Exploration of the immune cell infiltration-related gene signature in the prognosis of melanoma

Yangyang Zeng^{1,2}, Yulan Zeng³, Hang Yin^{1,2}, Fengxia Chen^{1,2}, Qingqing Wang^{1,2}, Xiaoyan Yu^{1,2}, Yunfeng Zhou^{1,2}

¹Hubei Key Laboratory of Tumor Biological Behaviors, Zhongnan Hospital of Wuhan University, Wuhan 430071, Hubei, China

²Department of Radiation and Medical Oncology, Zhongnan Hospital of Wuhan University, Wuhan 430071, Hubei, China

³Cancer Center, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei 430022, China

Correspondence to: Yunfeng Zhou; email: yfzhouwhu@163.com, https://orcid.org/0000-0003-4328-6646Keywords: melanoma, immune cell infiltration, prognostic signature, biomarkersReceived: May 21, 2020Accepted: November 10, 2020Published: January 10, 2021

Copyright: © 2021 Zeng et al. This is an open access article distributed under the terms of the <u>Creative Commons Attribution</u> <u>License</u> (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Melanoma is a life-threatening form of skin cancer with an elevated risk of metastasis and high mortality rates. The prognosis and clinical outcomes of cancer immunotherapy in melanoma patients are influenced by immune cell infiltration in the tumor microenvironment (TME) and the expression of genetic factors. Despite reports suggesting that immune-classification may have a better prediction of prognosis compared to the American Joint Committee on Cancer/Union for International Cancer Control (AJCC/UICC) TNM-classification, the definition of Immunoscore in melanoma is becoming a difficult challenge. In this study, we established and verified a 7-gene prognostic signature. Melanoma patients from the Cancer Genome Atlas (TCGA) were separated into a low-risk group and a high-risk group using the median risk score. Receiver operating characteristic (ROC) analysis for overall survival (OS) showed that the area under the curve (AUC) was 0.701 for 1 year, 0.726 for 3 years, and 0.745 for 5 years, respectively. Moreover, a nomogram was constructed as a practical prognostic tool, and the AUC was 0.829 for 3 years, and 0.803 for 5 years, respectively. Furthermore, we validated the above results in two datasets from the Gene Expression Omnibus (GEO) database and the relationship between 7-gene prognostic signature and immune infiltration estimated.

INTRODUCTION

Melanoma is a highly aggressive skin cancer that causes about 55,500 deaths annually [1]. Metastatic melanoma represents the cause of death in the vast majority of cases and has a 5-year relative survival rate of approximately 25% [2, 3]. Based on clinical and pathological features, there are different therapeutic options for melanoma, including surgical treatment, chemotherapy, radiotherapy, immunotherapy, or targeted therapy [4, 5]. Previous studies have reported that the immuneclassification may have a more superior prognostic value compared to the AJCC/UICC TNM-classification, which reveals the importance of immunological biomarkers in melanoma [6–9]. Despite the encouraging clinical outcomes of immunotherapy over the past decade, a significant proportion of patients develop resistance, which is a major obstacle to successful immunotherapy [3, 10]. In spite of considerable advances in genetic approaches that provide diagnostic, prognostic, or therapeutic information, the gene-based biomarkers have not yet been used in routine clinical practice [1, 11, 12]. Therefore, it is imperative to explore the immune-related genomic signature to supplement conventional prognostic factors and to improve the effects of immunotherapeutic drugs.

Increasing evidence shows that immune cell infiltration of the TME influences the prognosis and clinical outcome of cancer immunotherapy [13-16]. In melanoma, tumor-infiltrating lymphocytes are a favorable prognosticator and predictive biomarker for treatment response [17–19]. Moreover, increasing evidence suggests that an intense infiltration of B cells, activated T cells and mature dendritic cells could be associated with a positive prognosis, while an increased amount of macrophages, mast cells, and neutrophils predict worse survival in malignant melanoma [16, 20-22]. However, due to the difficulties experienced in the evaluation and the complex intratumoral immune reactions, the definition of Immunoscore is becoming a difficult challenge in melanoma [9, 23]. Although programmed death-ligand 1 (PD-L1) expression in tumor-infiltrating immune cells (TIICs) is related to a better prognosis of cancer, PD-L1 expression does not influence treatment decisions [1, 24]. Therefore, we identified potential immune-related prognostic genes for melanoma prognosis and therapeutic evaluation. A series of immune-related genes as potential prognostic predictors from TCGA was used to establish a 7-gene prognostic signature in melanoma. Furthermore, we validated the results in two GEO datasets and estimated the relationship between the established prognostic signature and immune infiltration.

RESULTS

Construction and verification of immune-related groups

A total of 471 melanoma samples from the TCGA dataset were classified into two immune-related groups using the single-sample gene set enrichment analysis (ssGSEA): 203 samples (43.1%) in the high immune cell infiltration group, and 268 samples (56.9%) in the low immune cell infiltration group (Figure 1A).

Moreover, we verified the reliability of the two immune-related groups classified by the above method. Firstly, we compared the degree of immune cell infiltration between immune-related groups by the Estimation of Stromal and Immune cells in Malignant Tumor tissues using the Expression data (ESTIMATE) algorithm. As expected, the Tumor Purity was higher in the low immune cell infiltration group, while the ESTIMATE Score, Immune Score, and Stromal Score were higher in the high immune cell infiltration group (Figure 1A). Figure 1B shows the significant associations between the scores and the two immunerelated groups (P < 0.001). We further found significantly higher expression level of the human leukocyte antigen (HLA) family genes and PD-L1 (Figure 1C, 1D, P < 0.001) and increased variety of immune cells among 22 TIICs subtypes in the high immune cell infiltration group compared with the other group (Figure 1E).

Furthermore, the results of gene set enrichment analysis (GSEA) suggested an intensive immune phenotype in the high immune cell infiltration group. The top 5 Gene ontology (GO) terms are shown in Supplementary Figure 1A, including the MHC class II protein complex and immunoglobulin complex. Moreover, 5 immunerelated Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were selected and shown in Supplementary Figure 1B, such as B cell and T cell receptor signaling pathway. As previously reported, we also confirmed that the immune cell infiltration was a favorable prognosticator in melanoma (Supplementary Figure 2), and was closely associated with the above results. Therefore, our immune cell infiltration groups could be used as immune-related groups for further analysis.

Screening of differentially expressed genes (DEGs) and functional analysis

A total of 659 DEGs were identified from a comparison of high vs. low immune cell infiltration groups in the TCGA dataset, and comprised of 607 up-regulated genes and 52 down-regulated genes. Functional analysis was performed using Metascape. The top GO terms and KEGG included lymphocyte activation and cytokinecytokine receptor interaction (Figure 2A, 2B). The network of GO and KEGG enriched terms is shown in Figure 2C. Moreover, the 659 DEGs were analyzed using the Search Tool for the Retrieval of Interacting Genes database (STRING) and were used to construct the protein-protein interaction (PPI) network. Several central modules were established by Molecular Complex Detection (MCODE) software (Figure 2D, Module 1). Furthermore, we determined the prognostic value of the 659 DEGs in TCGA and obtained 509 DEGs associated with OS (Supplementary Table 1).

Identification of immunity-related module and intersecting DEGs

The weighted gene co-expression network analysis (WGCNA) detected 7 co-expression modules and their association with immune-related groups were analyzed. The cluster dendrogram was established (Figure 3A) and the blue module was the most correlated module of



Figure 1. Construction and verification of immune-related groups in melanoma samples. (A) The high and low immune cell infiltration groups (Immunity_H and Immunity_L) were constructed by ssGSEA and further verified by the ESTIMATE algorithm. (B) The correlation between the scores calculated by the ESTIMATE algorithm and the two immune cell infiltration groups (P < 0.001). (C) The expression of HLA family genes. (D) The expression of PD-L1. (E) The difference of TIIC subtypes between the two immune cell infiltration groups.



Figure 2. Screening of DEGs and functional analysis. (A) Heatmap of the GO enriched terms. (B) Heatmap of the KEGG enriched terms. (C) A network of GO and KEGG enriched terms. (D) The top one module. The color of nodes represents the log (FC) value and the size reflects the number of interacting proteins with the designated protein.

the immune cell infiltration groups ($|\mathbf{r}| = 0.74$, P = 5e-83) (Figure 3B). A total of 841 genes in the selected module were extracted for further study (Supplementary Table 2), and 333 intersecting DEGs were obtained from the module with the highest correlation coefficient and 509 DEGs with prognostic value (Figure 3C).

Functional analysis of the 333 DEGs was carried out by the clusterProfiler package. Top GO terms included T cell activation in biological processes (BP), T cell receptor complex in cellular components (CC), and cytokine binding in molecular functions (MF) (Figure 3D). Moreover, the top 10 enriched pathways are shown in Figure 3E, including cytokine-cytokine receptor interaction and B cell receptor signaling pathway.

Establishment and validation of the 7-gene prognostic signature

The 333 DEGs were subjected to univariate Cox regression analysis and 294 DEGs were selected (P < 0.05) for further study. Using the Least Absolute Shrinkage and Selection Operator (Lasso) Cox regression analysis, 16 genes were filtered for stepwise multivariate Cox regression analysis (Figure 4A, 4B). Finally, 7 key genes (CYTL1, CCL8, FCGR2C, OAS1, HAPLN3, WIPF1, CLIC2) were used to construct the prognostic signature (Supplementary Figure 3): Risk score = 0.007* CYTL1 - 0.033* CCL8 - 0.021* FCGR2C - 0.015* OAS1 - 0.037* HAPLN3 - 0.014* WIPF1 - 0.023* CLIC2 (Supplementary Table 3). The distribution of the risk score, survival status, and gene expression profiles between the two groups are displayed in Figure 4C–4E.

Melanoma patients from TCGA were separated into a low-risk group (n = 224) and a high-risk group (n = 223) based on the median risk score as a cutoff value. The Kaplan-Meier analysis showed that the highrisk group had a poorer OS compared with the low-risk group (Figure 5A). Moreover, the time-dependent ROC analysis showed that the AUC was 0.701 for 1 year, 0.726 for 3 years, and 0.745 for 5 years, respectively, indicating that our prognostic signature had a strong predictive capacity (Figure 5B). Similar results were observed for DSS and PFS. The high-risk group had worse DSS (Figure 5C) and PFS (Figure 5E) compared with the low-risk group. The AUCs for 1-year, 3-year, and 5-year DSS rates using the 7-gene prognostic signature were 0.733, 0.757, and 0.766, respectively (Figure 5D). The AUCs for predicting 1-year, 3-year, and 5-year PFS were 0.588, 0.619, and 0.634, respectively (Figure 5F). These results demonstrated that the 7-gene prognostic signature was an effective prognostic indicator of OS, DSS, and PFS.

Furthermore, GSE54467 (n = 79) and GSE65904 (n =207) were utilized to validate the predictive value of OS and DSS. Based on the median risk score, the high-risk group (n=39) and low-risk group (n=40) were identified in the GSE54467 dataset. Likewise, patients in the highrisk group had a shorter OS (Supplementary Figure 4A). The AUC of our prognostic signature was 0.555 at 1 year, 0.682 at 3 years, and 0.761 at 5 years (Supplementary Figure 4B). In the GSE65904 dataset, the Kaplan-Meier analysis also showed that the highrisk group (n=103) had a worse DSS compared with the low-risk group (n=104) (Supplementary Figure 4C). The time-dependent ROC analysis also showed that our prognostic signature had a good predictive capacity in predicting 1-year, 3-year, and 5-year DSS (AUC = 0.641, 0.629, 0.590, respectively) (Supplementary Figure 4D).

Further, we performed a stratification analysis for OS to determine the prognostic power of our prognostic signature in subgroups of melanoma. The high-risk patients had a worse prognosis compared with the lowrisk patients in each stratum (Figure 6A–6J). Moreover, the prognostic signature had a better predictive ability compared with other clinical factors (Supplementary Figure 5A–5C). The pathologic T stage, N stage, and risk score were identified as independent prognostic factors (P < 0.001) (Supplementary Figure 5D, 5E). Eventually, we incorporated these prognostic factors to formulate an OS nomogram (Figure 6K). The ROC curves and calibration curves showed a favorable predictive capacity and stability in predicting the 3-year (AUC = 0.829) or 5-year (AUC = 0.803) OS (Supplementary Figure 6).

Immune infiltration score and immune cells infiltration analysis

The box chart showed that there were higher immune and stromal scores in the high-risk group compared with the low-risk group (Figure 7A and Supplementary Figure 7A, P < 0.001). The results in GSE54467, an independent dataset, also confirmed a similar correlation (Figure 7B and Supplementary Figure 7B, P < 0.001). Additionally, the expression of each gene (CCL8, FCGR2C, OAS1, HAPLN3, WIPF1, CLIC2) was significantly positively correlated to the immune score (Figure 7D–7I) and stromal score (Supplementary Figure 7D–7I). But CYTL1 had the opposite result in the immune score (Figure 7C) and had no relevance to the stromal score (Supplementary Figure 7C).

Furthermore, the risk score was significantly negatively correlated to B cells (Cor = -0.279), CD4+ T cells (Cor = -0.332), CD8+ T cells (Cor = -0.507), dendritic cells (Cor = -0.549), macrophages (Cor = -0.315), and



Figure 3. Identification of highly immune-related DEGs with prognostic value and functional analysis. (A) A hierarchical clustering dendrogram is built to detect co-expressed genes in modules in the TCGA dataset of melanoma. (B) A heatmap showing the relationships of consensus module-trait in different modules under the low and high immune cell infiltration groups. (C) Identification of 333 common DEGs from 509 DEGs with prognostic value and the blue module using the Venn diagram software. (D) Top 15 GO terms. (E) Top 10 KEGG pathways.

neutrophils (Cor = -0.619) (Figure 8, P < 0.001). Meanwhile, 6 prognostic genes (CCL8, FCGR2C, OAS1, HAPLN3, WIPF1, CLIC2) were associated with the abundance of B cell, CD4+T cell, CD8+T cell, dendritic cells, neutrophils and macrophages (Supplementary Figure 8B–8G, P < 0.05). The only gene (CYTL1) was unrelated to the 6 types of immune cell infiltration (Supplementary Figure 8A, P > 0.05).



Figure 4. Establishment and assessment of immune-related prognostic signature. (A, B) 294 DGEs determined using (A) LASSO regression analysis and (B) 10-fold cross-validation. (C–E) The distribution of (C) risk score, (D) survival status, and (E) immune-related gene expression profiles.

GSEA analysis

The GSEA analysis was performed for functional annotation of the 7-gene prognostic signature. We found that more than 10 immune-related GO terms and KEGG

pathways were visibly enriched in the low-risk group, while there were few enriched in the high-risk group. The top 5 immune-related GO terms are shown in Figure 9A, such as positive regulation of cytokine production. Besides, the top 5 immune-related pathways



Figure 5. Kaplan–Meier survival analysis and time-dependent ROC analysis of the 7-gene prognostic signature in the TCGA dataset. Kaplan-Meier survival curves and ROC curves of (A, B) OS, (C, D) DSS, and (E, F) PFS in patients with melanoma.

are displayed in Figure 9B, such as natural killer cell mediated cytotoxicity.

DISCUSSION

In this work, two immune-related groups were classified based on the degree of immune cell infiltration in the TCGA dataset of melanoma. A total of 333 DEGs with prognostic value were obtained by WGCNA. A novel 7gene prognostic signature in melanoma was established and its predictive value and accuracy validated in GSE54467 and GSE65904 datasets. Finally. the interaction between the immune 7-gene and prognostic signature was explored using the ESTIMATE algorithm, TIMER database and GSEA. Figure 10 displays a detailed flowchart of the current study.

Although studies have reported that the immuneclassification may have a better prognostic value than the AJCC/UICC TNM-classification, the definition of Immunoscore is a challenge in melanoma [23]. Previous studies suggest that the immune infiltration of the TME can be used to evaluate melanoma survival [25–27]. Therefore, we identified potential immune-related prognostic genes for melanoma prognosis. In our study, we established immune-related groups by ssGSEA and identified them using the ESTIMATE algorithm. We verified the reliability of the two immune-related groups by comparing different expressions of HLA family genes and PD-L1, which influence the human immune system and immunotherapy [28, 29], and using the CIBERSORT algorithm. Furthermore, the results of GSEA also showed different immune phenotypes in immune-related groups.

Based on the high heterogeneity of melanoma, there are more than 51,000 biomarkers, including tissuebased tumor cell and TME biomarkers [30, 31]. In this study, we obtained 659 DEGs, and 333 DEGs with prognostic value were identified and validated in the TCGA dataset by WGCNA. Functional analysis was performed and PPI networks for these genes were constructed. The results of the GO analysis in both 659



Figure 6. Stratification analysis and construction of a prognostic nomogram in patients. (A–J) Kaplan–Meier analysis of the subgroups including (A) \leq 60 years, (B) > 60 years, (C) female, (D) male, (E) stage I-II, (F) stage III-IV, (G) NO, (H) N1-3, (I) T1-2, and (J) T3-4. (K) OS predictive nomogram.

DEGs and 333 DEGs were closely related to the TME and immune function, such as lymphocyte activation and T cell activation. Besides, KEGG analysis showed that the B cell receptor signaling pathway was the overlapping enriched pathways. In the top one PPI module of the 659 DEGs, the main hub nodes were reported to affect the TME and immunotherapy, including HLA and guanine-binding protein (GBP) family induced by interferon and C-X-C motif chemokine ligand, like HLA-A [32] and GBP4 [33]. In this study, 7 genes (CYTL1, CCL8, FCGR2C, OAS1, HAPLN3, WIPF1, CLIC2) were identified as prognostic signatures for melanoma. CYTL1 was found to be highly expressed in neuroblastoma and a potential therapeutic target and diagnosis biomarker [34]. The expression of CCL8 in nude mice has been reported to inhibit human cervical cancer [35]. Moreover, in melanoma, previous research suggests that a local CCL8-rich environment promotes the selection of metastatic tumor cells, while a high CCL8 concentration



Figure 7. Relationship between the prognostic signature and immune score. (A, B) The high-risk group has lower immune scores compared with the low-risk group in (A) the TCGA dataset and (B) GSE54467. (C–I) The association between immune score and the expression of each gene in a 7-gene prognostic signature: (C) CYTL1, (D) CCL8, (E) FCGR2C, (F) OAS1, (G) HAPLN3, (H) WIPF1, (I) CLIC2.

inhibits their migration [36]. OAS1 induced by interferons is reported to be one gene in prognostic signature for patients with bladder cancer, despite a lack of sufficient experimental studies [37]. A previous study showed that WIPF1 knockdown inhibited natural killer cell cytotoxicity [38, 39]. CLIC2 is identified as a novel gene related to immune checkpoint proteins based on TCGA gene expression data [40], while another study suggests that it inhibits the hematogenous spread of tumor cells [41]. However, there is a need for further research to understand the function of FCGR2C and HAPLN3 in cancer. The role of the 7 genes in melanoma also requires more in-depth investigation.

Additionally, our prognostic signature was found to be an effective and stable prognostic indicator, since it demonstrated favorable prognostic value in predicting OS, DSS, and PFS outcomes in the TCGA dataset of melanoma. Two GEO datasets were used to validate the OS and DSS results. Moreover, stratification analysis in different subgroups was performed and a nomogram was constructed by integrating each independent clinical prognostic factor. Both ROC curves and calibration plots showed a robust and reliable predictive ability of the nomogram for OS. Therefore, our prognostic signature could predict more clinical prognostic factors and provide a useful tool to supplement the traditional clinical prognostic factors and improve the therapeutic effect. Nevertheless, the limitations of this study included that some risk factors for melanoma like a family history of melanoma could not be collected in the TCGA dataset, and the two selected GEO datasets lacked some corresponding clinical features data.

As for the interaction between the immune and the established 7-gene prognostic signature, we found that the low-risk group had higher immune and stromal scores. Previous studies demonstrated that intense infiltration of activated T cells, B cells, or mature dendritic cells could be positive prognosis factors [16, 20, 21]. Infiltrating CD8 lymphocytes are also reported to be a good prognostic factor for melanoma [42]. We observed a significantly negative correlation between the abundance of six immune cell subtypes and the risk score. Initially, neutrophils showed the strongest negative correlation with the risk score. However, it could be explained that the prognostic signature is



Figure 8. Relationship between the prognostic signature and the TIICs subtypes. Association between risk score and six subtypes of TIICs: (A) B cells. (B) CD4+ T cell. (C) CD8+ T cell. (D) Dendritic cells. (E) Macrophage. (F) Neutrophil.

mainly related to T cell activity and cytokine function: on one hand, CD8+ plays a major anti-tumor function, while on the other hand, some related cytokines may recruit anti-neoplastic neutrophils in the early stage [43]. Besides, the proportion of various immune infiltrating cells need to be considered for specific risk grouping. Moreover, except for CYTL1, the expression of each gene in the 7-gene prognostic signature was also found to be significantly related to immunization, such as immune score, stromal score and immune cell subtypes. Finally, GSEA analysis showed that many immune-related GO terms and KEGG pathways were enriched in the low-risk group. Due to the limitations associated with many current experiments on finite cell lines or samples, we carried out a comprehensive and systematical analysis of the TME in melanoma using a larger volume of data. However, despite the high heterogeneity of cancer or the lack of computer algorithms, continued exploration of these findings are likely to lead to novel insights into the molecular mechanisms.



Figure 9. GSEA analysis based on the risk value. (A) The significant enrichment of the top 5 GO terms in the high-risk group and the top 5 immune-related GO terms in the low-risk group. (B) The significant enrichment of the top 5 pathways in the high-risk group and the top 5 immune-related pathways in the low-risk group.

In conclusion, this study identified and validated 333 immune infiltration related genes as potential immune predictors for melanoma. Furthermore, we constructed and verified a 7-gene prognostic signature associated with the TME as a prognostic indicator. Although previous studies have studied immune infiltrationrelated prognostic models, especially those on melanoma [25, 26, 44, 45], in contrast, we first established reliable immune-related groups and obtained core immune infiltration related genes with prognostic value. We then constructed a favorable prognostic model for use in predicting more clinical prognostic factors and established a nomogram as a practical prognostic tool. Finally, we explored the relationship between the model, including its genes, immune infiltration score, and immune cell infiltration. Taken together, this study provides a series of novel immune-related biomarkers for prognosis of melanoma and establishes a 7-gene prognostic signature to supplement traditional clinical prognostic factors and improve the therapeutic effect.



Figure 10. The flowchart of the current study.

MATERIALS AND METHODS

Database

The HTSeq-FPKM data and clinical information of melanoma were acquired from the TCGA database (https://portal.gdc.cancer.gov). The cohort of melanoma for identifying immune-related genes consisted of 471 samples in TCGA. DSS and PFS data for melanoma was obtained from cBioportal (https://www.cbioportal.org/). Two GEO datasets (GSE54467, GSE65904) were gained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). Only patients with overall survival days \geq 30 days were included in the study.

Construction of immune-related groups

The infiltration levels of immune cells in 471 melanoma samples were estimated by ssGSEA [46], and the immune-related groups were constructed using the hclust function in R software. The ssGSEA analysis was carried out in the R package Gene Set Variation Analysis (GSVA) [47].

Verification of immune-related groups

The related scores for individual samples were determined by the ESTIMATE algorithm in R software. Then, the different gene expression levels of the HLA family genes and PD-L1 were analyzed between the immune-related groups. Moreover, the composition of 22 immune cell subtypes was determined by the CIBERSORT package [48]. The immune-related functional annotation and signaling pathways that were differentially activated in the two groups were validated by GSEA.

Screening of differentially expressed genes

The overlapping genes and normalized gene expression profile data in the TCGA dataset and the two GEO datasets were identified for further analysis. The comparison of high vs. low immune cell infiltration groups was executed in R package limma [49]. |logFC|> 0.5 and false discovery rate (FDR) < 0.05 were considered as the cut-off criteria.

Functional analysis

Functional analysis was carried out by Metascape (<u>http://metascape.org</u>) [50] or R package clusterProfiler. The PPI network with the confidence score > 0.9 was established using the STRING website (<u>https://string-db.org/</u>) [51] and Cytoscape software [52], and proteins with degree > 1 were selected.

Survival analysis

The Kaplan-Meier survival analysis was carried out to evaluate the correlation between the survival probability of patients and different groups characterized by different levels of immune cell infiltration, expression of individual gene, or risk score.

WGCNA of immune-related groups

The WGCNA was used to determine the relationship between co-expression gene modules and immunerelated groups in the WGCNA package [53]. The WGCNA network and the co-expressed gene modules were established and detected using the soft threshold power of $\beta = 5$, topological overlap matrix (TOM), and minimal module size of 50. 841 genes in the blue module with the highest correlation coefficient were selected for further analyses. A total of 333 DEGs with prognostic value in the TCGA dataset were obtained using the Venn diagram software.

Identification of immune-related gene prognostic signature

A total of 294 DEGs (P < 0.05) were selected by univariate Cox regression analysis. Out of these, 16 genes were identified by LASSO Cox analysis with 10-round crossvalidation. The prognostic signature was selected and optimized by stepwise multiple Cox regression analysis. The formula for the risk score was as follows [54]: Risk core = $\sum_{i=1}^{n} \text{Coef}_i * \text{Exp}_i$; *n*, Coef_i, Exp_i represented the number of signature genes, the coefficient, and the gene expression level, respectively. Univariate and multivariate Cox regression analysis were performed to evaluate the independence of the prognostic signature from clinical factors. The time-dependent ROC analysis was performed using the survival package. The nomogram and calibration curves were plotted by R package rms.

TIMER database analysis

The deconvolution algorithm provided by Tumor Immune Estimation Resource (TIMER) database (https://cistrome.shinyapps.io/timer/) [55] was used to determine the relationship between risk scores and TIICs using Pearson correlation. Additionally, the association between the expression of 7 genes and TIICs was also analyzed.

Statistical analysis

Statistical analyses were carried out in R version 3.6.3 (Package: limma, sva, pheatmap, ggpubr, org.Hs.eg.db, clusterProfiler, enrichplot, survival, glmnet, surviner, survivalROC, beeswarm, rms, etc).

Abbreviations

TME: Tumor microenvironment: AJCC/UICC: American Joint Committee on Cancer/Union for International Cancer Control; TCGA: The Cancer Receiver Genome Atlas: ROC: operating characteristic; OS: Overall survival; AUC: Area under the curve; DSS: Disease-specific survival; PFS: Progression-free survival; GEO: Gene Expression Omnibus; PD-L1: programmed death-ligand 1; TIICs: Tumour infiltrating immune cells; ssGSEA: Singlesample gene set enrichment; ESTIMATE: Estimation of Stromal and Immune cells in Malignant Tumor tissues using the Expression data; HLA: Human leukocyte antigen; GSEA: Gene set enrichment analysis; GO analysis: Gene ontology analysis; KEGG: Kyoto Encyclopedia of Genes and Genomes; DEGs: Differentially expressed genes; STRING: Search Tool for the Retrieval of Interacting Genes database; PPI: Protein-protein interaction; MCODE: Molecular Complex Detection; WGCNA: Weighted gene co-expression network analysis; BP: Biological processes; CC: Cellular components; MF: Molecular functions; LASSO: Least Absolute Shrinkage and Selection Operator; GSVA: Gene Set Variation Analysis; FDR: False discovery rate; TOM: Topological overlap matrix; TIMER: Tumor Immune Estimation Resource; HR: Hazard ratio.

AUTHOR CONTRIBUTIONS

Conceptualization: Yangyang Zeng and Yunfeng Zhou; Methodology: Yangyang Zeng and Yulan Zeng; Software: Yangyang Zeng and Hang Yin; Verification: Yangyang Zeng and Fengxia Chen; Visualization: Yangyang Zeng, Qingqing Wang, and Xiaoyan Yu; Writing (original draft): Yangyang Zeng and Yunfeng Zhou; Writing (review and editing): Yangyang Zeng and Yunfeng Zhou.

ACKNOWLEDGMENTS

We are grateful for free access to the TCGA and GEO databases. And we would like to thank Ganiou Assani for critically reviewing the manuscript.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

FUNDING

This research was funded by the National Natural Science Foundation of China (81472799).

REFERENCES

- Schadendorf D, van Akkooi AC, Berking C, Griewank KG, Gutzmer R, Hauschild A, Stang A, Roesch A, Ugurel S. Melanoma. Lancet. 2018; 392:971–84. <u>https://doi.org/10.1016/S0140-6736(18)31559-9</u> PMID:<u>30238891</u>
- Strub T, Ballotti R, Bertolotto C. The "ART" of epigenetics in melanoma: from histone "alterations, to resistance and therapies". Theranostics. 2020; 10:1777–97. https://doi.org/10.7150/thno.36218 PMID:32042336
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin. 2020; 70:7–30. <u>https://doi.org/10.3322/caac.21590</u> PMID:<u>31912902</u>
- Amaria RN, Menzies AM, Burton EM, Scolyer RA, Tetzlaff MT, Antdbacka R, Ariyan C, Bassett R, Carter B, Daud A, Faries M, Fecher LA, Flaherty KT, et al, and International Neoadjuvant Melanoma Consortium members. Neoadjuvant systemic therapy in melanoma: recommendations of the International Neoadjuvant Melanoma Consortium. Lancet Oncol. 2019; 20:e378–89.

https://doi.org/10.1016/S1470-2045(19)30332-8 PMID:<u>31267972</u>

- McKean MA, Amaria RN. Multidisciplinary treatment strategies in high-risk resectable melanoma: role of adjuvant and neoadjuvant therapy. Cancer Treat Rev. 2018; 70:144–53. <u>https://doi.org/10.1016/j.ctrv.2018.08.011</u> PMID:30195813
- Keung EZ, Gershenwald JE. Clinicopathological features, staging, and current approaches to treatment in high-risk resectable melanoma. J Natl Cancer Inst. 2020; 112:875–85. https://doi.org/10.1093/jnci/djaa012 PMID:32061122
- Galon J, Pagès F, Marincola FM, Thurin M, Trinchieri G, Fox BA, Gajewski TF, Ascierto PA. The immune score as a new possible approach for the classification of cancer. J Transl Med. 2012; 10:1. <u>https://doi.org/10.1186/1479-5876-10-1</u> PMID:22214470
- Galon J, Lugli A, Bifulco C, Pages F, Masucci G, Marincola FM, Ascierto PA. World-wide immunoscore task force: meeting report from the "Melanoma Bridge", Napoli, November 30th-December 3rd, 2016. J Transl Med. 2017; 15:212. https://doi.org/10.1186/s12967-017-1310-9

PMID:29058595

 Ros-Martínez S, Navas-Carrillo D, Alonso-Romero JL, Orenes-Piñero E. Immunoscore: a novel prognostic tool. Association with clinical outcome, response to treatment and survival in several malignancies. Crit Rev Clin Lab Sci. 2020; 57:432–43. https://doi.org/10.1080/10408363.2020.1729692 PMID:32175789

- Jenkins RW, Fisher DE. Treatment of advanced melanoma in 2020 and beyond. J Invest Dermatol. 2021; 141:23-31. <u>https://doi.org/10.1016/j.jid.2020.03.943</u> PMID:<u>32268150</u>
- 11. Syn NL, Teng MW, Mok TS, Soo RA. De-novo and acquired resistance to immune checkpoint targeting. Lancet Oncol. 2017; 18:e731–41. https://doi.org/10.1016/S1470-2045(17)30607-1 PMID:29208439
- Teng MW, Ngiow SF, Ribas A, Smyth MJ. Classifying cancers based on t-cell infiltration and PD-L1. Cancer Res. 2015; 75:2139–45. <u>https://doi.org/10.1158/0008-5472.CAN-15-0255</u> PMID:<u>25977340</u>
- Becht E, Giraldo NA, Germain C, de Reyniès A, Laurent-Puig P, Zucman-Rossi J, Dieu-Nosjean MC, Sautès-Fridman C, Fridman WH. Immune contexture, immunoscore, and Malignant cell molecular subgroups for prognostic and theranostic classifications of cancers. Adv Immunol. 2016; 130:95–190. <u>https://doi.org/10.1016/bs.ai.2015.12.002</u> PMID:<u>26923001</u>
- Petitprez F, Vano YA, Becht E, Giraldo NA, de Reyniès A, Sautès-Fridman C, Fridman WH. Transcriptomic analysis of the tumor microenvironment to guide prognosis and immunotherapies. Cancer Immunol Immunother. 2018; 67:981–988. <u>https://doi.org/10.1007/s00262-017-2058-z</u> PMID:<u>28884365</u>
- Fridman WH, Zitvogel L, Sautès-Fridman C, Kroemer G. The immune contexture in cancer prognosis and treatment. Nat Rev Clin Oncol. 2017; 14:717–34. <u>https://doi.org/10.1038/nrclinonc.2017.101</u> PMID:<u>28741618</u>
- Ladányi A. Prognostic and predictive significance of immune cells infiltrating cutaneous melanoma. Pigment Cell Melanoma Res. 2015; 28:490–500. <u>https://doi.org/10.1111/pcmr.12371</u> PMID:<u>25818762</u>
- Lee N, Zakka LR, Mihm MC Jr, Schatton T. Tumourinfiltrating lymphocytes in melanoma prognosis and cancer immunotherapy. Pathology. 2016; 48:177–87. <u>https://doi.org/10.1016/j.pathol.2015.12.006</u> PMID:<u>27020390</u>
- 18. Gartrell RD, Marks DK, Hart TD, Li G, Davari DR, Wu A, Blake Z, Lu Y, Askin KN, Monod A, Esancy CL, Stack EC, Jia DT, et al. Quantitative analysis of immune infiltrates

in primary melanoma. Cancer Immunol Res. 2018; 6:481–93.

https://doi.org/10.1158/2326-6066.CIR-17-0360 PMID:<u>29467127</u>

- Fu Q, Chen N, Ge C, Li R, Li Z, Zeng B, Li C, Wang Y, Xue Y, Song X, Li H, Li G. Prognostic value of tumorinfiltrating lymphocytes in melanoma: a systematic review and meta-analysis. Oncoimmunology. 2019; 8:1593806. <u>https://doi.org/10.1080/2162402X.2019.1593806</u> PMID:31143514
- 20. Selitsky SR, Mose LE, Smith CC, Chai S, Hoadley KA, Dittmer DP, Moschos SJ, Parker JS, Vincent BG. Prognostic value of B cells in cutaneous melanoma. Genome Med. 2019; 11:36. <u>https://doi.org/10.1186/s13073-019-0647-5</u> PMID:<u>31138334</u>
- Tucci M, Passarelli A, Mannavola F, Felici C, Stucci LS, Cives M, Silvestris F. Immune system evasion as hallmark of melanoma progression: the role of dendritic cells. Front Oncol. 2019; 9:1148. <u>https://doi.org/10.3389/fonc.2019.01148</u> PMID:31750245
- Iglesia MD, Parker JS, Hoadley KA, Serody JS, Perou CM, Vincent BG. Genomic analysis of immune cell infiltrates across 11 tumor types. J Natl Cancer Inst. 2016; 108:djw144. https://doi.org/10.1093/jnci/djw144 PMID:27335052
- Galon J, Fox BA, Bifulco CB, Masucci G, Rau T, Botti G, Marincola FM, Ciliberto G, Pages F, Ascierto PA, Capone M. Immunoscore and immunoprofiling in cancer: an update from the melanoma and immunotherapy bridge 2015. J Transl Med. 2016; 14:273. https://doi.org/10.1186/s12967-016-1029-z

PMID:27650038

- 24. Zhao T, Li C, Wu Y, Li B, Zhang B. Prognostic value of PD-L1 expression in tumor infiltrating immune cells in cancers: a meta-analysis. PLoS One. 2017; 12:e0176822. <u>https://doi.org/10.1371/journal.pone.0176822</u> PMID:<u>28453554</u>
- Wang P, Zhang X, Sun N, Zhao Z, He J. Comprehensive analysis of the tumor microenvironment in cutaneous melanoma associated with immune infiltration. J Cancer. 2020; 11:3858–70. <u>https://doi.org/10.7150/jca.44413</u> PMID:32328190
- Huang R, Mao M, Lu Y, Yu Q, Liao L. A novel immunerelated genes prognosis biomarker for melanoma: associated with tumor microenvironment. Aging (Albany NY). 2020; 12:6966–6980. https://doi.org/10.18632/aging.103054 PMID:32310824

- 27. Balatoni T, Mohos A, Papp E, Sebestyén T, Liszkay G, Oláh J, Varga A, Lengyel Z, Emri G, Gaudi I, Ladányi A. Tumor-infiltrating immune cells as potential biomarkers predicting response to treatment and survival in patients with metastatic melanoma receiving ipilimumab therapy. Cancer Immunol Immunother. 2018; 67:141–151. https://doi.org/10.1007/s00262-017-2072-1 PMID:28988380
- 28. McGranahan N, Rosenthal R, Hiley CT, Rowan AJ, Watkins TB, Wilson GA, Birkbak NJ, Veeriah S, Van Loo P, Herrero J, Swanton C, Jamal-Hanjani M, et al, and TRACERx Consortium. Allele-specific HLA loss and immune escape in lung cancer evolution. Cell. 2017; 171:1259-71.e11. https://doi.org/10.1016/j.cell.2017.10.001 PMID:29107330
- 29. Topalian SL, Taube JM, Pardoll DM. Neoadjuvant checkpoint blockade for cancer immunotherapy. Science. 2020; 367:eaax0182. https://doi.org/10.1126/science.aax0182 PMID:32001626
- 30. Villanueva J, Herlyn M. Melanoma and the tumor microenvironment. Curr Oncol Rep. 2008; 10:439–46. https://doi.org/10.1007/s11912-008-0067-y PMID:18706274
- 31. Donnelly D 3rd, Aung PP, Jour G. The "-OMICS" facet of melanoma: heterogeneity of genomic, proteomic and metabolomic biomarkers. Semin Cancer Biol. 2019; 59:165-74. https://doi.org/10.1016/j.semcancer.2019.06.014 PMID:31295564
- 32. Rodig SJ, Gusenleitner D, Jackson DG, Gjini E, Giobbie-Hurder A, Jin C, Chang H, Lovitch SB, Horak C, Weber JS, Weirather JL, Wolchok JD, Postow MA, et al. MHC proteins confer differential sensitivity to CTLA-4 and PD-1 blockade in untreated metastatic melanoma. Sci Transl Med. 2018; 10:eaar3342. https://doi.org/10.1126/scitranslmed.aar3342
- PMID:30021886 33. Xu L, Pelosof L, Wang R, McFarland HI, Wu WW, Phue
- JN, Lee CT, Shen RF, Juhl H, Wu LH, Alterovitz WL, Petricon E, Rosenberg AS. NGS evaluation of colorectal cancer reveals interferon gamma dependent expression of immune checkpoint genes and identification of novel IFNy induced genes. Front Immunol. 2020; 11:224. https://doi.org/10.3389/fimmu.2020.00224

PMID:32265897

34. Wen M, Wang H, Zhang X, Long J, Lv Z, Kong Q, An Y. Cytokine-like 1 is involved in the growth and metastasis of neuroblastoma cells. Int J Oncol. 2012; 41:1419-24.

https://doi.org/10.3892/ijo.2012.1552 PMID:22797702

35. Struyf S, Proost P, Vandercappellen J, Dempe S, Noyens B, Nelissen S, Gouwy M, Locati M, Opdenakker G, Dinsart C, Van Damme J. Synergistic up-regulation of MCP-2/CCL8 activity is counteracted by chemokine cleavage, limiting its inflammatory and anti-tumoral effects. Eur J Immunol. 2009; 39:843-57.

https://doi.org/10.1002/eji.200838660 PMID:19224633

- 36. Barbai T, Fejős Z, Puskas LG, Tímár J, Rásó E. The importance of microenvironment: the role of CCL8 in metastasis formation of melanoma. Oncotarget. 2015; 6:29111-28. https://doi.org/10.18632/oncotarget.5059 PMID:26320180
- 37. Qiu H, Hu X, He C, Yu B, Li Y, Li J. Identification and validation of an individualized prognostic signature of bladder cancer based on seven immune related genes. Front Genet. 2020; 11:12. https://doi.org/10.3389/fgene.2020.00012

PMID:32117435

- 38. Myers JA, Miller JS. Exploring the NK cell platform for cancer immunotherapy. Nat Rev Clin Oncol. 2020. [Epub ahead of print]. https://doi.org/10.1038/s41571-020-0426-7 PMID:32934330
- 39. Krzewski K, Chen X, Strominger JL. WIP is essential for lytic granule polarization and NK cell cytotoxicity. Proc Natl Acad Sci USA. 2008; 105:2568-73. https://doi.org/10.1073/pnas.0711593105 PMID:18258743
- 40. Choy CT, Wong CH, Chan SL. Embedding of genes using cancer gene expression data: biological relevance and potential application on biomarker discovery. Front Genet. 2019; 9:682. https://doi.org/10.3389/fgene.2018.00682 PMID:30662451
- 41. Ueno Y, Ozaki S, Umakoshi A, Yano H, Choudhury ME, Abe N, Sumida Y, Kuwabara J, Uchida R, Islam A, Ogawa K, Ishimaru K, Yorozuya T, et al. Chloride intracellular channel protein 2 in cancer and noncancer human tissues: relationship with tight junctions. Tissue Barriers. 2019; 7:1593775. https://doi.org/10.1080/21688370.2019.1593775 PMID:30929599
- 42. Piras F, Colombari R, Minerba L, Murtas D, Floris C, Maxia C, Corbu A, Perra MT, Sirigu P. The predictive value of CD8, CD4, CD68, and human leukocyte antigen-D-related cells in the prognosis of cutaneous malignant melanoma with vertical growth phase. Cancer. 2005; 104:1246-54.

https://doi.org/10.1002/cncr.21283 PMID:16078259

- 43. Giese MA, Hind LE, Huttenlocher A. Neutrophil plasticity in the tumor microenvironment. Blood. 2019; 133:2159–67. <u>https://doi.org/10.1182/blood-2018-11-844548</u> PMID:<u>30898857</u>
- 44. Yang S, Liu T, Nan H, Wang Y, Chen H, Zhang X, Zhang Y, Shen B, Qian P, Xu S, Sui J, Liang G. Comprehensive analysis of prognostic immune-related genes in the tumor microenvironment of cutaneous melanoma. J Cell Physiol. 2020; 235:1025–1035. https://doi.org/10.1002/jcp.29018 PMID:31240705
- Zhao Y, Schaafsma E, Gorlov IP, Hernando E, Thomas NE, Shen R, Turk MJ, Berwick M, Amos CI, Cheng C. A leukocyte infiltration score defined by a gene signature predicts melanoma patient prognosis. Mol Cancer Res. 2019; 17:109–19. <u>https://doi.org/10.1158/1541-7786.MCR-18-0173</u>

PMID:30171176

- 46. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, Angell H, Fredriksen T, Lafontaine L, Berger A, Bruneval P, Fridman WH, Becker C, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity. 2013; 39:782–95. https://doi.org/10.1016/j.immuni.2013.10.003 PMID:24138885
- 47. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013; 14:7. <u>https://doi.org/10.1186/1471-2105-14-7</u> PMID:<u>23323831</u>
- Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015; 12:453–57. <u>https://doi.org/10.1038/nmeth.3337</u> PMID:<u>25822800</u>
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015; 43:e47.

https://doi.org/10.1093/nar/gkv007 PMID:25605792

 Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019; 10:1523.

https://doi.org/10.1038/s41467-019-09234-6 PMID:<u>30944313</u>

 Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, von Mering C. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015; 43:D447–52.

https://doi.org/10.1093/nar/gku1003 PMID:25352553

- 52. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13:2498–504. <u>https://doi.org/10.1101/gr.1239303</u> PMID:14597658
- 53. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; 9:559. <u>https://doi.org/10.1186/1471-2105-9-559</u> PMID:19114008
- 54. Zhang M, Wang X, Chen X, Zhang Q, Hong J. Novel immune-related gene signature for risk stratification and prognosis of survival in lower-grade glioma. Front Genet. 2020; 11:363. <u>https://doi.org/10.3389/fgene.2020.00363</u> PMID:<u>32351547</u>
- Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B, Liu XS. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. Cancer Res. 2017; 77:e108–10.

https://doi.org/10.1158/0008-5472.CAN-17-0307 PMID:29092952

SUPPLEMENTARY MATERIALS

Supplementary Figures



Supplementary Figure 1. GSEA analysis comparing the immune phenotype between the two immune-related groups. (A, B) In high immune cell infiltration group, (A) the significant enrichment of top 5 GO terms, and (B) the significant enrichment of 5 immune-related KEGG pathways.



Survival curve (p=1.507e-04)

Supplementary Figure 2. Association between two immune-related groups and OS for melanoma patients.



Supplementary Figure 3. The forest plot for the 7-gene prognostic signature.



Supplementary Figure 4. Validation of the 7-gene prognostic signature. (A, B) Kaplan-Meier survival curves and ROC curves of OS in GSE54467. (C, D) Kaplan-Meier survival curves and ROC curves of DSS in GSE65904.



Supplementary Figure 5. Assessment of the independent prognostic value of the prognostic signature. The (A) 1-year ROC curves, (B) 3-year ROC curves, and (C) 5-years ROC curves for the signature and clinical features. (D) The univariate Cox regression analysis. (E) The multivariate Cox regression analysis.



Supplementary Figure 6. Evaluation of the OS nomogram. (A, B) The (A) 3-year and (B) 5-year ROC curves. (C, D) The calibration curves for the probability of (C) 3-year and (D) 5-year OS.



Supplementary Figure 7. Relationship between the prognostic signature and stromal score. (A, B) The high-risk group has lower stromal scores compared with the low-risk group in (A) the TCGA dataset and (B) GSE54467. (C–I) The association between stromal score and the expression of each gene in the 7-gene prognostic signature: (C) CYTL1, (D) CCL8, (E) FCGR2C, (F) OAS1, (G) HAPLN3, (H) WIPF1, (I) CLIC2.



Supplementary Figure 8. Association between the gene expression level and six subtypes of TIICs. (A) CYTL1, (B) CCL8, (C) FCGR2C, (D) OAS1, (E) HAPLN3, (F) WIPF1, (G) CLIC2.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1, 2.

Supplementary Table 1. Kaplan-Meier survival analysis.

Supplementary Table 2. The genes in the selected module identified by WGCNA.

Supplementary Table 3. Multivariate Cox regression analysis.

ID	Coef	HR	HR.95L	HR.95H	p-Value
CYTL1	0.007408	1.007435	1.002858	1.012033	0.00143
CCL8	-0.033	0.967543	0.929431	1.007217	0.107561
FCGR2C	-0.02117	0.979052	0.960903	0.997543	0.026581
OAS1	-0.0147	0.985407	0.968316	1.0028	0.099617
HAPLN3	-0.03669	0.963978	0.937684	0.991009	0.009321
WIPF1	-0.01375	0.98634	0.975503	0.997297	0.01468
CLIC2	-0.02314	0.97713	0.949366	1.005705	0.115682

Coef: Coefficient; HR: hazard ratio.