


ORIGINAL RESEARCH

Single-cell RNA-seq and bulk RNA-seq explore the prognostic value of exhausted T cells in hepatocellular carcinoma

Xiaolong Tang^{1,2} | Yandong Miao^{1,3} | Lixia Yang⁴ | Wuhua Ha¹ | Zheng Li⁵ | Denghai Mi^{1,4} 

¹The First Clinical Medical College, Lanzhou University, Lanzhou, China

²The Second Department of Gastrointestinal Surgery, Affiliated Hospital of North Sichuan Medical College, Nanchong, China

³Department of Oncology, Yantai Affiliated Hospital of Binzhou Medical University, The Second Clinical Medical College of Binzhou Medical University, Yantai, China

⁴Gansu Academy of Traditional Chinese Medicine, Lanzhou, China

⁵Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou, China

Correspondence

Denghai Mi and Zheng Li.

Email: mi.dh@outlook.com and lizhenglys@126.com

Funding information

Special Plan for Condition Construction of [Gansu Provincial Scientific Research Institutes], Grant/Award Number: 20JR10RA432; China Postdoctoral Science Foundation, Grant/Award Number: 2019M663860

Abstract

Hepatocellular carcinoma (HCC) remains a worldwide health problem. Mounting evidence indicates that exhausted T cells play a critical role in the progress and treatment of HCC. Therefore, a detailed characterisation of exhausted T cells and their clinical significance warrants further investigation in HCC. Based on the GSE146115, we presented a comprehensive single-cell Atlas in HCC. Pseudo-time analysis revealed that tumour heterogeneity progressively increased, and the exhausted T cells gradually appeared during tumour progression. Functional enrichment analysis revealed that the evolutionary process of exhausted T cells mainly contained the pathway of cadherin binding, proteasome, cell cycle, and T cell receptor regulation of apoptosis. In the International Cancer Genome Consortium database, we divided patients into three clusters with the T cell evolution-associated genes. We found that the exhausted T cells are significantly related to poor outcomes through immunity and survival analysis. In The Cancer Genome Atlas database, the authors enrolled weighted gene co-expression network analysis, univariate Cox analysis, and Lasso Cox analysis, then screened the 19 core genes in T cells evolution and built a robust prognostic model. This study offers a fresh view on evaluating the patients' outcomes from an exhausted T cells perspective and might help clinicians develop therapeutic systems.

KEYWORDS

bioinformatics, tumours

1 | INTRODUCTION

Primary liver cancer is the sixth most commonly diagnosed cancer and the third leading cause of cancer-associated death worldwide in 2020, with approximately 906,000 new cases and 830,000 deaths [1]. Hepatocellular carcinoma (HCC) accounts for more than 90% of liver cancer cases [2]. Conventional

treatments include hepatic resection, liver transplantation, local ablation with radiofrequency, and hepatic artery intervention, such as transarterial chemoembolisation etc., [3–5]. Currently, systemic therapies including tyrosine kinase inhibitor, immune checkpoint inhibitors, and monoclonal antibodies have challenged the use of conventional therapies in the treatment of liver cancer [6, 7]. However, as therapy advanced, the

Xiaolong Tang and Yandong Miao was contributed equally to this work.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *IET Systems Biology* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

prognosis of HCC patients is still poor, and 5-year overall survival rate is only about 10% [8]. Therefore, we need more strategies to reveal the pathological process of tumour development to help clinicians formulate therapeutic procedures [9, 10].

Tumours have a high degree of heterogeneity, resulting in people responding differently to the same therapy [11]. Besides, multilesion HCC is frequent and may result from concurrent carcinogenesis, intrahepatic metastases, or clonal spread [12]. Thus, lesions within the same tumour may exhibit unique genetic changes, biological behaviours, and microenvironmental characteristics and react differentially to treatments. Even tumour cells within the same lesion may have a variety of somatic mutations [13]. Some researchers indicated that an optimal strategy for multilesional HCC would manage all subtype lesions into account [14]. Therefore, exploration of intratumoral heterogeneity is a significant issue in HCC.

Although much work has been spent explaining tumour heterogeneity, our knowledge is still primarily confined to tumour cells [15]. Recent evidence suggests that stromal cells, including tumour-infiltrating immune cells, exhibit heterogeneity as well [16]. In addition, the tumour microenvironment (TME) is becoming a more prominent target of pharmacological therapy methods [17]. The fact that naturally occurring T cells with antitumour activity exist in human cancer has rationalised immunotherapy's use in oncology. The correlation between intratumoral T cell increase and better survival in cancer indicates that tumour-specific T cell activity plays a critical anticancer function [18]. With the development of single-cell transcriptome analysis, T cells can be annotated and analysed at a more detailed level in HCC progress. Notably, exhausted T cells have previously been observed as a unique dysfunctional cell lineage that develops during persistent infections and malignancies [19]. The exhausted T cells exhibit several distinct characteristics, including decreased effector cytotoxicity, decreased cytokine production, and upregulated multiple inhibitory molecular receptors, such as PDCD-1, LAG3, CTLA4 [20]. Moreover, many researchers suggested that the exhausted T cells may be a crucial factor in the success of immunotherapy and patient outcomes [21, 22]. In addition to the classical markers to define exhausted T cells, some scholars have proposed that in a different field, exhausted T cells show many phenomena beyond the traditional definition [23]. Seo, Hyungseok et al. reported that by changing the phenotype and transcriptional profile of chimaeric antigen receptor (CAR) T cells, basic leucine zipper ATF-like transcription factor (BATF) could improve their anti-tumour response [24]. Recently, Grebinoski, Stephanie et al. indicated that LAG3 might be a core gene regulating the function of CD8⁺ T cells, and the deficiency of LAG3 can reduce their exhaustion phenotype, which reveals the possibility of LAG3 as a target for immunotherapy [25]. Therefore, a thorough understanding of exhausted T cells and further exploration of their new crucial markers in oncology is critical for developing novel immunotherapies for HCC patients.

In this study, we obtained both single-cell RNA-seq data and bulk RNA-seq data from public databases and analysed

them through various computational biology techniques. The single-cell RNA-seq data was performed by Su, Xianbin et al. [26], which contained four HCC patients. We further described the tumour heterogeneity and explored the characterisation of exhausted T cells. By correlating the single-cell RNA-seq data with the ICGC database, we found that a high level of exhausted T cells was significantly associated with a low clinical outcome. Moreover, in the TCGA database, we formed a new prognostic model through the crucial genes in T cells evolution. The model is a stable measure to estimate the prognosis and an accurate indicator for evaluating the exhausted T cells' level in HCC patients. This investigation offers a fresh view on evaluating the HCC patients' outcome from an exhausted T cells perspective and might help clinicians develop therapeutic systems. The flow chart of the study design and analysis is shown in Figure 1.

2 | MATERIALS AND METHODS

2.1 | Data acquisition

The single-cell RNA-seq dataset of HCC (GSE146115) was obtained from the Gene Expression Omnibus database (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) and has been described by Su, Xianbin et al. [26] in their supplemental data. An additional single-cell RNA-seq dataset of HCC (GSE151530) described by Ma, Lichun et al. [27] was utilised for validation. Bulk RNA-seq data and the relevant patients' information of HCC were obtained from The Cancer Genome Atlas database (TCGA, <https://portal.gdc.cancer.gov/>, v29.0) and International Cancer Genome Consortium database (ICGC, <https://dcc.icgc.org/>). The matrix files of RNA-seq data for each sample were collated and annotated onto the genome by R software (version 4.0.2).

2.2 | Quality control and single-cell RNA-seq analysis

To explore the tumour cell heterogeneity in HCC cells, we enrolled single-cell RNA-seq data from GSE146115. The quality control (QC) process was conducted through R-package “Seurat” (version 3.0.1) [28]. It was considered low-quality and removed with less than 50 unique molecular identifiers (UMIs) or more than 5% mitochondrion-derived UMI count for single cells. The differences among the patients' batch were normalised with the Integrate function of “Seurat”. The top 1500 variable genes were selected for further analysis. Subsequently, principal component analysis (PCA) and t-distributed stochastic neighbour embedding (TSNE) were enrolled to process the data, and the major cell clusters were visualised through 2D TSNE plots [29]. The TSNE analysis represents a non-linear dimensionality reduction technique, conceived by Laurens van der Maaten et al. in 2008. This method is specifically designed for the visualisation of high-dimensional data [30]. We employed the FindAllMarkers

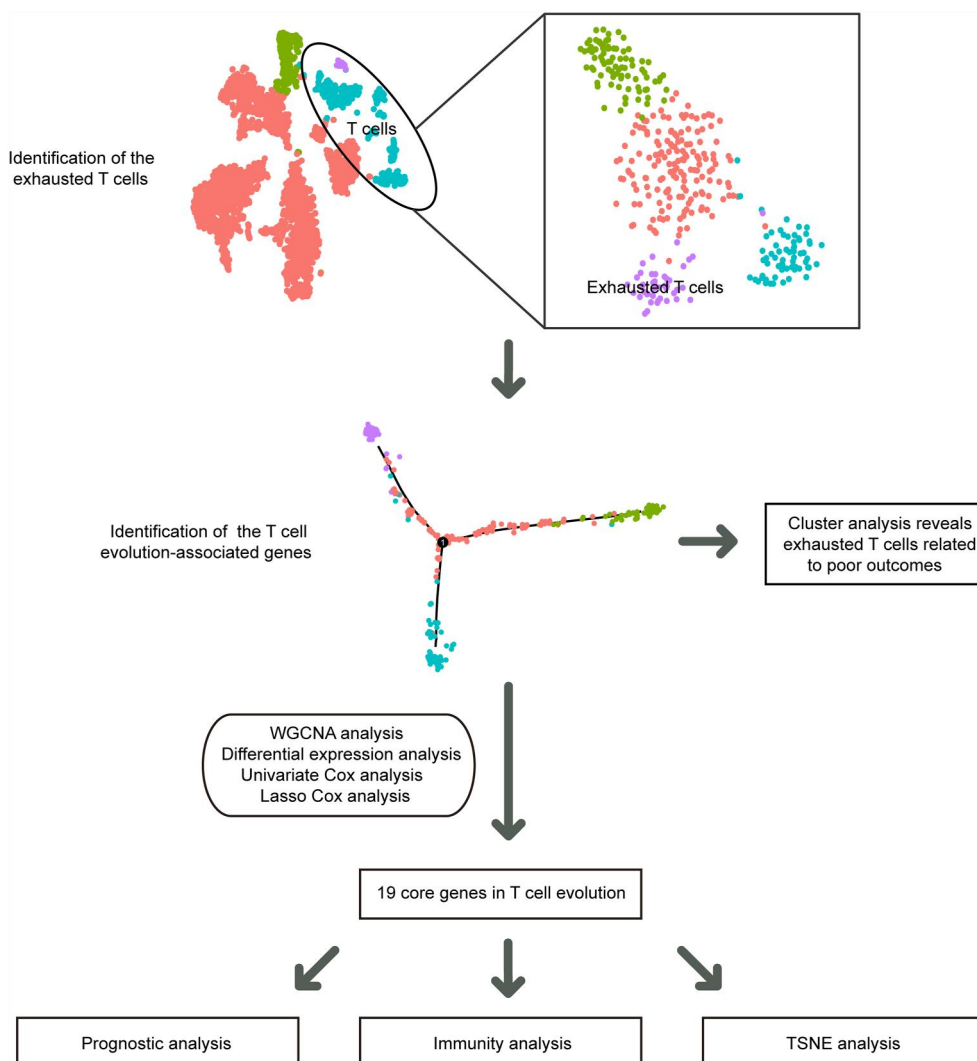


FIGURE 1 The flow chart of the study design and analysis.

function to identify the markers of each cell cluster. The primary cell characters were then recognised through the markers from the CellMarker database (<http://biocc.hrbmu.edu.cn/CellMarker/index.jsp>) [31].

2.3 | Pseudo-time trajectory analysis

To discover the evolutionary process of tumour cells and T cells during the tumour advancement, we visualised the trajectories through 2D TSNE plots by the R-package “monocle” [32]. Then, we analysed the distribution of time, branching, and cell clustering along these trajectories.

2.4 | Functional enrichment analysis

To understand how the function of exhausted T cells changes during tumour progression, we performed Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes

(KEGG) analysis based on the T cell evolution-associated genes through R-package “clusterProfiler”, “enrichplot”, “ggplot2”, and “org.Hs.for example,db” [33, 34]. Then, we used the Bioplanet model in Enrichr (<https://maayanlab.cloud/Enrichr/>) [35] and ClueGO's immunoassay model in Cytoscape (<http://apps.cytoscape.org/apps/cluego>) [36] to explore the immunity function further.

2.5 | Exploration of the relationship between subtypes of T cells and prognosis of HCC patients

To further explore the relationship between various subtypes of T cells and outcomes of HCC patients, we clustered the HCC patients in the ICGC database through T cell evolution-associated genes by R-package “ConsensusClusterPlus” [37]. Then, Kaplan–Meier (K-M) survival analysis, TME score analysis, and immune checkpoints analysis were used to reveal the clinicopathologic feature of each cluster of HCC patients

through R-package “survival”, “survminer” [38], “estimate”, and “limma” [39]. The TME score encompasses four distinct metrics, namely Stromal Score, Immune Score, ESTIMATE Score, and Tumour Purity. This algorithm primarily relies on the “estimate” R package and employs ESTIMATE (Estimation of Stromal and Immune cells in Malignant Tumour tissues using Expression data), a methodology predicated on gene expression data, for the purpose of calculating cellular composition and purity within tumour tissues [40].

2.6 | Screening for the core genes in T cell evolution and building the prognostic model

Weighted gene co-expression network analysis (WGCNA) was used on the basis of T cell evolution-associated genes to screen the core genes in T cells evolution. We constructed a scale-free co-expression network in the TCGA database and built some crucial modules associated with clinical features through R-package “WGCNA” [41]. An appropriate power of $\beta = 6$ was set to ensure a signed scale-free co-expression gene network. Then, based on the result of WGCNA, we further selected the differentially expressed genes (DEGs) between tumour and adjacent tissues using the Wilcoxon Test with R-package “limma”. We set the criteria as $|\text{Log Fold Change}| \geq 1$ and false discovery rate (FDR)-adjusted p -value ≤ 0.05 . Subsequently, we enrolled univariate COX analysis and least absolute shrinkage and selection operator (LASSO) regression analysis in screening the core genes with overall survival (OS). With the coefficients from the multivariate regression analysis and the expression level, we calculated the signature of T cell evolution (TCESig) of each patient by the following formula: $\text{TCESig}(\text{patient}) = \sum_{j=1}^n \text{Coeff}_j * X_j$, TCESig represents the prognostic risk score for each HCC patient, Coeff_j represents the coefficient, and X_j represents the expression level of every gene in the signature. We treated the TCGA database as a train set and treated the ICGC database as a test set. We adopted the median TCESig in TCGA as a cutoff value and separated HCC patients into the high-risk and low-risk groups in the TCGA and ICGC database. Subsequently, we arrange patients based on their risk scores, positioning them from left to right in ascending order of their scores. In accordance with this sequence, we generate a heatmap of TCESig's gene, risk curve, and survival state. By superimposing these three graphical representations, we aim to more effectively elucidate the relationship between TCESig and patient prognosis in both the training and testing sets. The R-package “time-ROC” was used to draw the Receiver Operating Characteristic (ROC) curve for investigating the sensitivity and specificity of the survival prediction by the TCESig [42]. Area Under Curve (AUC) delivered as an index of prognostic accuracy. K-M analysis was used to assess the survival rate for each group. Through univariate and multivariate Cox analysis, we verified whether TCESig was an independent prognosis factor in HCC.

2.7 | Validation of the relationship between the TCESig and immunity characteristics

To explore the relationship between the TCESig and subtypes of T cells, we examined the expression of TCESig's genes in different T cell groups. TSNE analysis was used to show the high expression region of TCESig's genes. TME score analysis, immune cell infiltration analysis, and immune checkpoints analysis were used to reveal the immunity characteristics of TCESig.

2.8 | Exploration of the TCESig' function and prognostic value

Several exhausted T cells-related gene sets were enrolled to illustrate the TCESig's function through Gene Set Enrichment Analysis (GSEA) (version 4.1.0) analysis [43]. Then, we examined the distribution of the TCESig in different clusters in the ICGC database. Since the CD8A expression can reflect the CD8T cells' level, and the TCESig may reflect the degree of accumulation of exhausted T cells. The combination of CD8A expression and TCESig may better reflect the degree of effective cytotoxic T cell deficiency in HCC patients. Hence, we performed survival analysis combining CD8A expression and the TCESig through the R-package “survival” and “survminer”.

2.9 | Statistical analysis

We used the R software (version 4.0.2) and strawberry-Perl (version 5.30.0) to perform data processing. In all hypothesis tests, p -values ≤ 0.05 were considered significant.

3 | RESULTS

3.1 | Tumour cell heterogeneity in HCC cells

Based on GSE146115, a total of 16 samples from four HCC patients were involved in this study, which contained 3200 single cells. Details of the patients are shown in Supplementary Table 1 (Table S1). The QC standards are described in Materials and Methods, and 3198 single cells matched the QC standards (Figure S1A-C). According to the CellMarker database, these cells were divided into 14 subgroups (Figure S1D). In addition, the 14 subgroups were mainly annotated into four main types, including hepatocytes, T cells, NK cells, and monocyte (Figure 2a–b). Then, through the function of monocle, we performed pseudo-time analysis to infer the co-expression module and visualise the cell differentiation trajectory. We found that the tumour heterogeneity gradually increased with the advance of tumour, and various subtypes of tumour cells and immune cells gradually appeared (Figure 2c–e). The results indicated that tumour cells and related immune

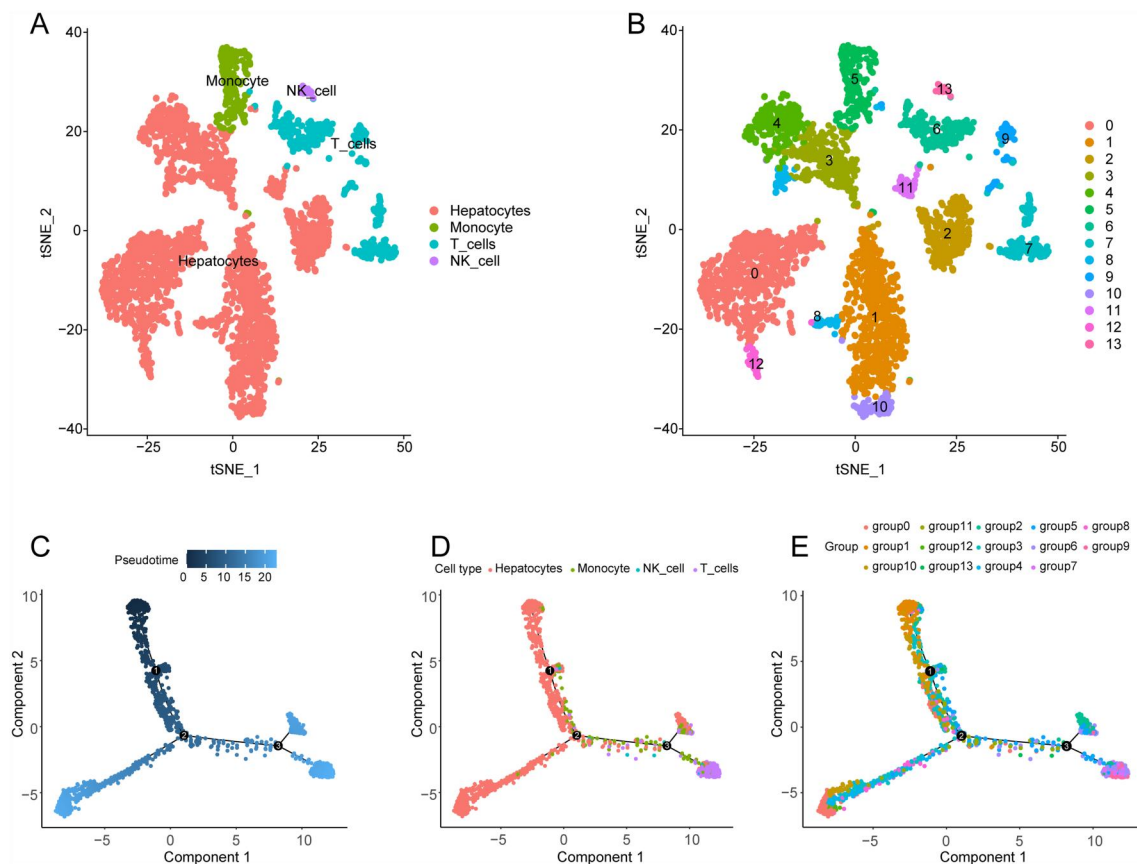


FIGURE 2 Overview of single-cell analysis in HCC tissues. (a) TSNE clustered all single tumour cells into four main types: hepatocytes, T cells, NK cells, and monocytes. (b) TSNE clustered all single tumour cells into 14 subgroups. (c-e) Pseudo-time analysis revealed the differentiation trajectory of all single cells in HCC, with each colour-coded for pseudo-time, cell types and subgroups. TSNE, t-distributed stochastic neighbour embedding.

cells are constantly evolving during tumour progression. However, the differences in functionality between the various divisions of single-cell subtypes remain unclear.

3.2 | T cells tend to be exhausted in the TME in HCC

As we all know, T cells play an essential role in tumour progression. Since we found that T cells were divided into various subtypes (Figure 2a–b), we isolated and analysed them again. Through the QC process, 354 single T cells matched the QC standards (Figure S2A–C). Based on the known markers mentioned in the CellMarker database, these T cells were divided into four subgroups (Figure 3a; Figure S2D). We found that the expression of PDCD1, an important inhibitory immune checkpoint, was significantly increased in group 3 (Figure 3b). Therefore, we suspected that the T cells of group 3 might be involved in exhausted T cells. Subsequently, we investigated the crucial immune checkpoints in the four cell groups (Figure 3c). Notably, we found that some vital inhibitory checkpoints were also upregulated in group 3, such as CTLA4, TIGIT, JAK1, LAG3, YTHDF1, LDHA, and IFNG. Since these genes are the markers of T cell exhaustion, the result indicated that T

cells from group 3 were exhausted in the TME. Then, we used monocle to perform pseudo-time analysis to infer the co-expression module and visualise the cell differentiation trajectory (Figure 3d–f). We found that the various subtypes of T cells gradually appeared with the advance of time, and the T cells from group 3, which represent the exhausted T cells, were distributed at the end of branch 2. These results suggested that in TME of HCC, Exhausted T cells gradually appeared with tumour progression.

3.3 | Functional enrichment analysis

In order to understand how the function of exhausted T cells changes during tumour progression, we performed a series of functional enrichment analyses by the DEGs of branch 2 in T cells differentiation, which contained the exhausted T cells. The results of GO analysis contained cadherin binding, cell adhesion molecule binding, and actin binding (Figure 4a). The results of KEGG analysis contained regulation of actin cytoskeleton, proteasome, and cell cycle (Figure 4b). In addition, the results of the Bioplanet model of Enrichr mainly contained cell cycle and T cell receptor regulation of apoptosis (Figure 4c). The results of Clue GO's immunoassay module in cytoscape mainly contained positive regulation of leucocyte

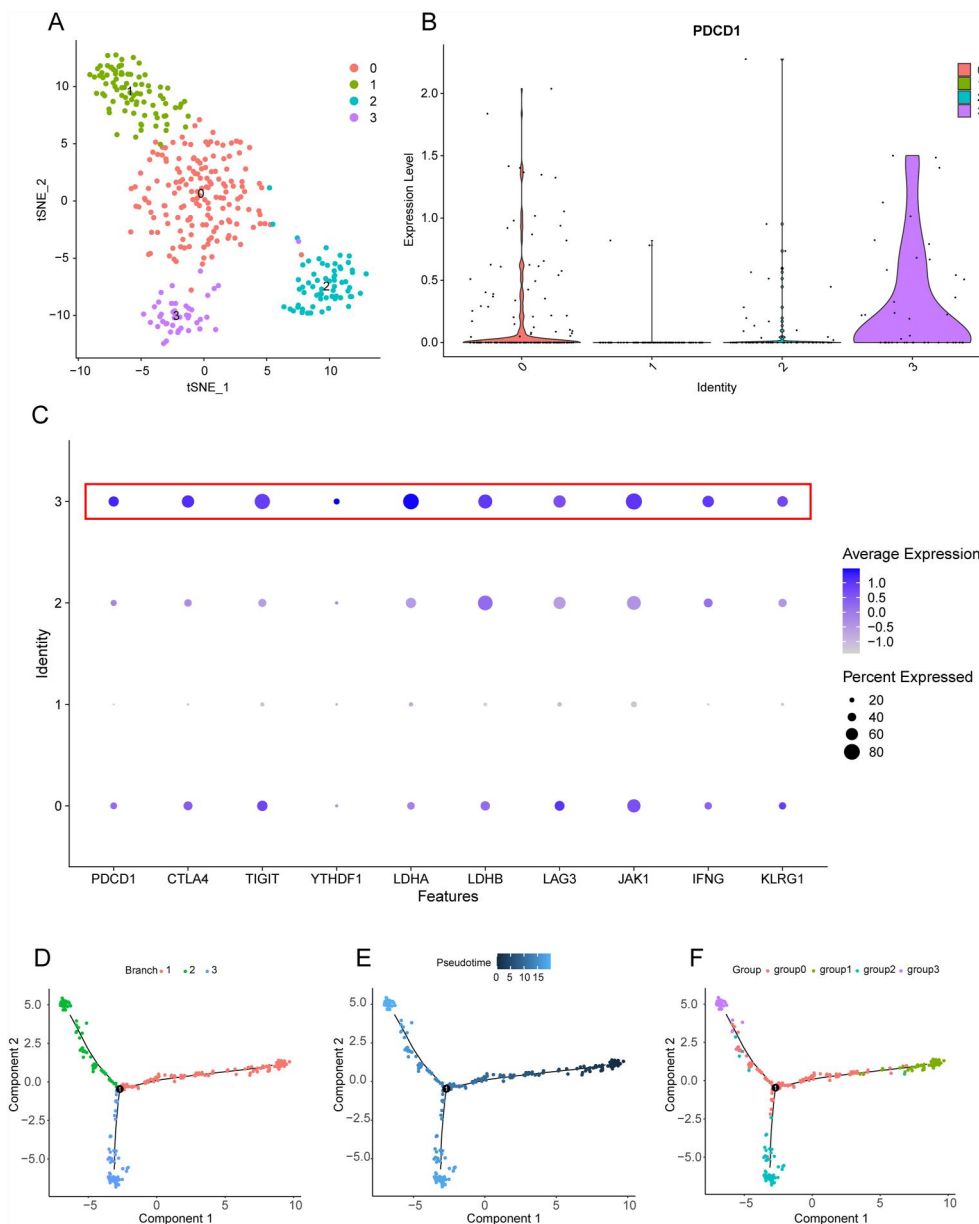


FIGURE 3 T cells tend to be exhausted in the TME of HCC. (a) tSNE clustered T cells into four groups. (b) Violin plot showed that PDCD1 expression was significantly increased in group 3. (c) Bubble plot revealed that the vital inhibitory checkpoints were upregulated in group 3. The red square represents group 3. (d-f) Pseudo-time analysis revealed the differentiation trajectory of T cells in HCC, with each colour-coded for branches, pseudo-time and subgroups. The T cells from group 3 were distributed at the end of branch 2.

chemotaxis, positive regulation of neutrophil chemotaxis, and antigen processing and presentation of exogenous peptide antigen via MHC class I (Figure 4d). These results shed light on the potential biological behaviours that T cells may experience during differentiation in tumour progression. During the tumour progression, antigen presentation and immune function were gradually enhanced with the increase of tumour-associated antigens. However, with the increase of tumour mutation burden, T cells were gradually overwhelmed, leading to the emergence of exhausted T cells, and finally, T cell apoptosis receptors were gradually activated, increasing T cell apoptosis.

3.4 | Exhausted T cells are associated with poor survival in HCC patients

To further investigate the relationship between various subtypes of T cells and outcomes of HCC patients, we clustered the HCC patients in the ICGC database using T cell evolution-associated genes. The patients fitted into three clusters since the results showed that when the number of sets is three, the difference between clusters is obvious, and the difference within clusters is inconspicuous (Figure 5a). K-M survival analysis revealed statistically significant differences among the three clusters (Figure 5b). The analysis suggested that the

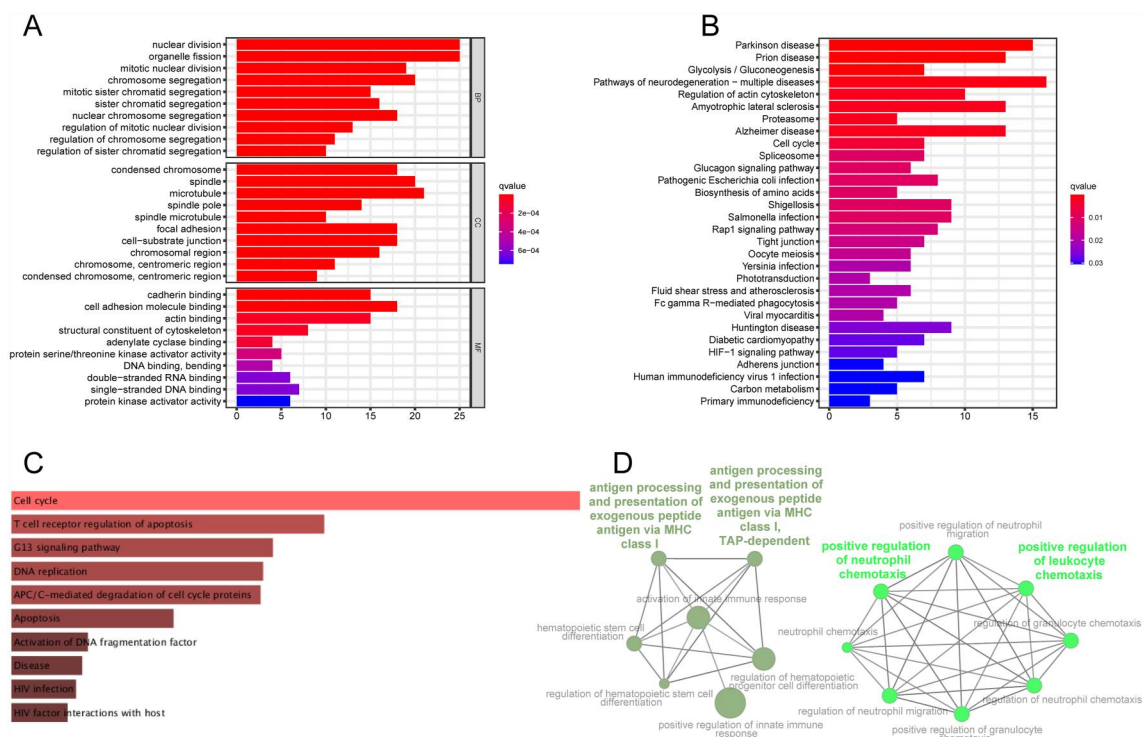


FIGURE 4 Functional enrichment analysis of branch 2. (a) GO enrichment analysis revealed the primary biological function and processes (BP), cellular component (CC), and molecular function (MF). (b) KEGG enrichment analysis revealed the primary pathways. (c) The Bioplanet model of Enrichr showed the primary pathways. (d) Clue GO's immunoassay module revealed the main pathways of immunity.

patients from cluster 1 (C1) had a worse prognosis than cluster 2 (C2) and cluster 3 (C3), and C2 had the best prognosis. We further analysed the TME score of the three clusters. Using an analysis of variance (ANOVA), we validated the differences in the TME scores between the clusters. The result showed that the ESTIMATE score, Immune score, and Stromal score of C2 were the highest, and the tumour purity of C2 was the lowest (Figure 5c–f). The variation in overall survival time across distinct clusters reflects the disparate prognosis associated with each cluster. As we all know, a favourable TME score predicts a better outcome. According to the C2's best prognosis, this result of C2's TME score analysis was consistent with our expectations. Interestingly, the ESTIMATE score, Immune score, and Stromal score of C1 were significantly higher than C3, and the tumour purity of C1 was significantly less than C3 (Figure S3A–D). However, the prognosis of C1 was significantly worse than C3 (Figure 5g). In order to further explore the reasons behind this unusual phenomenon, we conducted survival analysis on ESTIMATE score, Immune score, Stromal score, and tumour purity. We found that the Stromal score was significantly related to prognosis, and a higher Stromal score presents a better prognosis (Figure 5h). Although there was no statistical difference in Immune score, ESTIMATE score, and tumour purity, it could be seen that the prognosis with high Immune and ESTIMATE score was better than that with a low score, and the prognosis with low tumour purity was better than that with high tumour purity (Figure S3E–G). In addition, CD8A is a crucial molecule of T cell activation signal transduction and is considered a

favourable prognosis predictor. Likewise, we found that high expression of CD8A was significantly associated with the favourable prognosis of HCC patients in the ICGC database (Figure S3I). However, we found that CD8A was significantly higher in C1 than in C3 (Figure S3H). Therefore, we suspected that other vital factors are affecting patient outcomes of C1. Then, we further analysed the immune checkpoints of C1 and C3 (Figure 5j), and it can be seen that many critical inhibitory immune checkpoints, such as PDCD1, TIGIT, CTLA4, and HAVCR2, were higher in C1 than C3, indicating that there were more exhausted T cells in C1 than in C3, which may be the reason for the poor prognosis of C1. Besides, through GSEA analysis, we found that the CTLA4 pathway and PD 1 signalling were significantly enriched in C1 (Figure 5j), further suggesting a more accumulation of exhausted T cells in C1 than in C3.

3.5 | Searching the core genes in T cell evolution and building the prognostic model

We design the TCGA database as a train set and the ICGC database as a test set. In the TCGA database, we enrolled WGCNA to process the T cell evolution-associated genes and constructed some modules associated with clinical features through a scale-free co-expression network (Figure S4). A total of four modules were generated (Figure 6a). Both blue and turquoise modules were significantly associated with tumour grade and stage. Hence, we chose the genes of blue and

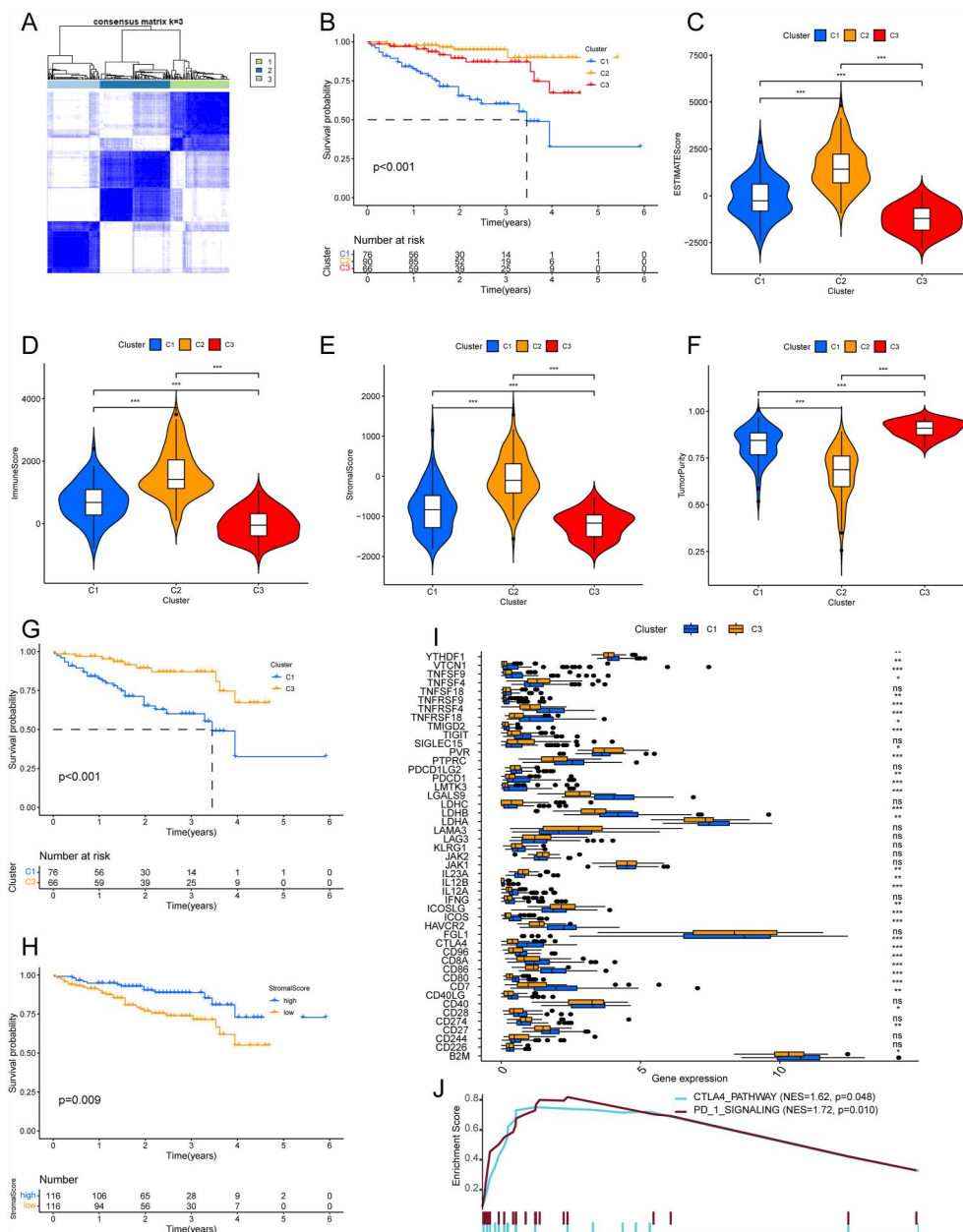


FIGURE 5 Exhausted T cells are associated with poor survival in HCC patients from the ICGC database. (a) The consensus matrix showed that patients fitted into three clusters through T cell evolution-associated genes. (b) K-M survival curve revealed that the three clusters' prognosis were significantly different. (c-f) Violin plot of tumour micro-environment (TME) score analysis showed that the three clusters' ESTIMATE score, Immune score, Stromal score, and Tumour purity were significantly different. (g) K-M survival curve revealed that the prognosis of patients in cluster 1 was significantly worse than that of cluster 3. (h) K-M survival curve revealed that the prognosis of patients with a high Stromal score was significantly better than patients with a low Stromal score. (i) The bar graph revealed that many critical inhibitory immune checkpoints were higher in cluster 1 than in cluster 3. (j) GSEA analysis displayed that the CTLA4 pathway and PD 1 signaling were significantly enriched in cluster 1. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

turquoise modules for further analysis. Then, using the Wilcoxon Test, we screened out the DEGs in tumour tissue and adjacent tissue from the two modules (Figure 6b-c). By univariate COX analysis, we further identified the genes associated with prognosis (Figure 6d). To further validate the accuracy of our data, we verified these genes using an additional independent single-cell RNA-seq dataset of HCC (GSE151530). As anticipated, we observed that many of these genes were highly expressed in exhausted T cells, reinforcing the reliability of our

initial results (Figure S5). Subsequently, according to Lasso regression analysis, we finally found 19 genes. Then, using the multivariate Cox analysis, the coefficients of the 19 genes were calculated (Table S2). Finally, a T cell evolution signature (TCESig) was established based on the 19 genes' expression level and coefficients to calculate each patient's risk. We found that there was a significant association between the TCESig and the clinical stage in the TCGA database (Table S3) and in the ICGC database (Table S1). Then, we used the risk score of

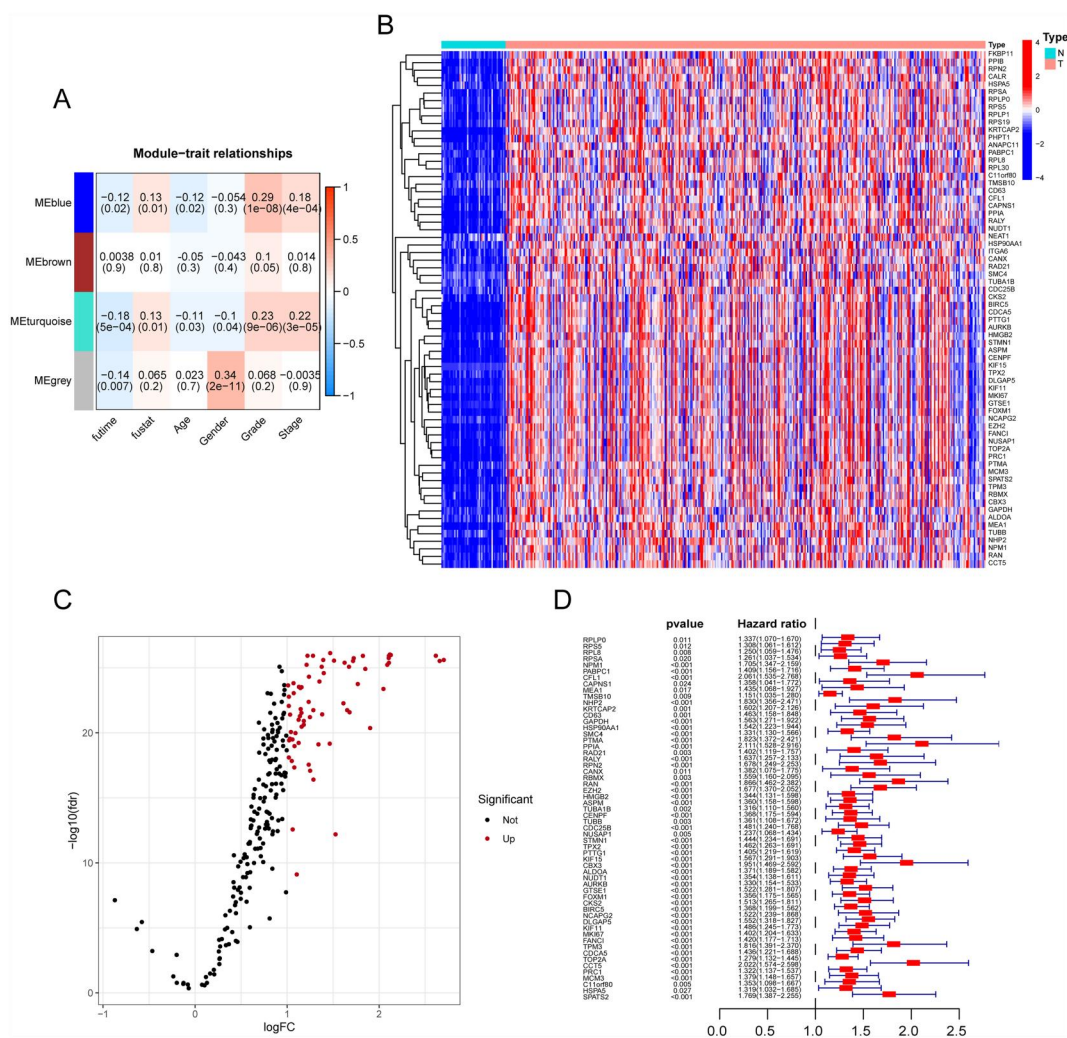


FIGURE 6 Searching the essential genes in T cell evolution and building the prognostic model in the TCGA database. (a) WGCNA analysis showed that four modules were generated, and blue and turquoise modules were significantly associated with tumour grade and stage. (b-c) Heatmap and volcano plot showed the DEGs in tumour tissue and adjacent tissue from blue and turquoise modules. (d) Univariate COX analysis further identified the genes potentially associated with OS.

TCESig to predict prognosis in both TCGA and ICGC databases. The distribution of risk scores and survival time suggested that patients' survival time decreased as the TCESig increased (Figure 7a-c; Figure 7f-h). In the TCGA database, the time-dependent ROC of TCESig showed that the AUC of 1-, 3-, and 5-year were 0.818, 0.775, and 0.761, respectively (Figure 7d). In the ICGC database, the time-dependent ROC of TCESig showed that the AUC of 1-, 3-, and 5-year were 0.695, 0.718, and 0.278, respectively (Figure 7i). K-M cumulative curves demonstrated that the overall survival time of patients with the high-risk score was significantly shorter than that of patients with low-risk score (Figure 7e; Figure 7j). In addition, compared with clinicopathology, univariate and multivariate Cox analysis showed that the TCESig was an independent prognostic predictor in the TCGA database (Figure 8a-b) and in the ICGC database (Figure 8c-d). These results suggested that the TCESig may be a reliable prognostic indicator in HCC.

3.6 | Validation of the relationship between the TCESig and exhausted T cells

We hypothesised that the TCESig is closely related to exhausted T cells. Hence, we analysed the expression level of the TCESig's genes in the four subtypes of T cells. The results showed that the TCESig's genes were highly expressed in group 3, which represent the exhausted T cells (Figure 9a), suggesting a close relationship between TCESig and exhausted T cells. In addition, through TSNE analysis, we identified the location of exhausted T cells (Figure 3a) and found that TCESig's genes were mainly highly expressed in the region of exhausted T cells (Figure 9b-c). Immune cell infiltration analysis showed that most immune cell infiltrations did not differ statistically between high and low-risk groups (Figure 9d). TME score analysis also showed no significant difference between the high and low-risk groups (Figure S6A-D). Then we analysed the relationship

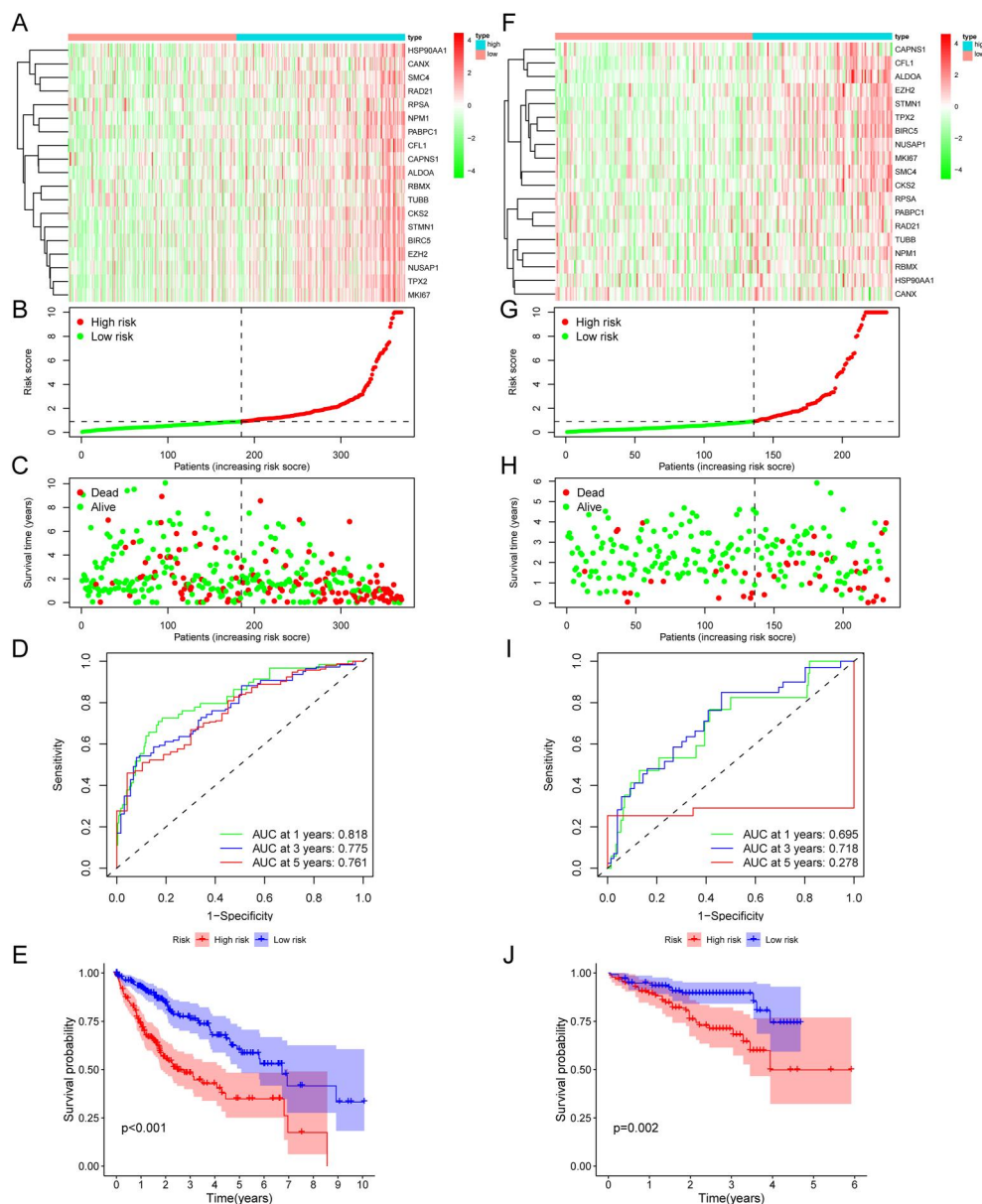
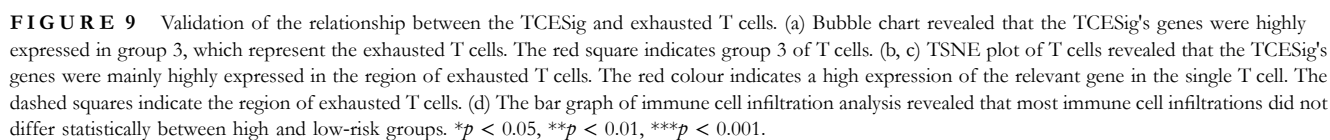
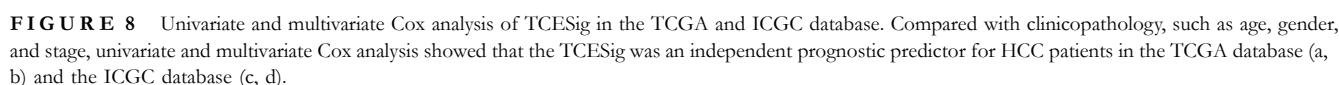


FIGURE 7 survival analysis of TCESig in the TCGA and ICGC database. Heatmap, distribution of risk scores, survival time and status revealed that the patients' survival time decreased as the TCESig increased in the TCGA database (a-c) and the ICGC database (f-h). The dashed lines present the cutoff value, which divided HCC patients into low-risk and high-risk groups. Time-dependent ROC curve showed that the AUC of TCESig of 1-, 3-, and 5-year were 0.818, 0.775, 0.761 in the TCGA database (d) and were 0.695, 0.718, 0.278 in the ICGC database (i). The K-M survival curve revealed that the prognosis of patients with high risk was significantly worse than those with low risk in the TCGA database (e) and ICGC database (j).

between the TCESig and the vital inhibitory immune checkpoints. The vital inhibitory immune checkpoints refer to essential molecular mechanisms that play a crucial role in regulating the immune response. These checkpoints act as negative regulators or “brakes” to prevent excessive or unintended activation of immune cells, ensuring the maintenance of self-tolerance and preventing autoimmunity. By inhibiting immune cell activation, these checkpoints can also contribute to immune evasion by cancer cells. We conducted a Wilcoxon Test to analyse the differences between high and low-risk groups and used Spearman's correlation to

investigate the association between patient risk scores and the expression of inhibitory immune checkpoints. The results showed that the vital inhibitory immune checkpoints were almost significantly higher in the high-risk group than in the low-risk group (Figure 10a–g) and positively correlated with the risk scores (Figure 10h–n). These results indicated that the TCESig was also an accurate indicator for evaluating the exhausted T cells' level in hepatocellular carcinoma. These results suggested that the genes in TCESig that led to poor outcomes may be primarily associated with exhausted T cells, not immune cell infiltration.



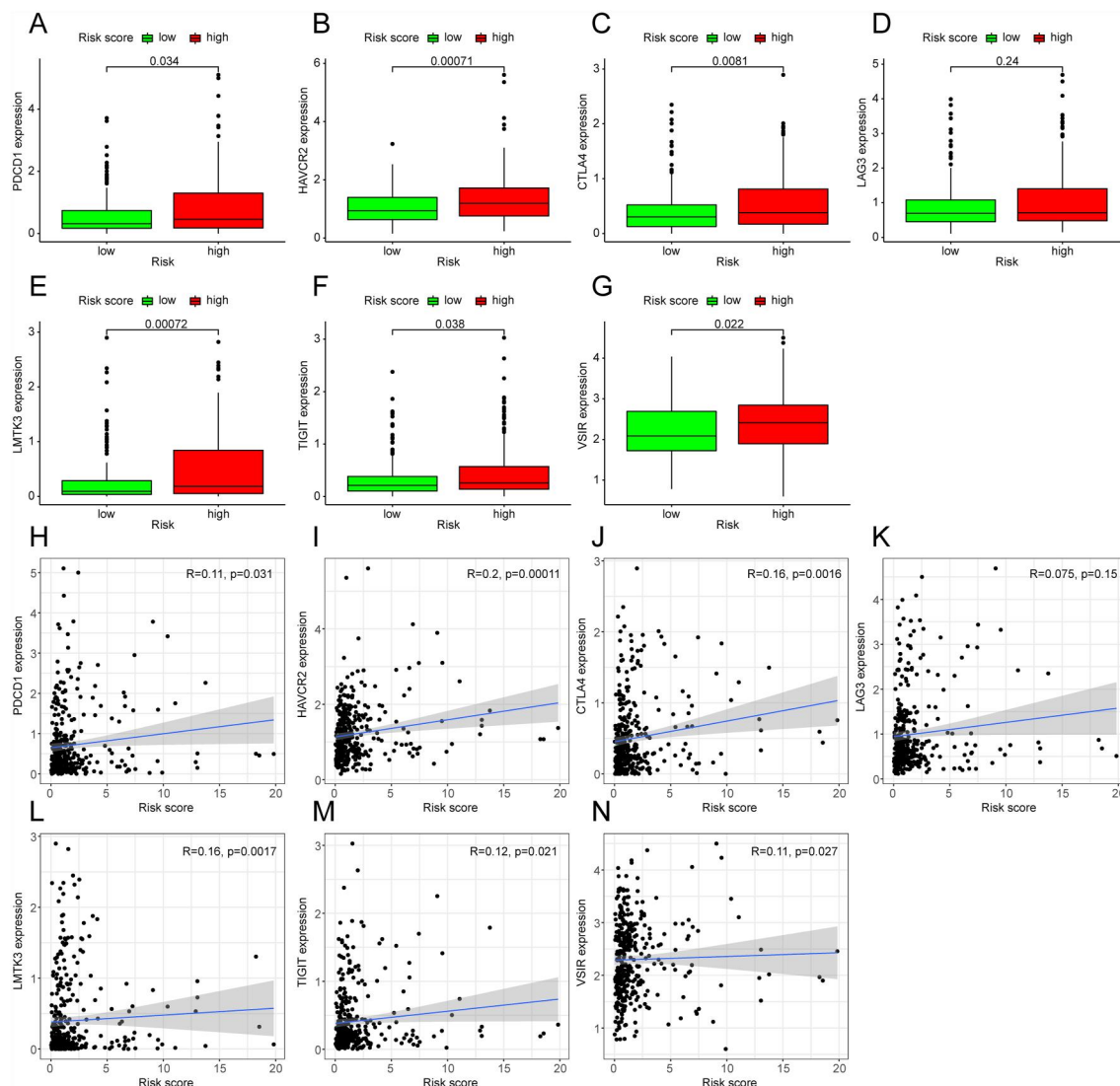


FIGURE 10 The relationship between the TCESig and the vital inhibitory immune checkpoints. (a-g) Bar graphs indicated that except for LAG3, almost vital inhibitory immune checkpoints were significantly higher in the high-risk group than in the low-risk group. (h-n) Correlation analysis revealed that except for LAG3, almost vital inhibitory immune checkpoints also positively correlated with the risk scores of TCESig.

3.7 | Exploration of the TCESig' function and prognostic value

To further validate the relationship between TCESig and exhausted T cells, we illustrated several exhausted T cells-related gene sets via GSEA analysis. The analysis showed that the exhausted T cells-related gene sets were significantly enriched in the high-risk score group of TCESig (Figure 11a). Consistent with the previous survival analysis result (Figure 5b), an ANOVA analysis of the ICGC database revealed that the risk-score in C1 was significantly higher than in C2 and C3 (Figure 11b). CD8⁺ T cells represent the primary subtype of cytotoxic T cells. CD8A, as a marker for CD8⁺ T cells, provides a reliable measure of the quantity of cytotoxic T cells. Because the combination of CD8A expression and TCESig may better reflect the degree of effective cytotoxic T cell deficiency in HCC patients, we performed survival analysis

combining CD8A expression and the TCESig. K-M survival analysis revealed CD8A Low/High risk group, which may have the least effective cytotoxic T cells, and has a significantly worse prognosis with a median survival rate of fewer than 2 years in the TCGA database (Figure 11c). These results further verified the relationship between the TCESig and exhausted T cells and revealed the potential of TCESig in predicting patient outcomes.

4 | DISCUSSION

With the advancement of cell sequencing technology, more information regarding the differentiation map and transcriptional heterogeneity has been revealed [44]. Here, we found that exhausted T cells are gradually present during HCC progression and may be associated with poor patient

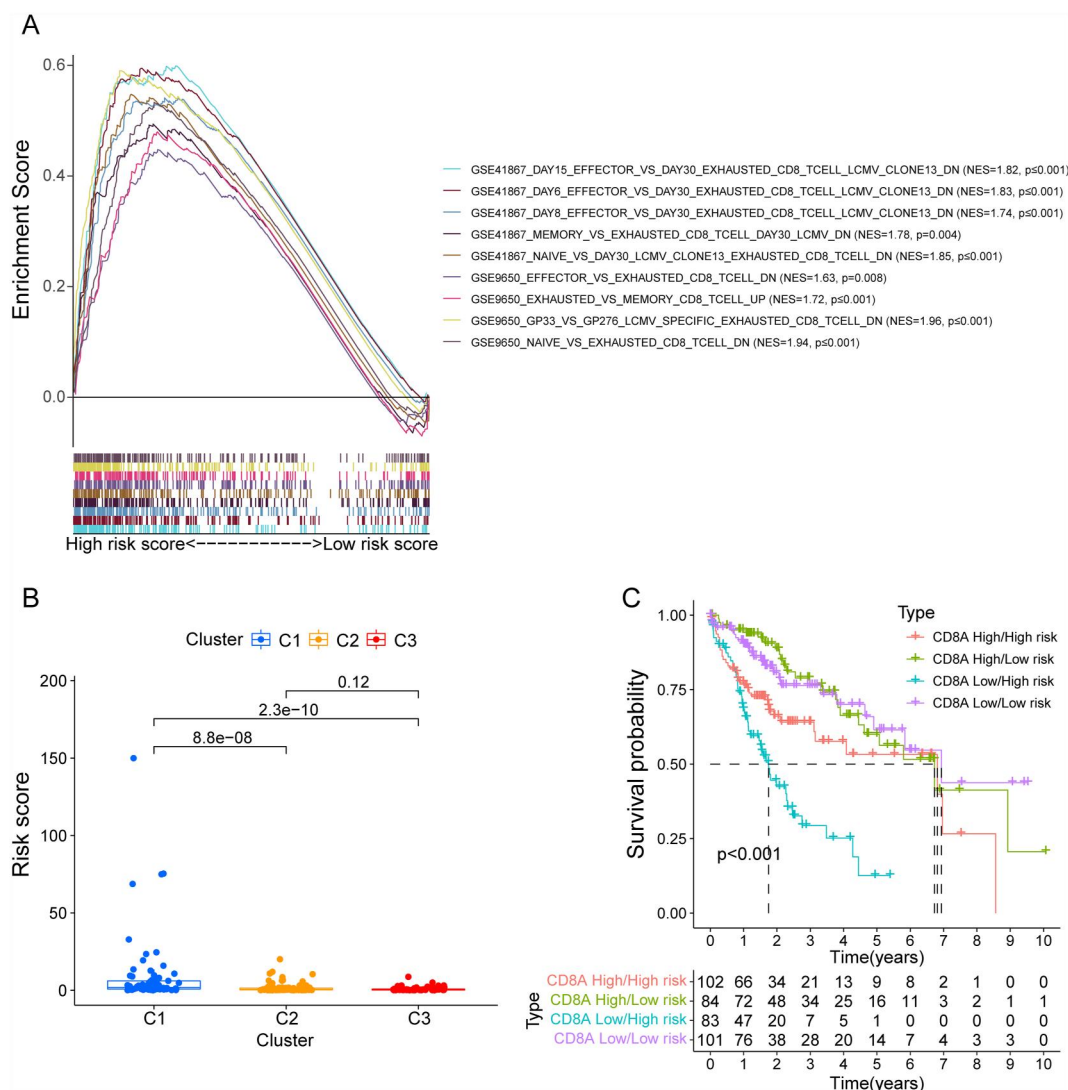


FIGURE 11 Exploration of the TCESig function and prognostic value. (a) GSEA analysis suggested that TCESig was associated with the gene set of exhausted T cells elevation. (b) The risk score of TCESig in C1 was significantly higher than C2 and C3 in the ICGC database. (c) K-M survival analysis revealed CD8A Low/High risk group has a significantly worse prognosis with a median survival rate of fewer than 2 years in the TCGA database.

outcomes through single-cell RNA-seq. We collected single-cell and bulk RNA-seq data from public databases and analysed them using various computational biology methods. We investigated the heterogeneity of tumour and characterised the exhausted T cells in HCC. The presence of exhausted T cells was shown to be strongly linked with poor clinical outcomes. Additionally, we developed a novel prognostic model using critical genes involved in T cell evolution. The model provides a consistent predictor of prognosis for patients with HCC. Interestingly, all model genes are significantly expressed in exhausted T cells, further indicating that exhausted T cells are associated with a poor prognosis for HCC patients. At last, we revealed that the model might also be an accurate indicator for evaluating the exhausted T cells' level. Collectively, this study provides a novel viewpoint on assessing the outcome of HCC patients from an exhausted T cell perspective, which may aid physicians in developing treatment approaches.

Through TSNE and pseudo-time plot, we found that the tumour heterogeneity gradually increased with HCC progression, and various subtypes of immune cells gradually appeared. The results indicated that tumour cells and related immune cells are constantly evolving during tumour progression. These findings corroborated with recent research that demonstrated the utility of single-cell data analysis in dissecting tissue into infiltration and tumour cells, elucidating the critical roles of humoral immunity infiltration in the landscape of HCC progress [45]. The pseudo-time analysis provided the trajectory of T cell differentiation to the exhaustion phenotype in HCC and revealed the genes and related pathways involved in this process. Blocking these genes or pathways may provide the possibility to interrupt T cell dysfunction and restore T cell activity and may contribute to an opportunity for finding new immunotherapies for HCC. We found that exhausted T cells gradually appeared during tumour progression and performed different gene-expression profiling. These results agree with

previous findings which showed that accumulation of exhausted T cells increased in tumour progression, and expression of various inhibitory immune checkpoints was enhanced in exhausted T cells [46]. Likewise, Zheng, Bo et al. indicated that tumour-associated T cells performed a robust function in HCC, and the existence of T cells with high PD-1 was confirmed by multiplex immunofluorescence tissue staining, and its enrichment showed definite prognostic value [47]. These results contribute to our understanding of the dynamic framework of T cells' function in HCC and may help develop rational cancer immunotherapies.

The GO and KEGG analysis explored the primary function and pathways of the evolutionary branch of exhausted T cells, which contains cadherin binding, proteasome, and cell cycle. These findings may reveal the potential functional changes during the T cells differentiation. Consistent with the literature, Banh, Cindy et al. confirmed that endogenous cadherin can inhibit T cell functions [48]. These results also corroborate the ideas of Widjaja, Christella E et al., who indicated that proteasome activity plays a critical role in determining the fate of T cells [49]. In addition, He, Tiansheng et al. suggested that proteasome activity regulates T cell activation and apoptosis [50]. The analysis of the Bioplanet model of Enrichr mainly contained cell cycle and T cell receptor regulation of apoptosis, suggesting that the cell cycle of T cells in exhausted T branch was significantly changed, and the cell receptors regulating apoptosis were elevated. The results of Clue GO's immunoassay module mainly contained positive regulation of leucocyte chemotaxis, positive regulation of neutrophil chemotaxis, and antigen processing and presentation of exogenous peptide antigen via MHC class I, revealing the progress that during the tumour advance, with the addition of tumour-associated antigens, antigen presentation and immunological response are steadily increased. However, as the tumour mutation burden increased, T cells became increasingly overwhelmed, resulting in the appearance of exhausted T cells. Finally, T cell apoptosis receptors were gradually activated, resulting in increased T cell apoptosis. These results are consistent with the commonly accepted definition of exhausted T cells [23], confirming the accuracy of our analysis.

All HCC patients in the ICGC database were divided into three clusters through the T cell evolution-associated genes. Comparing prognosis and TME score in each cluster, we were surprised to find a contradiction in these results that C1 had a worse prognosis than C3, while C1 had a better score of TME and a higher CD8A level than C3. Many studies have proved that a low TME score or CD8A level is associated with poor prognosis in cancer patients [51–54]. Likewise, we also found that patients with a high TME score or a high CD8A level have a better prognosis than patients with a low TME score or a low CD8A level in the ICGC database. Through detailed analysis, numerous crucial inhibitory immunological checkpoints were discovered to be higher in C1 than in C3, and the CTLA4 pathway and PD-1 signaling were significantly enriched in C1. Many studies have proved that exhausted T cells are significantly associated with the poor prognosis of patients in a variety of tumours [55–57]. Recently, Barsch, Maryam et al.

reported that in HCC patients, exhausted CD8⁺ T cells indicate a worse prognosis than tissue-resident memory T cells [58]. Likewise, Hung, Man Hsin et al. also reported that exhausted T cells are related to the poor prognosis of HCC patients, and the exhausted T cells are driven by tumour methionine metabolism [59]. These findings revealed that C1 had a higher proportion of exhausted T cells than C3, which may account for C1's poor prognosis. Our data are consistent with the report of Yang, Yanying et al. that the accumulation of exhausted T cells leads to a worse prognosis in HCC patients [46]. Besides, consistency with other studies ensured the reliability of subsequent analyses.

Through WGCNA analysis, differential expression analysis and COX analysis in the TCGA database, we found 19 core genes in T cell evolution and constructed a prognostic model as TCESig. We design the TCGA database as a train set and the ICGC database as a test set. TCGA and ICGC are two independent datasets without any common samples between them. The result of survival analysis showed that TCESig was a robust prognostic indicator in both training and test sets. In addition, our model's 1-, 3-, and 5-year AUC values were more accurate than the model of Zhang, Fapeng et al. built by eight TME-related genes in HCC [60]. Recently, Qin, Ge et al. demonstrated that NPM1, a gene of our model, selectively binds to the PD-L1 promoter in triple-negative breast cancer cells and activates its transcription, thus suppressing the T cell function *in vitro* and *in vivo* [61]. Ternet, Nicola et al. reported that CFL1's potential as a target in breast cancer immunotherapy should be further investigated [62]. Song, Kwon Ho et al. found that HSP90AA1 inhibition sensitised immune-refractory tumours to adoptive T cell transfer and PD-1 blocking, hence re-energising tumour-reactive T cells' immunological cycle [63]. Zhao, Ende et al. reported that EZH2, a methyltransferase, activated the Notch pathway by inhibiting the Notch repressors Numb and Fbxw7 by trimethylation of histone H3 at Lys27, which resulted in increased polyfunctional cytokine production and survival of T cells via Bcl-2 signaling. Furthermore, he suggested that EZH2 (+)CD8 (+) T cells were correlated with increased survival of patients [64]. Although the results of Zhao, Ende et al. differ from the study of Zhou, Liye et al. that EZH2 may be a therapeutic target since it has been shown to increase tumour cell antigen presentation and sensitize refractory tumours to anti-PD-1 treatment [65]. They are consistent with that EZH2 is an important gene for patients' outcomes through affecting the function of T cells. However, the relationship between some other genes in TCESig and T cells has not been fully discovered yet. These genes may also be involved in the differentiation progress of exhausted T cells. Further analysis of these genes may provide new targets for cancer immunotherapy.

Our results showed that the TCESig's genes were highly expressed in subgroup 3, which represents the exhausted T cells, indicating that the genes of TCESig are closely related to exhausted T cells. Tumour-infiltrating cytotoxic T cells can inhibit tumour development specifically. However, they often go into a condition of “exhaustion” or “dysfunction” [66]. Hence, in recent years, immuno-checkpoint blocking treatment

to correct T cell malfunction and exhaustion has gotten much interest [67]. Unfortunately, only a tiny percentage of cancer patients benefit from it in clinical trials [68]. Only 15%–20% of individuals with HCC responded to PD1 blocking [69]. One possible explanation is that anti-PD1 therapy alone only helps the less exhausted T cells [70], indicating that inhibiting a single immune checkpoint is inefficient in practice. Therefore, it is essential to identify more potential therapeutic targets and conduct comprehensive evaluation and treatment. The genes in our model may provide a new chance to get some novel therapeutic targets.

Our results showed that the TCESig was also an accurate indicator for evaluating the exhausted T cells' level in hepatocellular carcinoma. As we all know, the therapeutic window has been further widened by immunotherapy, including nivolumab alone or in combination with ipilimumab or pembrolizumab. Immunotherapy based on the use of immune checkpoint inhibitors, as single agents or in combination with kinase inhibitors, anti-angiogenic drugs, chemotherapeutic agents, and locoregional therapies, offers great promise in the treatment of HCC paralleling a lack of patients' selection criteria [71, 72]. Previous study indicated that preexisting exhausted T cells may have limited the ability to reinvigorate post-therapy [73]. The response of T cells to checkpoint blockade is from a specific subset of T cell clones that recently entered the tumour, and T cells clonal replacement was shown to be more common in exhausted T cells [73]. Given that our model performed well in predicting the accumulation of exhausted T cells and considering that exhausted T cells often show an inadequate response to checkpoint blockade, The model might suggest potential avenues for those involving the use of immune checkpoint inhibitors in HCC patients.

There are certain flaws in the current study that should be mentioned. The results of this study may have an extraordinary divergence due to the limited number of patients included in the analysis. Another vulnerability source assumes that we should manage more functional experiments to indicate the potential molecular mechanisms for predicting the effect of the genes in TCESig.

5 | CONCLUSIONS

In conclusion, we used single-cell RNA-seq and bulk RNA-seq to show that exhausted T cells are present during HCC progression and are correlated with poor patient outcomes. We investigated the heterogeneity of tumour and characterised the exhausted T cells in HCC. What is more, we created a new prediction model based on crucial genes involved in T cell differentiation. For patients with HCC, the model is a reliable predictor of prognosis and accumulation of exhausted T cells. This work offers a unique perspective on analysing the outcome of HCC patients from the standpoint of exhausted T cells, which may help physicians create treatment strategies and provide new therapeutic targets for further research.

AUTHOR CONTRIBUTION

Xiaolong Tang contributed to the conceptualization, data curation, formal analysis, software, visualization, and writing—original draft. Yandong Miao and Wuhua Ha performed the validation of this study. Yandong Miao wrote the writing—review & editing. Lixia Yang and Denghai Mi conducted the resources and supervision. Zheng Li contributed to the funding acquisition, investigation, and methodology. Denghai Mi conducted the project administration.

ACKNOWLEDGEMENTS

We thank the GEO, TCGA, ICGC, Ensembl, David, KEGG, and CellMarker databases for the availability of the data. This work was supported by the Special Plan for Condition Construction of [Gansu Provincial Scientific Research Institutes] (Grants No. 20JR10RA432) and the China Postdoctoral Science Foundation (Grant No. 2019M663860).

CONFLICT OF INTEREST STATEMENT

No potential conflicts of interest were disclosed.

DATA AVAILABILITY STATEMENT

The data used to support the findings of this study can be found in Single-cell RNA-seq data and were collected from the GEO database (GSE146115 and GSE151530) (<https://www.ncbi.nlm.nih.gov/geo/>). Bulk RNA-seq data and clinical data were collected from TCGA (<https://portal.gdc.cancer.gov/v29.0>) and ICGC (<https://dcc.icgc.org/>) databases.

ORCID

Denghai Mi  <https://orcid.org/0000-0002-8643-4496>

REFERENCES

1. Sung, H., et al.: Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA A Cancer J. Clin.* 71(3), 209–249 (2021). <https://doi.org/10.3322/caac.21660>
2. Kulik, L., El-Serag, H.B.: 'Epidemiology and management of hepatocellular carcinoma. *Gastroenterology* 156(2), 477–491.e1 (2019). <https://doi.org/10.1053/j.gastro.2018.08.065>
3. Wang, Y., et al.: Hepatic resection versus transarterial chemoembolization in infiltrative hepatocellular carcinoma: a multicenter study. *J. Gastroenterol. Hepatol.* 35(12), 2220–2228 (2020). <https://doi.org/10.1111/jgh.15060>
4. Samuel, D., Coilly, A.: Management of patients with liver diseases on the waiting list for transplantation: a major impact to the success of liver transplantation. *BMC Med.* 16(1), 113 (2018). <https://doi.org/10.1186/s12916-018-1110-y>
5. Kim, N., et al.: Stereotactic body radiation therapy vs. radiofrequency ablation in Asian patients with hepatocellular carcinoma. *J. Hepatol.* 73(1), 121–129 (2020). <https://doi.org/10.1016/j.jhep.2020.03.005>
6. Xu, F., et al.: Immune checkpoint therapy in liver cancer. *J. Exp. Clin. Cancer Res.* 37(1), 110 (2018). <https://doi.org/10.1186/s13046-018-0777-4>
7. Gordan, J.D., et al.: Systemic therapy for advanced hepatocellular carcinoma: ASCO guideline. *J. Clin. Oncol.* 38(36), 4317–4345 (2020). <https://doi.org/10.1200/jco.20.02672>
8. Wei, L., et al.: 'Genome-wide CRISPR/Cas9 library screening identified PHGDH as a critical driver for Sorafenib resistance in HCC. *Nat. Commun.* 10(1), 4681 (2019). <https://doi.org/10.1038/s41467-019-12606-7>

9. Tsilimigras, D.I., et al.: Hepatocellular carcinoma tumour burden score to stratify prognosis after resection. *Br. J. Surg.* 107(7), 854–864 (2020). <https://doi.org/10.1002/bjs.11464>
10. Khan, A.R., Wei, X., Xu, X.: Portal vein tumor thrombosis and hepatocellular carcinoma - the changing tides. *J. Hepatocell. Carcinoma* 8, 1089–1115 (2021). <https://doi.org/10.2147/jhc.s318070>
11. Dagogo-Jack, I., Shaw, A.T.: Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* 15(2), 81–94 (2018). <https://doi.org/10.1038/nrclinonc.2017.166>
12. Feo, F., Pascale, R.M.: Multifocal hepatocellular carcinoma: intrahepatic metastasis or multicentric carcinogenesis? *Ann. Transl. Med.* 3(1), 4 (2015)
13. Zhang, Q., et al.: Integrated multiomic analysis reveals comprehensive tumour heterogeneity and novel immunophenotypic classification in hepatocellular carcinomas. *Gut* 68(11), 2019–2031 (2019). <https://doi.org/10.1136/gutjnl-2019-318912>
14. Liu, F., et al.: Identification of FABP5 as an immunometabolic marker in human hepatocellular carcinoma. *J. Immunother. Cancer* 8(2), e000501 (2020). <https://doi.org/10.1136/jitc-2019-000501>
15. Patel, A.P., et al.: ‘Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344(6190), 1396–1401 (2014). <https://doi.org/10.1126/science.1254257>
16. Papalex, E., Satija, R.: Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat. Rev. Immunol.* 18(1), 35–45 (2018). <https://doi.org/10.1038/nri.2017.76>
17. Roma-Rodrigues, C., et al.: Targeting tumor microenvironment for cancer therapy. *Int. J. Mol. Sci.* 20(4), E840 (2019). <https://doi.org/10.3390/ijms20040840>
18. O'Donnell, J.S., Teng, M.W.L., Smyth, M.J.: Cancer immunoediting and resistance to T cell-based immunotherapy. *Nat. Rev. Clin. Oncol.* 16(3), 151–167 (2019). <https://doi.org/10.1038/s41571-018-0142-8>
19. Alfei, F., et al.: ‘TOX reinforces the phenotype and longevity of exhausted T cells in chronic viral infection. *Nature* 571(7764), 265–269 (2019). <https://doi.org/10.1038/s41586-019-1326-9>
20. Zeng, Z., Wei, F., Ren, X.: Exhausted T cells and epigenetic status. *Cancer Biol Med* 17(4), 923–936 (2020). <https://doi.org/10.20892/j.issn.2095-3941.2020.0338>
21. Kallies, A., Zehn, D., Utzschneider, D.T.: Precursor exhausted T cells: key to successful immunotherapy? *Nat. Rev. Immunol.* 20(2), 128–136 (2020). <https://doi.org/10.1038/s41577-019-0223-7>
22. Hsu, C.-L., et al.: Exploring markers of exhausted CD8 T cells to predict response to immune checkpoint inhibitor therapy for hepatocellular carcinoma. *Liver Cancer* 10(4), 346–359 (2021). <https://doi.org/10.1159/000515305>
23. Blank, C.U., et al.: ‘Defining “T cell exhaustion. *Nat. Rev. Immunol.* 19(11), 665–674 (2019). <https://doi.org/10.1038/s41577-019-0221-9>
24. Seo, H., et al.: BATF and IRF4 cooperate to counter exhaustion in tumor-infiltrating CAR T cells. *Nat. Immunol.* 22(8), 983–995 (2021). <https://doi.org/10.1038/s41590-021-00964-8>
25. Grebinoski, S., et al.: Autoreactive CD8+ T cells are restrained by an exhaustion-like program that is maintained by LAG3. *Nat. Immunol.* 23(6), 868–877 (2022). <https://doi.org/10.1038/s41590-022-01210-5>
26. Su, X., et al.: Clonal evolution in liver cancer at single-cell and single-variant resolution. *J. Hematol. Oncol.* 14(1), 22 (2021). <https://doi.org/10.1186/s13045-021-01036-y>
27. Ma, L., et al.: Single-cell atlas of tumor cell evolution in response to therapy in hepatocellular carcinoma and intrahepatic cholangiocarcinoma. *J. Hepatol.* 75(6), 1397–1408 (2021). <https://doi.org/10.1016/j.jhep.2021.06.028>
28. Butler, A., et al.: Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* 36(5), 411–420 (2018). <https://doi.org/10.1038/nbt.4096>
29. Kobak, D., Berens, P.: ‘The art of using t-SNE for single-cell transcriptomics. *Nat. Commun.* 10(1), 5416 (2019). <https://doi.org/10.1038/s41467-019-13056-x>
30. Maaten, L.V.D., Hinton, G.: ‘Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2625 (2008)
31. Zhang, X., et al.: CellMarker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res.* 47(D1), D721–D728 (2019). <https://doi.org/10.1093/nar/gky900>
32. Qiu, X., et al.: ‘Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* 14(10), 979–982 (2017). <https://doi.org/10.1038/nmeth.4402>
33. Walter, W., Sánchez-Cabo, F., Ricote, M.: ‘GOpilot: an R package for visually combining expression data with functional analysis. *Bioinformatics* 31(17), 2912–2914 (2015). <https://doi.org/10.1093/bioinformatics/btv300>
34. Kanehisa, M., et al.: KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45(D1), D353–D361 (2017). <https://doi.org/10.1093/nar/gkw1092>
35. Kuleshov, M.V., et al.: Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 44(W1), W90–97 (2016). <https://doi.org/10.1093/nar/gkw377>
36. Bindea, G., et al.: ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25(8), 1091–1093 (2009). <https://doi.org/10.1093/bioinformatics/btp101>
37. Wilkerson, M.D., Hayes, D.N.: ‘ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26(12), 1572–1573 (2010). <https://doi.org/10.1093/bioinformatics/btq170>
38. Blaser, N., et al.: Gems: an R package for simulating from disease progression models. *J. Stat. Software* 64(10), 1–22 (2015). <https://doi.org/10.18637/jss.v064.i10>
39. Ritchie, M.E., et al.: Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43(7), e47 (2015). <https://doi.org/10.1093/nar/gkv007>
40. Yoshihara, K., et al.: Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4(1), 2612 (2013). <https://doi.org/10.1038/ncomms3612>
41. Langfelder, P., Horvath, S.: ‘WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* 9(1), 559 (2008). <https://doi.org/10.1186/1471-2105-9-559>
42. Blanche, P., Dartigues, J.-F., Jacqmin-Gadda, H.: Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat. Med.* 32(30), 5381–5397 (2013). <https://doi.org/10.1002/sim.5958>
43. Subramanian, A., et al.: GSEA-P: a desktop application for gene set enrichment analysis. *Bioinformatics*, 23, (23), pp. 3251–3253, (2007). <https://doi.org/10.1093/bioinformatics/btm369>
44. Kiselev, V.Y., Andrews, T.S., Hemberg, M.: Challenges in unsupervised clustering of single-cell RNA-seq data. *Nat. Rev. Genet.* 20(5), 273–282 (2019). <https://doi.org/10.1038/s41576-018-0088-9>
45. Zhang, S., et al.: Single-cell RNA-seq analysis reveals microenvironmental infiltration of plasma cells and hepatocytic prognostic markers in HCC with cirrhosis. *Front. Oncol.* 10, 596318 (2020). <https://doi.org/10.3389/fonc.2020.596318>
46. Yang, Y., et al.: Analysis of single-cell RNAseq identifies transitional states of T cells associated with hepatocellular carcinoma. *Clin. Transl. Med.* 10(3), e133 (2020). <https://doi.org/10.1002/ctm2.133>
47. Zheng, B., et al.: Trajectory and functional analysis of PD-1high CD4+CD8+ T cells in hepatocellular carcinoma by single-cell cytometry and transcriptome sequencing. *Adv. Sci.* 7(13), 2000224 (2020). <https://doi.org/10.1002/advs.202000224>
48. Banh, C., Fugère, C., Brossay, L.: ‘Immunoregulatory functions of KLRG1 cadherin interactions are dependent on forward and reverse signaling. *Blood* 114(26), 5299–5306 (2009). <https://doi.org/10.1182/blood-2009-06-228353>
49. Widjaja, C.E., et al.: Proteasome activity regulates CD8+ T lymphocyte metabolism and fate specification. *J. Clin. Invest.* 127(10), 3609–3623 (2017). <https://doi.org/10.1172/jci90895>
50. He, T.-S., et al.: ALG-2 couples T cell activation and apoptosis by regulating proteasome activity and influencing MCL1 stability. *Cell Death Dis.* 11(1), 5 (2020). <https://doi.org/10.1038/s41419-019-2199-4>

51. Zeng, D., et al.: Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures. *Cancer Immunol Res* 7(5), 737–750 (2019). <https://doi.org/10.1158/2326-6066.cir-18-0436>
52. Li, Y., et al.: Identification of new therapeutic targets for gastric cancer with bioinformatics. *Front. Genet.* 11, 865 (2020). <https://doi.org/10.3389/fgene.2020.00865>
53. Ling, B., et al.: Microenvironment analysis of prognosis and molecular signature of immune-related genes in lung adenocarcinoma. *Oncol. Res.* 28(6), 561–578 (2021). <https://doi.org/10.3727/096504020x15907428281601>
54. Xu, D., et al.: Identification of immune subtypes and prognosis of hepatocellular carcinoma based on immune checkpoint gene expression profile. *Biomed. Pharmacother.* 126, 109903 (2020). <https://doi.org/10.1016/j.biopha.2020.109903>
55. Wagner, J., et al.: A single-cell atlas of the tumor and immune ecosystem of human breast cancer. *Cell* 177(5), 1330–1345.e18 (2019). <https://doi.org/10.1016/j.cell.2019.03.005>
56. Liu, F., et al.: Heterogeneity of exhausted T cells in the tumor microenvironment is linked to patient survival following resection in hepatocellular carcinoma. *OncoImmunology* 9(1), 1746573 (2020). <https://doi.org/10.1080/2162402x.2020.1746573>
57. Nakano, M., et al.: PD-1+ TIM-3+ T cells in malignant ascites predict prognosis of gastrointestinal cancer. *Cancer Sci.* 109(9), 2986–2992 (2018). <https://doi.org/10.1111/cas.13723>
58. Barsch, M., et al.: T-cell exhaustion and residency dynamics inform clinical outcomes in hepatocellular carcinoma. *J. Hepatol.* 77(2), 397–409 (2022). <https://doi.org/10.1016/j.jhep.2022.02.032>
59. Hung, M.H., et al.: Tumor methionine metabolism drives T-cell exhaustion in hepatocellular carcinoma. *Nat. Commun.* 12(1), 1455 (2021). <https://doi.org/10.1038/s41467-021-21804-1>
60. Zhang, F.-P., et al.: Construction of a risk score prognosis model based on hepatocellular carcinoma microenvironment. *World J. Gastroenterol.* 26(2), 134–153 (2020). <https://doi.org/10.3748/wjgv26.i2.134>
61. Qin, G., et al.: NPM1 upregulates the transcription of PD-L1 and suppresses T cell activity in triple-negative breast cancer. *Nat. Commun.* 11(1), 1669 (2020). <https://doi.org/10.1038/s41467-020-15364-z>
62. Ternette, N., et al.: Immunopeptidomic profiling of HLA-A2-positive triple negative breast cancer identifies potential immunotherapy target antigens. *Proteomics* 18(12), e1700465 (2018). <https://doi.org/10.1002/pmic.201700465>
63. Song, K.-H., et al.: HSP90A inhibition promotes anti-tumor immunity by reversing multi-modal resistance and stem-like property of immune-refractory tumors. *Nat. Commun.* 11(1), 562 (2020). <https://doi.org/10.1038/s41467-019-14259-y>
64. Zhao, E., et al.: Cancer mediates effector T cell dysfunction by targeting microRNAs and EZH2 via glycolysis restriction. *Nat. Immunol.* 17(1), 95–103 (2016). <https://doi.org/10.1038/ni.3313>
65. Zhou, L., et al.: Targeting EZH2 enhances antigen presentation, anti-tumor immunity, and circumvents anti-PD-1 resistance in head and neck cancer. *Clin. Cancer Res.* 26(1), 290–300 (2020). <https://doi.org/10.1158/1078-0432.ccr-19-1351>
66. Ma, J., et al.: PD1Hi CD8+ T cells correlate with exhausted signature and poor clinical outcome in hepatocellular carcinoma. *J Immunother Cancer* 7(1), 331 (2019). <https://doi.org/10.1186/s40425-019-0814-7>
67. Nishino, M., et al.: Monitoring immune-checkpoint blockade: response evaluation and biomarker development. *Nat. Rev. Clin. Oncol.* 14(11), 655–668 (2017). <https://doi.org/10.1038/nrclinonc.2017.88>
68. Chan, T.A., et al.: Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann. Oncol.* 30(1), 44–56 (2019). <https://doi.org/10.1093/annonc/mdy495>
69. El-Khoueiry, A.B., et al.: 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 (2017). [https://doi.org/10.1016/s0140-6736\(17\)31046-2](https://doi.org/10.1016/s0140-6736(17)31046-2)
70. Blackburn, S.D., et al.: Selective expansion of a subset of exhausted CD8 T cells by alphaPD-L1 blockade. *Proc. Natl. Acad. Sci. U. S. A.* 105(39), 15016–15021 (2008). <https://doi.org/10.1073/pnas.0801497105>
71. Solimando, A.G., et al.: Second-line Treatments for Advanced Hepatocellular Carcinoma: A Systematic Review and Bayesian Network Meta-analysis. *Clin Exp Med* (2021)
72. Leone, P., et al.: The evolving role of immune checkpoint inhibitors in hepatocellular carcinoma treatment. *Vaccines* 9(5), 532 (2021). <https://doi.org/10.3390/vaccines9050532>
73. Yost, K.E., et al.: Clonal replacement of tumor-specific T cells following PD-1 blockade. *Nat. Med.* 25(8), 1251–1259 (2019). <https://doi.org/10.1038/s41591-019-0522-3>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Tang, X., et al.: Single-cell RNA-seq and bulk RNA-seq explore the prognostic value of exhausted T cells in hepatocellular carcinoma. *IET Syst. Biol.* 17(4), 228–244 (2023). <https://doi.org/10.1049/syb2.12072>