The prediction of survival and analysis of microenvironment in patients with sarcoma

Jiarui Chen
The first affiliated hospital of Guangxi Medical University

Chong Liu
The first affiliated hospital of Guangxi Medical University

Tuo Liang
The first affiliated hospital of Guangxi Medical University

Zide Zhang
The first affiliated hospital of Guangxi Medical University

Zhaojun Lu
The first affiliated hospital of Guangxi Medical University

Shian Liao
The first affiliated hospital of Guangxi Medical University

Jie Jiang
The first affiliated hospital of Guangxi Medical University

Guoyong Xu
The first affiliated hospital of Guangxi Medical University

Hao Li
The first affiliated hospital of Guangxi Medical University

Tianyou Chen
The first affiliated hospital of Guangxi Medical University

Zhaojie Qin
The first affiliated hospital of Guangxi Medical University

Huijiang Liu
Nanning First People’ Hospital

Jiemei Cen (✉️ 935270023@qq.com)
The first affiliated hospital of Guangxi Medical University

Xinli Zhan
The first affiliated hospital of Guangxi Medical University

Research

Keywords: sarcoma, tumor microenvironment, immune gene, prognostic signature, TCGA
Abstract

Background:

Sarcomas were rare, aggressive, and heterogeneous group of tumors. The degree of tumor microenvironment cells, infiltrating immune cells and stromal cells in the tumor had an important impact on prognosis of sarcoma. The aim of this study was to identify the differentially expressed genes (DEGs) association with immune of sarcoma and potential prognostic immune biomarkers for predicting survival of sarcoma patients.

Methods:

The gene expression data and clinical data of sarcoma was downloaded from The Cancer Genome Atlas (TCGA) dataset. The immune scores and stromal scores were calculated by ESTIMATE algorithm. The limma package was used to identify the immune DEGs. ClusterProfiler package and STRING were further to analysis the immune DEGs. A prognostic signature was built based the immune gene and clinical data by univariate Cox regression analysis and multivariable Cox analysis. Finally, the prognostic signature was evaluated by functional assessment and the Gene Expression Omnibus (GEO) database.

Results:

The functional enrichment showed that the up immune DEGs was associated with immune cell activation, proliferation, and adhesion. A single prognostic signature was constructed with seven immune genes, and patient with high risk scores had a worse survival than those with low risk scores. Evaluating showed that the prognostic signature performed well and served as an independent factor in sarcoma.

Conclusions:

The chemokine receptors or chemokine ligands and immune related genes played a significant role in the mechanism of sarcoma, and the novel immune gene expression had a significant clinical guidance for the prognosis of patient in sarcoma.

Background

Sarcomas, derived from mesenchymal tissue, are malignant tumors comprising more than 50 different entities[1–3] with high histological heterogeneity within each subtype[4]. According to statistics reported, more than 10 thousand patients diagnosed with sarcoma and 3800 would be dead in sarcoma per year in the United States[5, 6]. The patients in sarcoma is poor prognosis with a high rate of recurrence after initial resection. However, the pathological mechanism of sarcoma is still unclear, and we are still helpless in the treatment of it.
The tumor microenvironment (TME) is the cellular milieu consisting mesenchymal cells, endothelial cells, tumor infiltrating immune cells (TIIC), extracellular matrix (ECM) molecules and inflammatory mediators[7], which can provide the metabolites and controls factors of proliferation and diffusion for many tumors. Tumor microenvironment has also been reported to seriously affect gene expression in tumor tissues[8–11]. Immune cells and stromal cells are the two main types of non-tumor components in the tumor microenvironment, and are considered to have important significance for the diagnosis and prognostic evaluation of tumors[12]. It has been confirmed to be related to relapse and overall survival[13–15].

To expose the immune microenvironment of sarcoma and the underlying pathological mechanism, the ESTIMATE algorithm[10] was performed to calculate the immune score of sarcoma from the TCGA database(https://portal.gdc.cancer.gov/)[16, 17], and limma package was used to analyzed immune DEGs based the immune score. With the data from IMMPORT website(https://www.immport.org/), we further screen out the immune genes related sarcoma. Finally, an immune gene prognostic signature was constructed with the immune genes.

**Materials And Methods**

**Data collection**

The level 3 gene expression profile (level 3 data) and clinical information of sarcoma was downloaded from TCGA database. The gene expression RNAseq-HTSeq–Counts was from Illumina. Clinical data consisted age, gender, survival, outcome, and histological type, et al.

**Data processing**

Based on R software(version 4.0.0), the gene expression data were standardized and ESTIMATE algorithm[10] contributed to attained the immune scores and stromal scores. The differentially expressed genes(DEGs) was analyzed by limma package[18] with a level of llogFCI ≥1.5 and adjusted P value ≤0.05.

**Enrichment analysis of DEGs**

Gene Ontology (GO) is an ontology widely used in the field of bioinformatics for annotating large scale genes and gene products[19]. Kyoto Encyclopedia of Genes and Genomes (KEGG) can be used to predict which pathways a particular gene is enriched[20]. ClusterProfiler package[21], an R package for comparing biological themes, completed analyzed the GO enrichment and KEGG pathway enrichment.

**PPI network construction**

We uploaded all the up regulated DEGs and down regulated DEGs to the online tool, STRING database (https://string-db.org/) and set the level at a high confidence (0.700) to construct the protein-protein interaction (PPI) network. The data of PPI networks was download from STRING and was inputted into
Cytoscape software[22] (version 3.7.1) for reconstruction. The plug in Molecular Complex Detection (MCODE) analyzed the PPI sub networks, and the sub networks would be considered only with a level of MCODE scores ≥10 and number of nodes ≥10.

Analysis of immune genes and construction of prognostic model

IMMPORT shared data enables searching and downloading shared biomedical research data(https://www.immport.org/), and we could obtain the immune related genes from it. Based on the immune related genes from IMMPORT and immune DEGs we analyzed, we obtained the immune genes of sarcoma. With the Survival package, a univariate Cox regression analysis was to examine the correlation of the expression level of immune genes and overall survival (OS) of patients with sarcoma. Genes with a P-value <0.05 were considered prognostic immune genes with expression levels significantly related to OS. And the prognostic immune genes selected by the “step” function were used for construction of the prognostic model. By fitting the prognosis immune genes to the multivariate Cox regression analysis with OS as the dependent variable, the relative contribution of these prognosis immune genes in sarcoma survival prediction was evaluated. Through the linear combination of the expression levels of immune genes with the multivariate Cox regression coefficient(α) as the weight, an immune expression-based prognostic risk score model was constructed. And the risk score (RS) was calculated as follows: RS = expression of gene₁ × α₁gene₁+expression of gene₂ × α₂gene₂+expression of gene₃ × α₃gene₃+...expression of geneₙ × αₙgeneₙ. And the sarcoma patients could be divided into high risk group and low risk group by the median risk scores of the immune gene prognostic model.

Evaluation and verification of prognostic model

To test the reliability of the prognostic model we constructed, receiver operating characteristic (ROC) curve was generated by survivalROC package in the R software to evaluate it. The univariate Cox regression analysis and multivariate Cox regression analysis would further assess whether the prognostic model can independently assess the prognosis of patients. In addition, the prognostic model was performed in the sarcoma data from GEO Datasets (https://www.ncbi.nlm.nih.gov/geo/).

Statistical analysis

FDRs(False discovery rate, adjusted p value) in limma and clusterProfiler were adjusted for multiple testing with the Benjamini-Hochberg procedure to control FDR[23]. Univariate analysis was performed using the log-rank test. Coexpression relationships was performed by the Pearson correlation coefficient. P value <0.05 was considered statistically significant. And statistical analyses were completed by R software (version 4.0.0), SPSS22.0(IBM Corporation, Armonk, NY, USA), and GraphPad Prism 8.

Results

The association of clinical data with immune scores and stromal scores
We downloaded 263 samples gene expression profiles and clinical data from TCGA dataset. Among them, males and females counted 119 (45.25%) and 144 (54.75%), respectively. Age covered from 20 to 90 years old (60.56±14.63). Leiomyosarcoma histologic subtype included 67 (25.48%) cases of conventional leiomyosarcoma, 34 (12.93%) cases of poorly differentiated or pleomorphic or epithelioid leiomyosarcoma, 4 (1.52%) cases of well differentiated leiomyosarcoma, and 158 (60.08%) cases were of unknown histological result. Positive and negative tumor margins were found in 74 (28.14%) and 139 (52.85%) cases, respectively, and 50 (19.01%) cases were unknown. With the ESTIMATE algorithm, immune scores ranged from -2039.13 to 3295.15 (433.88±1116.81), and stromal scores were distributed between -1275.33 to 2519.30 (882.36±718.82). The average immune scores was higher than stromal scores obviously (P<0.001). From the result, the male had a higher immune score and stromal score than the female (Figure 1 A-B, p<0.05). However, there was no obvious association between subtype and immune or stromal scores (Figure 1 C-D, p>0.05). According to the level of the score, we divided the patients into two groups, namely, high immune scores or stromal scores group and low immune scores or stromal scores group. The high immune or stromal scores group of patients had an older age than low immune or stromal scores, though there was no statistical significance (Figure 1 E-F, p>0.05). And with Kaplan-Meier algorithm, a low immune scores or stromal indicated that the patient had a worse survival, though, there was no statistically significant difference between stromal scores and survival prognosis(Figure 1 G, p<0.05, Figure 1 H, p>0.05).

Identification of significantly differentially expressed genes

To find out the relation of gene expression and immune scores or stromal scores, we analyzed the differentially expressed genes between high immune scores group and low immune scores group, as well as between high stromal scores group and low stromal scores group by limma package in R software. The heatmaps of differentially expressed genes were conducted by heatmap package (Figure 2 A-B). With a level of \( \text{logFC} \geq 1.5 \) and adjust p value <0.05, the volcano of DEGs were drawn(Figure 2 C-D), there were 1299 genes differentially expressed in high immune scores group compared with low immune scores, including 732 up regulated expressed genes and 567 down regulated genes. A total of 1223 genes were identified differentially expressed in high stromal group compared with low stromal group, including 520 up regulated genes and 703 down regulated genes. Interestingly, a total of 410 genes belonged to both up immune DEGs and up stromal DEGs, and there were 498 genes were intersected in down regulated DEGs and down stromal DEGs (Figure 2 E-F).

The functional analysis of DEGs

To reveal the correlation of immune with sarcoma, we had a deeper analyzed the function of DEGs compared by high immune scores and low immune scores, we analyzed the function of 732 up regulated genes and 567 down regulated genes with ClusterProfiler package in R software. GO term were divided into three aspects, including biology: biological process (BP), molecular function (MF), and cellular component (CC). And a total of 1193 GO terms were enriched by high regulated DEGs, including 1015 BPs, 76 CCs, and 102 MFs. GO result showed that the up regulated genes enriched in immune cell
activation, proliferation, and adhesion (Figure 3 A). It was revealed immune process played a significant role in sarcoma. And the down regulated DEGs were enriched in 477 GO Terms, including 324 BPs, 105 CCs, and 48 MFs. Down regulated DEGs were related to various muscle contractions (Figure 3 B). KEGG results showed that up regulated DEGs enriched in eight pathways, including Cytokine-cytokine receptor interaction, Viral protein interaction with cytokine and cytokine receptor, Chemokine signaling pathway, Osteoclast differentiation, and Cell adhesion molecules (CAMs) (Figure 3 C). And down regulated DEGs enriched in ten pathways, including Calcium signaling pathway, cAMP signaling pathway, Adrenergic signaling in cardiomyocytes, Vascular smooth muscle contraction, and Cardiac muscle contraction (Figure 3 D).

The construction of PPI network

We constructed the PPI network of DEGs to better understand the intrinsic interaction between the identified DEGs. With the data constructed by STRING and reconstructed by Cytoscape, we got the PPI networks of up regulated DEGs and down regulated DEGs, respectively (Supplement 1 A-B). To further analyzed the PPI network with the plug-in MCODE, with a level of Score (Density*#Nodes) ≥10 and Nodes ≥10, we got 4 subnets and one subnet from up regulated DEGs PPI network and down regulated PPI network, respectively(Figure 4 A-D). In the subnet of down regulated DEGs PPI, GNG4 owned the highest degree 24 followed by ADCY5 and ADCY2(degree=17). According to the GO result, we knew up regulated DEGs were associated with immune strongly. And the subnet 1 of up regulated DEGs, we called the top one subnet, was the most prominent subnet for owning only 36 nodes but with the most edges and highest scores. Thus, we decided to focus on this subnet to analyzed deeper. In the top one subnet, FPR2 owned the highest degree of 72, followed by C3(degree=66), C3AR1(degree=66), CXCL8(degree=66), CCR5(degree=65), CCR7(degree=65), CCR2(degree=59), CXCL1(degree=59), CXCL10(degree=57), and CXCR3(degree=57). Among the top ten genes, the CCR gene family and the CXCL gene family account for three. And the CCR gene family (CCR1, CCR2, CCR4, CCR5, and CCR7) and the CXCL gene family (CXCL1, CXCL2, CXCL6, CXCL8, CXCL9, CXCL10, CXCL11, and CXCL13) account for five and eight in the top one subnet, respectively. In addition, the CCL gene family (CCL4, CCL4L1, CCL5, CCL13, CCL19, and CCL20) was also significantly enriched in this subnetwork, accounting for 6. A matrix graph of Pearson's correlation analysis indicated that the expression of CCR family genes, some of CXCL family genes and CCL family genes were closely related to each other internally with a correlation coefficient ≥ 0.5(Figure 5 A-C). The numbers in each grid represent the correlations between the corresponding genes. In addition, we found that the high expression of CCR family genes indicated a better survival in a short time, though, there were no statistical significance in CCR1 and CCR4(Figure 6).

Construction of the immune genes’ expression-based prognostic signature

Based on the immune scores, we analyzed the differentially expressed genes. To further search for immune related genes, we download the immune genes from IMMPORT, and took the intersection of DEGs we analyzed and immune related genes from IMMPORT. A total of 85 genes were the common genes, that was the immune genes related to sarcoma. With the survival package, a univariate Cox
Regression analysis was to examine the correlation of the expression level of immune genes and overall survival of patients with sarcoma. A total of ten genes were considered as prognostic immune genes and selected using the “step” function to screen for the optimal combination (Figure 7 A). The follow seven prognostic immune genes were used for the prognostic model construction: BTC, EDN3, GREM2, IL17B, PAK3, RAC3, and ZC3HAV1L (Table 1). Survival analysis indicated that the prognostic model could separate the sarcoma into high risk group and low risk group, and patient with a high risk scores had a worse survival than those with a low risk scores. (Figure 7 B-C). Wilcoxon rank sum test was used to test the difference between the seven prognostic immune genes in the high-risk group and the low-risk group, and we found that its expression in the two groups was significantly different (Figure 7 D).

Evaluation and verification of prognostic model

ROC curve showed that the risk score performed well for the survival prediction of sarcoma, and the AUC of ROC curve was 0.765 (Figure 8 A). After the univariate Cox regression analysis and multivariate Cox regression analysis, we found the risk score was an independent prognostic factor in sarcoma (Figure 8 B). In addition, GSE17674 was downloaded from GEO Datasets verified the prognostic models. GSE17674, gene expression profiling was from GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array, contained forty-four Ewing sarcoma patient clinical data. Background correction, normalization and standardization of gene expression data were performed by RMA - Robust Multichip Average algorithm[24]. The risk scores of patients from GSE17674 were calculated and the Kaplan-Meier curves of event-free survival and overall survival were drawled (Figure 9). The patients with high risk scores had a worse event-free survival and overall survival, though, p>0.05 in the event-free survival.

Discussion

Sarcoma, a malignancy tumor with a high recurrence rate, brings us big challenges in treatment. Analyzing the tumor microenvironment by CIBERSORT algorithm to calculate the immune scores and stromal scores, there have been some literatures reported that there were a significant relationship between tumor microenvironment and tumor prognosis[25, 26]. Though, Runzhi Huang el at[11]. have analyzed the relationship between sarcoma recurrence and tumor immune microenvironment with TCGA data, this is the first report to predict the prognosis of sarcoma patients by using TCGA data to calculate the immune score to construct a prognostic model. In our study, we analyzed the immune scores and stromal scores. To best understand the relationship between immune and sarcoma, we analyzed the function of immune DEGs. The high expressed immune DEGs were enriched in immune-related biological process aspects such as T cell activation, regulation of T cell activation, leukocyte cell-cell adhesion, regulation of lymphocyte activation, and leukocyte proliferation, mononuclear cell proliferation, lymphocyte proliferation, regulation of leukocyte proliferation, regulation of mononuclear cell proliferation, regulation of leukocyte cell-cell adhesion, et al. They also enriched in immune-related pathways, including Cytokine-cytokine receptor interaction, Viral protein interaction with cytokine and cytokine receptor, Chemokine signaling pathway, Osteoclast differentiation, Cell adhesion molecules
In the first subnet of high expressed immune DEGs analyzed by MCODE, CCR gene family, CXCL gene family, and CCL gene family were clearly clustered. They had a significant coexpression relationship, especially CCR genes family. And high expression of CCR family genes indicated a better survival in a short time. CCRs (C-C motif chemokine receptors), CXCLs (C-X-C motif chemokine ligands), and CCLs (C-C motif chemokine ligands) belongs to chemokine receptors or chemokine ligands. Chemokines, chemokine receptors, and chemokine ligands play a significant role in cell migration, immune surveillance, inflammation, and tumorigenesis [27–29]. Numerous studies reported that chemokines and chemokine receptors were association with the tumor growth[30] via numerous mechanisms[31]. Chemokine receptors can inhibit the proliferation of tumor cells. Such as CCR1 reduce the proliferation of human liver cancer cells[32], CCR5 inhibits breast cancer progression in a p53-dependent manner[33]. CCR7-mediated leukocyte migration is greatly significant in the normal immune response. CCR7 recruits natural T cells and activated dendritic cells to the lymph nodes and initiate an adaptive immune response[34]. CCL19 is one of CCR7 ligands[35, 36], and can activates T-cells[37]. Inhibition of the CCR7 gene causes T cells to home to secondary lymphoid organs with significantly impaired function[38]. In addition, CXCL1 and CXCL2 have been shown to play a role in the growth of gastric cancer, adenocarcinoma, pancreatic cancer, and melanoma, lung cancer[39, 40].

To further analyzed the relation of sarcoma and immune, we screened out the immune genes related to sarcoma by immune DEGs and immune genes from IMMPOTT, and performed the univariable analyzed and multivariable Cox analyses to construct a prognostic model for predicting the prognosis of patients of sarcoma. A total of seven immune genes were used to construct the prognostic model. BTC, GREM2, IL17B, and PAK3 were protective immune genes related to sarcoma, while EDN3, RAC3, and ZC3HAV1L were risk immune genes related to sarcoma. EDN3 (endothelin 3) is an important paracrine factor for embryonic melanocyte proliferation, migration, and survival during fetal development[41]. Its involvement was confirmed in the development of malignant melanoma[42] and choriocarcinoma[43]. And Liu Y, et al. reported that EDN3 was highly produced by glioblastoma stem cells and was an essential mitogen for NCC development and migration[44]. RAC3 (Rac family small GTPase 3), a member of the Rac family of small guanosine triphosphatases (GTPases), played a significant role in regulating tumor metastasis and the tumorigenic process[45, 46]. Liu TQ demonstrated that knockdown the RAC3 obviously suppressed the proliferation and colony formation of lung cancer cells[47]. However, Wang G performed that high expression of RAC3 is associated with a longer survival of lung adenocarcinoma patients[48]. Until now, there are still very few reports on the ZC3HAV1L (zinc finger CCCH-type containing, antiviral 1 like) gene, especially the relationship between ZC3HAV1L and tumors. And we are the first to report that ZC3HAV1L is a risk factor for sarcoma.

**Conclusion**
To sum up, we were the first to analyzed the differential expressed genes based the different microenvironment in sarcoma and further revealed the mechanism of sarcoma, especially the role of CCRs, CXCLs, and CCLs in sarcoma. Combined with the genes from the IMMPORT website, we obtained the sarcoma-related immune genes, and constructed immune gene prognostic model to predict the survival of patients in sarcoma well.

Abbreviations

| Abbreviation | Description |
|--------------|-------------|
| DEGs         | Differentially expressed genes |
| TCGA         | The Cancer Genome Atlas |
| GEO          | Gene Expression Omnibus |
| TME          | Tumor microenvironment |
| TIIC         | Tumor infiltrating immune cells |
| ECM          | Extracellular matrix |
| GO           | Gene Ontology |
| KEGG         | Kyoto Encyclopedia of Genes and Genomes |
| PPI          | protein-protein interaction |
| MCODE        | Molecular Complex Detection |
| OS           | overall survival |
| RS           | risk score |
| ROC          | receiver operating characteristic |
| CCR          | C-C motif chemokine receptor |
| CXCL         | C-X-C motif chemokine ligand |
| CCL          | C-C motif chemokine ligand |
| EDN3         | endothelin 3 |
| RAC3         | Rac family small GTPase 3 |
| ZC3HAV1L     | zinc finger CCCH-type containing antiviral 1 like |
| FDR          | False discovery rate |

Declarations

The authors have no conflicts of interest to declare.
Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors agree to publish this article.

Availability of data and material

All data can be obtained from open and free data on the Internet.

Competing interests

The authors have no conflicts of interest to declare.

Funding

Not applicable.

Authors' contributions

JC and CL designed the study, collected data, and wrote the manuscript. JC and XZ revised the manuscript. GX, ZZ, ZL translated and revised the manuscript. JJ, HL, TC, ZQ prepared all the tables and figures. HL, TL, SL analyzed the result. JC and TL contributed to the design and testing of the data processing. All authors have read and approved the manuscript.

Acknowledgements

Thanks to Miss. Luo Jinying, a member of Shu Lun Pan in Hong Kong, for supporting to download the data of patients from TCGA data(Email: 476265104@qq.com).

References

1. Jo VY, Fletcher CD: WHO classification of soft tissue tumours: an update based on the 2013 (4th) edition. Pathology 2014, 46(2):95–104.
2. Brooks JS, Lee S. Contemporary diagnostics: sarcoma pathology update. Journal of surgical oncology. 2015;111(5):513–9.
3. Ducimetière F, Lurkin A, Ranchère-Vince D, Decouvelaere AV, Péoc’h M, Istier L, Chalabreysse P, Muller C, Alberti L, Bringuier PP, et al. Incidence of sarcoma histotypes and molecular subtypes in a prospective epidemiological study with central pathology review and molecular testing. PloS one. 2011;6(8):e20294.
4. Massarweh NN, Dickson PV, Anaya DA. Soft tissue sarcomas: staging principles and prognostic nomograms. Journal of surgical oncology. 2015;111(5):532–9.
5. Stiller CA, Trama A, Serraino D, Rossi S, Navarro C, Chirlaque MD, Casali PG. Descriptive epidemiology of sarcomas in Europe: report from the RARECARE project. European journal of cancer (Oxford England: 1990). 2013;49(3):684–95.

6. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. Cancer J Clin. 2009;59(4):225–49.

7. Hanahan D, Coussens LM. Accessories to the crime: functions of cells recruited to the tumor microenvironment. Cancer cell. 2012;21(3):309–22.

8. Curry JM, Sprandio J, Cognetti D, Luginbuhl A, Bar-ad V, Pribitkin E, Tuluc M. Tumor microenvironment in head and neck squamous cell carcinoma. Seminars in oncology. 2014;41(2):217–34.

9. 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. Journal of translational medicine. 2012;10:1.

10. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nature communications. 2013;4:2612.

11. Huang R, Meng T, Chen R, Yan P, Zhang J, Hu P, Zhu X, Yin H, Song D, Huang Z. The construction and analysis of tumor-infiltrating immune cell and ceRNA networks in recurrent soft tissue sarcoma. Aging. 2019;11(22):10116–43.

12. Jia D, Li S, Li D, Xue H, Yang D, Liu Y. Mining TCGA database for genes of prognostic value in glioblastoma microenvironment. Aging. 2018;10(4):592–605.

13. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pagès C, Tosolini M, Camus M, Berger A, Wind P, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science. 2006;313(5795):1960–4.

14. Geiger J, Hutchinson R, Hohenkirk L, McKenna E, Chang A, Mulé J. Treatment of solid tumours in children with tumour-lysate-pulsed dendritic cells. Lancet. 2000;356(9236):1163–5.

15. Wang C, Zhou X, Li W, Li M, Tu T, Ba X, Wu Y, Huang Z, Fan G, Zhou G, et al. Macrophage migration inhibitory factor promotes osteosarcoma growth and lung metastasis through activating the RAS/MAPK pathway. Cancer letters. 2017;403:271–9.

16. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemporary oncology (Poznan Poland). 2015;19(1a):A68–77.

17. Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet. 2013;45(10):1113–20.

18. 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(7):e47.

19. Gene Ontology C. The Gene Ontology (GO) project in 2006. Nucleic acids research. 2006;34(Database issue):D322–6.
20. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic acids research. 2017;45(D1):D353–61.

21. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics. 2012;16(5):284–7.

22. 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(11):2498–504.

23. Reiner A, Yekutieli D, Benjamini Y. Identifying differentially expressed genes using false discovery rate controlling procedures. Bioinformatics. 2003;19(3):368–75.

24. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics (Oxford England). 2003;4(2):249–64.

25. Xu WH, Xu Y, Wang J, Wan FN, Wang HK, Cao DL, Shi GH, Qu YY, Zhang HL, Ye DW. Prognostic value and immune infiltration of novel signatures in clear cell renal cell carcinoma microenvironment. Aging. 2019;11(17):6999–7020.

26. Zhang C, Zheng JH, Lin ZH, Lv HY, Ye ZM, Chen YP, Zhang XY. Profiles of immune cell infiltration and immune-related genes in the tumor microenvironment of osteosarcoma. Aging. 2020;12(4):3486–501.

27. Van Damme J, Proost P, Lenaerts JP, Opdenakker G. Structural and functional identification of two human, tumor-derived monocyte chemotactic proteins (MCP-2 and MCP-3) belonging to the chemokine family. The Journal of experimental medicine. 1992;176(1):59–65.

28. Fujiwara H, Hamaoka T. Coordination of chemokine and adhesion systems in intratumoral T cell migration responsible for the induction of tumor regression. Int Immunopharmacol. 2001;1(4):613–23.

29. Raman D, Sobolik-Delmaire T, Richmond A. Chemokines in health and disease. Experimental cell research. 2011;317(5):575–89.

30. Sarvaiya PJ, Guo D, Ulasov I, Gabikian P, Lesniak MS. Chemokines in tumor progression and metastasis. Oncotarget. 2013;4(12):2171–85.

31. Lacalle RA, Blanco R, Carmona-Rodríguez L, Martín-Leal A, Mira E, Mañes S. Chemokine Receptor Signaling and the Hallmarks of Cancer. International review of cell molecular biology. 2017;331:181–244.

32. Lu P, Nakamoto Y, Nemoto-Sasaki Y, Fujii C, Wang H, Hashii M, Ohmoto Y, Kaneko S, Kobayashi K, Mukaida N. Potential interaction between CCR1 and its ligand, CCL3, induced by endogenously produced interleukin-1 in human hepatomas. Am J Pathol. 2003;162(4):1249–58.

33. Mañes S, Mira E, Colomer R, Montero S, Real LM, Gómez-Moutón C, Jiménez-Baranda S, Garzón A, Lacalle RA, Harshman K, et al. CCR5 expression influences the progression of human breast cancer in a p53-dependent manner. The Journal of experimental medicine. 2003;198(9):1381–9.
34. Villablanca EJ, Raccosta L, Zhou D, Fontana R, Maggioni D, Negro A, Sanvito F, Ponzoni M, Valentinis B, Bregni M, et al. Tumor-mediated liver X receptor-alpha activation inhibits CC chemokine receptor-7 expression on dendritic cells and dampens antitumor responses. Nature medicine. 2010;16(1):98–105.

35. Kim CH, Pelus LM, Appelbaum E, Johanson K, Anzai N, Broxmeyer HE. CCR7 ligands, SLC/6Ckine/Exodus2/TCA4 and CKbeta-11/MIP-3beta/ELC, are chemoattractants for CD56+(+)CD16(-) NK cells and late stage lymphoid progenitors. Cellular immunology. 1999;193(2):226–35.

36. Yoshida R, Nagira M, Imai T, Baba M, Takagi S, Tabira Y, Akagi J, Nomiyama H, Yoshie O. EBI1-ligand chemokine (ELC) attracts a broad spectrum of lymphocytes: activated T cells strongly up-regulate CCR7 and efficiently migrate toward ELC. Int Immunol. 1998;10(7):901–10.

37. Steinman RM, Pack M, Inaba K. Dendritic cells in the T-cell areas of lymphoid organs. Immunological reviews. 1997;156:25–37.

38. Gunn MD, Kyuwa S, Tam C, Kakiuchi T, Matsuzawa A, Williams LT. Nakano HJJ:EM: Mice Lacking Expression of Secondary Lymphoid Organ Chemokine Have Defects in Lymphocyte Homing and Dendritic Cell Localization. 1999, 189(3):451–460.

39. Richmond A, Thomas HG. Purification of melanoma growth stimulatory activity. Journal of cellular physiology. 1986;129(3):375–84.

40. Cheng ZH, Shi YX, Yuan M, Xiong D, Zheng JH, Zhang ZY. Chemokines and their receptors in lung cancer progression and metastasis. Journal of Zhejiang University Science B. 2016;17(5):342–51.

41. Tang L, Su M, Zhang Y, Ip W, Martinka M, Huang C, Zhou Y. Endothelin-3 is produced by metastatic melanoma cells and promotes melanoma cell survival. J Cutan Med Surg. 2008;12(2):64–70.

42. Spinella F, Rosanò L, Di Castro V, Decandia S, Nicotra MR, Natali PG, Bagnato A. Endothelin-1 and endothelin-3 promote invasive behavior via hypoxia-inducible factor-1alpha in human melanoma cells. Cancer research. 2007;67(4):1725–34.

43. Rauh A, Windischhofer W, Kovacevic A, DeVaney T, Huber E, Semlitsch M, Leis HJ, Sattler W, Malle E. Endothelin (ET)-1 and ET-3 promote expression of c-fos and c-jun in human choriocarcinoma via ET(B) receptor-mediated G(i)- and G(q)-pathways and MAP kinase activation. Br J Pharmacol. 2008;154(1):13–24.

44. Liu Y, Ye F, Yamada K, Tso JL, Zhang Y, Nguyen DH, Dong Q, Soto H, Choe J, Dembo A, et al. Autocrine endothelin-3/endothelin receptor B signaling maintains cellular and molecular properties of glioblastoma stem cells. Molecular cancer research: MCR. 2011;9(12):1668–85.

45. Chatterjee M, Sequeira L, Jenkins-Kabaila M, Dubyk CW, Pathak S, van Golen KL. Individual rac GTPases mediate aspects of prostate cancer cell and bone marrow endothelial cell interactions. Journal of signal transduction. 2011;2011:541851.

46. Gest C, Joimel U, Huang L, Pritchard LL, Petit A, Dulong C, Buquet C, Hu CQ, Mirshahi P, Laurent M, et al. Rac3 induces a molecular pathway triggering breast cancer cell aggressiveness: differences in MDA-MB-231 and MCF-7 breast cancer cell lines. BMC Cancer. 2013;13:63.
47. Liu TQ, Wang GB, Li ZJ, Tong XD, Liu HX. Silencing of Rac3 inhibits proliferation and induces apoptosis of human lung cancer cells. Asian Pacific journal of cancer prevention: APJCP. 2015;16(7):3061–5.

48. Wang G, Wang H, Zhang C, Liu T, Li Q, Lin X, Xie J, Liu H. Rac3 regulates cell proliferation through cell cycle pathway and predicts prognosis in lung adenocarcinoma. Tumour biology: the journal of the International Society for Oncodevelopmental Biology Medicine. 2016;37(9):12597–607.

Tables

Table 1. The detailed information of immune gene prognostic model in sarcoma.

| Genes   | Hazard ratio | P value       | Coefficient |
|---------|--------------|---------------|-------------|
| BTC     | 0.90675      | 0.00136839    | -0.05946    |
| EDN3    | 1.061255     | 0.02837185    | 0.057797    |
| GREM2   | 0.912368     | 0.00131673    | -0.09467    |
| IL17B   | 0.936947     | 0.04169136    | 0.06285     |
| PAK3    | 0.930598     | 0.03411862    | -0.08109    |
| RAC3    | 1.252358     | 6.6651E-05    | 0.215092    |
| ZC3HAV1L| 1.194909     | 0.03087486    | 0.270264    |

Notes: Hazard ratio and P value were from the univariate Cox proportional hazards regression analysis in 263 sarcoma patients. Coefficient was from the multivariate Cox proportional hazards regression analysis in 263 sarcoma patients.

Figures
Figure 1

The association between immune or stromal scores and clinical data in sarcoma
Figure 2

Comparison of gene expression profile with immune scores and stromal scores in sarcoma. The heatmaps (A and B), volcanos (C and D), and Venn diagrams (E and F) of differentially expressed genes between high and low immune scores or stromal scores group. Note: H, high scores. L, low scores. lfc, Log(fold chang). Imm, immune score. Str, stromal score. DEGs, differentially expressed genes.
Figure 3
The function analysis of DEGs. The bubble diagrams of top 20 GO Terms of up DEGs(A) and down DEGs(B). The bubble diagrams of KEGG pathways of up DEGs(C) and down DEGs(D). Note: DEGs, differentially expressed genes. GO, Gene Ontology. Kyoto Encyclopedia of Genes and Genomes.
Figure 4

The subnets of up and down regulated genes PPI networks. A to C were the top 3 subnets from up regulated DEGs, E was the subnet from down regulated DEGs PPI network. The node size represents the connected degree. The red and blue represent up and down regulated, respectively, the darker the color, the greater the fold change. And the thickness of the edge represents the combined score.
Figure 5

Matrix graphs of Pearson's correlation analysis of genes. A, CCR genes; B, CXCL genes; C, CCL genes.
Figure 6

The survival analysis of CCR genes in sarcoma.
Figure 7

Prognostic risk analysis of immune genes in sarcoma. A, univariate Cox regression analysis of immune genes. B, From top to bottom are the risk score, survival status distribution, and the expression heat map of seven prognostic immune genes in low and high groups. C, Kaplan–Meier curves for the low- and high-risk groups. D, the violin plot of the expression of seven prognostic immune genes expression level in the low- and high-risk groups.
Figure 8

The evaluation of prognostic model. A, ROC curve analysis for predicting survival in sarcoma patients according to the risk score. B, From left to right are univariate Cox regression analysis and multivariate Cox regression analysis for predicting survival in sarcoma patients according to the risk score. Note: his_subtype, Histological subtype.

Figure 9

Kaplan–Meier curves of OS and EFS for the low- and high-risk groups in GSE17674. Note: OS, Overall survival. EFS, Event-free survival.