In-depth characterization of the biomarkers based on tumor-infiltrated immune cells reveals implications for diagnosis and prognosis in hepatocellular carcinoma

Wenbiao Chen, Kefan Bi, Xujun Zhang, Jingjing Jiang, Hongyan Diao [*]

State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Center for Infectious Disease, Collaborative InnovatEion Center for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, 310003, China

A R T I C L E   I N F O

A B S T R A C T

Hepatocellular carcinoma (HCC) is an immune-related tumor, that the type and number of tumor-infiltrated immune cells can serve as biomarkers for the clinical application. In this study, we constructed the immune model for diagnostic and prognostic prediction of HCC based on the systematic bioinformatics analyses on the component of immune cells from large samples transcriptome. CIBERSORT analysis found that the component of immune cells between 513 HCC and 473 adjacent normal tissues was different. M0 macrophages and regulatory T cells were mainly enriched in tumor tissues, whereas the $CD8^{+}$ T cell and activated $CD4^{+}$ memory T cells were the most in normal tissues. Using random forest and LASSO analyses, eleven immune cell types were mined out to construct the immune diagnostic model (IDG), which showed high efficiency in distinguishing cancer from normal tissues both in testing and validation groups. In addition, the immune prognostic model (IPG) consisting of five types of immune cells was constructed using the LASSO-Cox algorithm. It showed that HCC patients of the high-risk group had a significantly shorter survival time than those of low-risk group in testing, validation, and entire cohorts. Besides, Nomogram plots and decision curve analyses revealed that the IPG was positively associated with the HCC clinical classification of the Barcelona Clinic Liver Cancer (BCLC) stage, and showing more accuracy of prediction than independent BCLC stage. Related analyses found that IDG positively correlated with epithelial-mesenchymal transition (EMT) and cytotoxic factor-related genes and negatively correlated with immune checkpoint regulators related genes. From the GSEA analysis of the biological function of genes related to IPG, it was found that the genes of the high-risk group were enriched in some tumorigenesis related pathways, such as DNA replication, cell cycle, and PPARA. Therefore, this study identified IDG and IPG as efficient biomarkers for the diagnosis and prognosis of HCC.

1. Introduction

Hepatocellular carcinoma (HCC) is one of the malignant tumors threatening human health. There has been an increasing trend in the incidences of HCC worldwide. It ranks sixth and third for the incidence and mortality of malignant tumors, respectively [1]. Current treatment strategies of HCC include surgical resection, chemoradiotherapy, and immunotherapy. However, since the insidious onset and the lack of early diagnostic and prognostic biomarkers, more than 80% of HCC patients are generally in the middle and late stage at the time of treatment. This results in poor prognosis and loss of the optimal treatment opportunity [2]. Therefore, efficient diagnostic and prognostic specific biomarkers are critical to improving the clinical management of HCC.

The liver is a vital immune organ with unique immune cells that helps in the maintenance of the normal immune function of the human body. Numerous studies have shown that the immune dysregulation of the liver, such as immune surveillance escape, immune microenvironment alteration, and immunosuppression, is an important mechanism for the development and distant metastasis of HCC [3]. Recent studies have found infiltrated immune cells as an important part of the immune microenvironment that influences the occurrence of HCC [4]. Moreover, it was previously revealed that tumor-infiltrating immune cells are closely related to the prognosis of HCC and the target of immunotherapy [5]. Lessons learned from these studies suggest that different types of immune cell infiltration could be novel potential biomarkers for the diagnosis and prognosis of HCC.

Infiltration of various cell types mainly made up of B-cell, plasma-cell, and T-cell, in tumor tissue were characterized for better prognosis [6]. On the contrary, the tumor tissues enriched with high infiltration of regulatory T cells (Tregs), macrophages, and mast cells could have a poor prognosis [7]. Using the HCC mouse model to analyze the tumor-associated antigens, Liu et al. found that CD8$^+$ T cells exhibited exhaustion-like phenotypes, while macrophages showed accumulation during the progression of HCC. Their findings indicated immune checkpoint inhibitors and immunotherapies based on infiltrating immune cells for HCC [8,9]. However, the traditional methods, such as flow cytometry and immunohistochemistry, can neither assess entirely different immune effector cell types nor clearly distinguish between groups of closely related cells. This is mainly due to the limited number of immune markers that current techniques can simultaneously measure. Furthermore, the targets of immune cells in small sample experiments lack specificity and are not representative enough to be used to guide the clinical application [10]. As an alternative, bioinformatics analysis based on big data transcriptome has made it possible for a comprehensive investigation of the tumor-infiltrated immune cells and mining out of the immune-related biomarkers for tumors.

To improve the diagnosis and prognosis of HCC, the present study has characterized the infiltrated immune cells composition of tumor tissues from gene expression profiles using the Estimating Relative Subsets of RNA Transcripts (CIBERSORT) algorithm. The random-forest analysis, least absolute shrinkage, and selection (LASSO) algorithm, logistic regression analyses were applied to found out the immune-related biomarkers. The immune diagnostic model (IDG) and immune prognostic model (IPG) that provided the potential biomarkers for the diagnosis and prognosis of HCC patients were then constructed. Further analyses of bioinformatics associations revealed the IPG was related to clinical characteristics and molecular subtypes. The research flow chart is presented in Fig. 1.

2. Materials and methods

2.1. Dataset acquisition

The data used in this study were obtained from the Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/) public database. It contains high-throughput gene expression data of all kinds of diseases submitted by research institutions around the world [11]. To avoid experimental bias, we downloaded HCC data from four different research institutions, including GSE20140 (GSE20140-GPL5474, GSE20140-GPL8432, GSE20140-GPL18461), GSE10186, GSE54236, and GSE76427. The four datasets selected had prognostic information of the HCC tissues and a sample size greater than 50. To obtain more accurate results, the preliminarily processed data were divided into three groups: testing, validation, and entire cohorts, for experimental verification. The chip probe was then mapped to the human gene symbol using the R Bioconductor package. A total of 80 tumor tissues in GSE20140 and GSE10186 datasets had duplicates. Ultimately, we obtained 986 tissues that included 513 tumors and 473 normal tissues (Supplementary Table 1). The dataset GSE76427 had 115 tumor and 52 normal tissues containing the clinical information, such as age and gender. Barcelona Clinic Liver Cancer (BCLC) clinicopathological stage was used for association analyses.

2.2. Estimation of infiltrated immune cells

CIBERSORT was used to assess the proportion of the infiltrated immune cells in the tumor and normal tissues. It is a method for transferring transcriptomics data to the immune cell component of complex tissues to analyze the profiling of immune cell fractions from bulk tissue. The standard genes expression data were uploaded to the CIBERSORT web portal. The LM22 genes signature and 1000 permutations were used to quantify the proportions of 22 subsets of infiltrated immune cells [12].
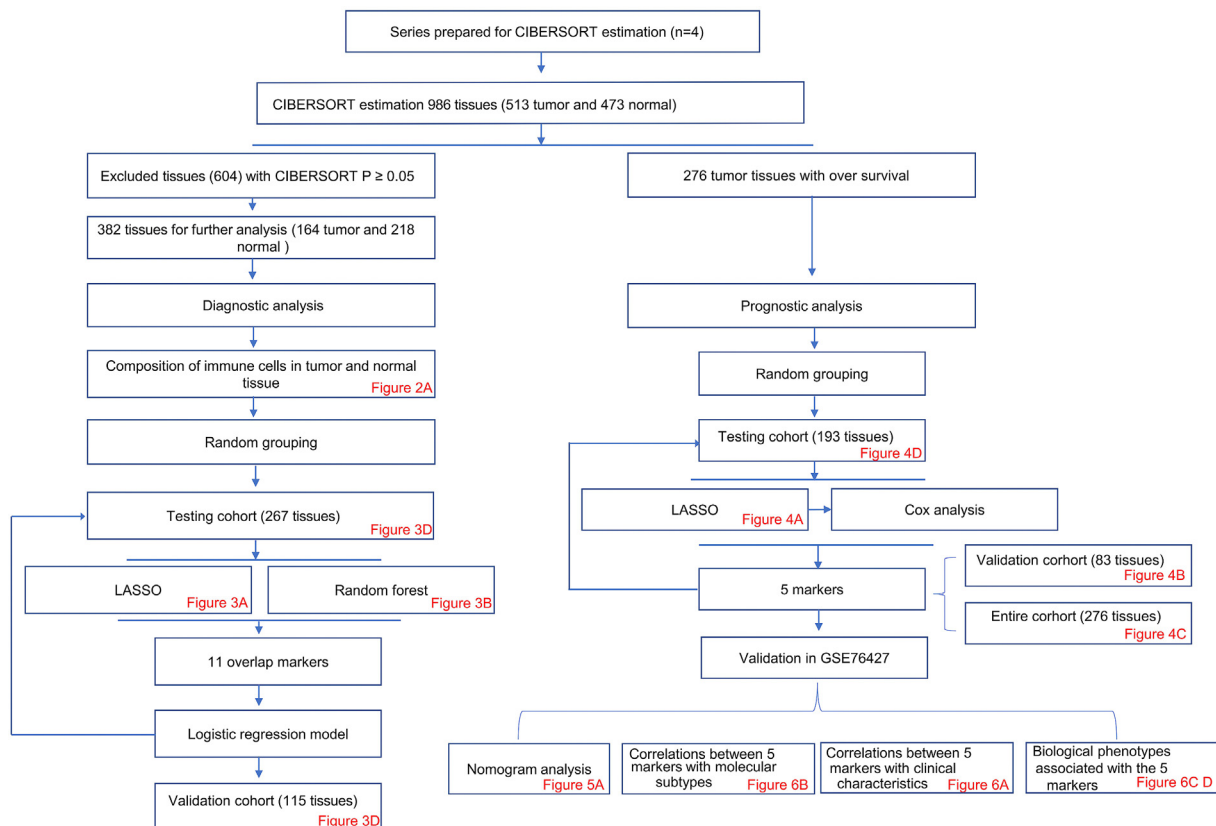

Fig. 1. Flow Chart of this study.

The infiltrated immune cells included T cells, B cells, macrophages, dendritic cells, NK cells, and myeloid subset cells. An empirically defined global *P*-value was determined to measure the confidence of deconvolution results of each tissue. The *P*-value of each tissue calculated by the CIBERSORT algorithm was <0.05, indicating that the population of the immune cell was correct. For each tissue, the value calculated by CIBERSORT was normalized, and immune cell type fractions summed up for direct comparison of cellular component across different phenotypes of immune cells and datasets.

2.3. Construction of the diagnostic and prognostic model

The tissues were randomly divided into testing and validation groups using the R Sample function package. The LASSO and random-forest analyses were conducted to select the suitable immune cell fractions. The overlapping markers were obtained between the two methods. The IDG was constructed based on the coefficients of each selected marker using a logistic regression algorithm. For the construction of IPG, we used the LASSO–Cox algorithm, an algorithm of estimation on parameters in high-dimensional models [13], to reduce the dimensionality and to select the most significant suitable immune cell fraction. The IDG was built based on the corresponding coefficients by using the Cox regression analysis.

2.4. Associated analysis and biological function on IPG

The relationship between IPG and gene signatures linked to tumor and immunity was analyzed by spearman's correlation test and the correlation diagram depicted using the Corrplot R package. The associations between the IPG and clinical characteristics were evaluated via Nomogram plots using the R package. The calibration curves were drawn to compare the ideal model and observed outcome. The decision curve was used to assess the discrimination of the IPG and BCLC stage. We conducted the Gene set enrichment analysis (GSEA) algorithm to analyze the biological functions of the genes using the JAVA program based on molecular signatures database of "c2. cp.kegg.v6.1. symbols" and "c2. cp.biocarta.v6.1. symbols".

2.5. Statistical analysis

All the statistical analyses in this study were performed using R software 3.5 and SPSS software 25. Student's *t*-test and Mann Whitney *U* test were used to make comparisons between groups for the normally distributed variables and nonnormally distributed variables, respectively. The Kaplan–Meier (KM) method was used to depict the survival curve to estimate the survival probability of patients. The statistical differences among means were tested using the log-rank test. The receiver operating characteristic (ROC) curves, generated by the pROC R package, were used to determine the sensitivity and specificity of KM analysis based on the area under the ROC curve (AUC). The violin plot obtained from the Vioplot R package was used to represent the different distribution of IPG among datasets. Combining with the clinical characteristic as concomitant variables, univariate and multivariate Cox regression analyses were undertaken to verified potentially prognostic value of IPG. All the statistical tests were two-tailed and statistical significance was defined as *P* value < 0.05.

3. Results

3.1. The difference of infiltrated immune cells between tumor and normal tissues

A total of 513 tumor and 473 normal tissues were subjected to CIBERSORT screening. Among these, 164 tumor and 218 normal tissues (P < 0.05) were used to investigate the composition of infiltrated immune cells in tumor and normal tissues using the CIBERSORT method.

The results showed a significant difference in the fractions of immune cells between tumor and normal tissues (Fig. 2A). The mast cells, macrophages, and T regs cells were the main component of tumor tissues, whereas the normal tissues consisted mainly of CD4$^+$ T and CD8$^+$ T cells. We also found that counts of the M0 macrophages and Tregs in the tumor tissues were consistently higher than in normal tissues. On the contrary, normal tissues had more CD8$^+$ T cells and activated CD4$^+$ memory T cells that tumor tissues.

The summary analysis of the distribution of infiltrated immune cells across the clinical characteristics in tumor tissues showed that plasma cells, M2 macrophages, activated CD4$^+$ memory T cells, M0 macrophages, and activated mast cells had the greatest enrichment (Fig. 2B). Even after regrouping the tumor tissues by clinical characteristics, such as the BCLC stage, the proportion of the five types of immune cells remained the same. This suggested their potential role for the biomarker of HCC. The results indicated a significant difference in the composition of infiltrated immune cells between tumor and normal tissues, suggesting the potential value of using infiltrated immune cells in the diagnostic and prognostic evaluation of HCC patients.

3.2. Construction of the immune diagnostic model

The 164 tumor and 218 normal tissues that satisfied the condition of P < 0.05 by CIBERSORT screening were randomly divided into testing (267 tissues) and validation (115 tissues) groups. We performed random forest and LASSO analyses to mine out the key biomarkers of infiltrated immune cells. Random forest analysis, an intelligent data mining algorithm [14], was applied to identify the most important immune cells that can be used to distinguish between tumors and normal tissues (Fig. 3A). A total of 14 immune cells that could clearly separate the tumor from the normal tissues well were obtained. (Supplementary Table 2). The LASSO analysis, a shrinkage estimation [15], was used to select the most significantly immune cells to differentiate tumor and normal tissues (Fig. 3B). It identified 15 types of immune cells with optimal values when lambda = 0.000001 (Supplementary Table 2). Eleven biomarkers that overlapped between the two analyses were retained (Supplementary Table 3). We then constructed the immune diagnostic model (IDG) based on the eleven biomarkers using a logistic regression algorithm. The selected immune cells were assessed as continuous variables in this model. When the IDG values were compared, tumor tissues had significantly higher values than normal tissues in testing, validation and entire sets of all datasets except GSE54236 (Fig. 3C). The ROC curves showed that the AUC value for the testing and validation sets were 0.905 and 0.852, respectively (Fig. 3D), indicating that IDG could be used to effectively distinguish the tumor from normal tissues.
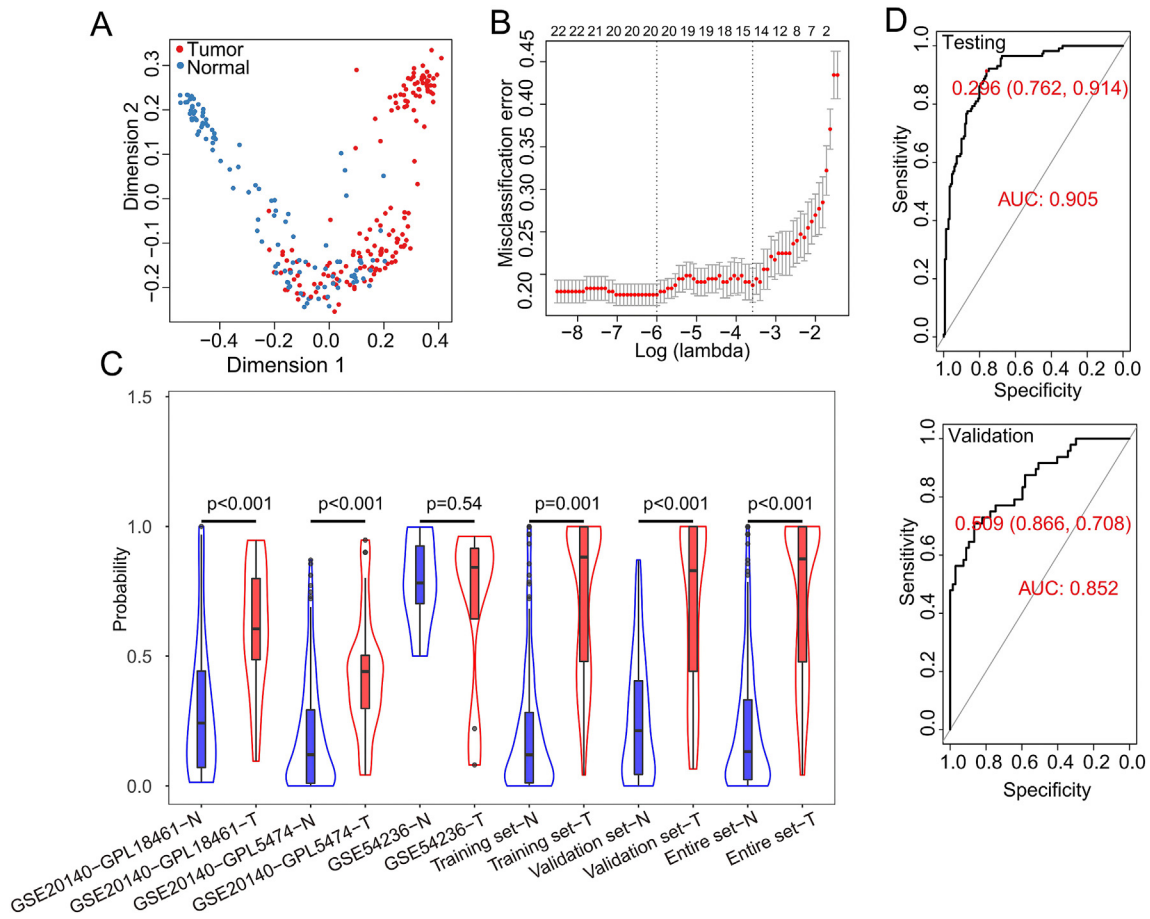
3.3. Construction of the immune prognostic model

A total of 276 tumor tissues that survival information and met the criteria for prognostic analysis were included in IPG construction. These tissues were separated into testing (193 patients) and validation (83 patients) groups. The LASSO-Cox algorithm was an algorithm of estimation on parameters in high-dimensional models [13], was performed to reduce the number of candidate immune cells and to select the most significant survival-associated immune cells (Fig. 4A). Five key biomarkers were identified for IPG construction (Supplementary Table 4). The IPG model was constructed through a Cox analysis in the testing groups. The AUC values were determined to assess the predictive ability of IPG (Supplementary Table 5). Based on the median value (0.411) of 5 key biomarkers of the testing cohort, patients were divided into high-risk and low-risk cohorts. Kaplan–Meier (KM) analyses showed that the prognosis of patients in the low-risk cohort was better than in the high-risk cohort (p < 0.0001) (Fig. 4B). Similar results were obtained in validation (p = 0.006), and entire (p < 0.0001) (Fig. 4C and D) cohorts.

To further verified the prognostic value of IPG, we drew the nomogram plots by combining the prognostic score of IPG on the GSE76427
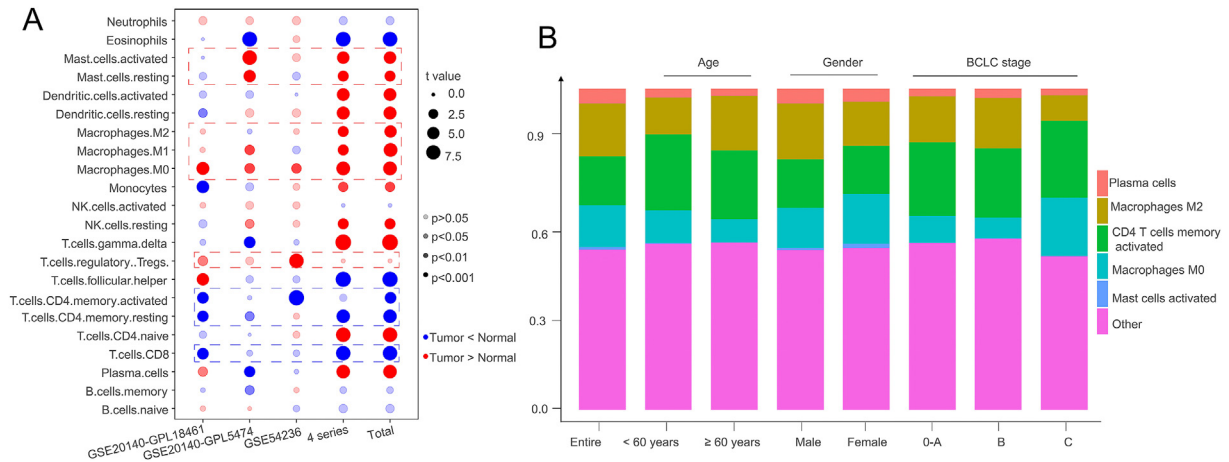
Fig. 2. The analysis of the composition of infiltrating immune cells. (A) Differences in immune cell fraction between tumor and normal control tissue. (B) The summary of immune cell fraction distributed across the clinical characteristics in tumor tissues. 4 series stand for 4 datasets, including GSE20140-GPL18461, GSE20140-GPL5474, GSE54236, and GSE76427; Total represents all the 6 datasets in this study; the red squares represent the key immune cells in the tumor tissues with higher proportions than normal tissues; the blue squares show the key immune cells in the normal tissues with higher proportions than tumor tissues.

Fig. 3. Construction of immune diagnostic model based on the infiltrating immune cells. (A) Multi-dimensional scale plot of an adjacent matrix using random forest analysis. (B) LASSO regression model used to reveal the misclassification error of different quantitative variables. The red dots and the grey lines show a misclassification error value and standard error (SE), respectively. The vertical dashed lines on the left and right represent the minimum criteria and 1-SE criteria, respectively. The lambda was randomly selected parameter. (C) Comparison of IDG value between normal and tumor tissues across different datasets. T: tumor tissues; N: normal tissues. (D) ROC curves of the IDG used to demonstrate the ability of IPG to differentiate between tumor and normal tissues.

dataset (Fig. 5A). The length of the line represented the degree of influence of different factors and the effect of different values of factors on the outcome. The nomogram plots showed that both the prognostic score of IPG and BCLC stage had a significant influence on the prognosis of HCC

patients. Moreover, the prognostic score of IPG contributed to the highest risk points (0–95) than other clinical characteristics. The nomogram plots used in our study performed relatively well when compared to the ideal model (Fig. 5B). Besides, the decision curve showed that the nomogram
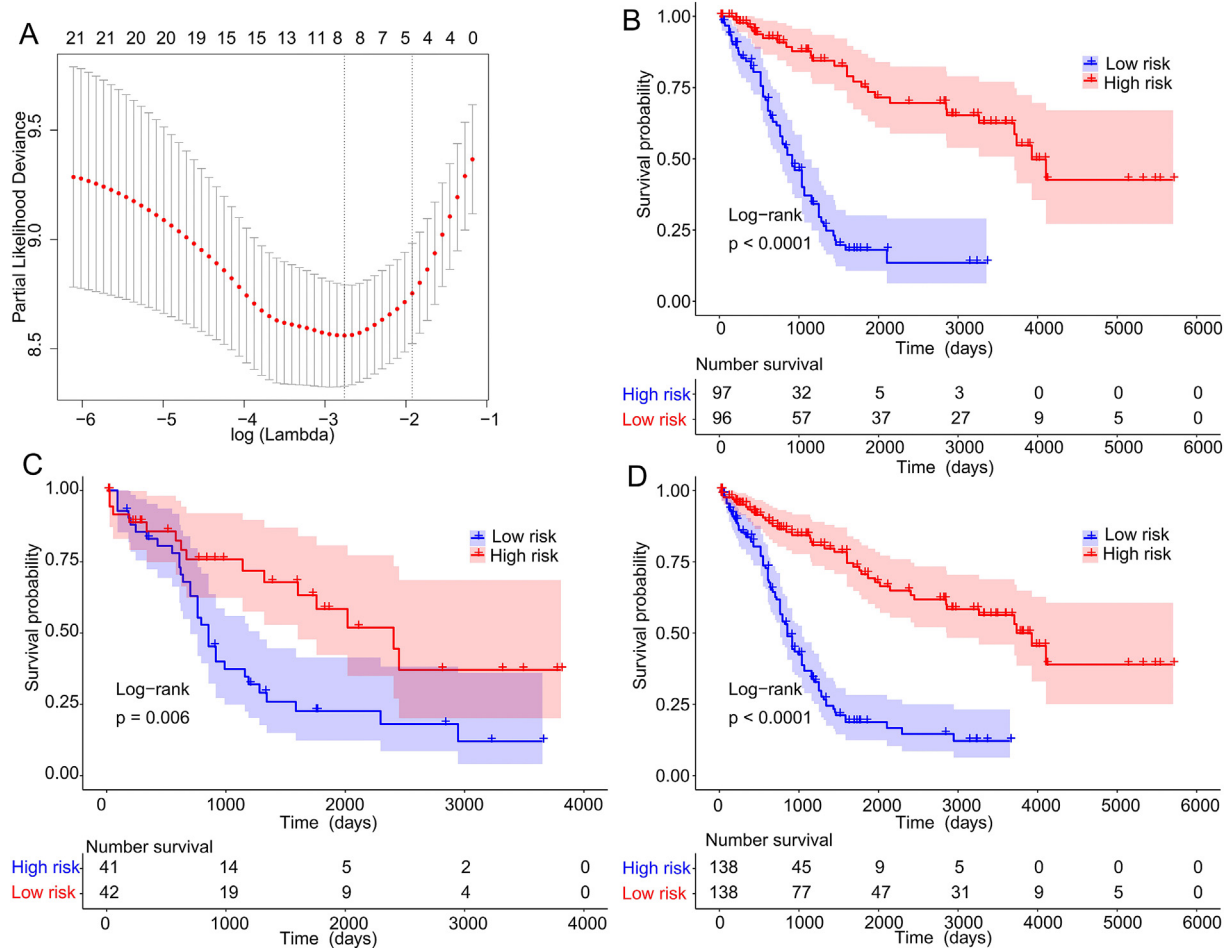
Fig. 4. Construction of immune prognostic model based on the infiltrating immune cells. (A) LASSO regression model used to reveal partial likelihood deviance of different quantitative variables. The red dots and the grey lines represent a misclassification error value and standard error (SE), respectively. The vertical dashed lines on the left and right represent the minimum criteria and 1-SE criteria, respectively. The lambda was randomly selected parameter. KM analysis for the overall survival of HCC patients based on IPG in the testing (B), validation (C), and entire cohorts (D).

IPG for the prognostic evaluation of HCC patients was more accurate and informing than the BCLC stage (Fig. 5C). To observe the relationship between IPG and clinical characteristics (age, gender, and BCLC stage), univariate (P = 0.037) and multivariate (P = 0.034) analyses were used to verify the independency of IPG regardless of adjusting the clinical characteristics (Supplementary Table 6). These results indicated that IPG was a robust biomarker for the prognostic evaluation of HCC patients.

3.4. Analysis of association of IPG with clinical characteristic, inflammation/related genes, and biological process

We investigated the relationship between IPG and clinical characteristics and found that the score of IPG varied across the BCLC stage and survival status. The score of IPG increased as the BCLC stage went up (p = 0.038), and scores of IPG in the dead group were higher than alive groups (p = 0.032) (Fig. 6A). We then investigated the association between IPG and inflammation-related genes, including epithelial-mesenchymal transition (EMT), cytotoxic factor, and immune checkpoint regulators. IPG was shown to positively correlate with all the genes of EMT and cytotoxic factor, and negatively correlate with several genes of immune checkpoint regulators, such as CTLA4 and LAG3 (Fig. 6B). Also, the biological process of IPG was analyzed through GSEA methods to elucidate its related functional pathway. Different analyses of genes between high-risk cohort and low-risk cohort were made. Both in molecular signatures database of "c2. cp.kegg.v6.1. symbols" and "c2. cp.biocarta.v6.1. symbols", the genes of the high-risk cohort were

enriched in some the tumorigenesis related pathways, including DNA replication, cell cycle, and PPARA pathway. The genes of the low-risk cohort were mainly associated with metabolism pathways, such as drug metabolism cytochrome p450, fatty acid metabolism, and tyrosine metabolism (Fig. 6C and D). The association analyses revealed that IPG correlated with some key clinical characteristics and genes and participated in several important biological functions. These findings showed the potential of IPG as a prognostic biomarker for HCC patients.

4. Discussion

The development of HCC is associated with the immune mechanism of the host [16]. Studies have shown that the liver immune microenvironment is reprogrammed under the influence of inflammation, which was closely related to HCC. HCC occurs mainly as a result of pathological changes in the long-term inflammatory response to chronic viral hepatitis [17]. The homeostasis of the immune microenvironment is the premise to maintain liver function [18]. However, the liver homeostasis of infiltrated immune cells is destroyed under chronic infection, prompting macrophages, neutrophils, mast cells, lymphocytes, and dendritic cells to release a large number of inflammatory factors to form inflammatory microenvironments. The inflammatory microenvironment promotes HCC induction, accelerates tumor progression, and contributes to the formation of new blood vessels [19,20]. Thus, much attention has been drawn to the component of infiltrated immune for its ability to be utilized as biomarkers for the clinical implication and the pathogenesis of HCC
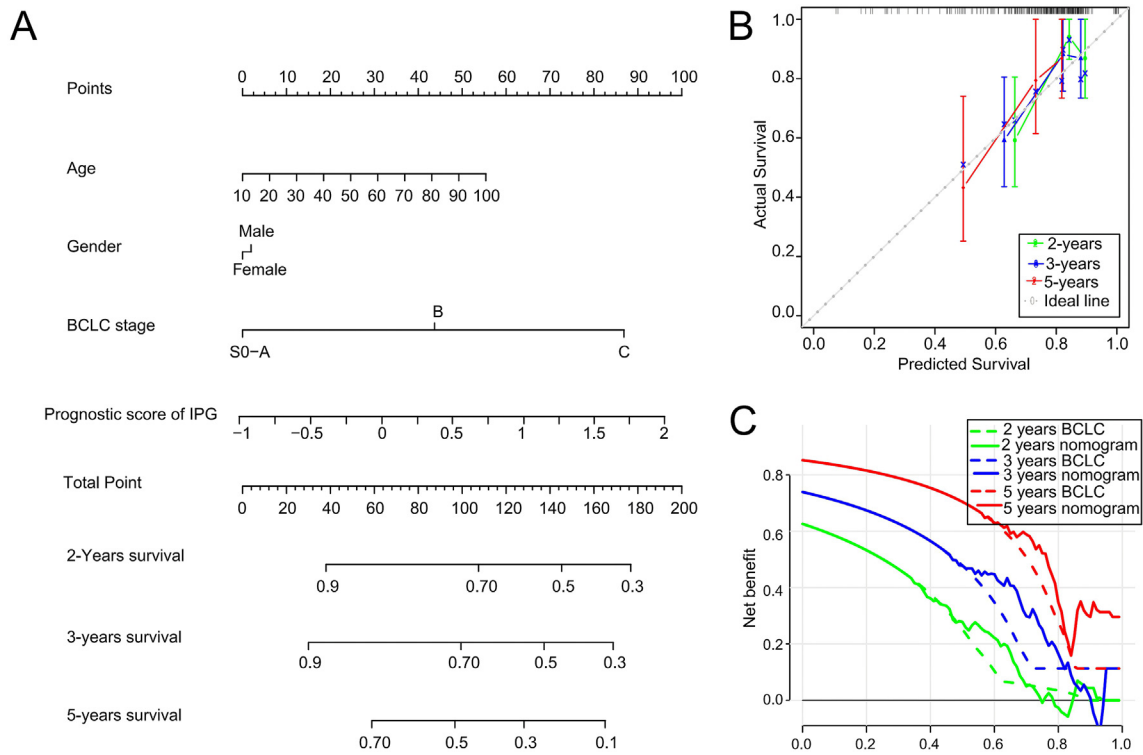
Fig. 5. Construction and validation of nomogram through a combination of IPG and clinical characteristics. (A) Nomogram plot for the prognostic prediction of HCC patients based on PG and clinical characteristics. (B) Calibration curves of nomograms for consistency between 2-, 3-, and 5-years survival predictions and observations. (C) Decision curve analyses for the prognostic prediction between nomogram and BCLC stage.
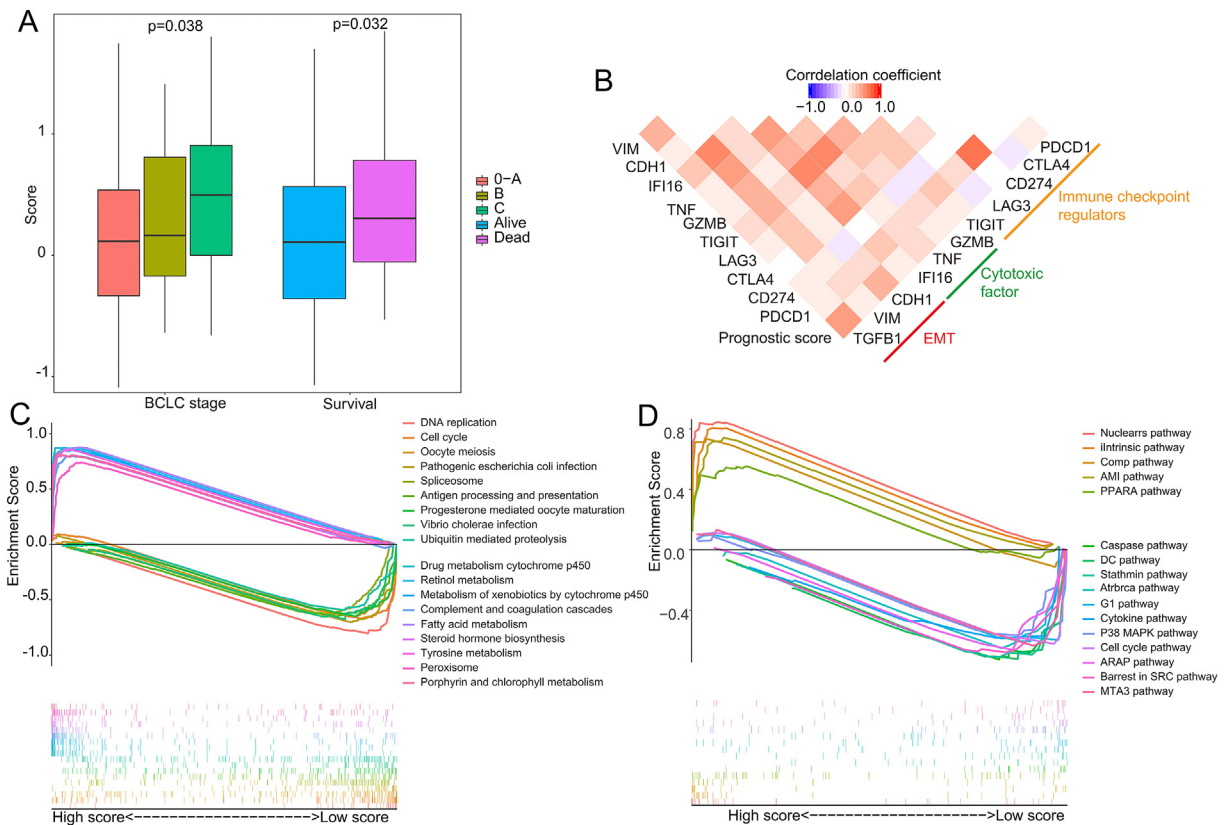
Fig. 6. Association analyses of IPG and clinical characteristics, genes, and biological processes. (A) The distribution of IPG value in different BCLC stages and survival status. (B) The association between IPG and genes linked to immune checkpoint regulators, cytotoxic factors, and EMT. (C) GSEA analyses on the biological pathways related to IPG for the database of (C) "c2. cp.kegg.v6.1. symbols" and (D) "c2. cp.biocarta.v6.1. symbols".

[21].

Although the insights of infiltrated immune cells have not benefited much the practical clinical application in HCC, they have provided a new strategy for the diagnosis and treatment. In this study, we used multi-dimensional bioinformatics analyses to examine the component of infiltrating immune cells in a large sample of HCC tissues and established IDG and IPG for the clinical implication. Our study not only demonstrated the IDG and IPG to be potential biomarkers for the HCC patients but also shown the infiltrating immune cells participating in its pathogenesis. Besides, our study overcame the difficulty of tissue acquisition, small sample size, and poor homogeneity. This research has solved the problem of single-dimensional analysis, lack of immune markers, and poor discrimination ability.

Early diagnosis and prognostic evaluation of HCC is the premise to improve the effect of treatment on patients. Therefore, specific markers for timely diagnosis and prognosis of HCC should be developed. A study on the predictive treatment efficacy of sorafenib in HCC patients using flow cytometry revealed that a high baseline $CD4^+$ T effector/Treg ratio could be used as a biomarker for significant prognosis in HCC patients with sorafenib treatment [22]. However, the simultaneous evaluation of effector immune cells, such as $CD8^+$, dendritic cells, and NK cells, in HCC patients, was not comprehensively performed on this platform. Similarly, using immunohistochemistry to investigate the potential biomarker of PD-L1/PD-1 for HCC, it was found that PD-L1 expression of macrophages cells was positively associated with the overall survival of the patient. This indicated that PD-L1$^+$ macrophages were characterized by immune activation with high levels of $CD8^+$ infiltration cells and immune-related gene expression [23]. However, the characterization of the effects of the immune cells on different cell types was limited, and the distinction of cells between tumor and normal tissue was weak. For these reasons, big data analytics on genes using high-throughput technology has been developed and verified for the screening of tumor biomarkers based on gene expression [24]. Therefore, the use of transcriptomic data to calculate tumor immune microenvironment could efficiently describe multiple immune cell phenotypes of different tissues in a large sample size patient cohort [25].

In this study, we investigated the composition of infiltrated immune cells through gene expression using the CIBEROSORT algorithm. We observed differences in the characteristics of the immune cell between tumor and normal tissues. For instance, macrophages and Tregs were more in the tumor tissues than in the normal tissues. These tumor-associated macrophages (TAM) not only prevented T cells from attacking tumor cells but also secreted growth factors to promote tumor angiogenesis [26]. The immunosuppressive cells in the tumor microenvironment, such as Tregs, plays an important role in the occurrence and development of tumor by inhibiting the function of immune effector cells through various mechanisms [27]. We hypothesized that the higher proportions of Tregs and macrophages in the tumor microenvironment contributed to tumorigenesis. On the other hand, we found that the proportion of $CD4^+$ T cells and $CD8^+$ T cells in the normal tissue were higher than tumor in tissues, which was attributed to the mechanism employed by T cells to find and kill cancer cells [28].

Our research showed the high AUC values of IDG for HCC diagnosis and significantly distinguished KM curve of prognosis in high- and low-risk groups by IPG. Thus, we strongly believe that our IDG and IPG could be potential biomarkers for the clinical implication of HCC patients. A general study used a high-throughput sequence to analyze the differential expression genes between tumor and normal and to identify the target for the pathogenesis and clinic of disease [29]. Nevertheless, data based on high-throughput sequencing might not be that clinically useful, hence the need to determine a suitable and effective scoring model to interpret these data clinically [30]. Our study enhanced the accuracy of diagnostic and prognostic prediction markedly by establishing the IDG and IPG based on a big sample size using systemic bioinformatics. Moreover, the IDG and IPG were based on the immune model, suggesting that they had a good practical clinical translation

value. Similarly, some studies have used genes database to establish an immune model score for the molecular classification and prognostic prediction on HCC. For example, Yutaka et al. classified the immune microenvironment of HCC into three distinct immunosubtypes [6], and Daniela et al. characterized molecular features of immune cells that infiltrate HCC [7]. By building IDG and IPG, our study has provided novel methods to screen out the potential biomarker for HCC clinical application.

The traditional HCC classification, such as the BCLC stage, TNM, and differentiation grade, is still the most common clinical guidelines in practice [31]. However, it has now been established that there may be significant differences in the clinical outcomes of patients at the same classification. The current classification of HCC neither provides comprehensive prognostic information nor responds well to the therapeutic outcome [32]. Molecular typing can better subdivide liver cancer, such as immune phenotype, cell origin, molecular pathway, mutation status, and gene expression, to distinguish different subtypes of HCC [33]. In this study, we also analyzed the association between IPG and clinical characteristics. The prognostic ability of IPG in this study was similar to that of the traditional BCLC stage. Our findings did not show our IPG to be superior to the BCLC stage in practical clinical application. Maybe the implementation of the IPG as a new component model for the prognosis and the integration of the IPG and BCLC stage could improve the prognostic evaluation of HCC patients.

The association analysis of IPG and molecular subtypes found that IPG negatively correlated with the immune checkpoint regulators related genes, such as CTLA4 and LAG3. This was in agreement with the immune checkpoint inhibitors in cancer immunotherapy that showed CTLA4 and LAG3, as the common biomarkers for designing the immunosuppressive site [34]. Furthermore, the GSEA analysis on the functional pathway related to IPG supported the genes engaged in the mechanism of tumorigenesis. Noteworthily, the GSEA result of low-risk cohort associated with multiple metabolism signatures. A growing number of studies have shown the involvement of immunometabolism in immune cell activation [35]. In particular, pathogen-mediated immune metabolism is involved in the pathogenesis of diseases such as HIV and intestinal flora [36,37].

Even though this study overcame the difficulty of tissue acquisition, small sample size, and poor homogeneity, there were some limitations. Our research was not comprehensive enough since it constructed the model based on the analysis of the composition of the immune cells in the HCC microenvironment only. The combination of the infiltrating of immune cells, pathological type, genes classification, and single-cell sequencing would be the best method to screen for HCC biomarkers. Besides, our study did not contrast between the cases with immune-checkpoint inhibitors and conventional therapy. Therefore, the utilization of our IDG and IPG in diagnostic and prognostic of HCC is still limited and needs further investigations. Furthermore, the component of infiltrating immune cells was derived from the transcriptomes of microarray using the CIBERSORT algorithm, which cannot be adopted in clinical practice immediately. The use of gene expression to predict the composition of immune cells remains to be explored.

5. Conclusion

This study has given a comprehensive analysis of the infiltrating immune cells in the utility of diagnosis and prognosis and constructed the IDG and IPG for the potential clinical application. The IDG and IPG may be used as an effective biomarker for improving the diagnosis and prognosis of HCC patients.

Credit author statement

manufacture of figures: Wenbiao Chen and Kefan Bi; writing of manu-
script: Wenbiao Chen.

Declaration of competing interest

The authors declare that they have no known competing financial
interests or personal relationships that could have appeared to influence
the work reported in this paper.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https
://doi.org/10.1016/j.jtauto.2020.100067.

References

[1] J. Bruix, G.J. Gores, V. Mazzaferro, Hepatocellular carcinoma: clinical frontiers and perspectives, Gut 63 (2014) 844–855.
[2] A. Forner, M. Reig, J. Bruix. Hepatocellular carcinoma, Lancet (London, England) 391 (2018) 1301–1314.
[3] O.V. Makarova-Rusher, J. Medina-Echeverz, A.G. Duffy, T.F. Greten, The yin and yang of evasion and immune activation in HCC, J. Hepatol. 62 (2015) 1420–1429.
[4] X. Tang, Z. Shu, W. Zhang, L. Cheng, J. Yu, M. Zhang, et al., Clinical significance of the immune cell landscape in hepatocellular carcinoma patients with different degrees of fibrosis, Ann. Transl. Med. 7 (2019) 528.
[5] B. Li, E. Severson, J.C. Pignon, H. Zhao, T. Li, J. Novak, et al., Comprehensive analyses of tumor immunity: implications for cancer immunotherapy, Genome Biol. 17 (2016) 174.
[6] Y. Kurebayashi, H. Ojima, H. Tsujikawa, N. Kubota, J. Maehara, Y. Abe, et al., Landscape of immune microenvironment in hepatocellular carcinoma and its additional impact on histological and molecular classification, Hepatology 68 (2018) 1025–1041.
[7] D. Sia, Y. Jiao, I. Martinez-Quetglas, O. Kuchuk, C. Villacorta-Martin, M. Castro de Moura, et al., Identification of an immune-specific class of hepatocellular carcinoma, based on molecular features, Gastroenterology 153 (2017) 812–826.
[8] Y.T. Liu, T.C. Tseng, R.S. Soong, C.Y. Peng, Y.H. Cheng, S.F. Huang, et al., A novel spontaneous hepatocellular carcinoma mouse model for studying T-cell exhaustion in the tumor microenvironment, Journal for immunotherapy of cancer 6 (2018) 144.
[9] G. Zhou, D. Sprengers, P.P.C. Boor, M. Doukas, H. Schutz, S. Mancham, et al., Antibodies against immune checkpoint molecules restore functions of tumor-infiltrating T cells in hepatocellular carcinomas, Gastroenterology 153 (2017), 1107-19.e10.
[10] M.X. Tian, W.R. Liu, H. Wang, Y.F. Zhou, L. Jin, X.F. Jiang, et al., Tissue-infiltrating lymphocytes signature predicts survival in patients with early/intermediate stage hepatocellular carcinoma, BMC Med. 17 (2019) 106.
[11] T. Barrett, S.E. Wilhite, P. Ledoux, C. Evangelista, I.F. Kim, M. Tomashevsky, et al., NCBI GEO: archive for functional genomics data sets–update, Nucleic Acids Res. 41 (2013) D991–D995.
[12] A.M. Newman, C.L. Liu, M.R. Green, A.J. Gentles, W. Feng, Y. Xu, et al., Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (2015) 453–457.
[13] I. Sohn, J. Kim, S.H. Jung, C. Park, Gradient lasso for Cox proportional hazards model, Bioinformatics 25 (2009) 1775–1781.
[14] X. Chen, H. Ishwaran, Random forests for genomic data analysis, Genomics 99 (2012) 323–329.
[15] Z. Li, M.J. Sillanpää, Overview of LASSO-related penalized regression methods for quantitative trait mapping and genomic selection, TAG Theoretical and applied genetics Theoretische und angewandte Genetik 125 (2012) 419–435.
[16] Y. He, Q. Dang, J. Li, Q. Zhang, X. Yu, M. Xue, et al., Prediction of hepatocellular carcinoma prognosis based on expression of an immune-related gene set, Aging 12 (2020) 965–977.
[17] Y.M. Yang, S.Y. Kim, E. Seki, Inflammation and liver cancer: molecular mechanisms and therapeutic targets, Semin. Liver Dis. 39 (2019) 26–42.
[18] M.W. Robinson, C. Harmon, C. O'Farrelly, Liver immunology and its role in inflammation and homeostasis, Cell. Mol. Immunol. 13 (2016) 267–276.
[19] D. Capece, M. Fischietti, D. Verzella, A. Gaggiano, G. Cicciarelli, A. Tessitore, et al., The inflammatory microenvironment in hepatocellular carcinoma: a pivotal role for tumor-associated macrophages, BioMed Res. Int. 2013 (2013) 187204.
[20] V. Chew, C. Tow, M. Teo, H.L. Wong, J. Chan, A. Gehring, et al., Inflammatory tumour microenvironment is associated with superior survival in hepatocellular carcinoma patients, J. Hepatol. 52 (2010) 370–379.
[21] E. Cariani, G. Missale, Immune landscape of hepatocellular carcinoma microenvironment: implications for prognosis and therapeutic applications, Liver Int. 39 (2019) 1608–1621.
[22] S.G. Kalathil, A. Hutson, J. Barbi, R. Iyer, Y. Thanavala, Augmentation of IFN-γ+ CD8+ T cell responses correlates with survival of HCC patients on sorafenib therapy, JCI Insight 4 (2019), e130116.
[23] C.-Q. Liu, J. Xu, Z.-G. Zhou, L.-L. Jin, X.-J. Yu, G. Xiao, et al., Expression patterns of programmed death ligand 1 correlate with different microenvironments and patient prognosis in hepatocellular carcinoma, Br. J. Canc. 119 (2018) 80–88.
[24] J. Zucman-Rossi, A. Villanueva, J.-C. Nault, J.M. Llovet, Genetic landscape and biomarkers of hepatocellular carcinoma, Gastroenterology 149 (2015), 1226-39.e4.
[25] C. Klijn, S. Durinck, E.W. Stawiski, P.M. Haverty, Z. Jiang, H. Liu, et al., A comprehensive transcriptional portrait of human cancer cell lines, Nat. Biotechnol. 33 (2015) 306–312.
[26] A.J. Petty, Y. Yang, Tumor-associated macrophages: implications in cancer immunotherapy, Immunotherapy 9 (2017) 289–302.
[27] Y. Takeuchi, H. Nishikawa, Roles of regulatory T cells in cancer immunity, Int. Immunol. 28 (2016) 401–409.
[28] V. Sasidharan Nair, E. Elkord, Immune checkpoint inhibitors in cancer therapy: a focus on T-regulatory cells, Immunol. Cell Biol. 96 (2018) 21–33.
[29] J. Xuan, Y. Yu, T. Qing, L. Guo, L. Shi, Next-generation sequencing in the clinic: promises and challenges, Canc. Lett. 340 (2013) 284–295.
[30] D.G. Altman, L.M. McShane, W. Sauerbrei, S.E. Taube, Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration, BMC Med. 10 (2012) 51.
[31] M. Roncalli, Y.N. Park, L. Di Tommaso, Histopathological classification of hepatocellular carcinoma, Dig. Liver Dis. : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver 42 (Suppl 3) (2010) S228–S234.
[32] C.-C. Chia, I. Rubinfeld, B.M. Scirica, S. McMillan, H.S. Gurm, Z. Syed, Looking beyond historical patient outcomes to improve clinical models, Sci. Transl. Med. 4 (2012), 131ra49-ra49.
[33] D. Sia, A. Villanueva, S.L. Friedman, J.M. Llovet, Liver cancer cell of origin, molecular class, and effects on patient prognosis, Gastroenterology 152 (2017) 745–761.
[34] P. Darvin, S.M. Toor, V. Sasidharan Nair, E. Elkord, Immune checkpoint inhibitors: recent progress and potential biomarkers, Exp. Mol. Med. 50 (2018) 165.
[35] J. Van den Bossche, L.A. O'Neill, D. Menon, Macrophage immunometabolism: where are we (going)? Trends Immunol. 38 (2017) 395–406.
[36] J.E. Belizário, J. Faintuch, M. Garay-Malpartida, Gut Microbiome Dysbiosis and Immunometabolism: New Frontiers for Treatment of Metabolic Diseases, vol. 2018, Mediators of inflammation, 2018, p. 2037838.
[37] L. Tarancon-Diez, E. Rodríguez-Gallego, A. Rull, J. Peraire, C. Viladés, I. Portilla, et al., Immunometabolism is a key factor for the persistent spontaneous elite control of HIV-1 infection, EBioMedicine 42 (2019) 86–96.