Journal of Cancer Research and Clinical Oncology
https://doi.org/10.1007/s00432-022-04325-2

RESEARCH

Discovery and characterization of tumor antigens in hepatocellular carcinoma for mRNA vaccine development

Jiantao Fu[1,2] · Feng Chen[3] · Yuanji Lin[4] · Jin Gao[1,2] · Anna Chen[1,2] · Jin Yang[1,2]

Received: 20 July 2022 / Accepted: 24 August 2022

Abstract

Background mRNA vaccines are emerging as new targets for cancer immunotherapy. However, the potential tumor antigens for mRNA vaccine design in hepatocellular carcinoma (HCC) remain to be elucidated.

Methods Genetic and RNA-Seq data were obtained from TCGA and ICGC. Tumor-specific antigens (TSAs) were identified by differential expression, mutation status, HLA binding, antigen-presenting cell (APC) correlation, immune checkpoint (ICP) relevance and prognosis. Consensus clustering was used for patient classification. The molecular and immune status of TSAs and clustered patients, including prognostic ability, tumor microenvironment, tumor-related signature and tumor immune dysfunction and exclusion (TIDE), were further characterized.

Results Five dysregulated and mutated TSAs were identified in HCC (TSA5): FXYD6, JAM2, GALNT16, C7, and CCDC146. Seven immune gene modules and five immune subtypes (IS1–IS5) of HCC were identified. The immune subtypes and TSA5-related modules showed distinct molecular, cellular and clinical characteristics. According to our study, IS1 patients may be suitable for vaccination.

Keywords mRNA vaccine · Hepatocellular carcinoma · Immune subtypes · Tumor antigens · Immunology

Abbreviations

mRNA	Messenger RNA
HCC	Hepatocellular carcinoma
TAAs	Tumor-associated antigens
TSAs	Tumor-specific antigens
TCGA	The Cancer Genome Atlas
ICGC	International Cancer Genome Consortium
APC	Antigen-presenting cell
ICP	Immune checkpoint
TIDE	Tumor immune dysfunction and exclusion
IS	Immune subtype
ICI	Immune checkpoint inhibitors
NSCLC	Non-small cell lung cancer
CNV	Copy number variation
FGA	Fraction genome altered
MSI	Microsatellite instability
WGCNA	Weighted gene coexpression network analysis
TOM	Topological overlap measure
TIMER	Tumor immune microenvironment
TIICs	Tumor immune infiltrating cells
DCs	Dendritic cells
ssGSEA	Single-sample gene set enrichment analysis
TILs	Tumor-infiltrating lymphocytes
ImmuCellAI	Immune cell abundance identifier
TIS	T cell-inflamed score
TTN	Titin
TP53	Tumor protein p53
PDCD1 PD-1	Programmed cell death 1
FXYD6	FXYD domain containing ion transport regulator 6

Jiantao Fu and Feng Chen contributed equally to this work.

✉ Jin Yang
20171129@hznu.edu.cn

[1] Department of Translational Medicine Center, The Affiliated Hospital of Hangzhou Normal University, Hangzhou 310015, Zhejiang, China

[2] Institute of Hepatology and Metabolic Diseases, Hangzhou Normal University, Hangzhou 310015, Zhejiang, China

[3] Department of Blood Bank, Blood Center of Zhejiang Province, Hangzhou 310052, Zhejiang, China

[4] Department of Research, Hangzhou MC Life Sciences Co., Ltd, Hangzhou 311500, Zhejiang, China

Published online: 29 August 2022

JAM2	Junctional adhesion molecule 2
GALNT16	Polypeptide N-acetylgalactosaminyltransferase 16
C7	Complement C7
CCDC146	Coiled-coil domain containing 146
OS	Overall survival
PFI	Progression-free survival
AFP	Alpha fetoprotein
GPC3	Glypican 3
GOLM1	Golgi membrane protein 1
CTLA4	Cytotoxic T lymphocyte-associated antigen-4
LAG3	Lymphocyte-activation gene 3
TIM-3 or HAVCR2	T cell immunoglobulin domain and mucin domain containing protein 3
IDO1	Indoleamine 2,3-dioxygenase 1
TIGIT	T cell immunoreceptor with Ig and ITIM domain
ICD	Immunogenic cell death
MDSCs	Myeloid-derived suppressor cells
Tregs	Regulatory T cells
BMP4	Bone morphogenetic protein 4

Introduction

Hepatocellular carcinoma (HCC) is characterized by complicated reprogramming of both genomic and cytological changes in cancerous hepatocytes (Farazi and DePinho 2006) and represents the major cause of cancer mortality globally. The current treatment modalities, including surgery, chemotherapy, kinases and immune checkpoint inhibitors (ICIs), are still unsatisfactory, especially for patients with the late stage exhibiting mutation heterogeneities in the majority of tumors (Bray et al. 2018), which cause the development of drug resistance and worse outcome (Vasan et al. 2019).

The (neo)antigenic profile of a tumor determines the effect of anticancer immunity (Vitale et al. 2019), as demonstrated by the fact that certain cancers develop resistance to ICI treatment as a result of antigen presentation abnormalities (Kalbasi and Ribas 2020). Thus, tumor-specific antigens (TSAs) have attracted much attention since they can be used as biomarkers to predict the therapeutic effects of ICI therapy and as targets for cancer immunotherapy (Wu et al. 2018). The last decade has seen an increase in the use of messenger RNA (mRNA) treatments in combination with chemotherapy and immunotherapy, and this area of study has become a hotbed for the creation of novel cancer therapies (Khan et al. 2021).

mRNA vaccines targeting multiple tumor-specific antigens can elicit broad immune responses and decrease tumor escape (Pardi et al. 2018), and multiple preclinical and clinical trials have been thus tested. For example, recent interim results from a multinational clinical study (NCT02410733) validated the immunogenicity of an mRNA-based melanoma vaccine (Sahin et al. 2020). For late-stage non-small cell lung cancer (NSCLC), a clinical trial showed the benefits of immunotherapy consisting of mRNA encoding six NSCLC-associated antigens, with minor side effects (Papachristofilou et al. 2019). These early results have fueled optimism for the creation of novel vaccines, and the emphasis is now on identifying TSAs that may act as antigenic targets for cancers (Roberts et al. 2020). Several studies have suggested potential tumor antigens aimed at mRNA vaccine development for pancreatic adenocarcinoma and cholangiocarcinoma (Huang et al. 2021a, b; Huang et al. 2021a, b). However, no mRNA vaccine for HCC has been developed thus far.

The purpose of this study was to explore novel HCC antigens for the mRNA vaccines progression and for choosing right therapy for patients. We identified five TSAs with vaccine potential, as well as five robust immune subtypes and seven functional modules of HCC based on immune-related gene clustering, thus providing support for a suitable vaccination program consisting of mRNA vaccines and appropriate recipients.

Materials and methods

Data collection and processing

We obtained patient transcriptome data and associated clinical information from the Cancer Genome Consortium (ICGC, https://www.icgc-argo.org) and The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov). The cBioPortal database (http://www.cbioportal.org/) was used to get the somatic mutation count, copy number variation (CNV), fraction genome altered scores (FGA) (Chin et al. 2007) and microsatellite instability score (MSI). For liver cancer-specific neoantigen analysis, all peptides with 8–11 amino acids resulting from mutations in expressed genes were extracted and predicted by NetMHCpan2.8 in the tsnadb database (Wu et al. 2018). A total of 2002 immune-related genes (ImmPort portal, http://www.immport.org/), including cytokines, chemokines, receptors, antigen-presenting cell (APC)-related, and immune-stimulating or immune-suppressing molecules, were prepared from both the discovery and validation cohorts.

WGCNA

To identify gene modules and hub genes, weighted gene coexpression network analysis (WGCNA) was performed (Langfelder and Horvath 2008). Briefly, the similarity

matrix was constructed by computing the Pearson correlation coefficient between two genes. It was then transformed into an adjacency matrix with a soft threshold of $\beta = 3$ and finally into a topological matrix using the topological overlap measure (TOM) to quantify the degree of association between genes. To cluster the genes, the 1-TOM distance was employed, and then a dynamic pruning tree was constructed to identify the modules. The hub genes in the network were those with the highest K in rankings.

Immune subtyping

The 2002 immune-related genes were grouped based on their expression profiles, and a consistency matrix was constructed to identify immunological subtypes (Wilkerson and Hayes 2010). The consensus clustering method used 1000 repeats of the 1-Pearson correlation distance measure, and the optimal partition was determined based on the high intracluster consensus and low intercluster consensus.

Cancer immunity

The TIMER database (https://cistrome.shinyapps.io/timer/) was used to assess the quantity of tumor immune infiltrating cells (TIICs), which include T cells, B cells, macrophages, neutrophils, and dendritic cells (DCs) (Li et al. 2020). To further subdivide more specific immune cell types, we used single-sample GSEA (ssGSEA) to analyze the immune cell infiltration score of each sample (Hänzelmann et al. 2013) by applying the gene signatures of 28 tumor-infiltrating lymphocytes (TILs) (Ru et al. 2019). Immune Cell Abundance Identifier (ImmuCellAI) was utilized for estimation of the 24 immune cell types focusing on 18 T cell subsets (Miao et al. 2020). Gene markers of exhausted CD8+ T cells were retrieved from the study of single-cell sequencing of infiltrating T cells in HCC (Zheng et al. 2017).

The activity of the cancer immunity cycle, which includes cancer cell antigen release (Step 1), cancer antigen presentation (Step 2), antigen priming and activation (Step 3), immune cell trafficking to tumors (Step 4), immune cell infiltration into tumors (Step 5), cancer cells recognized by T cells (Step 6), and killing of cancer cells (Step 7) (Chen and Mellman 2013), was evaluated using the ssGSEA method (Xu et al. 2018).

Comprehensive analysis of immune and molecular features

The pancancer T cell-inflamed score (TIS) was used to define the pre-existing immune response to cancer (Ayers et al. 2017). We computed the TIS in HCC as a weighted linear combination of the scores from the T cell-inflamed gene expression signature. Tumor immune dysfunction and exclusion (TIDE, http://tide.dfci.harvard.edu/) scores were calculated as previously described (Fu et al. 2020) and were used to predict immunotherapy responsiveness.

To further characterize the immunological and molecular activities of the various groups, we used ssGSEA to compare the results for specific gene signatures. Kaplan–Meier survival curves were utilized to investigate prognostic differences. Correlation analysis was used to determine the relationship between immune subtypes and other markers. P values < 0.05 were considered statistically significant.

Results

Identification of HCC-related tumor antigens

Figure 1 summarizes the workflow of this study (Fig. 1). We identified 3522 protein-coding genes that most likely encode HCC-associated antigens by comparing the aberrantly expressed genes between tumors and normal tissues. Meanwhile, 4387 mutant genes probably encoding tumor-specific antigens were filtered by analyzing altered genome fractions and mutation counts in each sample (Fig. 2A, B). Titin (TTN) and tumor protein p53 (TP53) were identified as the most commonly altered genes in both FGA and mutation counts. Next, HLA binding of the predicted neoantigens was filtered by NetMHCpan2.8. Overall, 675 dysregulated and frequently mutated HCC-specific genes were identified.

Due to the high connection between programmed cell death 1 (PDCD1, PD-1) expression and CD8+ T cell genes and signatures across malignancies (Paré et al. 2018), we eliminated genes having a significantly positive correlation to PD-1 expression while retaining genes with a positive correlation to CD8A expression. In addition, prognosis-associated tumor antigens were selected from the aforementioned genes as potential candidates for mRNA vaccine generation (Fig. 2C). Finally, five TSAs were detected (named TSA5 hereafter), including FXYD domain containing ion transport regulator 6 (FXYD6), junctional adhesion molecule 2 (JAM2), polypeptide N-acetylgalactosaminyltransferase 16 (GALNT16), Complement C7 (C7), and coiled-coil domain containing 146 (CCDC146) (Fig. 2D). Enhanced expression levels of FXYD6, JAM2, GALNT16, C7 and CCDC146 were significantly associated with macrophage, DC, B and/or T cell infiltration, as well as a better prognosis (Fig. 3A, B).

Identification of HCC-related immune subtypes

Tumors and their microenvironment have unique immunological statuses in various immune subtypes, which aids in the identification of vaccine-eligible individuals. For this, consensus clustering was performed using the expression

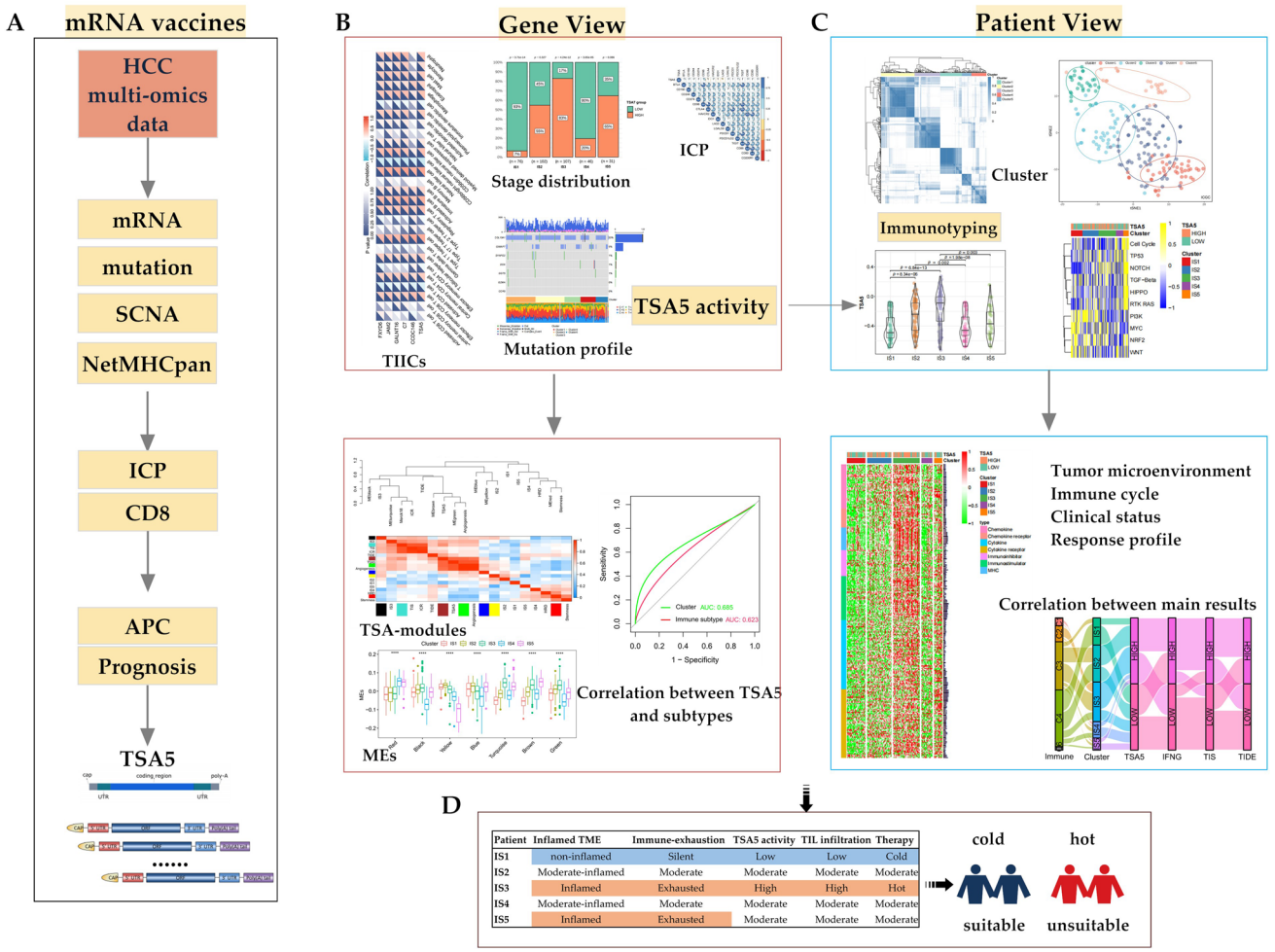

Fig. 1 Schematic diagram of the study. **A** Factors such as differential expression, mutation status, HLA binding prediction, immune relevance, and the prognosis algorithm were combined to screen for promising candidates. Five genes (TSA5) that have potential immunity-inducing effects were determined as candidates for the HCC mRNA vaccine. **B** Characterization of TSA5 activity. Correlation between TSA5 and immunological status, which includes TIICs, immunomodulators, immune checkpoints, cancer immunity cycle, and *T* cell-inflamed score. Modules were determined by the WGCNA method, and the significance of TSA5 in predicting molecular subtypes of HCC was established. **C** Identification of patients suitable for mRNA vaccines. By clustering immune genes in the HCC patients, five distinctive HCC subtypes were identified with cold or hot features. The specific immune status of the different subtypes was revealed. The Sankey diagram indicates the main results of this study. **D** Patients with IS1 are suitable for TSA5 vaccination

patterns of 2022 immune-related genes from 371 HCC cases in the TCGA cohort. The cumulative distribution function and functional delta area were used to estimate $k = 5$, the number of immune-related genes that seemed to cluster stably (Fig. 4A, B), resulting in the identification of five immunological subtypes (IS1–5) (Fig. 4C). Among them, IS3 tumors were associated with better survival probability, while IS4 tumors had the worst prognosis (Fig. 4D). The clinical stage distribution in each immune subtype is shown in Fig. 4E. Similar to the TCGA cohort, the immune subtype was predictive in the ICGC cohort (Fig. 4G–J) and significantly varied throughout stages (Fig. 4K).

Increased expression of alpha fetoprotein (AFP), glypican 3 (GPC3), and Golgi membrane protein 1 (GOLM1) are well-established prognostic and diagnostic markers of HCC associated with cancer progression (Piñero et al. 2020). Both the TCGA and ICGC cohorts demonstrated substantially varied levels of expression of these biomarkers across immunological subtypes in this research. IS4 exhibited increased expression of AFP and GPC3, while IS1 showed reduced expression of both markers (Fig. 4F, L). Overall, immune subtyping encompasses typical clinical characteristics, such as conventional staging or AFP expression, for predicting HCC patient prognosis.

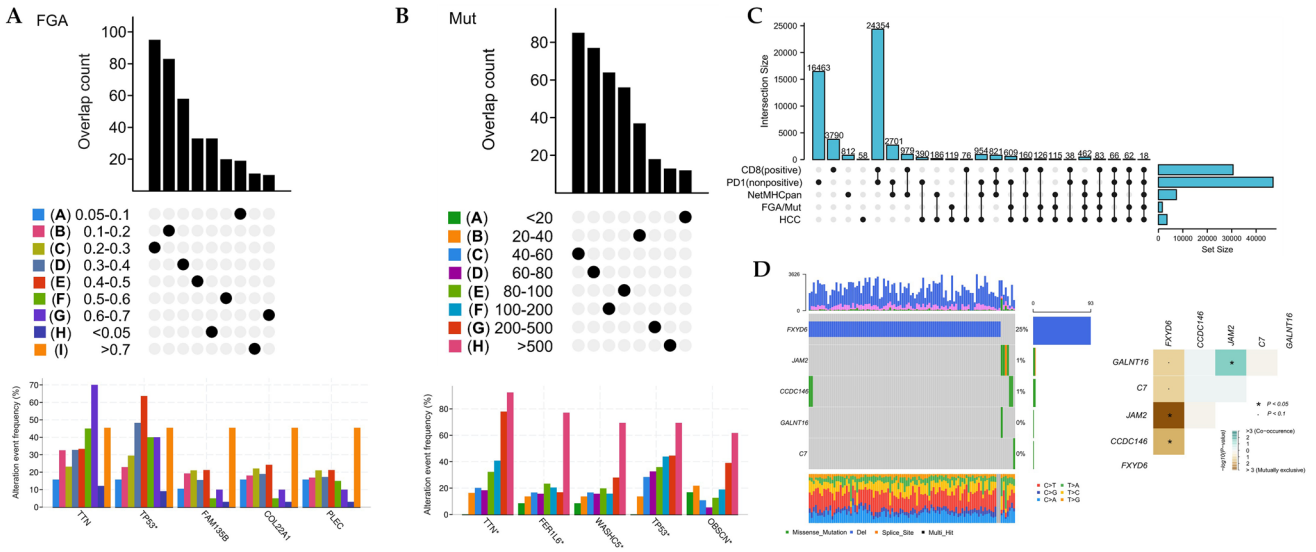

Fig. 2 HCC-related tumor antigen identification. **A** Samples overlapping in the group of fraction genome altered. The lower panel indicates the top 5 genes with the highest frequency in the FGA groups. **B** Samples that overlap in mutation count groups and the top 5 most often occurring genes in mutation count groups. **C** Overlapping plot of filtering steps. **D** Oncoplot of TSA5. The right panel shows the mutational interactions among TSA5

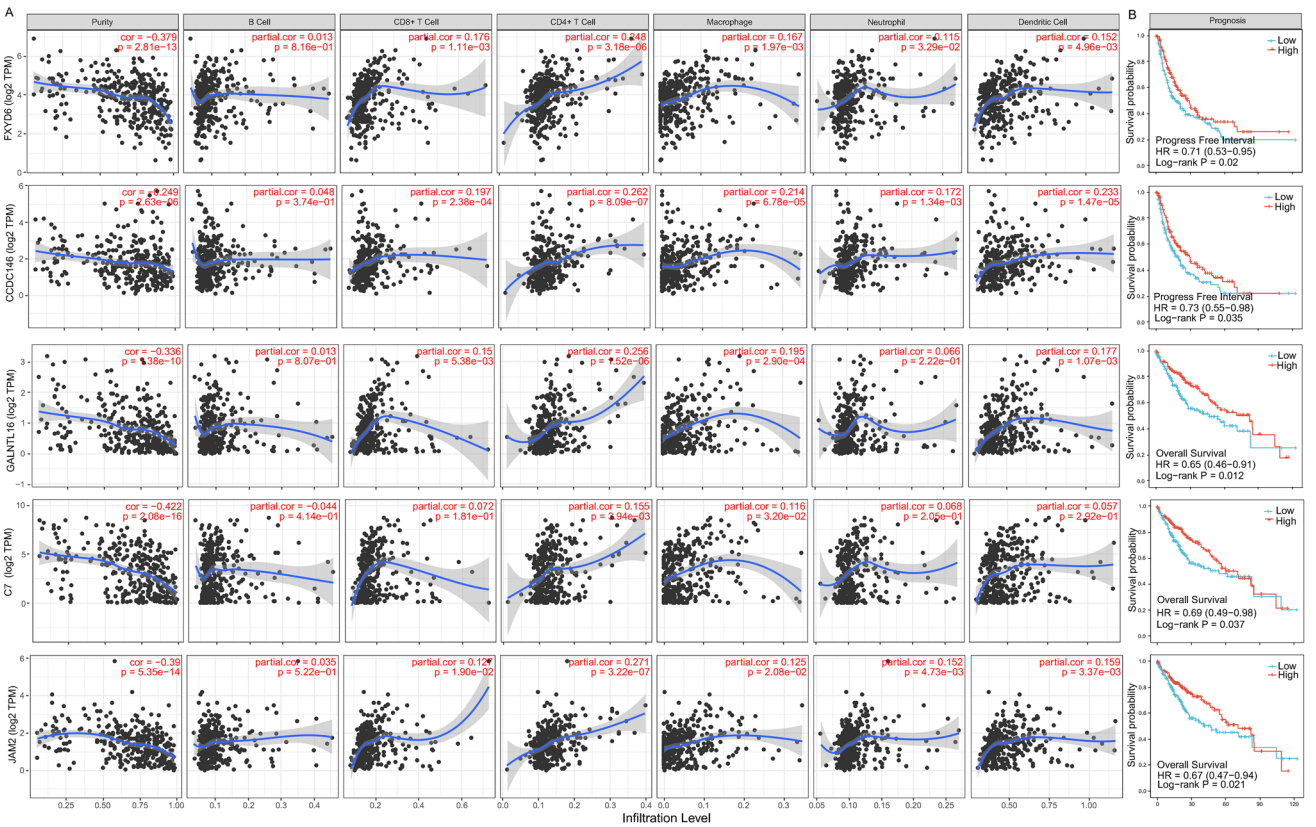

Fig. 3 Association between TSA5 and APCs and prognosis. **A** Correlations between the expression levels of FXYD6, JAM2, GALNT16, C7, and CCDC146 and infiltration of macrophages, dendritic cells, B cells and T cells in HCC. **B** Kaplan–Meier curves illustrating the overall survival (OS) or progression-free survival (PFI) of HCC patients stratified by TSA5 expression levels, using the median as the cutoff

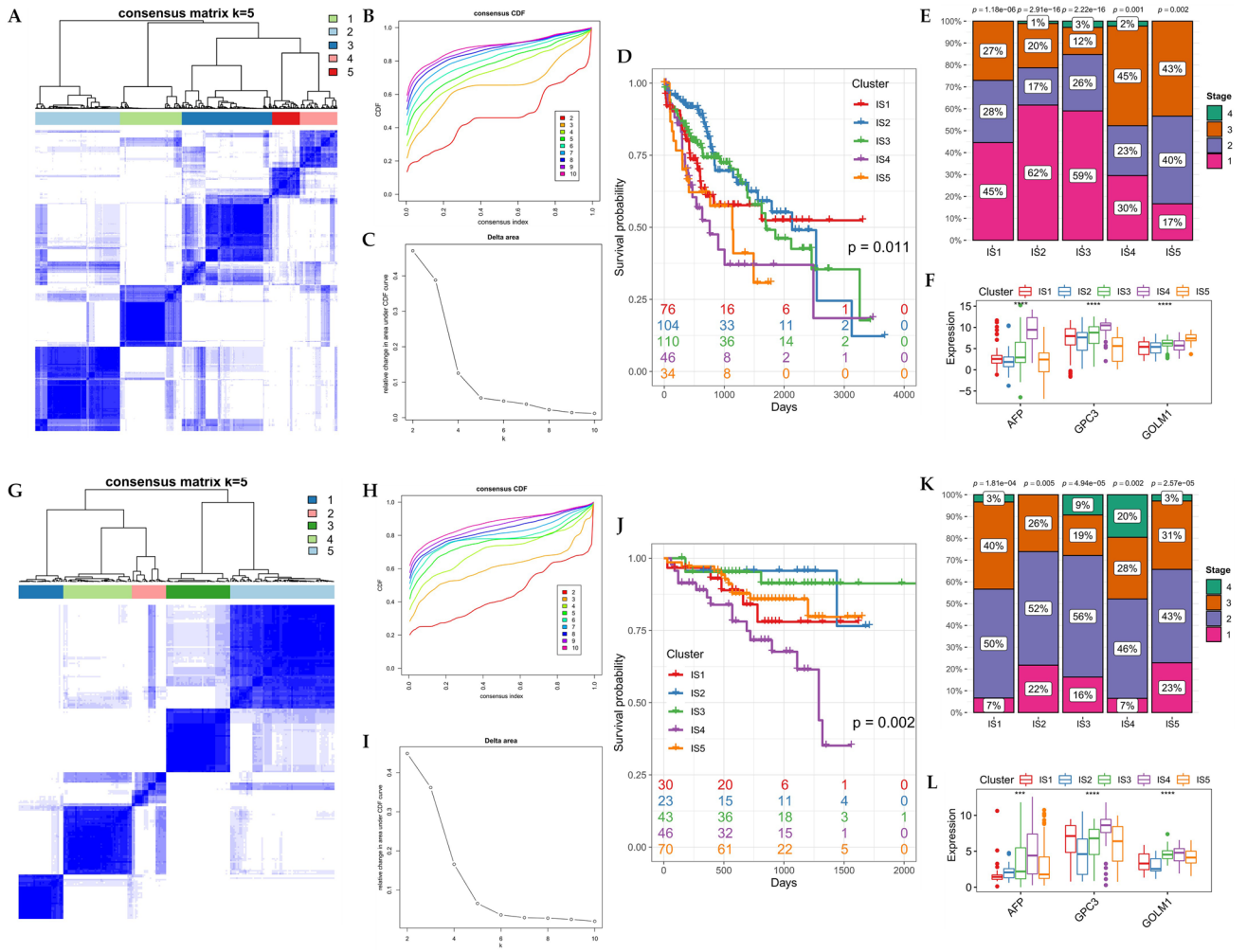

Fig. 4 HCC immune subtypes. **A** Heat map of sample clustering in the TCGA cohort. **B** Cumulative distribution function curve of immune-related genes in the TCGA cohort. **C** Delta area of immune-related genes. **D** Kaplan–Meier curves depicting the OS of different subtypes in the TCGA cohort. **E** The TCGA cohort's distribution of HCC stages by subtype. **F** Associations between immune subtypes and biomarkers in the TCGA cohort. **G–L** Consistent findings validated in the ICGC cohort

Characterization of TSA5 activity in HCC

Among the different subtypes, the IS1 subtype exhibited the lowest TSA5 activity, while the IS3 subtype had the greatest (Fig. 5A). Notably, decreased TSA5 activity was linked with a poor outcome in HCC patients (Fig. 5B).

TSA5 activity was shown to be significantly correlated with the infiltration of memory-effector CD8+ T cells, Th1 T cells, activated B cells, NK cells, macrophages, and dendritic cells (Fig. 5C). TSA5 also exhibited a favorable association with the TIICs' effector genes (Fig. 5E).

Immune checkpoint inhibitors were shown to be underexpressed in the noninflamed tumor microenvironment (TME). In this study, the majority of immune checkpoint inhibitors, including PD-1, CD274 (PD-L1), CTLA4, lymphocyte activation gene 3 (LAG3), T cell immunoglobulin domain and mucin domain containing protein 3 (TIM-3; also known as HAVCR2), indoleamine 2,3-dioxygenase 1 (IDO1), and T cell immunoreceptor with Ig and ITIM domain (TIGIT), were found to have negative correlations with TSA5 activity. These findings have been confirmed in the ICGC cohort.

Accordingly, the inflamed phenotype exhibited the greatest incidence of CD8 positivity. TSA5 activity was shown to be modestly associated with CD8A and TIS expression (Fig. 5F). These findings suggested that TSA5 activation was compatible with effective anticancer immunity.

Six different molecular subtypes (C1–C6) have been consistently identified in previous large-scale pancancer genomic profiling investigations, with C1, C2, C3, and C4 comprising the predominant types in HCC (Thorsson et al. 2019). IS1 and IS2 were mostly overlapping with C3 and C4, IS3 with C2, C3 and C4, and IS4 and IS5 with C1, C2, C3 and C4 (Fig. 5I). C3 and C4 were shown to be linked with improved prognoses, while C1 and C2 were found to

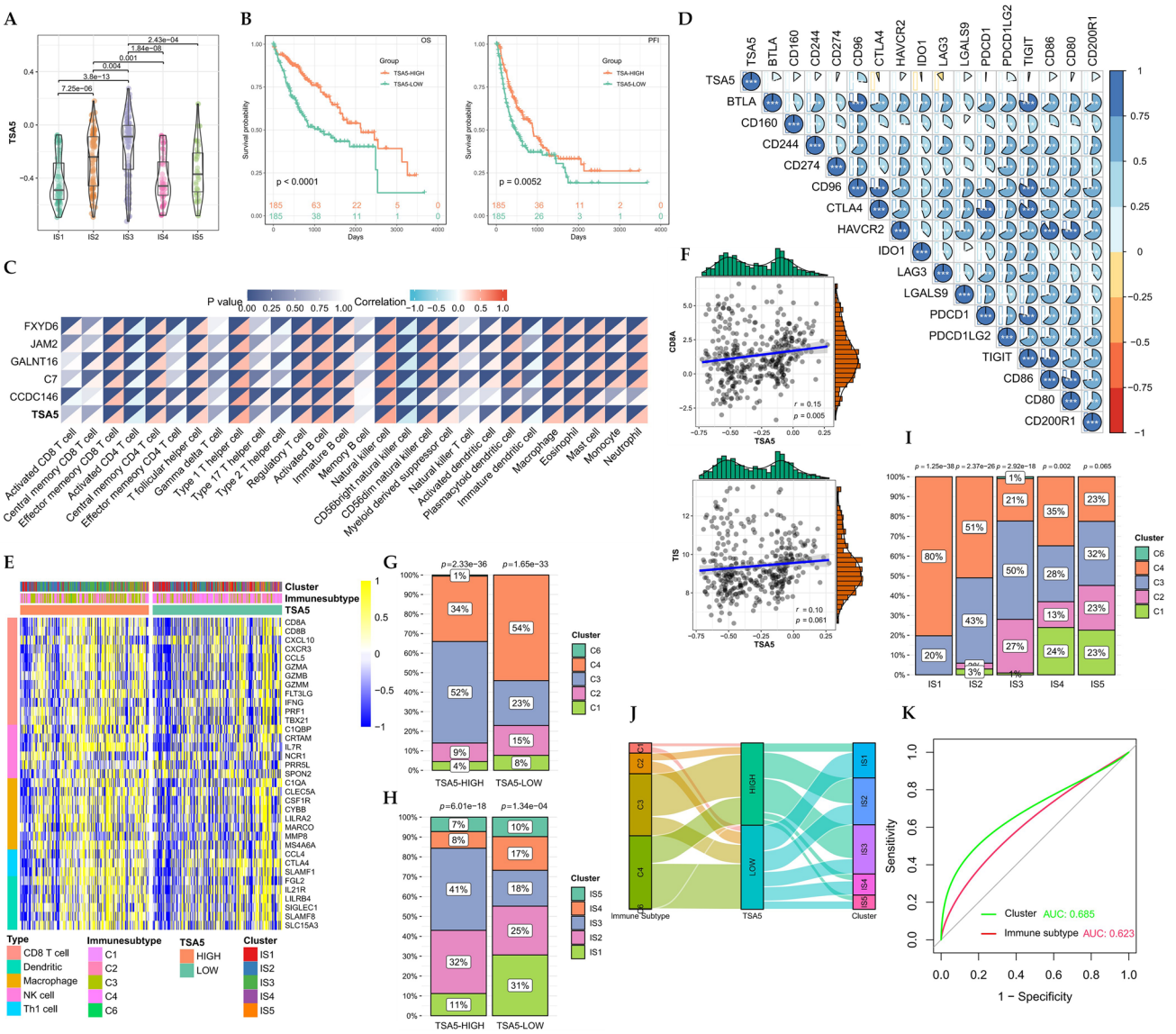

Fig. 5 Characterizing the immune activity of TSA5 in IS1–5. **A** Differences in TSA5 activity across the IS1–IS5 subtypes in HCC. **B** High activity of TSA5 results in better OS or PFI in HCC. **C** Correlation between TSA5 activity and the infiltration levels of 28 types of TILs. **D** Correlations between TSA5 and 16 inhibitory immune checkpoints. **E** Differences in effector gene expression between high- and low-TSA5 tumor-associated immune cells (NK cells, macrophages, dendritic cells, CD8+T cells, and Th1 cells). **F** Correlation between TSA5 and CD8A expression and the T cell-inflamed score. **G–I** Subtype differences between high- and low-TSA5 groups. **J** Relationship between TSA5 and the immune subtyping groups. **K** Predictive accuracy of TSA5 for HCC subtypes

be connected with poor prognoses. These findings correlated with increased survival in IS1 and IS2 cancers vs IS4 and IS5 tumors.

As shown in Fig. 5H, the IS2 and IS3 subtypes were more common in the high-TSA5 group, and there was a more even distribution of IS1–IS5 subtypes in the low-TSA5 group. Similarly, the low-TSA5 group had more C2 and C4 samples and less C3 samples than the high-TSA5 group (Fig. 5I). TSA5's predictive accuracy for IS varied from 0.623 to 0.685 (Fig. 5K). These results establish the relevance of our TSA5 definition to immunogenicity and the reliability of the immunotyping method.

Associations between HCC subtypes and immune modulators

The activities associated with cancer immunity are controlled by the functions of immunomodulators (Chen and Mellman 2013). Globally, most immunomodulators were highly expressed in the IS3 and IS5 groups (Fig. 6A).

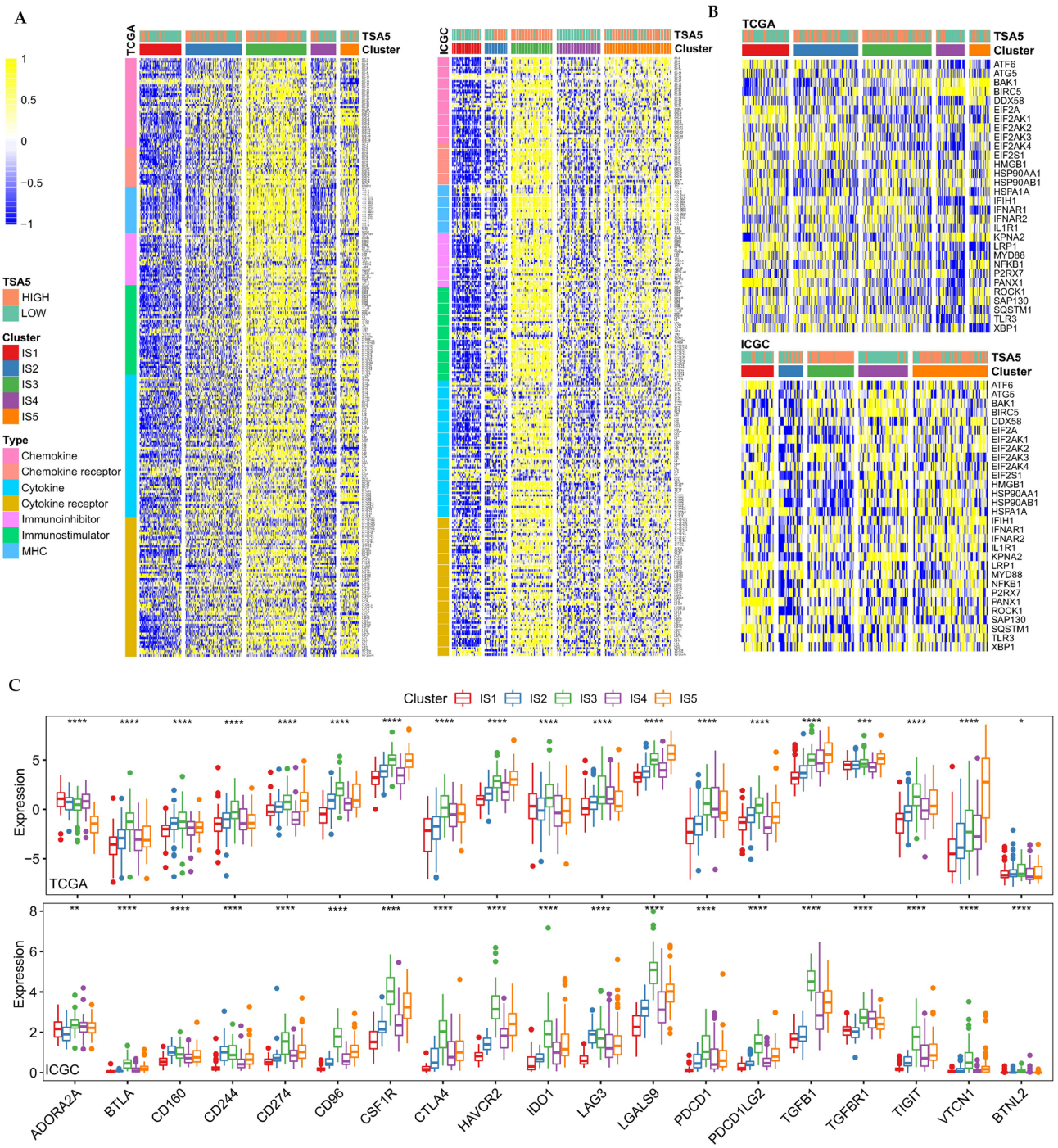

Fig. 6 Associations between immunomodulators and immunological subtypes. **A** Differences in the expression of immunomodulators among IS1–5 in the TCGA and ICGC cohorts. **B** Expression of ICD-related genes is differentially expressed among HCC immune subtypes in the TCGA and ICGC cohorts. **C** Differential expression of ICP-related genes among the HCC immune subtypes in the TCGA and ICGC cohorts

IS3 and IS5 patients had higher levels of CCR8, CCR10, CXCR3, and CXCR6, but lower levels of CCR2, CCR5, and CXCR4 (Ma et al. 2019), indicating that they have a distinct immune-exhausted state.

IS1 was found to be negatively correlated with a large number of immunomodulators (Fig. 6A). In particular, the loss or low expression of CXCR3 and its chemokine ligands CXCL9 and CXCL10, as well as CCR5 (Dangaj et al. 2019)

and its ligands CCL3, CCL4 and CCL5 in the IS1 group, corresponded to TIL desertification (González-Martín et al. 2011). Furthermore, bone morphogenetic protein 4 (BMP4) and IL27 expression was elevated in the IS1 group. IL27 promotes IFNγ responses in HCC cells (Rolvering et al. 2017). BMP4 augments the survival of HCC cells under hypoxic and hypoglycemic conditions by promoting the glycolysis pathway (Zhong et al. 2021). In addition, the levels of cytokines such as CSF1, IL-11, IL18 and IL1β were decreased in the IS1 group. IL1β promotes HCC metastasis by upregulating CD274 and CSF1 (He et al. 2021). Elevation of IL-11 levels triggers HCC outgrowth (D. Wang et al. 2019).

The majority of immunological checkpoint (ICP)-related genes were differentially expressed among subtypes (Fig. 6C). CD274, CD96, CSF1R, CTLA4, HAVCR2, IDO1, LAG3, LGALS9, PDCD1, PDCD1LG2 (PDL2), TGFBR1, TIGIT, and VTCN1 were substantially elevated in IS3 or IS5 tumors but considerably downregulated in IS1 tumors.

Immunogenic cell death (ICD) can elicit anticancer immune responses that reinforce the therapeutic effects of conventional anticancer chemotherapies and radiotherapy (Deutsch et al. 2019). Thirty ICD genes showed substantial variation between subtypes (Fig. 6B). For example, EIF2A, EIF2Ak1, HMGB1, HSPA1A, LRP1, and PANX1 were increased in IS1 tumors, whereas BAK1, BIRC5, and KPNA2 were upregulated in IS4 tumors. As a result, immunotyping may serve as a biomarker for mRNA vaccines by reflecting the expression levels of ICP and ICD modulators.

HCC-immunity of immune subtypes

Given the important role of the TME in the response to mRNA vaccines, we next examined the immune cell composition in IS1–5. IS3 and IS5 showed similar immune patterns, including the infiltration of both antitumor immune cells, such as activated CD4 T cells, central memory CD4 T cells, effector memory CD8+ T cells, natural killer T cells, and activated DCs, and protumor immune cells, such as myeloid-derived suppressor cells (MDSCs) and regulatory T cells (Tregs). Nevertheless, all these subtypes have notably varying immune cell compositions. For example, IS1 had greater memory B cell scores than the other groups, while IS2 had higher eosinophil and Th17 cell scores (Fig. 7A, B).

Across the seven-step cancer immunity cycle (Xu et al. 2018), the activities of the majority of the steps in the cycle were shown to be downregulated in the IS1 group, including cancer cell antigen release (Step 1), priming and activation (Step 3), and immune cell trafficking to tumors (Step 4). Interestingly, the activity of cancer antigen presentation (Step 2) and recognition of cancer cells by T cells (Step 6) were upregulated in the IS1 group. The activity of Step 7 (killing of cancer cells) was relatively high in the IS1 group (Fig. 7C).

Solid tumors lacking TILs are considered to be immunologically cold (Galon and Bruni 2019). Thus, IS1 and IS2 have a "cold" immunological phenotype, whereas IS3 and IS5 have a "hot" immunological phenotype. Similar results were observed in the ICGC cohort (Fig. 7B). Collectively, the immunological subtypes represent the cellular and molecular immune status of HCC patients and serve as predictive markers for mRNA vaccination.

Associations between immune subtypes and cancer characteristics

A higher burden of mutation load and higher somatic mutation rates are correlated with stronger anticancer immunity (Rooney et al. 2015). As shown in Fig. 8A, IS1 showed an elevated TMB compared with IS2 and IS3 in the ICGC cohort. Furthermore, there was a significantly higher MSI score in the IS1 group in both the ICGC and TCGA cohorts (Fig. 8B). In addition, IS1 and IS2 showed a significantly lower TIS and exhausted CD8+ T signatures than the IS3, IS4 and IS5 groups (Fig. 8C, D).

The distribution of oncogenic hallmark pathways is shown in Fig. 8E. WNT-β catenin signaling has been linked to an immunosuppressive phenotype (Roelands et al. 2020), while PI3K-Akt signaling was associated with CD8+ T cell infiltration in the HCC microenvironment (Zhang et al. 2019). Both pathways were more active in the IS1 group (Fig. 8E).

We then used the TIDE and immunologic constant of rejection (ICR) algorithms (Roelands et al. 2020) to assess the potential clinical efficacy of immunotherapy in different subgroups. A higher TIDE prediction score represented a higher potential for immune evasion, which suggested that the patients were less likely to benefit from immunotherapy. The IS1 subgroup had a lower TIDE and ICR score than the other subgroups, as well as lower T cell exclusion and dysfunction scores, implying that IS1 patients could benefit more from ICI therapy than other patients (Fig. 8F).

Correlations between immunological subtypes and previously established molecular signatures were also examined. As shown in Fig. 8G, IS3 and IS5 had the highest scores for angiogenesis, leukocyte fraction, stromal fraction, macrophage regulation, lymphocyte infiltration, IFN-gamma response and TGF-β response. Thus, IS3 and IS5 were considered an overall favorable immune-activated phenotype and had high levels of TILs and immunomodulators. IS1 was characterized by low scores for TP53, leukocyte fraction, stromal fraction, macrophage regulation, lymphocyte infiltration and TGF-β response signatures. IS4 exhibited high stemness, TP53, and proliferation scores but low scores

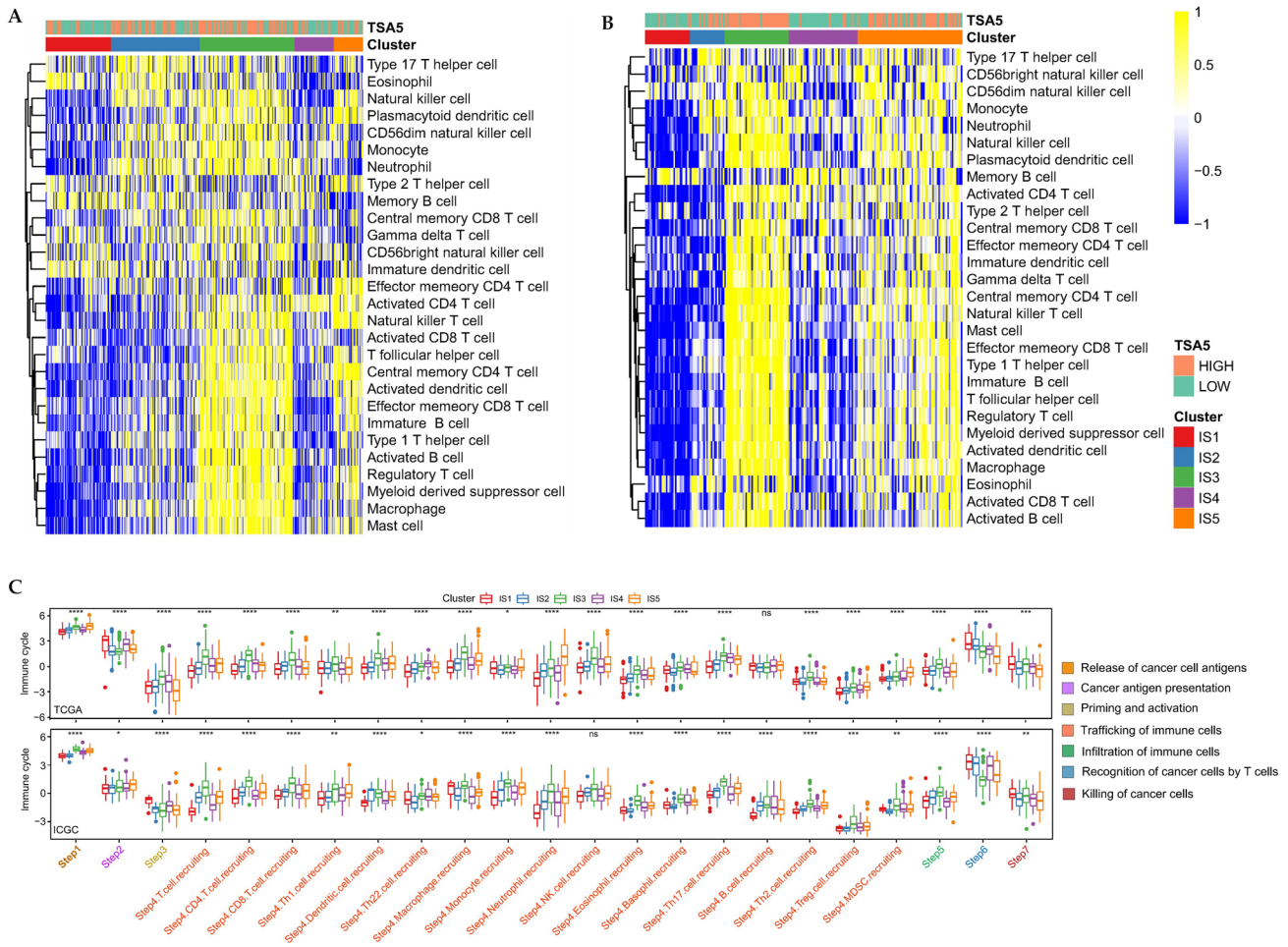

Fig. 7 The tumor immunity of HCC subtypes. **A**, **B** Scores of differential enrichment for 28 immune cell signatures in HCC immune subtypes from the TCGA and ICGC cohorts. **C** Differential enrichment scores of immune cell signatures related to the seven-step cancer immunity cycle in the TCGA and ICGC cohorts

for angiogenesis and IFN-gamma responsiveness, indicating both an immune-cold and an immunosuppressive feature.

These findings suggest that the immune subtype presenting the unique characteristics of HCC patients and those with IS1 may be more responsive to mRNA vaccination.

Characterization of TSA5-related immune modules of HCC

Next, we used the WGCNA method to find the relationship between immune gene coexpression modules and immune subtypes or TSA5 activity. As shown in Fig. 9A, we obtained 7 coexpression modules, except the gray colors indicate genes outside of any functional modules (Fig. 9A). TSA5 activity was highly correlated with the green and brown modules, which are associated with the PI3K-Akt and MAPK signaling pathways, respectively (Fig. 9B, F).

We next examined the distribution of modules' eigengenes across five ISs and discovered that seven modules had substantially different distributions (Fig. 9D). For example, IS1 demonstrated that the turquoise and brown modules had low eigengenes whereas the yellow module had high eigengenes. IS4 has the greatest number of eigengenes in the red module and the fewest in the black, blue, and green modules (Fig. 9B, D). The analysis of survival-relevant genes within the red module revealed that increased expression ratings were associated with a poor prognosis, and vice versa in the turquoise module (Fig. 9D, E). In addition, the red module was associated with cell cycle progression and T cell activation. Similarly, the turquoise module was associated with cytokine–cytokine receptor interactions (Fig. 9C). Furthermore, we utilized hub genes in each module to assess immune cell infiltration in various modules (Fig. 9G), and the results corroborated the above-mentioned findings (Fig. 7A, B).

The therapeutic potential of mRNA vaccines in cancer patients with certain immunological subtypes is primarily determined by T cell infiltration and activation in the

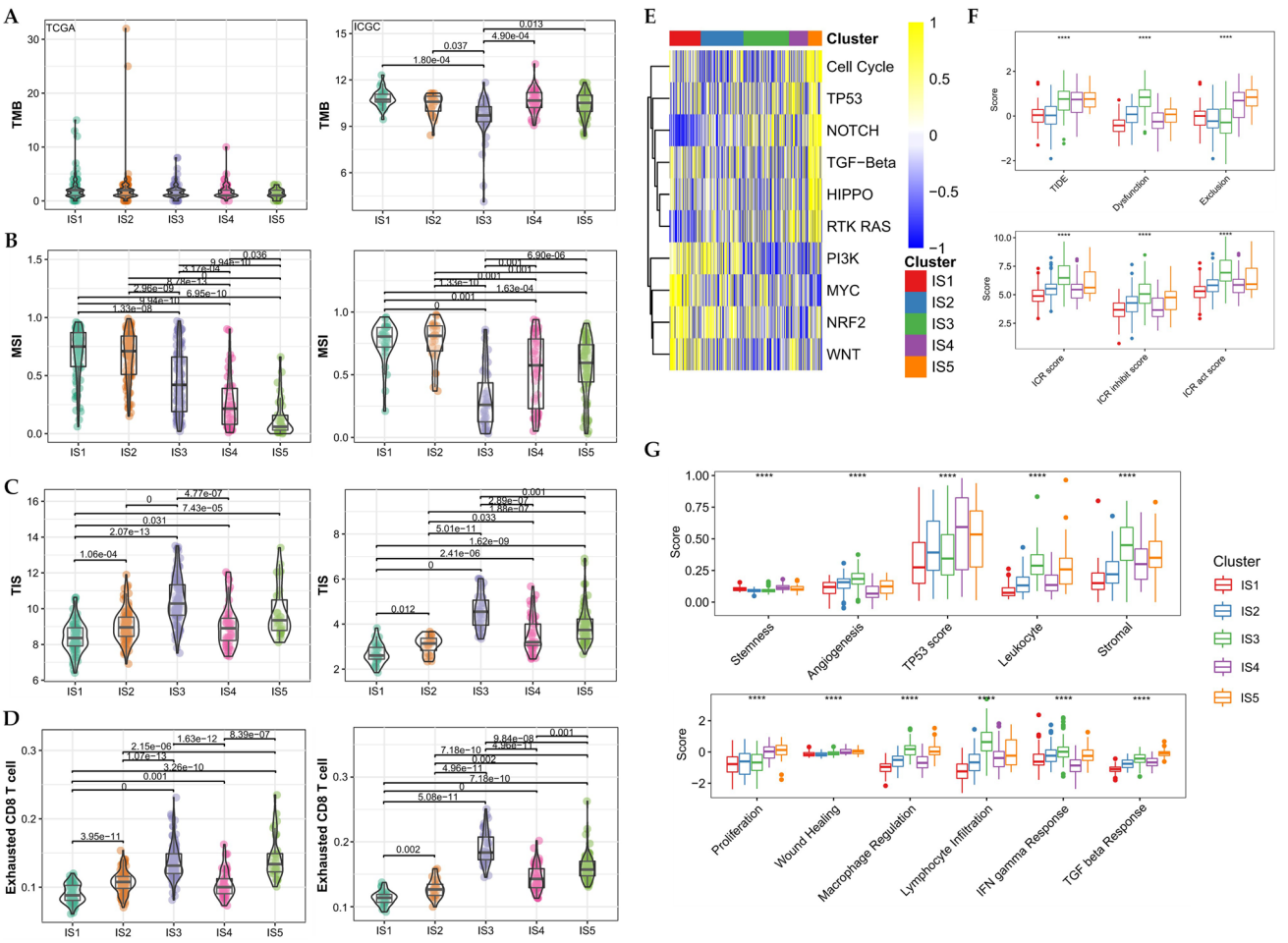

Fig. 8 Associations between immune subtypes and molecular signatures. **A**–**D** TMB, MSI, TIS, and exhausted CD8+T score in the TCGA and ICGC cohorts. **E** Changes in 10 oncogenic pathways in HCC IS1–IS5. **F** TIDE and ICR scores in HCC IS1–IS5. **G** Differential enrichment scores of the molecular signatures among HCC IS1–IS5

TME, as well as suppression of immune-suppressive cells. Upregulation of hub genes in the red module, such as NUF2, TTK, and KIF11, was associated with a negative correlation with the immune response, as measured by cytotoxic and exhausted T cell scores (Fig. 9G), implying that patients who expressed high levels of the genes clustered into red modules may have a poor response to the mRNA vaccine. Thus, the red module's hub genes may serve as indicators for the mRNA vaccination response.

Discussion

Over the past 20 years, clinical studies of TAA-based cancer vaccines have been mostly disappointing, largely due to low immunogenicity. Interest in mRNA-based cancer vaccines was rekindled by advances in immune checkpoint inhibition and the initial success of RNA vaccinations (Pardi et al. 2018).

To target HCC TSAs, several factors affecting tumor immunogenicity were considered in this study. First, we screened the downregulated neoantigens with poor prognosis to ensure that the dose escalation of the mRNA vaccines might increase the therapeutic effect. Second, the candidate TSAs were distributed across the different FGA/mutation groups in HCC patients with a high degree of heterogeneity, thus increasing the chance of finding responders. Indeed, an RNA vaccine targeting multiple nonmutated TAAs prevalent in melanoma recently induced a durable objective response in ICI-experienced patients (Sahin et al. 2020). Third, APC relevance was determined, indicating that the mutated peptides of TSA had HLA binding affinity and may be processed and presented to CD8+T cells in the presence of sufficient lymphocyte infiltration to elicit an immunological response. Fourth, the candidates have weak or negative correlations with ICPs. While cytotoxic CD8+T cells are essential for antitumor immunity, when combined with a suppressive tumor microenvironment and extended antigen exposure,

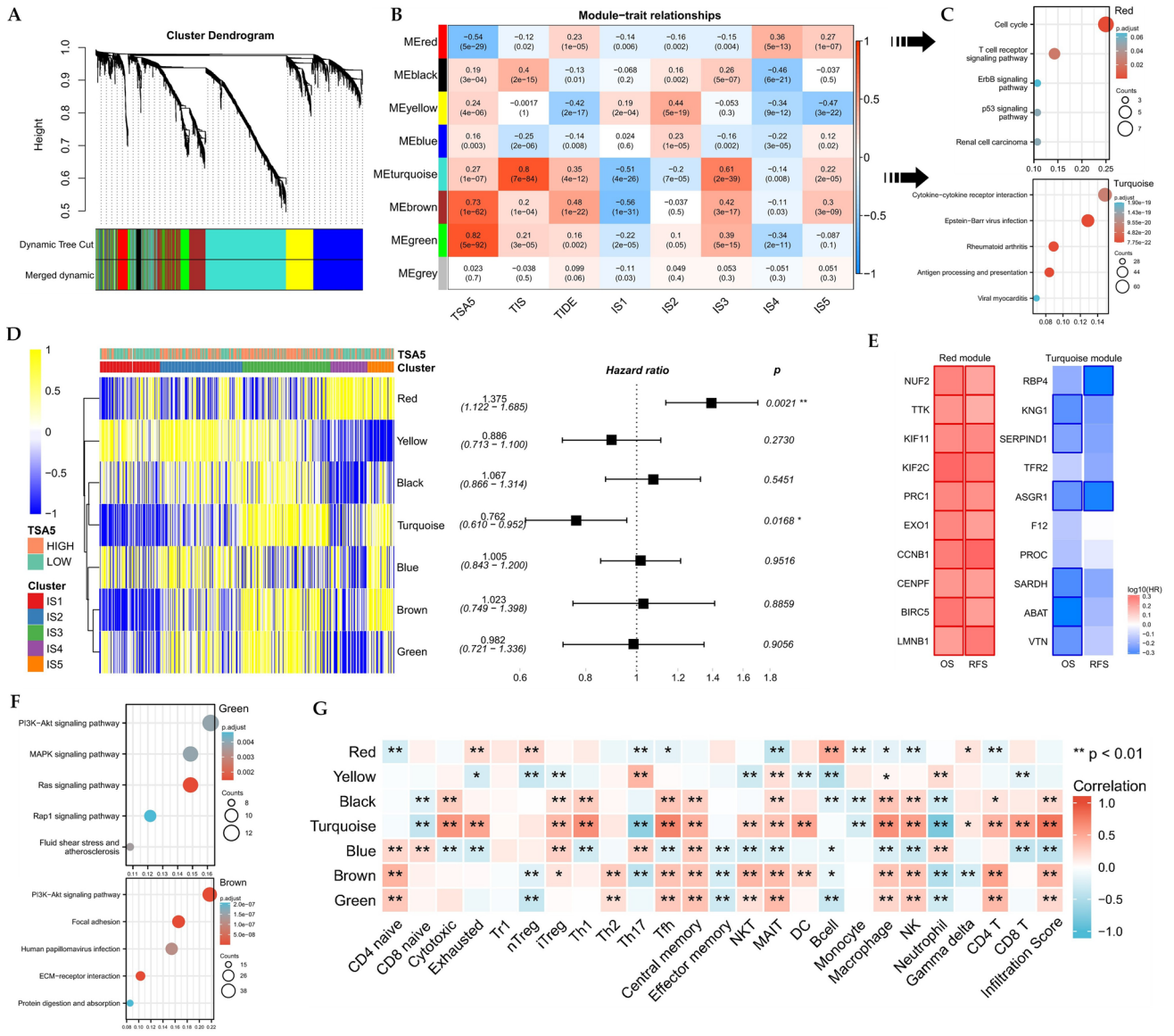

Fig. 9 Immune gene coexpression modules of HCC. **A** Dendrogram of the immune-related coexpression network. **B** Module–trait relationship. **C** An enrichment map highlighting the most frequently occurring KEGG keywords in the red and turquoise modules. The dot size and color intensity reflect the number of genes and their degree of enrichment, respectively. **D** Differential distribution of each module's eigengenes among HCC subtypes. Forest maps illustrating the single-factor survival analysis of seven HCC modules. **E** Survival map of the top 10 hub genes of the red and turquoise modules. **F** Dot plot illustrating the green module's enriched terms. **G** Correlation between the GVSA score of each module and immune cells

tumor-specific effector CD8+ T cells are prone to enter a T cell exhaustion stage [13]. Exhausted CD8+ T cells show a hierarchical decline in their ability to produce cytokines and kill, in contrast to functioning effector and memory T cells. Through these steps, a series of targetable antigens were identified, of which FXYD6, JAM2, GALNT16, C7, and CCDC146 are promising mRNA vaccine candidates.

Although these candidates need to be functionally verified, prior research shows that they have the potential to be used in mRNA vaccines. For example, FXYD6 has been reported to affect the activity of Na(+)/K(+)-ATPase as well as the downstream Src-ERK signaling pathway in HCC (Gao et al. 2014). JAMs function in leukocyte recruitment to inflamed tissue, and in mice suffering from chronic liver injury, such as autoimmune hepatitis, endothelial cells trigger increased expression of JAM2 (Hintermann et al. 2018). GALNT16 encodes a N-acetylgalactosaminyltransferase that modifies protein O-glycosylation, a process involved in tumor development (Raman et al. 2012; Wu et al. 2019). As

a potential tumor suppressor, C7 is correlated with tumor progression and prognosis (Ying et al. 2016).

Although many studies have shown a change in local immune responses in an antitumor direction in HCC patients, such as enhanced infiltration of cytotoxic T, NK, and NKT cells(Ma et al. 2019), the overall response rate to checkpoint inhibitors in HCC patients is only 15–20% (El-Khoueiry et al. 2017). Consistently, HCC patients had higher expression of exhaustion-related inhibitory receptors on CD8+ T cells (Ma et al. 2019), and the intensity and quantity of ICPs expressed by exhausted T cells were positively correlated with the degree of exhaustion. Increased PD-L1 expression in cancer cells and/or APCs indicates that an otherwise efficient T cell response has been suppressed (Wang et al. 2020). As demonstrated in this study, TSA5 was negatively correlated with the expression of most ICPs. Next, TSA5 activity was highly correlated with the green and brown modules, representing overlap with some immune-inflamed features of the IS3 group (Fig. 9B). Consistently, TSA5 activity is positively correlated with CD8 and TIS, suggesting inherent immunogenicity. Third, TSA5 activity is positively correlated with multiple immune active cells, including CD8+ T cells, DCs, and macrophages. These results shed light on the feasibility of using TSA5 as a target for cancer vaccination.

Given that mRNA vaccines are beneficial for only a subset of patients, immunological subtypes were defined to identify the group suitable for cancer vaccination. Molecular, cellular, and clinical differences were distinct among the five immune subtypes. Both the ICGC and TCGA cohorts' prognostic predictions revealed that patients with IS3 tumors had a better prognosis, whereas those with IS4 tumors had the poorest. These findings imply that immunotyping may outperform established tumor markers such as AFP and GPC3 as well as traditional staging in predicting the prognosis of HCC patients. The high expression of ICPs in HCC, especially in IS3 tumors, presents an immune-exhausted tumor microenvironment, which may impede an efficient immune response from mRNA vaccination. In addition, the expression of ICD modulators was evenly distributed among the subtypes, but there were fewer ICD modulator-expressing samples in the IS4 subgroup than in the other subgroups. However, IS1 tumors have a high level of ICD modulator expression, indicating that IS1 tumors may have more promise for mRNA vaccination than other immune subtypes.

Due to the fact that the immunological state of the tumor influences the effectiveness of mRNA vaccines, we further identified the immune cell components in the various clusters. IS3 and IS5 had significantly higher levels of activated CD8 T cells, activated B cells, monocytes, and effector memory CD4 T cells than IS1 and IS2, whereas IS1, IS2, and IS4 had higher levels of CD56dim NK cells, supporting the notion that IS3 and IS5 are immunologically hot phenotypes and IS1, IS2, and IS4 are immunologically cold phenotypes. Furthermore, IS1 was characterized by low TP53, leukocyte and stromal fraction, macrophage regulation, lymphocyte infiltration and TGF-β response signatures, as well as TIS and TIDE scores. IS4 was associated with high scores for stemness, TP53 and wound healing and low scores for angiogenesis and IFN-gamma response and scored high for TIDE and TIS, indicating an immunologically cold and moderately inflamed phenotype. Correspondingly, patients with IS1 with higher TMB and MSI may have greater responsiveness to mRNA vaccines. Nevertheless, the vaccination antigens and other prognostic indicators discovered in this research merit further exploration with the goal of developing applications.

Conclusions

FXYD6, JAM2, GALNT16, C7, and CCDC146 are potential HCC antigens for mRNA vaccine design. Patients with immune subtype 1 are a suitable population for vaccination. This study provides a framework for developing anti-HCC mRNA vaccines and biomarkers for vaccination.

Acknowledgements We would like to thank ICGC (https://www.icgc-argo.org) and TCGA (https://portal.gdc.cancer.gov) for data collection, as well as The cBioPortal database (http://www.cbioportal.org/) and The TIMER database (https://cistrome.shinyapps.io/timer/) for the provision of data processing.

Author contributions JY conceived and designed this work. JF and FC conducted bioinformatic analyses. YL, JG, and AC helped with data collection, analysis, and interpretation. JF and JY wrote the manuscript. JY revised the manuscript. All the authors read and approved the final manuscript.

Funding This research was supported by grants from the National Natural Science Foundation of China (81772520), the Natural Science Foundation of Zhejiang Province (LGF19H030004), and Zhejiang medical and health technology project (2018PY039).

Data availability The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Declarations

Conflict of interest The authors declared no potential conflicts of interest in terms of the research, authorship, and/or publication of this article.

Ethical approval No application.

Consent to participate No application.

Consent for publication No application.

References

Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR et al (2017) IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. J Clin Invest 127(8):2930–2940. https://doi.org/10.1172/jci91190

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68(6):394–424. https://doi.org/10.3322/caac.21492

Chen DS, Mellman I (2013) Oncology meets immunology: the cancer-immunity cycle. Immunity 39(1):1–10. https://doi.org/10.1016/j.immuni.2013.07.012

Chin SF, Teschendorff AE, Marioni JC, Wang Y, Barbosa-Morais NL, Thorne NP et al (2007) High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. Genome Biol 8(10):R215. https://doi.org/10.1186/gb-2007-8-10-r215

Dangaj D, Bruand M, Grimm AJ, Ronet C, Barras D, Duttagupta PA et al (2019) Cooperation between constitutive and inducible chemokines enables *T* cell engraftment and immune attack in solid tumors. Cancer Cell 35(6):885-900.e810. https://doi.org/10.1016/j.ccell.2019.05.004

Deutsch E, Chargari C, Galluzzi L, Kroemer G (2019) Optimising efficacy and reducing toxicity of anticancer radioimmunotherapy. Lancet Oncol 20(8):e452–e463. https://doi.org/10.1016/s1470-2045(19)30171-8

El-Khoueiry AB, Sangro B, Yau T, Crocenzi TS, Kudo M, Hsu C et al (2017) Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. Lancet 389(10088):2492–2502. https://doi.org/10.1016/s0140-6736(17)31046-2

Farazi PA, DePinho RA (2006) Hepatocellular carcinoma pathogenesis: from genes to environment. Nat Rev Cancer 6(9):674–687. https://doi.org/10.1038/nrc1934

Fu J, Li K, Zhang W, Wan C, Zhang J, Jiang P et al (2020) Large-scale public data reuse to model immunotherapy response and resistance. Genome Med 12(1):21. https://doi.org/10.1186/s13073-020-0721-z

Galon J, Bruni D (2019) Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. Nat Rev Drug Discov 18(3):197–218. https://doi.org/10.1038/s41573-018-0007-y

Gao Q, Chen X, Duan H, Wang Z, Feng J, Yang D et al (2014) FXYD6: a novel therapeutic target toward hepatocellular carcinoma. Protein Cell 5(7):532–543. https://doi.org/10.1007/s13238-014-0045-0

González-Martín A, Gómez L, Lustgarten J, Mira E, Mañes S (2011) Maximal T cell-mediated antitumor responses rely upon CCR5 expression in both CD4(+) and CD8(+) T cells. Cancer Res 71(16):5455–5466. https://doi.org/10.1158/0008-5472.Can-11-1687

Hänzelmann S, Castelo R, Guinney J (2013) GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinform 14:7. https://doi.org/10.1186/1471-2105-14-7

He Q, Liu M, Huang W, Chen X, Zhang B, Zhang T et al (2021) IL-1β-induced elevation of solute carrier family 7 member 11 promotes hepatocellular carcinoma metastasis through up-regulating programmed death ligand 1 and colony-stimulating factor 1. Hepatology. https://doi.org/10.1002/hep.32062

Hintermann E, Bayer M, Conti CB, Fuchs S, Fausther M, Leung PS et al (2018) Junctional adhesion molecules JAM-B and JAM-C promote autoimmune-mediated liver fibrosis in mice. J Autoimmun 91:83–96. https://doi.org/10.1016/j.jaut.2018.05.001

Huang X, Tang T, Zhang G, Liang T (2021a) Identification of tumor antigens and immune subtypes of cholangiocarcinoma for mRNA vaccine development. Mol Cancer 20(1):50. https://doi.org/10.1186/s12943-021-01342-6

Huang X, Zhang G, Tang T, Liang T (2021b) Identification of tumor antigens and immune subtypes of pancreatic adenocarcinoma for mRNA vaccine development. Mol Cancer 20(1):44. https://doi.org/10.1186/s12943-021-01310-0

Kalbasi A, Ribas A (2020) Tumour-intrinsic resistance to immune checkpoint blockade. Nat Rev Immunol 20(1):25–39. https://doi.org/10.1038/s41577-019-0218-4

Khan P, Siddiqui JA, Lakshmanan I, Ganti AK, Salgia R, Jain M et al (2021) RNA-based therapies: a cog in the wheel of lung cancer defense. Mol Cancer 20(1):54. https://doi.org/10.1186/s12943-021-01338-2

Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinform 9:559. https://doi.org/10.1186/1471-2105-9-559

Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q et al (2020) TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res 48(W1):W509-w514. https://doi.org/10.1093/nar/gkaa407

Ma J, Zheng B, Goswami S, Meng L, Zhang D, Cao C et al (2019) PD1(Hi) CD8(+) *T* cells correlate with exhausted signature and poor clinical outcome in hepatocellular carcinoma. J Immunother Cancer 7(1):331. https://doi.org/10.1186/s40425-019-0814-7

Miao YR, Zhang Q, Lei Q, Luo M, Xie GY, Wang H et al (2020) ImmuCellAI: a unique method for comprehensive T-Cell subsets abundance prediction and its application in cancer immunotherapy. Adv Sci (weinh) 7(7):1902880. https://doi.org/10.1002/advs.201902880

Papachristofilou A, Hipp MM, Klinkhardt U, Früh M, Sebastian M, Weiss C et al (2019) Phase Ib evaluation of a self-adjuvanted protamine formulated mRNA-based active cancer immunotherapy, BI1361849 (CV9202), combined with local radiation treatment in patients with stage IV non-small cell lung cancer. J Immunother Cancer 7(1):38. https://doi.org/10.1186/s40425-019-0520-5

Pardi N, Hogan MJ, Porter FW, Weissman D (2018) mRNA vaccines—a new era in vaccinology. Nat Rev Drug Discov 17(4):261–279. https://doi.org/10.1038/nrd.2017.243

Paré L, Pascual T, Seguí E, Teixidó C, Gonzalez-Cao M, Galván P et al (2018) Association between PD1 mRNA and response to anti-PD1 monotherapy across multiple cancer types. Ann Oncol 29(10):2121–2128. https://doi.org/10.1093/annonc/mdy335

Piñero F, Dirchwolf M, Pessôa MG (2020) Biomarkers in hepatocellular carcinoma: diagnosis prognosis and treatment response assessment. Cells 9(6):1370. https://doi.org/10.3390/cells9061370

Raman J, Guan Y, Perrine CL, Gerken TA, Tabak LA (2012) UDP-N-acetyl-α-D-galactosamine:polypeptide N-acetylgalactosaminyltransferases: completion of the family tree. Glycobiology 22(6):768–777. https://doi.org/10.1093/glycob/cwr183

Roberts TC, Langer R, Wood MJA (2020) Advances in oligonucleotide drug delivery. Nat Rev Drug Discov 19(10):673–694. https://doi.org/10.1038/s41573-020-0075-7

Roelands J, Hendrickx W, Zoppoli G, Mall R, Saad M, Halliwill K et al (2020) Oncogenic states dictate the prognostic and predictive connotations of intratumoral immune response. J Immunother Cancer 8(1):e000617. https://doi.org/10.1136/jitc-2020-000617

Rolvering C, Zimmer AD, Kozar I, Hermanns HM, Letellier E, Vallar L et al (2017) Crosstalk between different family members: IL27 recapitulates IFNγ responses in HCC cells, but is inhibited by IL6-type cytokines. Biochim Biophys Acta Mol Cell Res 1864(3):516–526. https://doi.org/10.1016/j.bbamcr.2016.12.006

Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N (2015) Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell 160(1–2):48–61. https://doi.org/10.1016/j.cell.2014.12.033

Ru B, Wong CN, Tong Y, Zhong JY, Zhong SSW, Wu WC et al (2019) TISIDB: an integrated repository portal for tumor-immune system interactions. Bioinformatics 35(20):4200–4202. https://doi.org/10.1093/bioinformatics/btz210

Sahin U, Oehm P, Derhovanessian E, Jabulowsky RA, Vormehr M, Gold M et al (2020) An RNA vaccine drives immunity in checkpoint-inhibitor-treated melanoma. Nature 585(7823):107–112. https://doi.org/10.1038/s41586-020-2537-9

Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH et al (2019) The immune landscape of cancer. Immunity 51(2):411–412. https://doi.org/10.1016/j.immuni.2019.08.004

Vasan N, Baselga J, Hyman DM (2019) A view on drug resistance in cancer. Nature 575(7782):299–309. https://doi.org/10.1038/s41586-019-1730-1

Vitale I, Sistigu A, Manic G, Rudqvist NP, Trajanoski Z, Galluzzi L (2019) Mutational and antigenic landscape in tumor progression and cancer immunotherapy. Trends Cell Biol 29(5):396–416. https://doi.org/10.1016/j.tcb.2019.01.003

Wang D, Zheng X, Fu B, Nian Z, Qian Y, Sun R et al (2019) Hepatectomy promotes recurrence of liver cancer by enhancing IL-11-STAT3 signaling. EBioMedicine 46:119–132. https://doi.org/10.1016/j.ebiom.2019.07.058

Wang M, Wang S, Desai J, Trapani JA, Neeson PJ (2020) Therapeutic strategies to remodel immunologically cold tumors. Clin Transl Immunol 9(12):e1226. https://doi.org/10.1002/cti2.1226

Wilkerson MD, Hayes DN (2010) ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics 26(12):1572–1573. https://doi.org/10.1093/bioinformatics/btq170

Wu J, Zhao W, Zhou B, Su Z, Gu X, Zhou Z et al (2018) TSNAdb: a database for tumor-specific neoantigens from immunogenomics data analysis. Genom Proteom Bioinform 16(4):276–282. https://doi.org/10.1016/j.gpb.2018.06.003

Wu H, He G, Song T, Zhang Y, Chen X, Chen H et al (2019) Evaluation of GALNT16 polymorphisms to breast cancer risk in Chinese population. Mol Genet Genom Med 7(8):e848. https://doi.org/10.1002/mgg3.848

Xu L, Deng C, Pang B, Zhang X, Liu W, Liao G et al (2018) TIP: a web server for resolving tumor immunophenotype profiling. Cancer Res 78(23):6575–6580. https://doi.org/10.1158/0008-5472.Can-18-0689

Ying L, Zhang F, Pan X, Chen K, Zhang N, Jin J et al (2016) Complement component 7 (C7), a potential tumor suppressor, is correlated with tumor progression and prognosis. Oncotarget 7(52):86536–86546. https://doi.org/10.18632/oncotarget.13294

Zhang Q, Lou Y, Yang J, Wang J, Feng J, Zhao Y et al (2019) Integrated multiomic analysis reveals comprehensive tumour heterogeneity and novel immunophenotypic classification in hepatocellular carcinomas. Gut 68(11):2019–2031. https://doi.org/10.1136/gutjnl-2019-318912

Zheng C, Zheng L, Yoo JK, Guo H, Zhang Y, Guo X et al (2017) Landscape of infiltrating *T* cells in liver cancer revealed by single-cell sequencing. Cell 169(7):1342-1356.e1316. https://doi.org/10.1016/j.cell.2017.05.035

Zhong J, Kang Q, Cao Y, He B, Zhao P, Gou Y et al (2021) BMP4 augments the survival of hepatocellular carcinoma (HCC) cells under hypoxia and hypoglycemia conditions by promoting the glycolysis pathway. Am J Cancer Res 11(3):793–811