A novel panel based on Immune infiltration and tumor mutational burden for prognostic prediction in hepatocellular carcinoma

CURRENT STATUS: Posted

Hewei Wu, Xingrong Zheng, Tao Pan, Xiaoan Yang, Genglin Zhang, Yunwen Lian, Jiaxin Lin, Liang Peng, Chan Xie

Hewei Wu
Third Affiliated Hospital of Sun Yat-Sen University

Xingrong Zheng
Third Affiliated Hospital of Sun Yat-Sen University

Tao Pan
Third Affiliated Hospital of Sun Yat-Sen University

Xiaoan Yang
Third Affiliated Hospital of Sun Yat-Sen University

Genglin Zhang
Third Affiliated Hospital of Sun Yat-Sen University

Yunwen Lian
Third Affiliated Hospital of Sun Yat-Sen University

Jiaxin Lin
Third Affiliated Hospital of Sun Yat-Sen University

Liang Peng
Third Affiliated Hospital of Sun Yat-Sen University

✉ pzp33@hotmail.com Corresponding Author

Chan Xie
Third Affiliated Hospital of Sun Yat-Sen University

✉ happyxiechan@hotmail.com Corresponding Author

ORCiD: https://orcid.org/0000-0002-0225-5483
Subject Areas

**Oncology  Cancer Biology**

Keywords

*Tumor mutation burden; hepatocellular carcinoma; Immune infiltration; neoplasm recurrence; vascular invasion*
Abstract

Background: Tumor mutation burden (TMB) was associated with prognosis in various malignancies, but it remains to be elucidated in hepatocellular carcinoma (HCC). We aimed to investigate the prognostic effects of TMB and its relationship with immune infiltration so as to establish a panel model capable of predicting prognosis.

Methods: TMB was calculated in all 374 HCC patients from the Cancer Genome Atlas, and high TMB and low TMB groups were determined based on propensity score matching and X-tile analysis. WCGNA was utilized to screen out genes related to immune cells and TMB, followed by multivariate analysis to generated the final TMB-Infiltration (TMB-IF) panel. 453 HCC patients from ICGC-LIRI-JP, GSE76427 and GSE20017 verified the robustness and extensibility of TMB-IF. Additionally, whole-exome sequencing (WES) was performed in 8 follow-up patients to investigate the relationship between HCC recurrence and TMB.

Results: Patients in high TMB group had worse prognosis than those in low TMB group, with a cutoff TMB value of 4.9. Enrichment analysis demonstrated that differentially expressed genes were mainly related to T cell activation, cell membrane localization and matrix composition. Tumor immune infiltration analysis revealed the infiltrations of Th2, Th17, and Tgd were up-regulated in high TMB group, while those of Tr1, MAIT, and DC were up-regulated in low TMB group. And TMB-IF fit well with the actual survival observation, with a C-index 0.785 (0.700-0.870), which verified in ICGC-LIRI-JP was 0.670 (0.573-0.767), and 0.774 (0.670-0.877) in GSE76427. In addition, TMB-IF showed a good performance in predicting tumor vascular invasion with a C-index of 0.847 (0.778-0.916).

Conclusions: Higher TMB means worse prognosis in HCC patients. Patients with higher frequency of immune-related gene mutations and TMB are prone to relapse after radical treatment.

Introduction

Hepatocellular carcinoma (HCC) is the third most common digestive cancer in global incidence, and mortality ranks the fourth highest among all tumors. Rates of both incidence and mortality in men are 2 to 3 times higher than women. And the mortality of HCC ranks second only to that of lung cancer among men (1). Although the pace of increasing in woman has slowed from previous years, incidence in men has continued to grow at a rate of 2% to 3% annually for the past decade(2). Main risk factors for HCC are chronic infection with hepatitis B virus (HBV) or hepatitis C virus (HCV), alcohol consumption, aflatoxin-contaminated foodstuffs, smoking, obesity and type 2 diabetes(3). Besides, it is not easy to screen out for early stage, and middle or late stage found when clinical features emerge.

Treatment of HCC mainly includes surgical resection, radiofrequency ablation, liver transplantation, radiotherapy, transcatheter arterial chemoembolization (TACE) and targeted therapy(4). In the past few years, Immune checkpoint inhibitors (ICIs) was considered to be one of the most promising treatments in cancer immunotherapy. Currently, several clinical trials of programmed death-1/programmed death-ligand 1 (PD-1/PD-L1) inhibitors in HCC have been conducted to show patient benefit(5, 6). Tumor mutation burden (TMB) is one of the popular biomarkers for predicting the efficacy of PD-1 inhibitors. Previously studies demonstrated a correlation between high TMB and clinical benefits of PD-1 inhibitors in lung cancer(7). And HCC has a higher tumor mutation load than the average for other solid tumors(8). In immunotherapy for most tumors, overall survival (OS) of patients with high TMB was significantly longer than that of patients with low TMB, with a median overall survival difference of approximately 40%(9). In the 2019 guidelines for non-small cell lung cancer (NSCLC), a new item of TMB was added to identify lung cancer patients who were eligible for dual immunotherapy with "Nivolumab plus Ipilimumab" and "Nivolumab" single-drug immunotherapy(10, 11). However, the effect of TMB in HCC remains unclear.

The essential work of the immune system is to identify and eliminate the abnormal cells. Immunotherapy is
aimed to arouse and strengthen the body's immune system to kill the tumor cells through various methods(12). In this case, the higher TMB, the more personality of the tumor is different from normal tissue, and the easier it is to be the target of immune cells, so the more likely it is theoretically to response to immunotherapy(13). Surrounding tissues, immune cells, blood vessels and extracellular matrix together constitute the tumor microenvironment, which is the "fertile soil" to help the tumor expand and invade faster(14). Immune cells involved in tumor development are diverse and highly heterogeneous. Complex and subtle relationship between different types of immune infiltrating cells and cell receptors often related to vascular invasion and tumor escape(15). Previous studies have shown CD8+ T cells that express different levels of PD1 are enriched in HCC tissues(16), but the current immunotherapy for HCC is relatively limited compared with results of immunotherapy in melanoma and non-small cell lung cancer. Previously, a study of 5,000 T-cell sequencing data has described the landscape of infiltrating T cells in HCC, showing that large Numbers of dysfunctional, killing T cells and inhibiting T cells clustered in tumor tissues(17). In addition, depleted T cells may further potentially evolve into inhibitory T cells.

Vascular invasion (VI) has been widely demonstrated to be closely associated with poor prognosis in patients with HCC after surgery resection(18, 19). Microvascular invasion (MVI) was defined as the presence of metastatic HCC micro-emboli in hepatic vessels. Even in tumors of same stages receiving the same treatment, the prognosis is still different(20), attributing it to the biological behavior of the tumor. MVI prior VI and VI has attracted more attention, which proved to be high risk factor for early recurrence of HCC after operation(21). Studies have indicated that HCC with microvascular invasion are needed to be comprehensively evaluated for a larger resection range and a wider ablation margin, and should be considered as a candidate for liver transplantation more cautiously(22, 23). However, even accepting radical resection, almost half of the patients relapsed within 3 years(24). At present, few studies have been conducted on the relationship between VI and infiltrating immune cells, nor on the relationship between recurrence and TMB in HCC.

To sum up, TMB has a general adaptability in pan-cancerous patients, especially those who receive immunotherapy. By analogy with other tumor clinical findings, TMB does play a potential role in HCC-related immunity. But there is still lack of monographs on the relationship between HCC and TMB currently. The aim of our study is to explore the differences of immune infiltrating cells at different TMB levels, identify related genes and construct prognostic panels. In addition, we hope to further explore the potential role of TMB in vascular invasion and HCC recurrence.

Methods

1. Somatic mutation, Transcriptome, microarray data acquisition and pre-processing

Somatic mutation data of 375 LIHC patients was downloaded from the publicly available TCGA database through the GDC data portal (https://portal.gdc.cancer.gov/), corresponding data format was based on the VarScan software platform "Masked Somatic Mutation". MiRNA Expression profiling, clinical data including age, gender, tumor grade, pathological stage, AJCCTNM stages and survival outcomes were all download for uniqueness matching. Meanwhile, miRNA expression files and related clinical outcome data of 203 HCC patients from Japan were also downloaded from ICGC official website (https://dcc.icgc.org/). In addition, gene expression chip profiles of GSE20017 (n=135) and GSE76427 (n=115) were also derived from the Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/). Along with the microarray data, corresponding platforms files "GPL8432" and "GPL10558" were also downloaded.

Mutation Annotation Format (MAF) of somatic variants downloaded from TCGA were visualized using "maftools" R package. Uniqueness matching with transcriptome data, collating ICGC gene expression profiling and GEO microarray data were all done using the Perl script (Perl version 5.28). TMB was defined as the total number of somatic gene coding errors detected per million bases, including base substitution, insertion, or deletion. Mutation frequency with number of variants/the length of exons (38 million) of 374 HCC patients were all calculated. Patients with well-matched clinical outcomes were subjected to Kaplan-Meier analysis followed by X-tile analysis(25) to determine the best TMB cutoff. Then, according to age and gender, the two groups of high
and low TMB were matched by the propensity score matching (PSM) in SPSS software, and finally the training data set of the panel was selected.

2. Patients and samples involved

8 patients diagnosed with HCC who underwent radiofrequency ablation during November 2017 and August 2018 at the third Affiliated Hospital of Sun Yat-sen University were enrolled. The operation was performed by two attending physicians with more than five years' experience. A total of 8 samples of liquid nitrogen quick-frozen HCC tissue that had been histologically and clinically diagnosed were harvested in this current study. Normal liver tissue from 2cm past the tumor edge were collected for controls. Each sample was immediately immersed in liquid nitrogen after isolation and then transferred to -80 °C within 30 minutes for later whole-exome sequencing (WES). Prior patients' consent and approval of the Institutional Research Ethics Committee of the Third Affiliated Hospital were obtained for the regular follow-up. Postoperative tumor status and recurrence will be recorded.

3. Differentially expressed genes and functional enrichment analysis

According to the results of X-tile analysis and PSM, transcriptome data of 374 HCC patients were divided into high TMB group and low TMB group via R software. "limma" R package was utilized to identify differentially expressed genes (DEGs) between the two groups. To avoid missing potentially important immune-related differential genes, we set the screening criteria at |Fold Change (FC)| >1 and False Discovery Rate (FDR) <0.05, which was also conducive to further weighted gene co-expression network analysis (WCGNA). After screening out DEGs, we used "org.HS.Eg.db", "clusterProfiler", "enrichplot", "ggplot2" packages to implement the Gene Ontology (GO) enrichment analysis of DEGs. Similarly, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of the DEGs was also achieved by the above 4 R packages. Preliminary Protein-Protein interaction (PPI) network between the differential genes was explored by STRING online database (http://string-db.org). Subsequently, the file containing nodes information was imported into Cytoscape software (version 3.6.1) for PPI network plotting. Top 10 nodes ranked by degree were filtered out using CytoHubba plug-in.

4. ImmuneCellAI and prognostic analysis of immune cells

After comprehensive consideration of TIMER database(26) and CIBERSORT algorithm(27), we further conducted immune cell infiltration analysis in ImmuneCellAI(28), which demonstrated powerful and unique function in tumor immune infiltration estimation and immunotherapy response prediction (http://bioinfo.life.hust.edu.cn/ImmuneCellAI/). Transcriptome data from both groups were submitted to ImmuneCellAI, and all original matrices containing infiltrating immune cells information were downloaded. All boxplots were drawn with "ggpubr" R package, and bilateral Wilcoxon rank-sum test and P value were used to compare the differences between the two groups. Then, scores of all immune cells were normalized by percentage, a stack graph was drawn using the "ggplot2" R package. Meanwhile, Kaplan-Meier analysis for each type of immune cells showed differences in survival outcomes at different levels of infiltration. R package utilized was "survival", and the log-rank test of P <0.05 was considered statistically significant.

5. WCGNA and tumor immune related gene identification

After screening out potential key immune cells, they were included as independent immune traits for weighted gene co-expression network analysis (WCGNA). Background genes were differentially expressed genes, and the R package involved was "WGCNA". Power parameter was predicted through the pickSoftThreshold function, which can provide appropriate soft threshold power for network construction by calculating scale-free topology fitting index of multiple powers. CutreeDynamic function was used to trim the gene level cluster tree, and then the co-expression module was obtained. ModuleEigengenes function in R WGCNA package was used to calculate the differences between the feature genes of each module. Association between modules and immune traits was assessed by Pearson's correlation. We further extracted immune cell-related genes under the conditions of Module Membership > 0.8 and Gene Significance correlation > 0.4. Genes from statistically significant modules were further incorporated into the multivariate Cox regression to screen genes related to prognosis.

6. TMB-IF gene panel generation and verification
Both "LR forward" and "LR backward" were used for variable filtering, and all independent prognostic factors determined by multivariate Cox regression analysis were included to establish the TMB-Infiltration (TMB-IF) panel to investigate the probability of 0.5-, 1- and 1.5- overall survival (OS). Concordance index (C-index) was calculated to assess the consistency between the actual observations and probability predicted by TMB-IF, with bootstrap method of 1000 resamples. ROC curves were performed to compared the identification of TMB-IF predictions of 0.5, 0.5-, 1-, 1.5- and 3- OS, using the “survivalROC” R package. Then, the 1.5-OS predictions of TMB-IF and TMN stages were compared, ROC curves were drawn, and AUC values were calculated. The remaining three validation cohorts (ICGC-JP, GSE76427, GSE20017) were used similar approach to validate the model. Nomogram and calibration diagrams involved were drawn using the “rms” R package.

7. Statistical analysis

Statistical significance of differences among variables with normal distribution was estimated by Student's t-test, while non-normal distribution variables were analyzed by Mann-Whitney U test. Qualitative variables were analyzed by Pearson χ2 test or Fisher’s exact test. Correlation was calculated using Spearman's and distance correlation analysis. Survival probability was calculated by Kaplan–Meier method, and Log-rank test was used to test the significance of differences in survival curves. Multivariate analysis adopted Cox proportional risk model, and methods of variable filtering were likelihood ratio test of maximum partial likelihood estimation (forward: LR and backward: LR). Accuracy of survival prediction was evaluated by receiver operating characteristic curve (ROC) analysis and Harrell's concordance index (C-index) analysis. All statistical analyses were performed using R software (version 3.6.2) and SPSS software (version 26.0). Double-tailed P < 0.05 was considered statistically significant.

Sample collection, DNA extraction and Whole exome sequencing (WES) of 8 follow-up patients were available in Supplemental Materials and Methods.

Results

1. characteristics of included patients and Landscape of mutation profiles in HCC.

Detailed characteristics of patients in the low TMB and high TMB group were presented in Table 1. There was a significant difference in TMB level between the two independent sets (P < 0.01). After matching with propensity score, composition of pathology grade in the two groups was basically similar (P = 0.63). As for gender and age, there was no paramount difference between the two groups (P = 0.23, P = 0.61). It could be seen that age > 60 and male had sufficient advantages in the proportion of population, especially in the high TMB group (67.57%, 72.97%). Notably, a considerable proportion of people in high TMB group were infected with the virus (40.54% vs 10.81%, including HBV and HCV infection), suggesting that influence of potential virus DNA integration. No significant differences found in alcohol consumption and family history of cancer in our study cohort (P = 0.80, P = 0.76). 72.97% of patients in the high TMB group were in stage I-II of TNM staging, and accounted for the majority of patients in ACJJ-T staging (72.97%), compared to patients in low TMB group (both 32.43%).

Somatic mutation profiles with VCF format were visualized after PSM and X-tile analysis (Fig. 1A, Figure S1). Genetic information for the Top 20 cumulative mutations were TP53, TTN, MUC16, CTNNB1, PCLO, HMCN1, OBSCN, ALB, LRPI, MUC4, RYR2, SPTA1, SYNE1, AHNAK2, ARID1A, CSMD3, DNAH2, DNAH5, PTPRO and WDR87. C > T was the most common single nucleotide mutation (SNV) in HCC (Fig. 1B). Base changes in each sample and top 10 mutated genes in HCC were summarized in Fig. 1E, and the frequency of single nucleotide polymorphism (SNP) was higher than insertion or deletion. Consistent and exclusive association among the mutant genes were shown in Fig. 1D, where green represented the co-occurrence relationship and brown represented the exclusion relationship. Figure 1C showed the ranking TMB of HCC among all tumors included in TCGA, which was above average.

All 8 patients followed up were male patients with chronic hepatitis B (CHB), and Barcelona Clinic Liver Cancer (BCLC) grades were in A ~ B. Child-Pugh score ranged from 5 to 7, which indicated these patients who received radiofrequency ablations were generally in good condition. P7 and P11 recurred during follow-up, and the lesions
were more numerous than before. Clinicopathologic characteristics of the 8 patients followed up are summarized in Table S1. WES was applied to the HCC tissues and adjacent normal tissues. Genetic information for the Top 100 cumulative mutations in each sample was showed in Fig. 2A. The frequency of SNP was higher than insertion or deletion, and C > A was the more common in these patients (Fig. 2B). Time points of P7 and P11 recurrence were shown in Fig. 2C, and they were accompanied by higher TMB and more mutations in TMB-IF (Fig. 2D).

2. Differential abundance of infiltrating immune cells in low- and high- TMB group

ImmuneCellAI was adopted to estimate the immune infiltration between low- and high- TMB group. Percentage normalized stack graph showed that immune infiltration scores of the two groups were generally similar (Fig. 3A). Top 10 immune infiltration scores in low TMB group were Macrophage (58.72%), DC (38.57%), Th17 (38.09%), MAIT (36.80%), Tc (34.45%), Th2 (33.77%), iTreg (30.64%), CD8_T cell (29.25%) and NK (25.87%), respectively. Meanwhile, in the high TMB group, the top 10 immune cells were Macrophage (54.40%), MAIT (32.83%), Th17 (51.48%), Tc (37.17%), Th2 (27.98%), iTreg (31.17%), DC (30.24%), NK (29.48%), Neutrophil (32.51), and CD8_T cell(32.49%), respectively. Percentage table of all immune scores could be found in Table S2.

Figure 3B showed the immune cells with different infiltration in the two groups, including Tr1 (P = 0.02), Th2 (P = 0.03), Th17 (P = 0.02), MAIT (P = 0.05), DC (P = 0.02) and Tgd (P = 0.04). Th2, Th17 and Tgd showed higher immune infiltration scores in the high TMB group, while Tr1, MAIT and DC were higher in the low TMB group. In addition, the differences of the remaining 14 types of immune cells were all displayed in Figure S2A, but they were not significant. In order to further screen out immune cells with potential impact on the survival of patients, Kaplan-Meier survival curves for all immune cells were plotted based on the median infiltration fraction. Finally, the potential immune cells were: CD8_T cell (P = 0.03), B cell (P = 0.05) and Tfh (P = 0.02) (Fig. 3C). The remaining Kaplan-Meier survival curves of 20 types of immune cells were all shown in Figure S2B.

3. Comparison of gene expression profiles and functional pathways analysis between low- and high- TMB groups

All the DEGs were listed in Table S3, from which it can be seen that with the increase of TMB, the expression of most genes was down-regulated. Subsequently, we conducted GO enrichment analysis for all of them (Fig. 4C). In the part of Biological Process (BP), there were obvious enrichment in terms of T cell activation and regulation of lymphocyte activation. Other terms with potential critical roles included regulation of cell-cell differentiation, positive regulation of cell differentiation, and Regulation of response to cytokine stimulus. In the cellular component (CC) part, DEGs were mostly enriched in terms related to cell membrane, such as membrane raft, plasma membrane raft, membrane microdomain, and cell substrata adhesives junction. In the part of Molecular Function (MF), the most enriched terms were the cell adhesion molecule binding, followed by more on the cytokine receptor binding, cytokine receptor binding and the glycosaminoglycan binding. KEGG pathway analysis showed that DEGs play an important role in the TNF signaling pathway, which was often required for activated monocytes or macrophage cells to kill or inhibit tumor cells (Fig. 4D). Meanwhile, PPI network of DEGs was constructed, and the 10 hub genes were screened out: PTPRC, CYBB, CTSS, LCP2, FYB, VCAM1, SYK, CSF2RB, TLR5, IL1B (Fig. 4A, B). Subsequently, expression levels of hub genes in each patient were shown by heatmap (Fig. 3E). The red part of the figure represents high expression, composed of patients with low TMB.

4. Identification of immune-related genes modules and associations of overall survival

All the immunocytes with different infiltration and those with potential influence on survival were included in WGCNA analysis as immune traits. Then, the correlation matrix of 10 modules of different colors and 10 immune traits with potential influence on prognosis was established (Fig. 5A), in which most of the modules were not strongly related to immune traits. Macrophage, DC, MAIT and Infiltration Score were well correlated with the genes in the blue module (r = 0.62, 0.62, 0.59, 0.68; P<0.05). Genes screened from the blue module were shown in Fig. 5B, and the remaining modules such as red, green, pink and magenta were all exhibited in Figure S3. Names of the gene symbol screened from all the modules and their correlation P values were listed in Table S4. After selection of the genes with good correlation in WGCNA, the multivariate Cox regression were performed to
screen variables using the forward and backward likelihood ratio method. 15 genes with statistical significance were screened out: DCN, EVI2A, FPR3, DSE, FYB1, P2RY13, CSF2RB, GEM, PMP22, SLC9A9, CTSS, CYBB, VCAM1, DOCK8 and SYK (All P < 0.05) (Fig. 5C), which constituted the final TMB-IF panel.

5. Construction and assessment of TMB-IF for HCC

Patients were segmented to the high- and low- risk groups according to the increased risk score (Fig. 6A). Kaplan-Meier survival analysis was re-conducted with the novel panel, and the results showed a good differentiation for patients (P < 0.01) (Fig. 6B). To determine the predictive accuracy of the TMB-IF, receiver operating characteristic (ROC) curves was drawn (Fig. 6C), which demonstrated that the area under the curve (AUC) was 0.85 for 0.5-year survival, 0.89 for 1-year survival, 0.90 for 1.5-year survival, 0.91 for 3-year survival. In addition, we compared the superiority of the novel TMB-IF with the traditional TMN staging with the threshold of 1.5 years, and the results showed that the AUC value of TMN was 0.68 (0.90 vs 0.68, Fig. 6D).

203 HCC patients from ICGC database were introduced to verify the accuracy of TMB-IF, and the survival curve showed good discrimination (P < 0.01) (Fig. 6E). Increasing risk map and heatmap were attached to Fig. 6F. Prognostic nomogram that integrated all the panel genes derived from the training cohort was shown in Fig. 7A. Calibration curve for the probability of survival at 1 and 1.5-year showed optimal consistency between the prediction by TMB-IF and actual observations (Fig. 7B), with a C-index of 0.785 (95% CI, 0.700–0.870). And the calibration curve of 231 patients from ICGC database was shown in Fig. 7C, with a C-index of 0.670 (95%CI, 0.573-0.767). In addition, survival curve of 115 patients from GSE76427 dataset also demonstrated TMB-IF model had a good risk discrimination (Fig. 7E). The calibration curve was shown in Fig. 7D, calculated C-index was 0.774 (95% CI, 0.670–0.877). Risk distribution of each patient in GSE76427 was shown in Figure S4A. To further explore the extensibility of TMB-IF panel, we also tried to apply it to the GSE20017 dataset to verify the vascular invasion in HCC. ROC curve of GSE20017 showed the AUC value was 0.847 (95%CI, 0.778–0.916) (Fig. 7F). Heatmap and calibration curve were shown in Figure S4B and Figure S4C.

Discussion

Currently, PD-1 inhibitors are mainly approved for the second-line treatment of advanced HCC patients with poor response to Sorafenib, which based on two previously reliable clinical studies(29, 30). Because of the different treatment responses, researchers have been looking for the best biomarkers that can well predict the efficacy of immunotherapy. PD-1/PD-L1 expression level(31), tumor mutation burden(13, 32), high microsatellite instability (MSI-H)(33), tumor infiltrating T cell content(34), and neutrophil-lymphocyte ratio (NLR)(35) are currently more widely accepted indicators of response efficacy. Although the predictive value was questioned as well as PD-1/PD-L1 expression level, (36), TMB has been proven indeed valuable in predicting the efficacy of immunotherapy by subsequent studies and most clinical trials (37).

Although there are few large clinical data explaining the role of TMB in HCC, its value can be inferred from other cancers that are currently more mature in immunotherapy. In the clinical study code named CheckMate227, Nivolumab plus low-dose Ipilimumab significantly prolonged the 1-year progression free survival period in advanced non-small cell lung cancer (NSCLC) patients who had TMB ≥ 10 mut/MB, regardless of the PD-1 expression level(38). Meanwhile, CheckMate-568, which was carried out in parallel, verified the conclusion(39). However, the TMB > 10/MB threshold may not be suitable for HCC. NSCLC ranked second in TMB among all tumors in TCGA database, while HCC ranked 12th. According to the X-tile analysis in this study, the optimal cutoff value of TMB for the two groups of patients was 4.9, which was about half of the threshold of NSCLC. But the threshold of 4.9 only reflected the ability of neoantigens to be produced by the tumor in non-immunotherapy HCC patients, which has yet to be confirmed by new clinical trials. In fact, researchers are trying to circumvent TMB to find ways to enhance the production and expression of neoantigens in tumors, so that patients with lower TMB also can benefit from the immunotherapy(40). Suggestions proposed now are using the existing chemotherapeutic/targeted agents, radiotherapy, oncological viruses and disrupting endogenous pathways in tumor cells. There is no doubt that these approaches still need further verification.

According to the results of GO enrichment and KEGG analysis of the DEGs in our study, it is noteworthy that the
matrix and cell adhesion are involved, except the term of T cell activation that we are concerned with. Recently, osteopontin has been found to regulate the sensitivity of immune cells to PD-1 inhibitors in liver cancer, which was a protein widely distributed in the matrix and cells with function of cell adhesion, signal transduction and tumor metastasis(41). The colony stimulating factor – 1 signaling pathway induced by osteopontin can destroy the tumor-related macrophages and make HCC sensitive to anti-PD-L1 blockade. This suggests that when we pay more attention to the immune cells, perhaps matrix components in the background environment are subtly influencing the immunocyte migration. As for top 10 selected hub genes finally fit into TMB-IF, key genes were identified as FYB, CSF2RB, CTSS, CYBB, VCAM1 and SYK. CTSS (Cathepsin S) silencing has been proved to have a potential role of inducing apoptosis of HCC cells(42). VCAM1, a gene associated with vascular cellular cytokines molecule, have been suggested that it may endue the growth and infiltration capacity of HCC cells(43). And spleen tyrosine kinase (SYK), which plays an important role in immune cell signaling pathway, has been reported as a biomarker for HCC and a potential target role for liver fibrosis(44). However, there is no research on FYB, CSF2RB and CYBB genes in HCC at present. Other component genes of the TMB-IF panel included DCN, DSE, EVI2A, FPR3, P2RY13, GEM, PMP22, SLC9A9 and DOCK8. DCN has been reported to be associated with DNA methylation and hydroxy-methylation in HBV-related HCC(45). DSE dysregulation may contributed to the malignant behavior of HCC cells through suppression of CCL5 signaling(46). The remaining genes are rarely reported in HCC and still require further basic research.

One of the characteristics of HCC is the rich blood supply and early blood-borne metastasis(47). To further improve the response of PD-1 inhibitors in advanced HCC, it is necessary to understand the mechanism of abnormal PD-1/PD-L1 signaling and immune escape microenvironment. Tumor infiltrating lymphocytes (TILs) are representative components of host anti-tumor immune response(48). CD3, CD4, CD8 and Foxp3 positive T lymphocytes are the most common subsets of TILs. Among them, tumor infiltration of CD8 positive T cells plays the most important role in host immune defense against tumor progression. In fact, patients in the high TMB group of this study did not showed the up-regulation in cytotoxic T lymphocytes as we expected. Our study demonstrated the infiltration abundance of Th2, Th17 and Tgd increased in the high TMB group, while Tr1, MAIT and DC increased in the low TMB group. Survival analysis showed that patients in the Low-TMB group had better survival outcomes, which we speculated might be related to increased infiltration of DC and MAIT cells through previous studies. More DC cells mean a greater chance of exposure of killer T cells to neoantigens. Another immune cell enriched in the low TMB group were Tr1 cells, which induce and maintain immune tolerance and immune regulatory cells. However, studies have shown that plasmacytoid dendritic cells (pDCs) promote the role of Tr1 cells in tumor immunosuppression in HCC(49), which may interfere with the efficacy of immunotherapy. MAIT cells are naturally equipped with anti-tumor function, but if the function of HCC infiltrated MAIT cells are impaired or recoded, the direction of anti-tumor immunity will be shifted to tumor promotion(50). In the high TMB group, increased Th2 broke the balance of Th1/Th2, which is one of the poor prognostic factors in HCC patients (51). And Th17 is a subset of inflammatory T helper cells that has been showed to potentially stimulate the development of HCC(52). Gamma-delta T cells (Tgd) possess unique characteristics and antigen recognition ability, as well as unique tissue affinity and cytotoxicity, which enable these cells to induce long-lasting immunity in response to different pathological conditions(53). Tgd cells are more likely to play an important role when the immunosuppressive state of HCC broken by pd-1 inhibitors.

A recent study by Shinji Itoh et al. confirmed the presence of vessels that encapsulated the tumor cluster (VETC) was associated with a high positive rate of PD-1 in HCC tissues, successfully established a link of PD-1, tumor vascular and immunotherapy(54). Therefore, combination of PD-1 inhibitors and anti-angiogenic drugs was expected to significantly improve the response rate of immunotherapy in HCC patients(55). It is for this reason that we have tried to apply TMB-IF to predict vascular tumor invasion, and the good discrimination was a little bit unexpected. Then, encouraged by the above performance, we further explore the potential application of TMB-IF in the recurrence of early stage of HCC. In fact, this result was not satisfactory because not all the gene mutations of TMB-IF can be detected. Although limited by the small sample size, it also can be seen that relapsed patients have higher levels of TMB than those without recurrence, and indeed have more mutations in TMB-IF-related genes. Previous studies have confirmed a high rate of recurrence after radiofrequency ablation, with nearly half of people relapsed within three years (56). And recent researches suggested that it is related to a late dynamic of immature NK cells (CD56+) and altered myeloid DC (PDL1+)(57). Therefore, it is possible that PD-1/PD-L1 inhibitors will also play a role in preventing early recurrence of HCC, in the coming era of
"immunotherapy plus". To summarize, PD-1 inhibitors may have a broader indication than current clinical studies of advanced HCC, and more researches are needed in the future.

There are some limitations exist in this study: (I) basic experiments are needed to verify the correlation between gene signature and immune cell infiltration. (II) more clinical samples are needed to verify the prognostic effect of TMB-IF and its potential relationship with immune infiltration and vascular invasion. (III) no data on the broader application of PD-1 inhibitors, and we are trying to explore further.

Conclusions

Higher TMB was associated with poorer survival outcomes. The TMB-IF panel based on tumor mutation burden and immune cell infiltration has demonstrated the stability and scalability of discrimination. But the intricate relationship between immune cells, tumor cells and immunotherapy still need to be further sorted out. PD-1 inhibitors may have a wider application in liver cancer in the future, hoping to reduce the 3- and 5-year high recurrence, but a series of new clinical studies are still needed.

Abbreviations

TMB, tumor mutation burden; HCC, hepatocellular carcinoma; TCGA, the Cancer Genome Atlas; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEGs, differentially expressed genes; WCGNA, weighted gene co-expression network analysis; TMB-IF, TMB-Infiltration; ICGC, International Cancer Genome Consortium. AUC, area under the curve. ImmuCellAI, immune cell abundance identifier. PPI, protein protein interaction. MAIT, Mucosal associated invariant T cells; Tc, cytotoxic T cells; Tex, exhausted T cells; Tr1, type 1 regulatory cells; nTreg, Natural regulatory T cells; iTreg, Induced regulatory T cells; Th1, Type 1 T helper cells; Th2, Type 2 T helper cells; Th17, Type 17 T helper cells; Tfh, Follicular helper T cells; Tcm, Central memory T cells; Tem, Effector memory T cells; NKT, Natural killer T cells; MAIT, Mucosal-associated invariant T cells; DC, Dendritic cells; Tgd, Gamma-delta T cells; TMB, tumor mutational burden; Low_IF, low Infiltrating; High_IF, high Infiltrating.

Declarations

Ethics approval and consent to participate

This study complies with the Declaration of Helsinki. Prior patients’ consent and approval of the Institutional Research Ethics Committee of the Third Affiliated Hospital has been obtained for the use of the clinical materials.

Consent for publication

Not applicable

Availability of data and material

All data generated and analyzed during this study are available from public database (see Methods) and supplementary materials.

Competing interests

The authors declare that there are no competing interests regarding the publication of this paper.

Funding:

This study was supported by grants from the National major science and technology project for the prevention
and treatment of AIDS and viral hepatitis [2018ZX10302205-002, 2018ZX10302204], Natural Science Foundation of China [No. 81472259, 81570539 and 81873572]. Natural Science Foundation of Guangdong Province [2014A030313042]. Guangzhou Science and Technology Plan Projects [201904010442], Science and Technology Project of Guangdong Province [2015A020212007], Sun Yat-sen University Clinical Research 5010 Program [2018009], Research project on degree and postgraduate education reform in Guangdong province [2018JGXM04], Young teacher training program of Sun Yat-sen University [16ykpy40], Open project of Key Lab of Tropical Disease Control (Sun Yat-sen University), Ministry of Education [2019kft07].

Author Contributions

C.X., and L.P. participated in the design, conception, and coordination of studies and interpretation of the data. H.W. and X.Z. prepared the manuscript and participated in the acquisition and interpretation of data. T.P. contributed to the revision of the article. X.Y., G.Z., Y.L. and J.L. assisted in the acquisition and interpretation of data and performed the statistical analysis. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424. 2018-11-01.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics. 2020. CA: A Cancer Journal for Clinicians. 2020 2020-01-01;70(1):7–30.
3. Fujiwara N, Friedman SL, Goossens N, Hoshida Y. Risk factors and prevention of hepatocellular carcinoma in the era of precision medicine. J HEPATOL. 2018;68(3):526–49.
4. Galle PR, Forner A, Llovet JM, Mazzaferro V, Piscaglia F, Raoul J, et al. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. J HEPATOL. 2018;69(1):182–236.
5. El-Khoueiry AB, Sangro B, Yau T, Crocenzi TS, Kudo M, Hsu C, et al. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. The Lancet. 2017;389(10088):2492–502.
6. Bandese N. Factors influencing late presentation for health care among men with cancer esophagus attending Hospice Africa Uganda (HAU). ANN ONCOL. 2018;29:v1.
7. MD H, MK C, MM A, E C, PA A, A A, et al. Tumor Mutational Burden and Efficacy of Nivolumab Monotherapy and in Combination with Ipilimumab in Small-Cell Lung Cancer. CANCER CELL. 2018;33(5):853–61.
8. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Blankin AV, et al. Signatures of mutational processes in human cancer. NATURE. 2013;500(7463):415–21.
9. Samstein RM, Lee C, Shouzhtari AN, Hellmann MD, Shen R, Janjigian YY, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. NAT GENET. 2019;51(2):202–6.
10. Hellmann MD, Paz-Ares L, Bernabe CR, Zurawski B, Kim SW, Carcereny CE, et al. Nivolumab plus Ipilimumab in Advanced Non-Small-Cell Lung Cancer. N Engl J Med. 2019;2019-11-21(21):2020–31. 381(.
11. Carbone DP, Reck M, Paz-Ares L, Creelan B, Horn L, Steins M, et al. First-Line Nivolumab in Stage IV or Recurrent Non–Small-Cell Lung Cancer. NEW ENGL J MED. 2017 2017-06-22;376(25):2415–26.
12. Wei SC, Duffy CR, Allison JP. Fundamental Mechanisms of Immune Checkpoint Blockade Therapy. CANCER DISCOV. 2018;8(9):1069–86.
13. Yarchoan M, Hopkins A, Jaffee EM. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. NEW ENGL J MED. 2017 2017-12-21;377(25):2500–1.
14. Wu T, Dai Y. Tumor microenvironment and therapeutic response. CANCER LETT. 2017;387:61–8.
15. Jiang X, Wang J, Deng X, Xiong F, Ge J, Xiang B, et al. Role of the tumor microenvironment in PD-L1/PD-1-mediated tumor immune escape. MOL CANCER. 2019;18(1):10.
16. HD K, GW S, MK SP, MH J, HJ K, K, et al. Association Between Expression Level of PD1 by Tumor-Infiltrating CD8 T Cells and Features of Hepatocellular Carcinoma. GASTROENTEROLOGY. 2018;155(6):1936–50.
17. Zheng C, Zheng L, Yoo J, Guo H, Zhang Y, Guo X, et al. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. CELL. 2017;169(7):1342–56.
18. XF X, ZL HXJH, YH LWYL. Z, et al. Risk Factors, Patterns, and Outcomes of Late Recurrence After Liver Resection for Hepatocellular Carcinoma: A Multicenter Study From China. JAMA SURG. 2019;154(3):209–17.
19. Pommergaard H, Rostved AA, Adam R, Thygesen LC, Salizzoni M, Gómez Bravo MA, et al. Vascular invasion and survival after liver transplantation for hepatocellular carcinoma: a study from the European Liver Transplant Registry. HPB. 2018;20(8):768–75.
20. TW SL, KD K, MW S, HK LHR L, et al. Effect of Microvascular Invasion Risk on Early Recurrence of Hepatocellular Carcinoma After Surgery and Radiofrequency Ablation. ANN SURG. 2019.
21. Erstad DJ, Tanabe KK. Prognostic and Therapeutic Implications of Microvascular Invasion in Hepatocellular Carcinoma. ANN SURG ONCOL. 2019;26(5):1474–93.
22. Vitale A, Peck-Radosavljevic M, Giannini EG, Vibert E, Sieghart W, Van Poucke S, et al. Personalized treatment of patients with very early hepatocellular carcinoma. J HEPATOL. 2017;66(2):412–23.
23. Lee S, Kang TW, Song KD, Lee MW, Rhim H, Lim HK, et al. Effect of Microvascular Invasion Risk on Early Recurrence of Hepatocellular Carcinoma After Surgery and Radiofrequency Ablation. ANN SURG. 2019;1.
24. P T. G J, B S, M S, S R. Recurrence of hepatocellular cancer after resection: patterns, treatments, and prognosis. ANN SURG. 2015;261(5):947–55.
25. Camp RL. X-Tile. A New Bio-Informatics Tool for Biomarker Assessment and Outcome-Based Cut-Point Optimization. CLIN CANCER RES. 2004 2004-11-01;10(21):7252–9.
26. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. NAT METHODS. 2015;12(5):453–7.
27. Miao YR, Zhang Q, Lei Q, Luo M, Xie GY, Wang H, et al. ImmuCellAI: A Unique Method for Comprehensive T-Cell Subsets Abundance Prediction and its Application in Cancer Immunotherapy. ADV SCI. 2020;7(7):1902880.
28. El-Khoueiry AB, Sangro B, Yau T, Crocenzi TS, Kudo M, Hsu C, et al. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. The Lancet. 2017;389(10088):2492–502.
29. Finn RS, Ryoo B, Merle P, Kudo M, Bouattour M, Lim HY, et al. Pembrolizumab As Second-Line Therapy in Patients With Advanced Hepatocellular Carcinoma in KEYNOTE-240: A Randomized, Double-Blind, Phase III Trial. J CLIN ONCOL. 2020 2020-01-20;38(3):193–202.
30. Balar AV, Castellano D, O'Donnell PH, Grivas P, Vuky J, Powles T, et al. First-line pembrolizumab in cisplatin-ineligible patients with locally advanced and unresectable or metastatic urothelial cancer (KEYNOTE-052): a multicentre, single-arm, phase 2 study. The Lancet Oncology. 2017;18(11):1483–92.
31. ZR C, CF C, SM DFLG. A, R E, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. GENOME MED. 2017;9(1):34.
37. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. ANN ONCOL. 2019;30(1):44-56.
38. Hellmann MD, Ciuleanu T, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, et al. Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden. NEW ENGL J MED. 2018 2018-05-31:378(22):2093-104.
39. Ready N, Hellmann MD, Awad MM, Otterson GA, Gutierrez M, Gainor JF, et al. First-Line Nivolumab Plus Ipilimumab in Advanced Non–Small-Cell Lung Cancer (CheckMate 568): Outcomes by Programmed Death Ligand 1 and Tumor Mutational Burden as Biomarkers. J CLIN ONCOL. 2019 2019-04-20;37(12):992-1000.
40. Lu L, Jiang J, Zhan M, Zhang H, Wang QT, Sun SN, et al. Targeting Neoantigens in Hepatocellular Carcinoma for Immunotherapy: A Futile Strategy? HEPATOLOGY. 2020 2020-04-16.
41. Zhu Y, Yang J, Xu D, Gao XM, Zhang Z, Hsu JL, et al. Disruption of tumour-associated macrophage trafficking by the osteopontin-induced colony-stimulating factor-1 signalling sensitises hepatocellular carcinoma to anti-PD-L1 blockade. GUT. 2019 2019-09-01;68(9):1653-66.
42. X W, L X, G Y, D L, T P, D L, et al. Cathepsin S silencing induces apoptosis of human hepatocellular carcinoma cells. AM J TRANSPL RES. 2015;7(1):100-10.
43. Huang J, Deng Q, Wang Q, Li K, Dai J, Li N, et al. Exome sequencing of hepatitis B virus-associated hepatocellular carcinoma. NAT GENET. 2012;44(10):1117–21.
44. Qu C, Zheng D, Li S, Liu Y, Lidofsky A, Holmes JA, et al. Tyrosine kinase SYK is a potential therapeutic target for liver fibrosis. HEPATOLOGY. 2018;68(3):1125-39.
45. Ye C, Tao R, Cao Q, Zhu D, Wang Y, Wang J, et al. Whole-genome DNA methylation and hydroxymethylation profiling for HBV-related hepatocellular carcinoma. INT J ONCOL. 2016;49(2):589–602.
46. HR WCL, CK Y, TJ L, CT T, L, CH L. DSE regulates the malignant characters of hepatocellular carcinoma cells by modulating CCL5/CCR1 axis. AM J CANCER RES. 2019;9(2):347-62.
47. Llovet JM, Zucman-Rossi J, Pikarsky E, Sangro B, Schwartz M, Sherman M, et al. Hepatocellular carcinoma. NAT REV DIS PRIMERS. 2016 2016-04-14;2:16018.
48. C L, D R, B Z, W Z, X W, Z C, et al. Current perspectives on the immunosuppressive tumor microenvironment in hepatocellular carcinoma: challenges and opportunities. MOL CANCER. 2019;18(1):130.
49. Pedroza-Gonzalez A, Zhou G, Vargas-Mendez E, Boor PP, Mancham S, Verhoef C, et al. Tumor-infiltrating plasmacytoid dendritic cells promote immunosuppression by Tr1 cells in human liver tumors. ONCOIMMUNOLOGY. 2015;4(6):e1008355. 2015-06-03.
50. Duan M, Goswami S, Shi J, Wu L, Wang X, Ma J, et al. Activated and Exhausted MAIT Cells Foster Disease Progression and Indicate Poor Outcome in Hepatocellular Carcinoma. Clinical cancer research: an official journal of the American Association for Cancer Research. 2019 2019-01-01;25(11):3304-16.
51. HL L, JW J, SW L, SH Y, JH K, SW N, et al. Inflammatory cytokines and change of Th1/Th2 balance as prognostic indicators for hepatocellular carcinoma in patients treated with transarterial chemoembolization. SCI REP-UK. 2019;9(1):3260.
52. YM Y, SY K. E S. Inflammation and Liver Cancer: Molecular Mechanisms and Therapeutic Targets. SEMIN LIVER DIS. 2019;39(1):26-42.
53. M L, J D, MC D. Key Features of Gamma-Delta T-Cell Subsets in Human Diseases and Their Immunotherapeutic Implications. FRONT IMMUNOL. 2017;8:761.
54. Itoh S, Yoshizumi T, Yugawa K, Imai D, Yoshiya S, Takeishi K, et al. Impact of Immune Response on Outcomes in Hepatocellular Carcinoma: Association with Vascular Formation. HEPATOLOGY. 2020 2020-02-29.
55. G K, S I, R K. The biology of Hepatocellular carcinoma: implications for genomic and immune therapies. MOL CANCER. 2017;16(1):149.
56. L JH, Z Y, Du CHW. L, J W, et al. A randomized trial comparing radiofrequency ablation and surgical resection for HCC conforming to the Milan criteria. ANN SURG. 2010;252(6):903-12.
57. JC PR, AS NFM. C, N B, AJ G, et al. Dynamic of systemic immunity and its impact on tumor recurrence after radiofrequency ablation of hepatocellular carcinoma. ONCOIMMUNOLOGY. 2019;8(8):1615818.
Table 1. Clinical baseline of 74 HCC patients included in study from TCGA cohort

| Characteristics          | Low TMB, n=37(%) | High TMB, n=37(%) | P value |
|--------------------------|------------------|-------------------|---------|
| Age                      |                  |                   |         |
| 60                       | 20(54.05%)       | 25(67.57%)        | 0.23    |
| 60                       | 17(45.95%)       | 12(32.43%)        |         |
| Gender                   |                  |                   |         |
| Male                     | 25(67.57%)       | 27(72.97%)        | 0.61    |
| Female                   | 12(32.43%)       | 10(27.03%)        |         |
| Viral infection          |                  |                   |         |
| Positive                 | 4(10.81%)        | 15(40.54%)        | 0.00    |
| Negative                 | 33(89.19%)       | 22(59.46%)        |         |
| Alcohol consumption      |                  |                   |         |
| Yes                      | 12(32.43%)       | 11(29.73%)        | 0.80    |
| No                       | 25(67.57%)       | 26(70.27%)        |         |
| Family history           |                  |                   |         |
| Yes                      | 7(18.92%)        | 6(16.22%)         | 0.76    |
| No                       | 30(81.08%)       | 31(83.78%)        |         |
| Pathology grade          |                  |                   |         |
| G1/G2                    | 25(67.57%)       | 23(62.16%)        | 0.63    |
| G3/G4                    | 12(32.43%)       | 14(37.84%)        |         |
| TNM Stage                |                  |                   |         |
| Stage I & II             | 12(32.43%)       | 27(72.97%)        | 0.00    |
| Stage III & IV           | 25(67.57%)       | 10(27.03%)        |         |
| AJCC-T                   |                  |                   |         |
| T1-T2                    | 13(35.14%)       | 27(72.97%)        | 0.00    |
| T3-T4                    | 24(64.86%)       | 10(27.03%)        |         |
| TMB level                | 3.12(1.24)       | 6.88(1.82)        | 0.00    |

Abbreviation: TMB, tumor mutation burden; AJCC, American Association of Cancer.

**Figure Legends Of Supplementary Figures**

**Figure S1. Results of X-tile analysis of 374 HCC patients in TCGA database**

(A,D) Survival curve after the optimal cutoff value was calculated by X-tile analysis, P<0.001. (B, C) Risk distribution, and the determination of best cutoff value of tumor mutational burden (TMB).

**Figure S2. The remaining types of immune infiltrating cells between different groups and survival analysis**

(A) 14 types of immune cell infiltration, in addition to the immune cells listed in Figure 2. (B) Survival analysis of other immune infiltrating cells, in addition to the five immune cells shown in Figure 2.

**Figure S3. Immune related gene modules of WCGNA analysis**

(A~E) Immune related genes in red, green, pink and magenta modules.

**Figure S4. Additional descriptions for GSE20017 and GSE76427**

(A) Risk distribution of 115 patients of GSE76427 dataset. (B) Heatmap of selected key genes in each patient in GSE20017. (C) Calibration curve of TMB-IF in predicting tumor vascular invasion.
Altered in 71 (95.95%) of 74 samples.
Figure 1

Landscape of mutation profiles in involved hepatocellular carcinoma tissue samples. (A) Mutation information of top 20 genes in each sample. Different colors and notes at the bottom represented types of mutations. (B) Overall distribution of the six different base substitution mutation frequency (above) and conversion ratio in each sample (below). (C) Tumor mutational burden ranking of HCC among all tumors in the Cancer Genome Atlas. (D) Co-occurrence associations across top 25 mutated genes. (E) Summary of the mutation information in all HCC tissue. TMB, tumor mutational burden; SNP, single nucleotide polymorphism; SNV, single nucleotide variants. HCC, hepatocellular carcinoma.
C

E
Landscape of mutation profiles in involved hepatocellular carcinoma tissue samples. (A) Mutation information of top 20 genes in each sample. Different colors and notes at the bottom represented types of mutations. (B) Overall distribution of the six different base substitution mutation frequency (above) and conversion ratio in each sample (below). (C) Tumor mutational burden ranking of HCC among all tumors in the Cancer Genome Atlas. (D) Co-occurrence associations across top 25 mutated genes. (E) Summary of the mutation information in all HCC tissue. TMB, tumor mutational burden; SNP, single nucleotide polymorphism; SNV, single nucleotide variants. HCC, hepatocellular carcinoma.
Landscape of mutation profiles of 8 patients with recurrence HCC. (A) Mutation information of Top 100 genes in each sample. Different colors and notes at the bottom represented types of mutations. (B) Summary of variant classification, variant type and SNV class information in 8 HCC tissue. (C) Time points of the recurrence of P7 and P11, and the location of the lesions. (D) TMB distribution of each patients and somatic mutation in TMB-IF among these HCC patients. TMB, tumor mutational burden; IF, Infiltration.
Landscape of mutation profiles of 8 patients with recurrence HCC. (A) Mutation information of Top 100 genes in each sample. Different colors and notes at the bottom represented types of mutations. (B) Summary of variant classification, variant type and SNV class information in 8 HCC tissue. (C) Time points of the recurrence of P7 and P11, and the location of the lesions. (D) TMB distribution of each patients and somatic mutation in TMB-IF among these HCC patients. TMB, tumor mutational burden; IF, Infiltration.
Infiltrating immune cells between different TMB groups and the effect on overall survival. (A) 24 types of infiltrating cells estimated by ImmuneCellAI. Different colors represent different types of immune cells, and the column length has been normalized by percentage. (B) 7 kinds of immune cells with significant infiltration in different TMB groups, and 3 others may affect the survival and prognosis of patients. (C) Kaplan-Meier survival analysis identified 4 types of immune cells that may affect the survival status. MAIT, Mucosal associated invariant T cells; Tc, cytotoxic T cells; Tex, exhausted T cells; Tr1, type 1 regulatory cells; nTreg, Natural regulatory T cells; iTreg, Induced regulatory T cells; Th1, Type 1 T helper cells; Th2, Type 2 T helper cells; Th17, Type 17 T helper cells; Tfh, Follicular helper T cells; Tcm, Central memory T cells; Tem, Effector memory T cells; NKT, Natural killer T cells; MAIT, Mucosal-associated invariant T cells; DC, Dendritic cells; Tgd, Gamma-delta T cells; TMB, tumor mutational burden; Low_IF, low Infiltrating; High_IF, high Infiltrating.
Figure 3

Infiltrating immune cells between different TMB groups and the effect on overall survival. (A) 24 types of infiltrating cells estimated by ImmuneCellAI. Different colors represent different types of immune cells, and the column length has been normalized by percentage. (B) 7 kinds of immune cells with significant infiltration in different TMB groups, and 3 others may affect the survival and prognosis of patients. (C) Kaplan-Meier survival analysis identified 4 types of immune cells that may affect the survival status. MAIT, Mucosal associated invariant T cells; Tc, cytotoxic T cells; Tex, exhausted T cells; Tr1, type 1 regulatory cells; nTreg, Natural regulatory T cells; iTreg, Induced regulatory T cells; Th1, Type 1 T helper cells; Th2, Type 2 T helper cells; Th17, Type 17 T helper cells; Tfh, Follicular helper T cells; Tcm, Central memory T cells; Tem, Effector memory T cells; NKT, Natural killer T cells; MAIT, Mucosal-associated invariant T cells; DC, Dendritic cells; Tgd, Gamma-delta T cells; TMB, tumor mutational burden; Low_IF, low Infiltrating; High_IF, high Infiltrating.

Figure 4

A
C

- chondrocyte development
- positive regulation of cytokine production
- positive regulation of cell adhesion
- extracellular matrix organization
- regulation of cytokine–mediated signaling pathway
- T cell activation
- regulation of cellular response to growth factor stimulus
- regulation of lymphocyte activation
- regulation of response to cytokine stimulus
- regulation of cell–cell adhesion
- collagen–containing extracellular matrix
- membrane raft
- membrane microdomain
- membrane region
- actin filament
- collagen trimer
- focal adhesion
- cell–substrate adherens junction
- cell–substrate junction
- plasma membrane raft
- extracellular matrix structural constituent
- cytokine receptor binding
- glycosaminoglycan binding
- enzyme inhibitor activity
- hydrolase activity, acting on carbon–nitrogen
- sulfur compound binding
- SH2 domain binding
- proteoglycan binding
Figure 4

Differentially expressed genes and functional enrichment analysis between the two groups. (A, B) Protein-protein interaction network of all Differentially expressed genes. Orange represents the screened hub genes, and the size of circle represents the level of gene expression. (C) Go enrichment analysis of all differentially expressed genes. (D) KEGG analysis of all differentially expressed genes: TNF signaling pathway, Rheumatoid arthritis and Platelet activation. (E) Heatmap of selected hub genes in each patient. From left to right in order of low TMB and high TMB group. MF, Molecular Function; BP, Biological Process; CC, Cellular Component; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.
Figure 4

Differentially expressed genes and functional enrichment analysis between the two groups. (A, B) Protein-protein interaction network of all Differentially expressed genes. Orange represents the screened hub genes, and the size of circle represents the level of gene expression. (C) Go enrichment analysis of all differentially expressed
genes. (D) KEGG analysis of all differentially expressed genes: TNF signaling pathway, Rheumatoid arthritis and Platelet activation. (E) Heatmap of selected hub genes in each patient. From left to right in order of low TMB and high TMB group. MF, Molecular Function; BP, Biological Process; CC, Cellular Component; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.
Figure 5
Weighted Gene Co-Expression Network Analysis and screening of immune infiltration and prognostic genes. (A) WGCNA modules related to immune infiltrating cells. Different color represented different modules and correlation was displayed by heatmap. (B) Screening of genes associated with immunity and prognosis in the blue module. (C) Forest plot shows multivariate regression results and genes that potentially incorporated into the TMB-IF panel. WGCNA, weighted gene co-expression network analysis; TMB-IF, TMB infiltration. Tr1, type 1 regulatory cells; Th2, Type 2 T helper cells; Th17, Type 17 T helper cells; Th, Follicular helper T cells; MAIT, Mucosal-associated invariant T cells; DC, Dendritic cells; Tgd, Gamma-delta T cells.

| Gene | (N=74) | 95% CI          | log2 Fold Change |\(p\) Value |
|------|--------|-----------------|-----------------|-------------|
| DCN  |        | 7.0e-01         | (5.7e-01 - 8.8e-01) |             |
| EVI2A|        | 5.9e-02         | (1.1e-02 - 3.3e-01) |             |
| FPR3 |        | 2.4e+00         | (1.2e+00 - 4.8e+00) |             |
| DSE  |        | 2.5e-05         | (3.8e-07 - 1.7e-03) |             |
| MNDA |        | 2.9e+00         | (9.7e-01 - 8.6e+00) |             |
| FYB1 |        | 5.3e+01         | (6.1e+00 - 4.6e+02) |             |
| P2RY13|       | 9.8e+00         | (1.1e+00 - 8.7e+01) |             |
| CSF2RB|       | 2.6e-05         | (3.2e-07 - 2.1e-03) |             |
| GEM  |        | 5.5e+00         | (2.4e+00 - 1.3e+01) |             |
| PMP22|        | 1.8e+00         | (1.3e+00 - 2.5e+00) |             |
| SLC9A9|       | 1.9e+00         | (1.1e+00 - 3.2e+00) |             |
| CTSS |        | 9.5e-01         | (9.1e-01 - 1.0e+00) |             |
| CYBB |        | 2.2e+00         | (1.5e+00 - 3.2e+00) |             |
| VCAM1|        | 1.1e+00         | (1.0e+00 - 1.2e+00) |             |
| DOCK8|        | 1.9e-01         | (4.1e-02 - 8.8e-01) |             |
| SYK  |        | 5.0e-01         | (3.4e-01 - 7.5e-01) |             |

# Events: 32; Global p-value (Log-Rank): 1.3232e-05
AIC: 213.71; Concordance Index: 0.83
Figure 3

A

Module-trait relationships

C

Hazard ratio

| Gene       | Hazard Ratio | p-value      |
|------------|--------------|--------------|
| HLA-DRA    | 9.9e-01      | (9.9e-01 - 1.0e+00) |
| DCN        | 7.0e-01      | (5.7e-01 - 8.8e-01) |
| EVI2A      | 5.9e-02      | (1.1e-02 - 3.3e-01) |
Figure 5

Weighted Gene Co-Expression Network Analysis and screening of immune infiltration and prognostic genes. (A) WGCNA modules related to immune infiltrating cells. Different color represented different modules and correlation was displayed by heatmap. (B) Screening of genes associated with immunity and prognosis in the blue module. (C) Forest plot shows multivariate regression results and genes that potentially incorporated into the TMB-IF panel. WGCNA, weighted gene co-expression network analysis; TMB-IF, TMB infiltration. Tr1, type 1 regulatory cells; Th2, Type 2 T helper cells; Th17, Type 17 T helper cells; Tfh, Follicular helper T cells; MAIT, Mucosal-associated invariant T cells; DC, Dendritic cells; Tgd, Gamma-delta T cells.

Figure 6
Figure 6

Construction and evaluation of the TMB-IF prognostic panel. (A) Overall distribution of patients identified with the threshold of high and low risk. Screened prognostic genes in the two groups was demonstrated by heatmap. (B) Kaplan-Meier survival analysis and survival curves of high and low risk patients. (C) ROC curve was utilized to evaluate accuracy of the prognostic model at 0.5-, 1-, 1.5- and 3-year overall survival. (D) Comparison of TMB-IF with TMN staging and Pathology Grade. (E) Kaplan-Meier survival analysis and survival curves of 203 patients of ICGC database. (F) Risk distribution of 203 patients of ICGC database. ICGC, International Cancer Genome Consortium. ROC, Receiver operating characteristic.
Figure 6

Construction and evaluation of the TMB-IF prognostic panel. (A) Overall distribution of patients identified with the threshold of high and low risk. Screened prognostic genes in the two groups was demonstrated by heatmap. (B) Kaplan-Meier survival analysis and survival curves of high and low risk patients. (C) ROC curve was utilized to evaluate accuracy of the prognostic model at 0.5-, 1-, 1.5- and 3-year overall survival. (D) Comparison of TMB-IF with TMN staging and Pathology Grade. (E) Kaplan-Meier survival analysis and survival curves of 203 patients of ICGC database. (F) Risk distribution of 203 patients of ICGC database. ICGC, International Cancer Genome Consortium. ROC, Receiver operating characteristic.

Figure 7

A
Figure 7

nomogram and validation of TMB-IF prognostic panel. (A) Nomogram of all genes that have significance for prognosis. (B) Internal calibration curve of the TMB-IF panel. Red points correspond to the fitting degree of 1-year survival probability, and the orange points correspond to the fitting degree of 1.5 years survival probability. (C) Calibration curve of 203 patients in ICGC database. (D) Calibration curve of 115 patients in GSE76427 dataset. (E) Kaplan–Meier survival analysis and survival curves of GSE76427 dataset. (F) ROC curve was utilized to evaluate the potential role of TMB-IF in predicting tumor vascular invasion. ROC, Receiver operating characteristic.
Figure 7

Nomogram and validation of TMB-IF prognostic panel. (A) Nomogram of all genes that have significance for prognosis. (B) Internal calibration curve of the TMB-IF panel. Red points correspond to the fitting degree of 1-year survival probability, and the orange points correspond to the fitting degree of 1.5 years survival probability. (C) Calibration curve of 203 patients in ICGC database. (D) Calibration curve of 115 patients in GSE76427 dataset. (E) Kaplan–Meier survival analysis and survival curves of GSE76427 dataset. (F) ROC curve was utilized to evaluate the potential role of TMB-IF in predicting tumor vascular invasion. ROC, Receiver operating characteristic.

Supplementary Files

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

- TableS1.docx
- TableS2.docx
- TableS3.docx
- TableS4.docx
- FigS2.pdf
- FigS4.pdf
- TableS1.docx
- TableS2.docx
- TableS3.docx
- TableS4.docx
- FigS2.pdf
- FigS4.pdf
