Assessing immune infiltration and the tumor microenvironment for the diagnosis and prognosis of sarcoma

Naiqiang Zhu
Chengde Medical University Affiliated Hospital https://orcid.org/0000-0002-9540-4408
Jingyi Hou (hjy_2016@126.com)
https://orcid.org/0000-0002-3311-4309

Primary research

Keywords: sarcomas, immune infiltration, prognosis, weighted gene co-expression analysis, tumor microenvironment

Posted Date: February 28th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-15489/v1

Version of Record: A version of this preprint was published on December 2nd, 2020. See the published version at https://doi.org/10.1186/s12935-020-01672-3.

Abstract

Background

Sarcomas, cancers originating from mesenchymal cells, are comprehensive tumors with poor prognoses, yet their tumorigenic mechanisms are largely unknown. In this study, we aimed to characterize infiltrating immune cells and genes associated with the immunologic response to sarcomas.

Method

The “CIBERSORT” algorithm was used to calculate the amount of L22 immune cell infiltration in sarcomas. Then, the “ESTIMATE” algorithm was used to assess the “Estimate,” “Immune,” and “Stromal” scores. Weighted gene co-expression network analysis (WGCNA) was utilized to identify the significant module related to the immune therapeutic target. Gene ontology (GO) enrichment and Kyoto Encyclopedia of Gens and Genomes (KEGG) analysis were applied using the “clusterProfiler” package in R for annotation and visualization.

Results

Macrophages were the most common immune cells infiltrating sarcomas. The number of CD8 T cells was negatively associated with that of M0 and M2 macrophages, and positively associated with M macrophages in sarcomas samples. The clinical parameters (disease type, gender) significantly increased with higher Estimate, Immune, and Stromal scores, and with a better prognosis. The blue module was significantly associated with CD8 T cells. Functional enrichment analysis showed that the blue module was mainly involved in chemokine signaling and the PI3K-Akt signaling pathway. CD48, P2RY10 and RASAL3 were identified and validated at the protein level.

Conclusion

Based on the immune cell infiltration and immune microenvironment, three key genes were identified, which suggest novel molecular mechanisms of sarcoma metastasis.

Background

Sarcomas are a widespread, heterogeneous group of tumors that occur on the skin, under the skin, in the periosteum, and on the ends of the long bones of adolescents and the elderly (overall incidence: 1–2 / 100,000 annually) [1], which is characteristic of cancers originating from mesenchymal cells [2]. Histopathologically, sarcomas are classified as either bone or soft tissue sarcomas [3]. To date, the etiology of sarcomas is not well characterized; however, their incidence appears to be associated with heredity [4], viral infection [5], trauma[6], environmental factors [7], and exposure to radiation [8]. Compared with other cancers, the degree of malignancy of sarcomas is relatively high [9], and hematogenous metastasis can spread to various organs, such as lung, brain, liver, and bone [10, 11].

Since sarcomas rarely display clinical manifestations in the initial stages of the disease, most sarcomas are diagnosed at more advanced stages.

Standard of care for patients with sarcoma is mainly comprised of local surgery, chemotherapy, and radiotherapy [12]. Among these, radical surgery or amputation is the most common [13]. Chemotherapy and radiotherapy are also included before and after surgery to prevent recurrence [14]; however, the success rate of treatment is currently low, and many patients still have poor prognosis and die of cancer-related cause. For some patients with distant metastases, palliative local treatments are chosen to control and delay disease progression [15]. With the enhanced study of the immune system, immunotherapy has emerged as very promising method to treat sarcomas after surgery and chemotherapy [16]. Researchers have adopted various immunotherapy strategies for different types of immune deficiency, but the main treatment obstacles are identifying the specific target antigen and dealing with the severe adverse effects of the selected treatment [17]. Additional research is urgently needed to identify ways to ameliorate the toxic response to immunotherapy, identify specific targets related to sarcoma, improve the effectiveness and safety of immunotherapy, and design new combinations of immune checkpoint inhibitors and other therapies [18]. Therefore, the characterization of sarcoma-specific biomarkers and the molecular mechanisms responsible for the transformation of normal cells to sarcoma are essential for the success of sarcoma immunotherapy.

Since immune infiltration and the tumor microenvironment may predict potential molecular mechanisms of sarcomas, we used the CIBERSORT (Cell-Type Identification By Estimating Relative Subsets Of RNA Transcripts) algorithm to characterize differential expression patterns of immune cell infiltration between sarcoma samples and normal samples in 22 subpopulations of immune cells, and we used the “Estimate” algorithm to analyze the Stromal and Immune scores of differential gene expression. Then, we used weighted gene co-expression network analysis (WGCNA, https://bmcbioinformatics.biomedcentral.com) to identify several key genes related to immune therapeutic target, to suggest novel molecular mechanisms responsible for the transformation and growth of sarcomas.

Materials And Methods

Sample Acquisition

The sarcoma transcription was downloaded from the TCGA database [19] via the Genomic Data Commons (GDC) data portal. Clinical information of SARC patients was also acquired from the TCGA. Filter criteria for eligible samples were as follows: 1) samples with both transcriptome data and clinical information were included; 2) samples with duplicated data or null values were excluded.

CIBERSOFT Evaluation

The “limma” package in R was used to normalized the data to estimate the percentage of infiltrating immune cell, and then standardized gene expression data were uploaded to CIBERSORT [20, 21]. Among them, the LM22 signature and 1,000 permutation were added [22], with CIBERSORT cases ($P < 0.05$) included in the survival analysis.

Estimation Evaluation

Estimate-, Immune- and Stromal scores of sarcoma samples were calculated with the ESTIMATE algorithm of the “estimate” package [23]. The “limma” package in R [24] was applied to identify differentially expressed genes (DEGs) with p-values < 0.05, and |log fold change (FC)| > 1. Then, the relationship between DEGs and overall survival of the patients from whom the sarcoma samples were acquired was analyzed with the “survival” package in R [25].

Co-expression Analysis

Samples identified by both CIBERSORT and ESTIMATE were included in the co-expression analysis. The co-expression analysis was performed using the WGCNA package in R language. [26, 27]. To maintain the gene network level off to the scale-free topology and enough connectivity, four were identified as the best soft power threshold when the degree of independence was 0.8. Then, gene modules were detected based on a topological matrix (TOM). Genes with high correlation were divided into one module (minimum module size = 20). To further evaluate the robust and reliability of the modules, the permutation test (50x) was performed with the “modulePreservation function.” [28]. As instructed, the modules with a Zsummary.qual < 5 were not considered stable in the co-expressed network. Modules with Zsummary < 2 indicated “no preservation,” 2 < Zsummary < 10 indicated “weak preservation,” and Zsummary > 10 indicated “strong preservation” (which is recommended to be used as the significant module). Modules with a Zsummary < 10, and a Zsummary.qual < 5 were excluded from the following analysis [29, 30].

Identification Of Significant Modules And Functional Analysis

In this study, gene significance (GS) [31] was used to calculate the correlation coefficients. For example, the module significance (MS) was indexed as the average GS for the genes in given module. When the modules with the highest MS values were regarded as the significant modules. The analysis of genes in significant modules were performed using gene ontology (GO) [32] enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) [33] pathway analysis, included in the “clusterProfiler” package [34] in R, with p-values < 0.05.

Gene Identification And Validation

Module connectivity (cor.geneModule Membership[MM]) and clinical trait relationship (cor.geneTraitSignificance) of genes in the significant module were calculated. In addition, the genes identified as significant were uploaded to the STRING database (http://string-db.org) (confidence > 0.1) [35], and Cytoscape was performed to establish the protein-protein interaction (PPI) network. The MCODE plug-in was utilized to identify the most significant sub-module, with a degree cut-off of 2, k-score of 2, and max. depth of 100. Then, we used the “Network Analyzer”to identify nodes in the network, in which the size and the color of the nodes represents MCODE scores. Genes with cor.geneModule > 0.8, cor.geneTraitSignificance > 0.2, and the largest MCODE sub-network were considered for further analysis. For further validation, identified genes were analyzed with the “survival” package in R, where $p < 0.05$ was considered statistically significant. In addition, the Human Protein Atlas Database (HPAD) [36] was used to validate the protein -level of these genes .

Genetic Alterations

CBio Cancer Genomics Portal [37, 38] is an open-access website used for the visualization and analysis of various cancers. In this study, this platform was utilized to investigate and compare the genetic alternations of the hub genes.

Results

Workflow Analysis and Data Description

The analytical workflow is shown in Fig. 1. First, we evaluated the differences in immune cell infiltration and tumor microenvironments of sarcoma tissues and normal, adjacent tissues. Next, we characterized the significant module associated with CD8 T cells by WGCNA, and hub genes were identified and validated. Samples from 263 patients with sarcomas and two samples of normal, adjacent tissues were obtained from the TCGA database for further analysis.

The Percent of Immune Cells in Sarcoma Samples and Clinical Correlation

Figure 2A and Supplementary Fig. 1 shown the percentage of 22 different immune cells calculated by the CIBERSORT algorithm. Macrophages were the predominant immune cell type in sarcoma tissues. The violin plot (Fig. 2B) shows that the fraction of resting Mest cells ($P = 0.096$) was significantly different in sarcoma tissue vs. normal, adjacent tissue. In addition, the fraction of CD8 T cells was negatively associated with M0 macrophages ($R = -0.43$) and M2 macrophages ($R = -0.41$), and positively associated with M1 macrophages ($R = 0.51$) and follicular, helper T cells ($R = 0.6$)(Fig. 2C).We performed a Kaplan-Meier survival analysis to evaluate the correlationship between the 22 immune cell subtypes and the prognosis of patients from whom the tumors were acquired. The fraction of activated dendritic cells (p =

9.293e-04), resting dendritic cells (p = 0.031), M2 marcophages (p = 1.668e-04), neutrophils (p = 0.046), and plasma cells (p = 0.031) were significantly different among disease types (Fig. 3A-E). The fractions of activated dendritic cells (p = 0.04) and M2 macrophages (p = 0.023) were significantly different among disease recurrence (Fig. 3F-G), and the fraction of activated dendritic cells (p = 0.021), resting NK cells (p = 0.007), and follicular helper T cells (p = 0.028) were also significantly different among total necrosis percent (Fig. 3H-J). In addition, the fractions of activated NK cells (p = 0.049), CD8 T cells (p < 0.001), regulatory T cells (p = 0.003), M0 marcophages (p = 0.002), and M1 marcophages (p = 0.038) were significantly correlated with overall survival (Fig. 4).

Immune-, Stroma- and Estimate- Scores Correlated with Clinical Parameters

The “ESTIMATE” package in R was performed to match and calculate the Immune-, Stromal-, and Estimate- scores of 263 patients. As shown in Table 1, the stromal scores ranged from − 1336.63 to 2476.22, the Immune scores ranged from − 1722.08 to 3499.2, and the Estimate scores ranged from − 2977.43 to 5336.41. The “survival” package in R was utilized to analyze the correlations of Estimate, Immune, and Stromal scores with overall survival (Fig. 5). The patients with tumors that had high Estimate, Immune, and Stromal scores had a significantly better prognosis than patients with tumors that had low Estimate, Immune, and Stromal score group (p = 0.004, p = 0.007, p = 0.017, respectively). Furthermore, the relationship among Estimate-, Immune-, Stromal- scores and clinical parameters were evaluated (Fig. 6). The clinical parameters clinical parameters (disease type, gender) significantly increased as Estimate-, Immune-, and Stromal- scores increased (p < 0.05).

Table 1

Estimate scores, Immune scores, Stromal scores, and clinical parameters of sarcoma samples

Characteristic	n	Estimate score (range)	Immune score (range)	Stromal score (range)
Age				
< 61.5	135	-2977.42 to 5336.40	-1722.08 to 3499.29	-1336.63 to 2476.22
≥ 61.5	128	-2142.73 to 5178.61	-1135.92 to 3382.39	-2142.73 to 5178.61
Gender				
Female	144	-2886.10 to 5336.41	-1722.09 to 3499.29	-1336.63 to 2032.58
Male	119	-2977.43 to 4985.82	-1646.75 to 3382.38	-1330.67 to 2476.22
Disease type				
Fibromatous	40	-504.94 to 5336.41	-962.99 to 3499.29	153.81 to 2157.00
Lipomatous	60	-1889.11 to 5187.61	-1110.77 to 3382.39	-778.34 to 2476.22
Myomatous	106	-2114.77 to 4221.40	-1520.77 to 2998.72	-727.47 to 1727.86
Nerve Sheath	10	-490.23 to 3520.43	-733.46 to 2271.22	-122.82 to 1249.22
Soft Tissue	37	-274.38 to 4985.82	-496.70 to 3348.72	106.11 to 1731.42
Synovial-like	10	-2977.43 to 726.05	-1722.08 to -174.09	-1336.63 to 900.13
Tumor total necrosis (%)				
0%	71	-2886.10 to 5336.41	-1722.08 to 3499.29	-1336.63 to 2230.57
< 10%	38	-1645.38 to 4600.14	-1067.28 to 3027.08	-628.27 to 1985.01
≥ 10, < 50%	62	-2977.43 to 4510.53	-1646.76 to 3140.17	-1330.67 to 1793.80
> 50%	12	-2114.77 to 3725.37	-1520.77 to 2293.27	-727.46 to 1707.18
unknown	80	-1889.11 to 5178.61	-1110.77 to 3199.34	-778.34 to 2476.22
Disease recurrence				
Yes	29	-2647.50 to 4104.67	-1722.08 to 2386.97	-925.41 to 1985.01
No	146	-2977.43 to 5336.41	-1646.76 to 3499.29	-1336.63 to 2230.57
Unknown	88	-1889.11 to 5178.61	-1110.77 to 3199.35	-778.34 to 2476.22

Identification Of DEGs And Clinical Correlation

Supplementary Fig. 2 demonstrated that the gene expression profiles can be used to differentiate the two groups (high vs. Low, for each). Using the Immune scores, a total of 2199 DEGs between the high-score and low-score group were identified in which 1070 genes were upregulated, and 1129 genes were downregulated. Similarly, for stromal scores, 995 DEGs were upregulated and 1498 DEGs were downregulated between the high-score and low-score groups. The genes differentially upregulated and downregulated in the high vs. low immune and stromal scores groups were further analyzed. Of these 1509 genes, 729 were upregulated and 780 were downregulated (Supplementary Fig. 3).

Weighted Gene Co-expression Network Construction And Module Preservation Analysis

After validation, the 1509 DEGs were used to create a co-expression network using WGCNA. The analysis was performed as previously described [39]. Briefly, the gene expression profile matrix of all samples was transformed into a Pearson's correlation coefficient matrix, and the distribution of connections among these genes indicated that the profile met the criteria of a scale-free network. Then we constructed network based on the scale-free network (Fig. 7A). We assigned a power value of 4 with a 0.8 degree of independent (Supplementary Fig. 4, Fig. 7B). The size of the seven modules ranged from 21 to 338 genes. Genes that were not co-expressed were assigned to the grey module, and were not further analyzed. Furthermore, the results of module stability analysis demonstrated that the Zsummary.qual for module preservation of gold modules was 5, and the Zsummary statistic for module preservation of the blue, green, yellow, turquoise, and brown modules was 10 (Fig. 7C, Supplementary List 1). Based on these data, the gold module was not considered stable in these analyses, and it was not used in subsequent analyses.

Identifying Significant Modules And Module Functional Annotation

The association of the six modules with types of immune cell infiltration was analyzed (Fig. 8A), and the blue module was identified as having the highest correlation with CD8 T cells compared with other modules (cor = 0.9, p = 7.7e-63). The eigengenes and adjacencies were calculated according to their correlation (Fig. 8B), and the five modules were divided into two main clusters. Furthermore, a heatmap was produced based on the interaction relationship of the six modules (Fig. 8C). Since CD8 T cells recognized and kill cancer cells by expressing cytokines and cytotoxic molecules, they have been identified as a key target for immunotherapy. The blue module had the highest correlation with CD8 T cells, which suggested that the genes in blue module were candidates for immunotherapy biomarkers of sarcoma. To further elucidate the functional of the significant module, all genes in the blue module were analyzed with the "clusterProfiler" package in R to identify representative KEGG pathways and GO terms. As shown in Fig. 9A, the most significantly enriched pathways of the blue module following KEGG pathway analysis were enhanced in chemokine signaling, the PI3K-AKT signaling pathway, and the JAK-

STAT signaling pathway (Supplementary List 2). GO enrichment analysis shown that the blue module contained biological processes mainly involved in chemokine-mediated, and complement receptor-mediated signaling pathways; and response to interferon-gamma, lipopolysaccharide, interferon-gamma, and chemokines. The cellular component (CC) was bent on the plasma, lysosomal, lytic vacuole membrane. Molecular function (MF) mainly enriched on the chemokine activity (Fig. 9 (B-D), Supplementary List 2).

Identification And Validation Of Hub Genes

Based on the following criteria (|MM| > 0.8, |GS| > 0.2, and the largest sub-network), four genes with high connectivity in the clinically significant modules were identified as hub genes (Fig. 10). The “survival” package in R was performed to calculate the survival analysis ($p < 0.05$ as statistically significant). Table 2 shows that the survival analysis of hub genes identified CD48 antigen (CD48), putative P2Y purinoceptor 10 (P2RY10), and RAS protein activator like-3 (RASAL3). Figure 11 shows that these three genes were significant and positively associated with overall survival. In addition, immunohistochemistry (IHC) staining data acquired from the HPAD confirmed the differential expression of the genes (Fig. 12).

Table 2
Hub genes in the significant module

Gene symbol	Co-expression analysis		MCODE analysis	
	GS	MM	Connectivity degree	MCODE_score
CD48	0.51	0.95	24	18
P2RY10	0.50	0.89	17	16.83
RASAL3	0.58	0.95	23	16.83

Genetic alterations

CBioPortal was used to evaluate genetic changes in CD48, P2RY10, and RASAL3. Figure 13A shows that the frequency of mutations in CD48 was 11% in P2RY10 it was 9%, and in RASAL3 it was 12%. Amplification, and mRNA was the main type. The three genes were altered in 22% (45/206) of the patients (Fig. 13B).

Discussion

Sarcomas, originating from mesenchymal cells, are comprehensive tumors with poor prognoses. There are no obviously clinical symptoms at the early stage of the disease, and distant metastasis occurs at the late stage. However, its potential molecular mechanism is still unclear. Among the available therapeutic strategies, immunotherapy is considered the safest and most effective. In this study, tumor-infiltrating

immune cells and the tumor-microenvironmental scores were used to construct a co-expression network to identify genes associated with the prognosis and metastasis of sarcomas.

We used the CIBERSORT algorithm to identify different patterns of immune cell infiltration in sarcoma samples, and characterize corresponding clinical traits. In the sarcoma tumors, 35–40% of the immune cells were macrophages. Macrophages are myeloid immune cells that are strategically positioned throughout body tissue and they can present antigens on their external membrane, which can promote the release of inflammatory cytokines by sarcoma tumors [40]. Macrophages are major components of the tumor microevironment and orchestrate various aspects of the immune response to sarcomas. They can affect tumorigenesis by either promoting cytokine release by immune cells or by enhancing the antitumor response [41, 42]. In addition, our Kaplan-Meier analysis suggests that marcophages are associated with disease type, overall survival, and percent necrosis in sarcoma tumors. CD8 T cells are the primary cell type responsible for immune responses. They can specifically recognize and kill cancer cells by secreting cytokines and cytotoxic molecules [43]. We found that the fraction of CD8 T cells was negatively associated with M0 macrophages and M2 macrophages and positively associated with M macrophages in sarcomas samples. These data suggest that CD8 T cells and macrophages might be potential markers for the prognosis of patients with sarcoma.

The tumor microenvironment (TME) is a complex, integrated system, which is different from the microenvironment established by normal cells and their surrounding tissue [44]. The TME plays a pivotal role in tumor progression and metastasis, and may significantly influence therapeutic response to cancer treatment [45, 46]. In this study, we calculated the Immune-, Stromal- and Estimate- scores for sarcoma samples by applying the ESTIMATE algorithm. The results shown that (disease type and gender) significantly increased with higher Estimate, Immune, and Stromal scores. These data suggest that the high scoring-groups have a better prognosis. Then, genes differentially expressed between low- and high-immune/stromal score groups were identified and characterized as having more or less DEGs. These data were used in the WGCNA analysis.

WGCNA is a systematic biological algorithm [26], which is used to reveal the association between genes and clinical phenotypes [47]. It has been widely used to identified potential biomarkers for Alzheimer's disease [48], breast cancer [49], and osteoarthritis [50]. In this paper, WGCNA identified 7 modules. Among these, the blue module was significantly associated with immune cell subtypes related to CD8 T cells. KEGG analysis discovered that the genes in blue module were mainly enriched for cytokine/cytokine receptor interaction, chemokine signaling pathways, the PI3K-AKT signaling pathway, and the JAK-STAT signaling pathway. Proinflammatory cytokines are involved in cancer progression, and cytokine/cytokine receptor interactions may be essential mediators of inflammation in the development and prognosis of sarcomas [44, 51, 52]. Our data indicate that the cytokine/cytokine receptor signaling pathway is involved in sarcoma progression[53]. A chemokine signaling helps coordinate cell migration.[54]. The JAK-STAT signaling pathway transfer signals from cell membrane receptors to the nucleus in sarcoma tumors [55]. It modulate the activity of immune system, especially the fate of helper T cells [56]. Zhang et al. found that downregulated expression of HGDF promotes tumor development and progression by coordinating

the PI3K-AKT signaling pathway in sarcomas. In addition, GO analysis has shown that the blue module is predominantly involved in chemokine-mediated and complement receptor-mediated signaling and chemokine activity. Previous researches has confirmed that the plasma, lysosomal, and lytic vacuole membranes may be potential targets for treating sarcomas [55, 57–60].

In the blue module, CD48, P2RY10, and RASAL3 were identified as differentially expressed. They were associated with survival analysis results and were validated at the protein level. CD48, a member of CD2 immunoglobulin superfamily (IgSF) participates in activation and differentiation pathways in CD84, CD150, CD229 and CD244 [61]. Liu et al. [62] suggested CD48 as a key gene for the induction of histiocytic sarcoma in mouse skeletal muscle. P2RY10 belongs to the family of G -protein -coupled receptors, which are activated by adenosine and uridine [63]. Wang et al. found that P2RY10 was potentially involved in the immune response and the development of sarcomas [64]. RASL3, a novel member of the RasGAP Rasal family, is predominantly expressed by T cells [65, 66], and RasGAP activity stimulates ERK phosphorylation. Previous studies have not linked RASL3 to the cancer [67]; therefore, it may be a novel immunotherapy target for the treatment of patients with sarcoma.

Conclusions:

In this study, the ESTIMATE algorithm-based immune score significant correlated, and the type and levels of L22 immune cells infiltrating sarcoma tumors were also assessed. We used the WGCNA algorithm to establish a co-expression network involving key genes related to immunotherapy and the development of sarcomas: CD48, P2RY10, and RASAL3. We propose that these three genes should be considered as potential prognostic biomarkers and/or therapeutic targets of immunotherapy.

Abbreviations

BP, biological process; CC, cellular component; CD48, CD48 antigen; DEGs, differentially expressed genes; FC, fold change; FDR, False discovery rate; GDC, Genomic Data Commons; GO, gene ontology; GS, gene significance; IHC, immunohistochemistry; KEGG, Kyoto Encyclopedia of Genes and Genomes; MCODE, ; MDS, multidimensional scaling; MEs, module eigengenes; MF, molecular function; MM, ; PPI, protein-protein interaction; P2RY10, Putative P2Y purinoceptor 10; RASAL3, RAS protein activator like-3; TCGA, The Cancer Genome Atlas; TOM, topological matrix; WGCNA, weighted gene co-expression network analysis;

Declarations

Ethic approval and consent to participate

Not applicable.

Consent for publication

Not applicable

Availability of data and materials

The data was from TCGA database in this manuscript.

Competing Interests

The authors declare that there are no conflicts of interest.

Funding

Project supported by the National Natural Science Foundation of China (Grant No.81703659) and Chengde city science and technology research and development projects (201904A103).

Author's Contributions

HJY conceived and designed the study. HJY and ZNQ performed the experiment and wrote the paper. All authors are responsible for reviewing data. All authors read and approved the final manuscript.

Acknowledgement:

Not applicable.

Supplementary Materials

Supplementary Excel List 1 The Zsummary.pres and Zsummary.qual of modules.

Supplementary Excel List 2 The GO enrichment and KEGG pathway enrichment in the blue module.

Supplementary Figure 1. Heatmap of immune cells estimated.

Supplementary Figure 2. Heatmap of DEGs in the groups with low- and high- scores. (A) Immune scores; (B) Stromal scores.

Supplementary Figure 3. Common DEGs in immune- and stromal- scores. (A) Commonly up-regulated genes; (B) Commonly down-regulated genes.

Supplementary Figure 4. (A) The scale-free fit index for soft-thresholding powers; (B)The mean connectivity for soft-thresholding powers.

Supplementary Figure 5. Multi-dimensional scaling (MDS) plot of the co-expression network.

References

1. Hornick JL. Subclassification of pleomorphic sarcomas: How and why should we care. Ann Diagn Pathol. 2018;37:118-24. https://doi.org/10.1016/j.anndiagpath.2018.10.006.
2. Ferrari A, Dirksen U, Bielack S. Sarcomas of Soft Tissue and Bone. Prog Tumor Res. 2016;43:128-41. https://doi.org/10.1159/000447083.
3. Hatina J, Kripnerova M, Houfkova K, Pesta M, Kuncova J, Sana J, et al. Sarcoma Stem Cell Heterogeneity. Adv Exp Med Biol. 2019;1123:95-118. https://doi.org/10.1007/978-3-030-11096-3_7.
4. Paumard-Hernández B, Calvete O, Inglada Pérez L, Tejero H, Al-Shahrour F, Pita G, et al. Whole exome sequencing identifies PLEC, EXO5 and DNAH7 as novel susceptibility genes in testicular cancer. Int J Cancer. 2018;143:1954-62. https://doi.org/10.1002/ijc.31604.
5. Golas G, Jang SJ, Naik NG, Alonso JD, Papp B, Toth Z. Comparative analysis of the viral interferon regulatory factors of KSHV for their requisite for virus production and inhibition of the type I interferon pathway. Virology. 2020;541:160-73. https://doi.org/10.1016/j.virol.2019.12.011.
6. Cheng Y, Yu C, Zhu S, Guo L, Zhang Y, Zhang Y, et al. Nonleukemic granulocytic sarcoma of orbit after blunt trauma: A case report and review of literature. Medicine (Baltimore). 2018;97:e0373. https://doi.org/10.1097/MD.0000000000010373.
7. Buchta CM, Boi SK, Miller BJ, Milhem MM, Norian LA. Obesity Does Not Exacerbate the Protumorigenic Systemic Environment in Sarcoma Subjects. Immunohorizons. 2017;1:20-8. https://doi.org/10.4049/immunohorizons.1700001.
8. Hui JY. Epidemiology and Etiology of Sarcomas. Surg Clin North Am. 2016;96:901-14. https://doi.org/10.1016/j.suc.2016.05.005.
9. McEvoy CR, Fox SB, Prall O. Emerging entities in NUTM1-rearranged neoplasms. Genes Chromosomes Cancer. 2020;https://doi.org/10.1002/gcc.22838.
10. Soyfer V, Corn BW, Shtraus N, Honig N, Meir Y, Kollender J, et al. Single-institution Experience of SBRT for Lung Metastases in Sarcoma Patients. Am J Clin Oncol. 2017;40:83-5. https://doi.org/10.1097/COC.0000000000000103.
11. Stevenson JD, Doxey R, Abudu A, Parry M, Evans S, Peart F, et al. Vascularized fibular epiphyseal transfer for proximal humeral reconstruction in children with a primary sarcoma of bone. Bone Joint J. 2018;100-B:535-41. https://doi.org/10.1302/0301-620X.100B4.BJJ-2017-0830.R1.
12. Requena C, Alsina M, Morgado-Carrasco D, Cruz J, Sanmartín O, Serra-Guillén C, et al. Kaposi Sarcoma and Cutaneous Angiosarcoma: Guidelines for Diagnosis and Treatment. Actas Dermosifiliogr. 2018;109:878-87. https://doi.org/10.1016/j.ad.2018.06.013.
13. Ghirardi V, Bizzarri N, Guida F, Vascone C, Costantini B, Scambia G, et al. Role of surgery in gynaecological sarcomas. Oncotarget. 2019;10:2561-75. https://doi.org/10.18632/oncotarget.26803.
14. Kirov G, Kazandzhiev P. [Retroperitoneal sarcomas]. Khirurgiia (Sofiia). 1991;44:14-7.
15. Dangoor A, Seddon B, Gerrand C, Grimer R, Whelan J, Judson I. UK guidelines for the management of soft tissue sarcomas. Clin Sarcoma Res. 2016;6:20. https://doi.org/10.1186/s13569-016-0060-4.

16. Shitara K, Ueha S, Shichino S, Aoki H, Ogiwara H, Nakatsura T, et al. First-in-human phase 1 study of IT1208, a defucosylated humanized anti-CD4 depleting antibody, in patients with advanced solid tumors. J Immunother Cancer. 2019;7:195. https://doi.org/10.1186/s40425-019-0677-y.
17. Tagliabue L, Capozza A, Maioli C, Luciani A, Ierardi AM, Carrafiello G. Immunotherapy treatment: an issue for metabolic response. Q J Nucl Med Mol Imaging. 2018;62:140-51. https://doi.org/10.23736/S1824-4785.17.03035-7.
18. Pandolfi F, Franza L, Todi L, Carusi V, Centrone M, Buonomo A, et al. The Importance of Complying with Vaccination Protocols in Developed Countries: "Anti-Vax" Hysteria and the Spread of Severe Preventable Diseases. Curr Med Chem. 2018;25:6070-81. https://doi.org/10.2174/0929867325666180518072730.
19. Blum A, Wang P, Zenklusen JC. SnapShot: TCGA-Analyzed Tumors. Cell. 2018;173:530. https://doi.org/10.1016/j.cell.2018.03.059.
20. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12:453-7. https://doi.org/10.1038/nmeth.3337.
21. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol Biol. 2018;1711:243-59. https://doi.org/10.1007/978-1-4939-7493-1_12.
22. Murray NP, Fuentealba C, Salazar A, Reyes E. Platelet-to-lymphocyte ratio and systemic immune-inflammation index versus circulating prostate cells to predict significant prostate cancer at first biopsy. Turk J Urol. 2020;1-7. https://doi.org/10.5152/tud.2020.19203.
23. Ciardullo S, Muraca E, Perra S, Bianconi E, Zerbini F, Oltolini A, et al. Screening for non-alcoholic fatty liver disease in type 2 diabetes using non-invasive scores and association with diabetic complications. BMJ Open Diabetes Res Care. 2020;8:https://doi.org/10.1136/bmjdrc-2019-000904.
24. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43:e47. https://doi.org/10.1093/nar/gkv007.
25. George B, Seals S, Aban I. Survival analysis and regression models. J Nucl Cardiol. 2014;21:686-94. https://doi.org/10.1007/s12350-014-9908-2.
26. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559. https://doi.org/10.1186/1471-2105-9-559.
27. Pei G, Chen L, Zhang W. WGCNA Application to Proteomic and Metabolomic Data Analysis. Methods Enzymol. 2017;585:135-58. https://doi.org/10.1016/bs.mie.2016.09.016.
28. Chen J, Wang X, Hu B, He Y, Qian X, Wang W. Candidate genes in gastric cancer identified by constructing a weighted gene co-expression network. PeerJ. 2018;6:e4692. https://doi.org/10.7717/peerj.4692.
29. Luo Y, Shen D, Chen L, Wang G, Liu X, Qian K, et al. Identification of 9 key genes and small molecule drugs in clear cell renal cell carcinoma. Aging (Albany NY). 2019;11:6029-52. https://doi.org/10.18632/aging.102161.

30. Wang J, Tian GG, Zheng Z, Li B, Xing Q, Wu J. Comprehensive Transcriptomic Analysis of Mouse Gonadal Development Involving Sexual Differentiation, Meiosis and Gametogenesis. Biol Proced Online. 2019;21:20. https://doi.org/10.1186/s12575-019-0108-y.
31. Li W, Liu J, Zhao H. Identification of a nomogram based on long non-coding RNA to improve prognosis prediction of esophageal squamous cell carcinoma. Aging (Albany NY). 2020;12:1512-26. https://doi.org/10.18632/aging.102697.
32. Gene Ontology Consortium. Gene Ontology Consortium: going forward. Nucleic Acids Res. 2015;43:D1049-56. https://doi.org/10.1093/nar/gku1179.
33. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28:27-30.
34. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16:284-7. https://doi.org/10.1089/omi.2011.0118.
35. Hunstad JP, Daniels MA, Crantford JC. Autologous Flap Gluteal Augmentation: Purse-String Technique. Clin Plast Surg. 2018;45:261-7. https://doi.org/10.1016/j.cps.2017.12.008.
36. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015;347:1260419. https://doi.org/10.1126/science.1260419.
37. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2:401-4. https://doi.org/10.1158/2159-8290.CD-12-0095.
38. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013;6:pl1. https://doi.org/10.1126/scisignal.2004088.
39. Lin J, Yu M, Xu X, Wang Y, Xing H, An J, et al. Identification of biomarkers related to CD8+ T cell infiltration with gene co-expression network in clear cell renal cell carcinoma. Aging (Albany NY). 2020;12:https://doi.org/10.18632/aging.102841.
40. Varol C, Mildner A, Jung S. Macrophages: development and tissue specialization. Annu Rev Immunol. 2015;33:643-75. https://doi.org/10.1146/annurev-immunol-032414-112220.
41. Mantovani A, Marchesi F, Malesci A, Laghi L, Allavena P. Tumour-associated macrophages as treatment targets in oncology. Nat Rev Clin Oncol. 2017;14:399-416. https://doi.org/10.1038/nrclinonc.2016.217.
42. Cassetta L, Pollard JW. Targeting macrophages: therapeutic approaches in cancer. Nat Rev Drug Discov. 2018;17:887-904. https://doi.org/10.1038/nrd.2018.169.
43. Duhen T, Duhen R, Montler R, Moses J, Moudgil T, de Miranda NF, et al. Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. Nat Commun. 2018;9:2724. https://doi.org/10.1038/s41467-018-05072-0.
44. Kim J, Bae JS. Tumor-Associated Macrophages and Neutrophils in Tumor Microenvironment. Mediators Inflamm. 2016;2016:6058147. https://doi.org/10.1155/2016/6058147.

45. Hui L, Chen Y. Tumor microenvironment: Sanctuary of the devil. Cancer Lett. 2015;368:7-13. https://doi.org/10.1016/j.canlet.2015.07.039.
46. Wu T, Dai Y. Tumor microenvironment and therapeutic response. Cancer Lett. 2017;387:61-8. https://doi.org/10.1016/j.canlet.2016.01.043.
47. Shi Y, Zhu S, Yang J, Shao M, Ding W, Jiang W, et al. Investigation of Potential Mechanisms Associated with Non-small Cell Lung Cancer. J Comput Biol. 2020;https://doi.org/10.1089/cmb.2019.0081.
48. Liang JW, Fang ZY, Huang Y, Liuyang ZY, Zhang XL, Wang JL, et al. Application of Weighted Gene Co-Expression Network Analysis to Explore the Key Genes in Alzheimer's Disease. J Alzheimers Dis. 2018;65:1353-64. https://doi.org/10.3233/JAD-180400.
49. Pei J, Wang Y, Li Y. Identification of key genes controlling breast cancer stem cell characteristics via stemness indices analysis. J Transl Med. 2020;18:74. https://doi.org/10.1186/s12967-020-02260-9.
50. Shui X, Xie Q, Chen S, Zhou C, Kong J, Wang Y. Identification and functional analysis of long non-coding RNAs in the synovial membrane of osteoarthritis patients. Cell Biochem Funct. 2020;https://doi.org/10.1002/cbf.3491.
51. Qian Z, Zhang Z, Wang Y. T cell receptor signaling pathway and cytokine-cytokine receptor interaction affect the rehabilitation process after respiratory syncytial virus infection. PeerJ. 2019;7:e7089. https://doi.org/10.7717/peerj.7089.
52. Xue K, Cao J, Wang Y, Zhao X, Yu D, Jin C, et al. Identification of Potential Therapeutic Gene Markers in Nasopharyngeal Carcinoma Based on Bioinformatics Analysis. Clin Transl Sci. 2019;https://doi.org/10.1111/cts.12690.
53. Rutkowski P, Kamińska J, Kowalska M, Ruka W, Steffen J. Cytokine and cytokine receptor serum levels in adult bone sarcoma patients: correlations with local tumor extent and prognosis. J Surg Oncol. 2003;84:151-9. https://doi.org/10.1002/jso.10305.
54. Gustavsson M. New insights into the structure and function of chemokine receptor:chemokine complexes from an experimental perspective. J Leukoc Biol. 2020;https://doi.org/10.1002/JLB.2MR1219-288R.
55. Kahn J, Deverapalli SC, Rosmarin D. JAK-STAT signaling pathway inhibition: a role for treatment of various dermatologic diseases. Semin Cutan Med Surg. 2018;37:198-208. https://doi.org/10.12788/j.sder.2018.041.
56. Seif F, Khoshmirsafa M, Aazami H, Mohsenzadegan M, Sedighi G, Bahar M. The role of JAK-STAT signaling pathway and its regulators in the fate of T helper cells. Cell Commun Signal. 2017;15:23. https://doi.org/10.1186/s12964-017-0177-y.
57. Ishida Y. Fine structure of primary reticulum cell sacroma of the brain. Acta Neuropathol Suppl. 1975;Suppl 6:147-53. https://doi.org/10.1007/978-3-662-08456-4_25.
58. Avnet S, Lemma S, Cortini M, Pellegrini P, Perut F, Zini N, et al. Altered pH gradient at the plasma membrane of osteosarcoma cells is a key mechanism of drug resistance. Oncotarget. 2016;7:63408-23. https://doi.org/10.18632/oncotarget.11503.

59. Liu Q, Li Z, Shang H, Zhang Q, Wang X, Zhang Y, et al. Scavenger Receptor Class B Type 1 (SR-B1) being a Potential Biomarker for the Diagnosis of Liposarcoma and Associated with the Degree of Differentiation of Liposarcomas. J Cancer. 2019;10:4326-32. https://doi.org/10.7150/jca.31730.
60. Yoshida A, Wakai S, Ryo E, Miyata K, Miyazawa M, Yoshida KI, et al. Expanding the Phenotypic Spectrum of Mesenchymal Tumors Harboring the EWSR1-CREM Fusion. Am J Surg Pathol. 2019;43:1622-30. https://doi.org/10.1097/PAS.0000000000001331.
61. Wang Z, Xiao Y, Guan W, Wang M, Chen J, Zhang L, et al. Acute myeloid leukemia immune escape by epigenetic CD48 silencing. Clin Sci (Lond). 2020;134:261-71. https://doi.org/10.1042/CS20191170.
62. Liu J, Hettmer S, Milsom MD, Hofmann I, Hua F, Miller C, et al. Induction of histiocytic sarcoma in mouse skeletal muscle. PLoS One. 2012;7:e44044. https://doi.org/10.1371/journal.pone.0044044.
63. Qin Y, Verdegaal EM, Siderius M, Bebelman JP, Smit MJ, Leurs R, et al. Quantitative expression profiling of G-protein-coupled receptors (GPCRs) in metastatic melanoma: the constitutively active orphan GPCR GPR18 as novel drug target. Pigment Cell Melanoma Res. 2011;24:207-18. https://doi.org/10.1111/j.1755-148X.2010.00781.x.
64. Wang LX, Li Y, Chen GZ. Network-based co-expression analysis for exploring the potential diagnostic biomarkers of metastatic melanoma. PLoS One. 2018;13:e0190447. https://doi.org/10.1371/journal.pone.0190447.
65. Muro R, Nitta T, Okada T, Ideta H, Tsubata T, Suzuki H. The Ras GTPase-activating protein Rasal3 supports survival of naive T cells. PLoS One. 2015;10:e0119898. https://doi.org/10.1371/journal.pone.0119898.
66. Saito S, Kawamura T, Higuchi M, Kobayashi T, Yoshita-Takahashi M, Yamazaki M, et al. RASAL3, a novel hematopoietic RasGAP protein, regulates the number and functions of NKT cells. Eur J Immunol. 2015;45:1512-23. https://doi.org/10.1002/eji.201444977.
67. Qi Y, Wang N, Pang LJ, Zou H, Hu JM, Zhao J, et al. Identification of potential mutations and genomic alterations in the epithelial and spindle cell components of biphasic synovial sarcomas using a human exome SNP chip. BMC Med Genomics. 2015;8:69. https://doi.org/10.1186/s12920-015-0144-7.

Figures

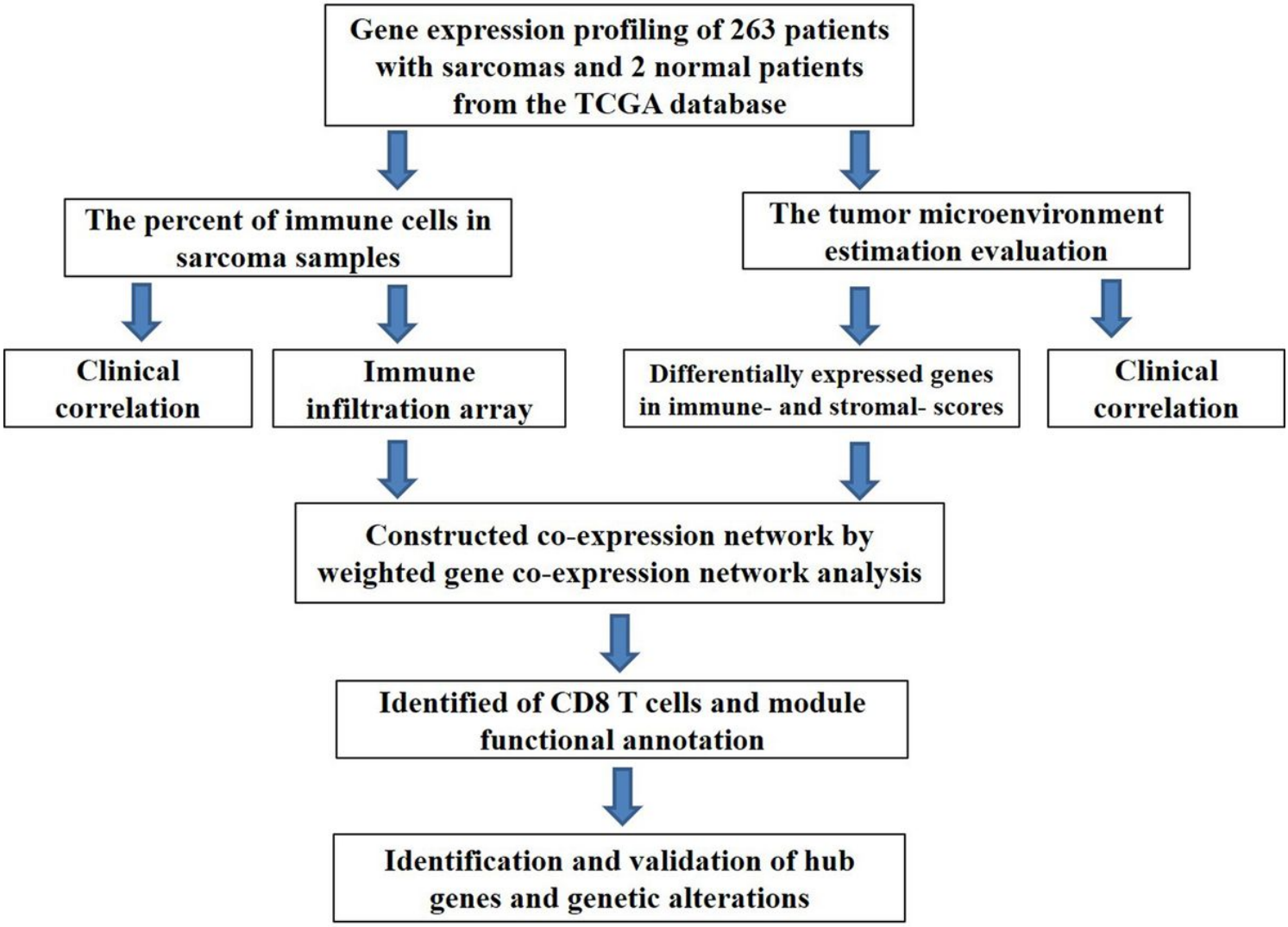

Figure 1

Flow chart of the analytical process.

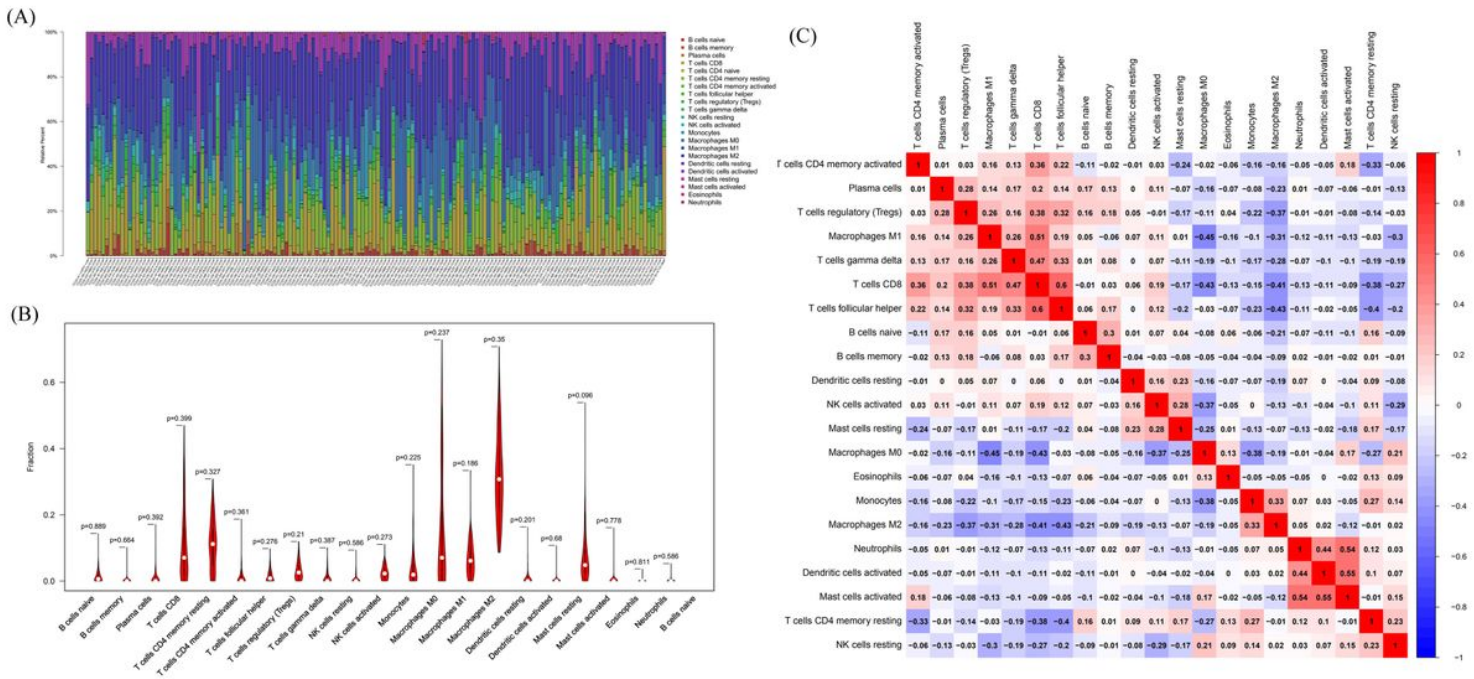

Figure 2

(A) Composition of immune cells. (B) Violin plot of immune cells. (C) Co-expression patterns among fractions of immune cells.

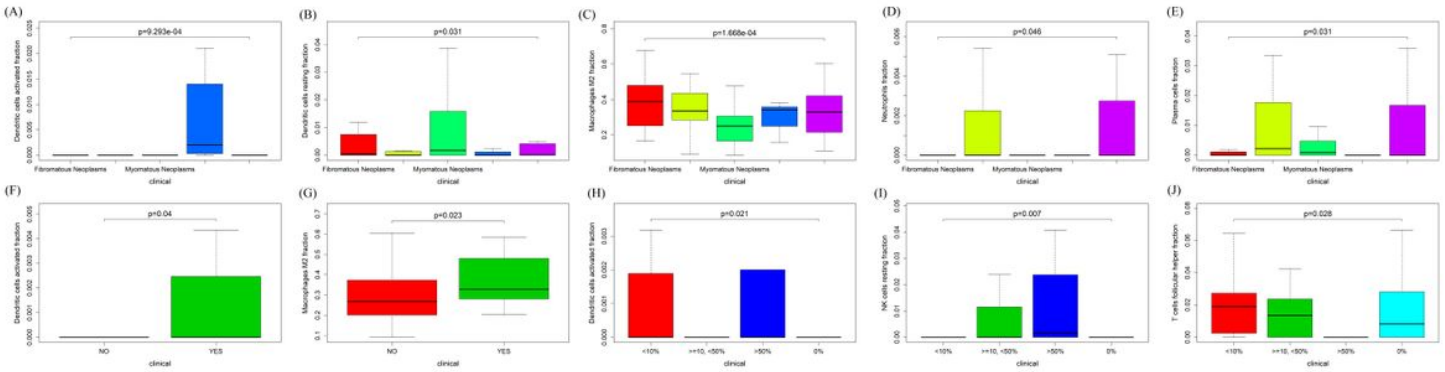

Figure 3

The fraction of (A) activated dendritic cells, (B) resting dendritic cells, (C) M2 macrophages, (D) neutrophils, and (E) plasma cells by disease type. The fraction of (F) activated dendritic cells and (G) M2 macrophages among disease recurrence. The fraction of (H) activated dendritic cells, (I) resting NK cells, and (J) follicular helper T cells among total necrosis percent.

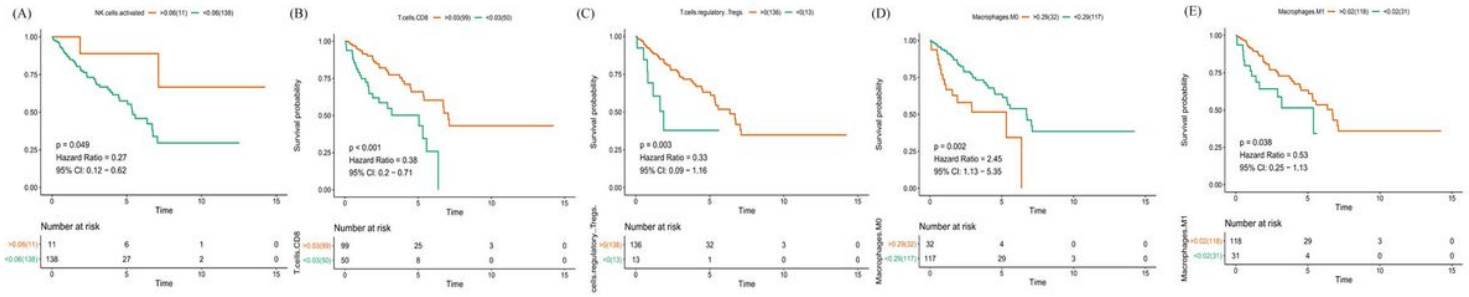

Figure 4

Survival curves of the fraction of (A) activated NK cells, (B) CD8 T cells, (C) regulatory T cells, (D) M0 macrophages, and (E) M1 macrophages.

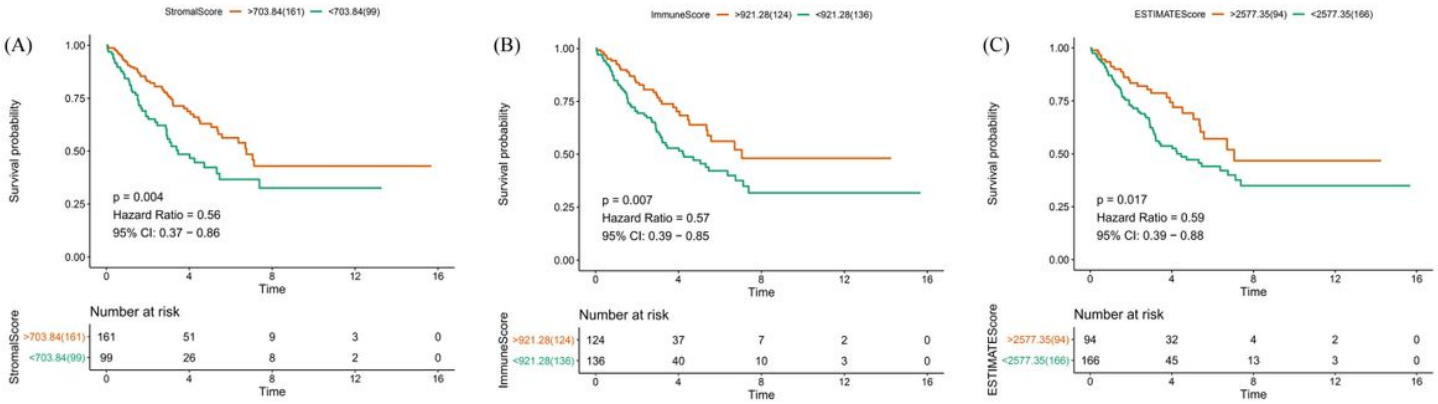

Figure 5

The association between (A) Estimate-, (B) Immune-, (C) Stromal scores of sarcoma tumors and the overall survival of patients harboring these tumors.

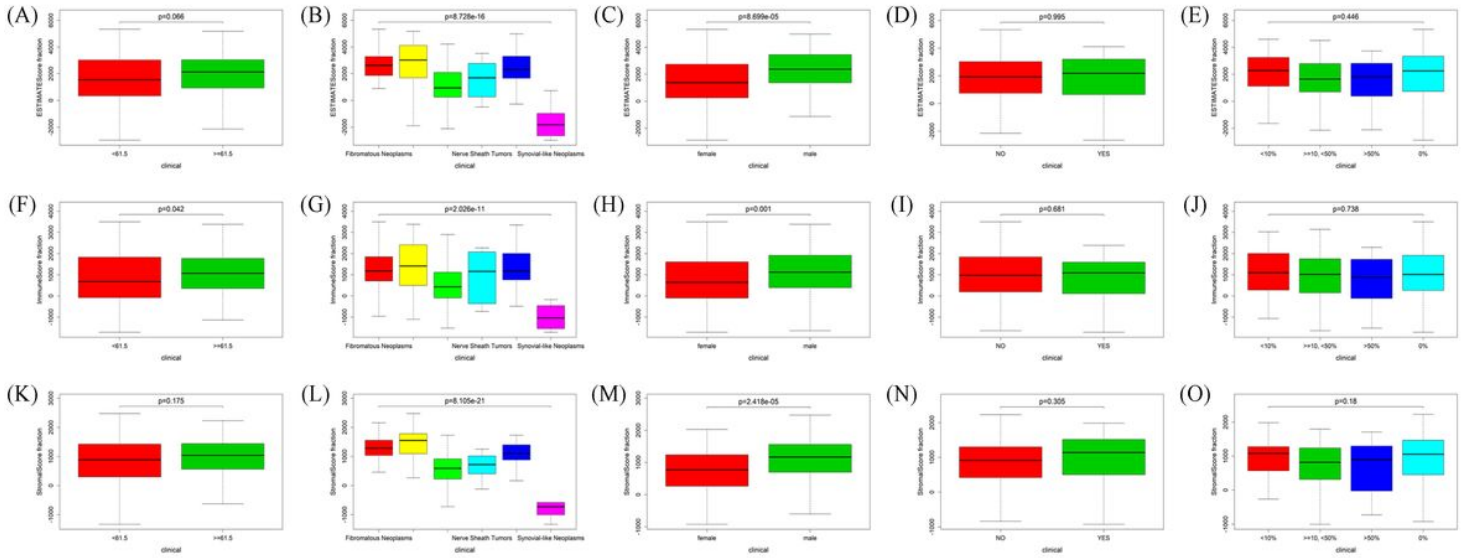

Figure 6

The relationship between Estimate-, Immune-, and Stromal scores of sarcoma tumors and clinical parameters of sarcoma patients. (A) Estimate scores and age; (B) Estimate scores and disease type ($p < 0.05$); (C) Estimate scores and gender ($p < 0.05$); (D) Estimate scores and disease recurrence; (E) Estimate scores and percent of necrosis; (F) Immune scores and age ($p < 0.05$); (G) Immune scores and disease type ($p < 0.05$); (H) Immune scores and gender ($p < 0.05$); (I) Immune scores and disease recurrence; (J) Immune scores and percent of necrosis; (K) Stromal scores and age; (L) Stromal scores and disease type ($p < 0.05$); (M) Stromal scores and gender ($p < 0.05$); (N) Stromal scores and disease recurrence; (O) Stromal scores and percent of necrosis.

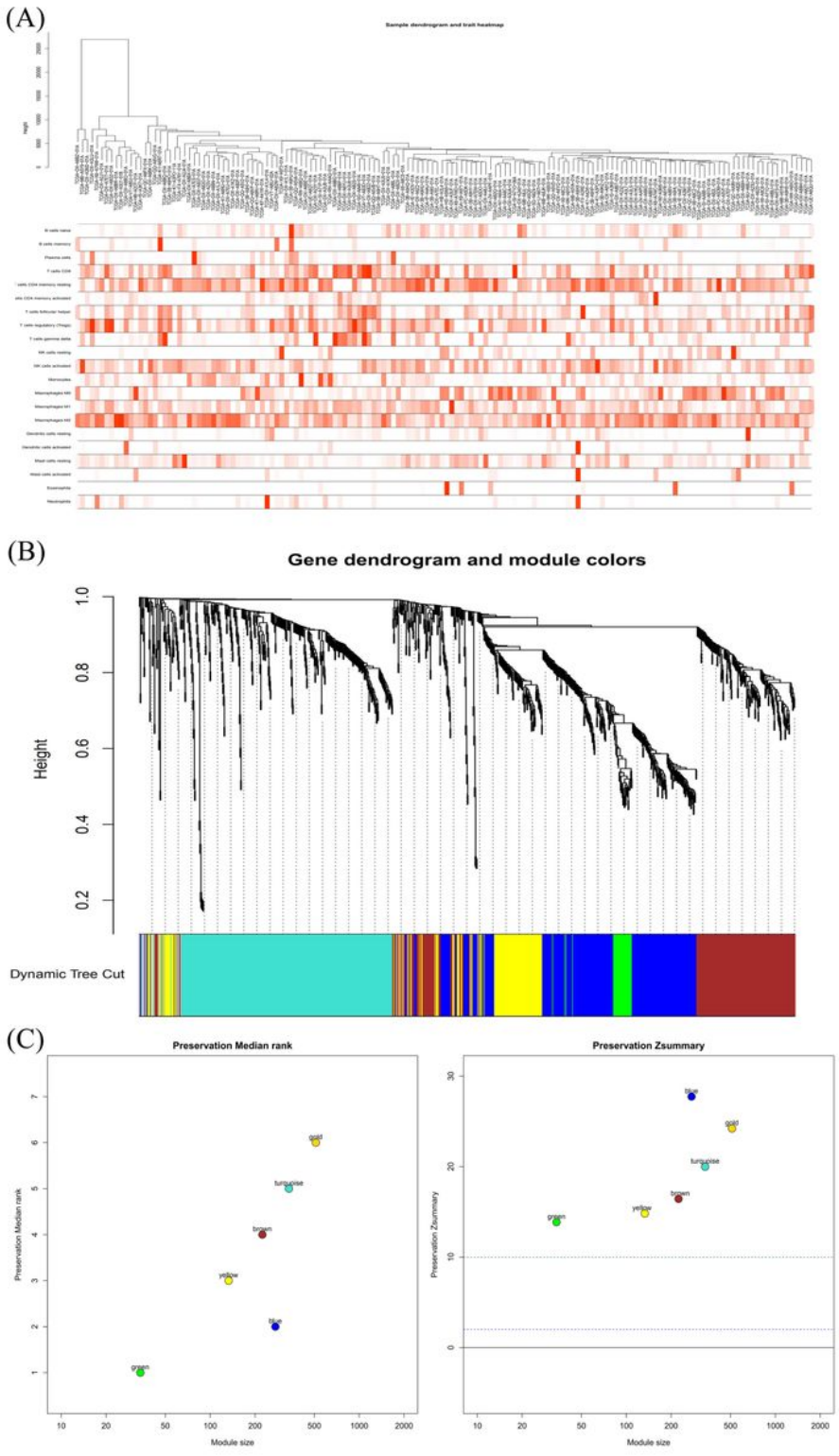

Figure 7

(A) Hierarchical clustering dendrogram of samples; (B) Co-expression network modules. (C) Median rank and Zsummary statistics of the most variant gene module preservation.

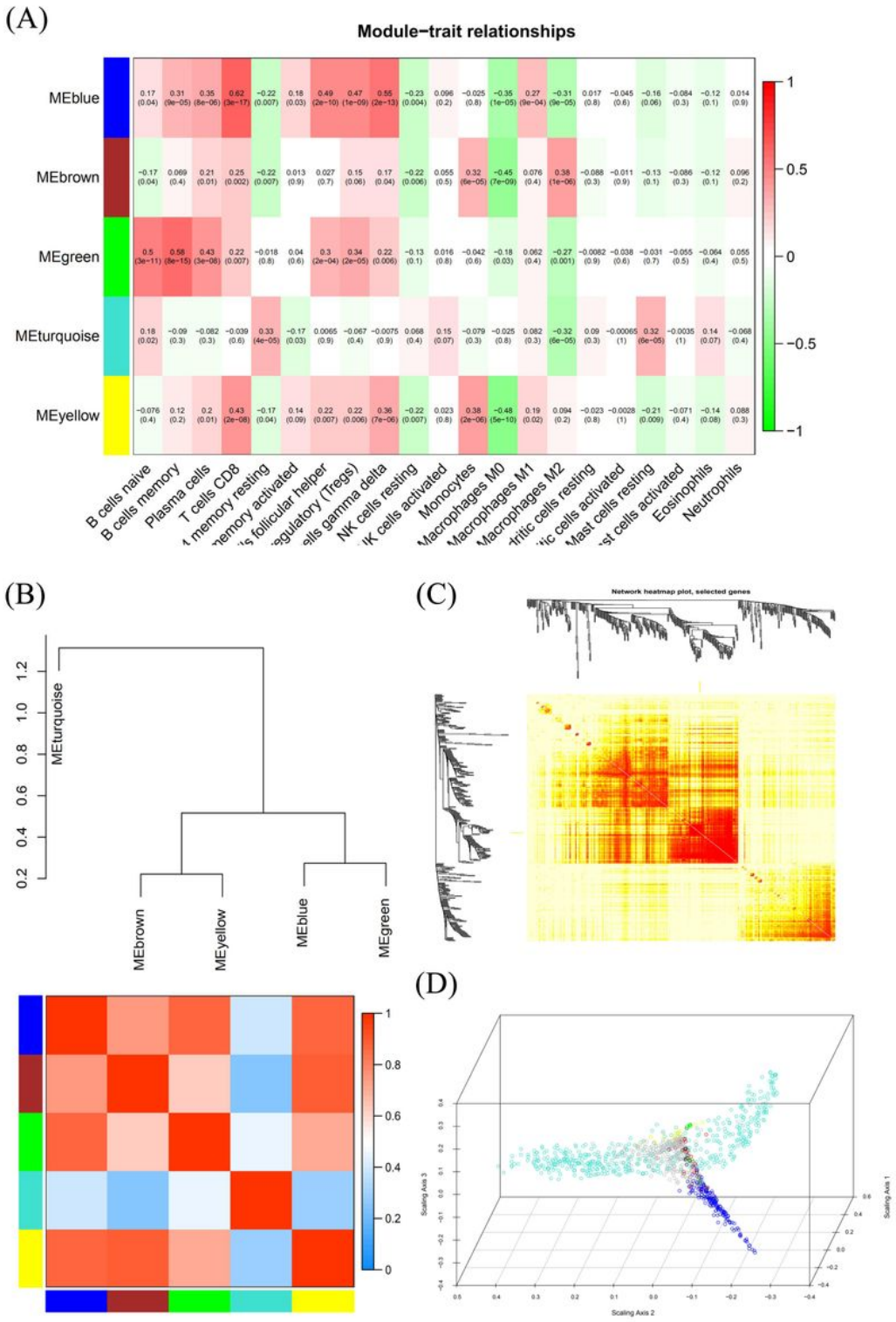

Figure 8

(A) Heatmap of the correlation between module eigengenes and immune cell types; (B) Network dendrogram and heatmap between the module and immune types; (C) Interaction of co-expressed genes; (D) Multidimensional scaling (MDS) plots representing the co-expression network.

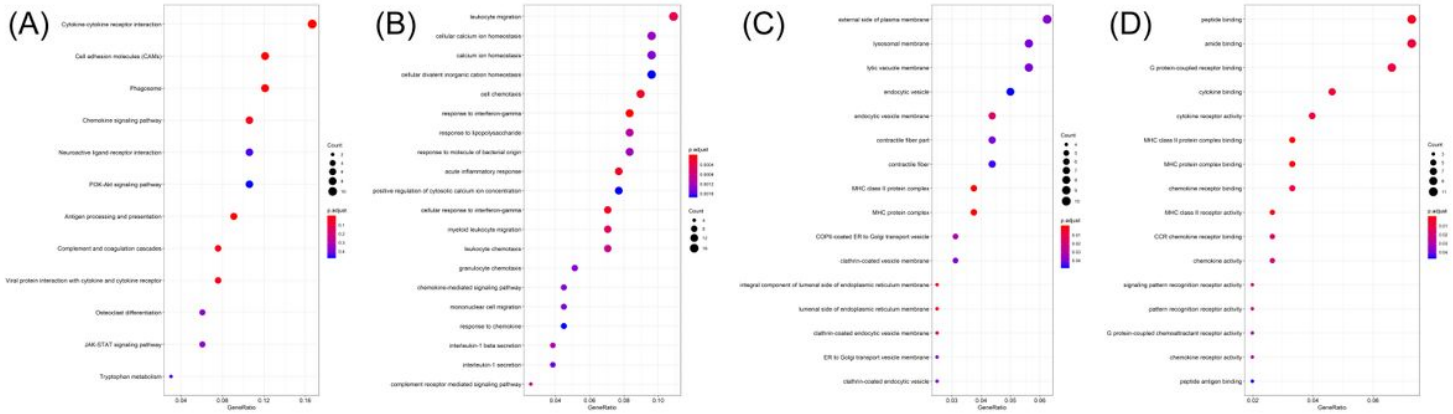

Figure 9

Pathway enrichment and GO analysis of the blue module. (A) KEGG pathway analysis; (B) Biological process (BP) analysis; (C) Cellular component (CC) analysis; (D) Molecular function (MF) analysis.

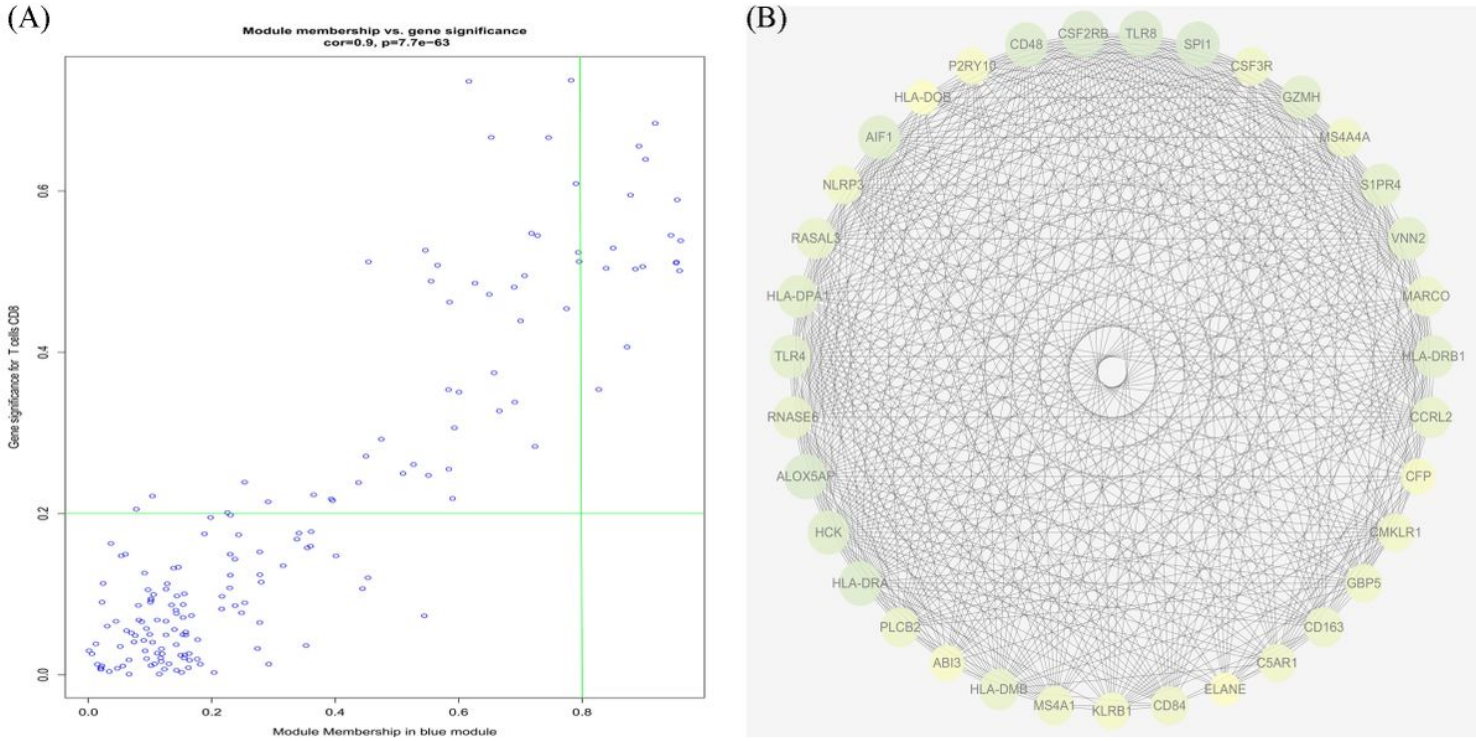

Figure 10

Gene identification. (A) Scatter plot of eigengenes in the blue module; (B) PPI network of genes in the MCODE sub-network.

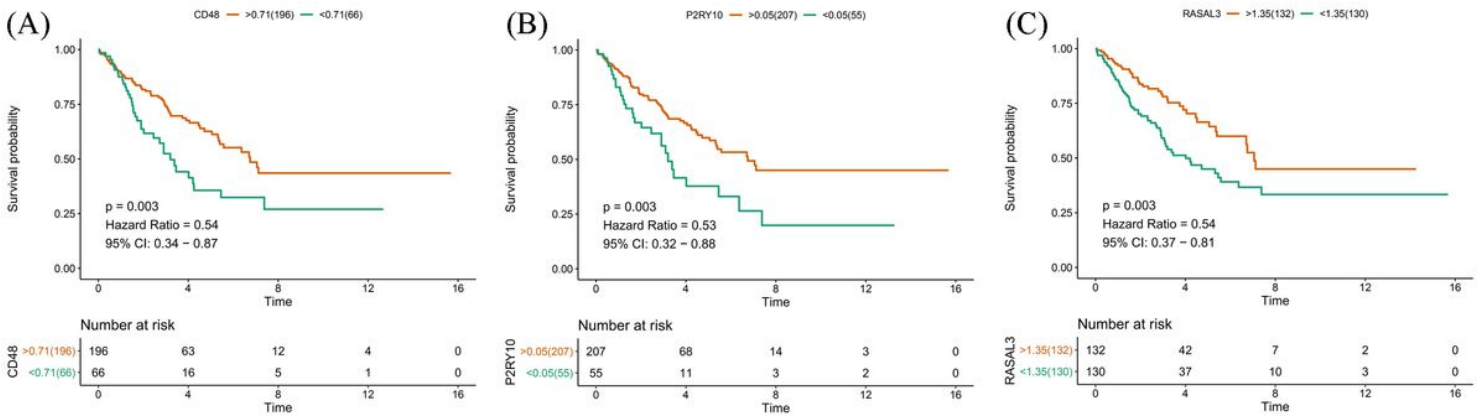

Figure 11

Overall survival of the three genes. (A) CD48; (B) P2RY10; (C) RASAL3.

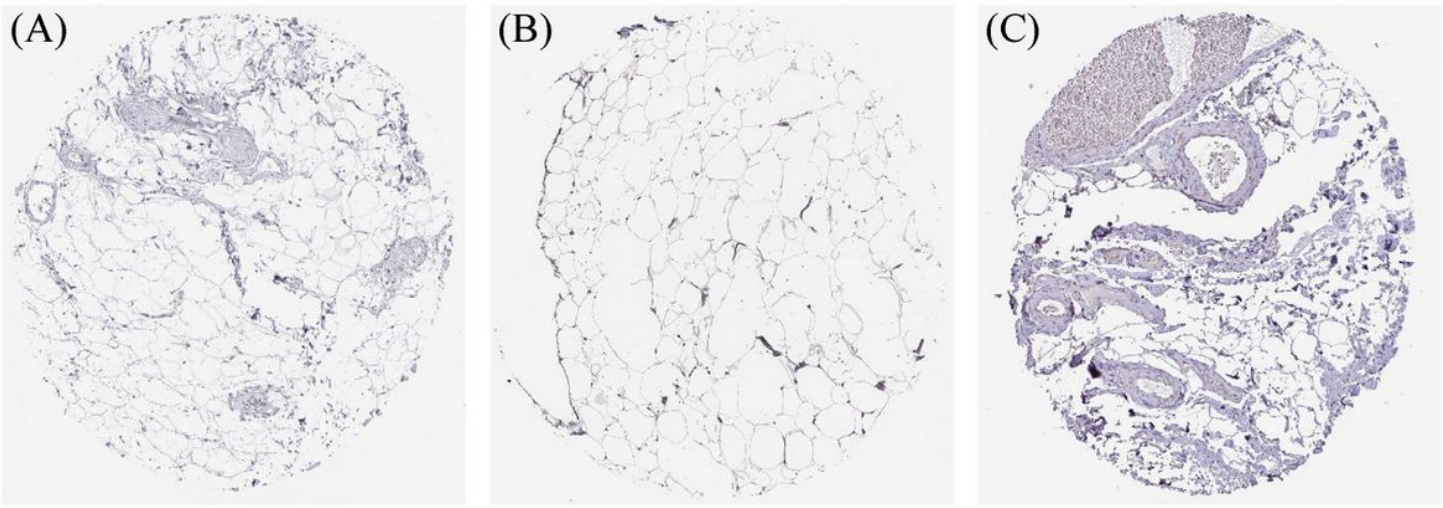

Figure 12

Immunohistochemical analysis of (A) CD48, (B) P2RY10, (C) RASAL3 expression.

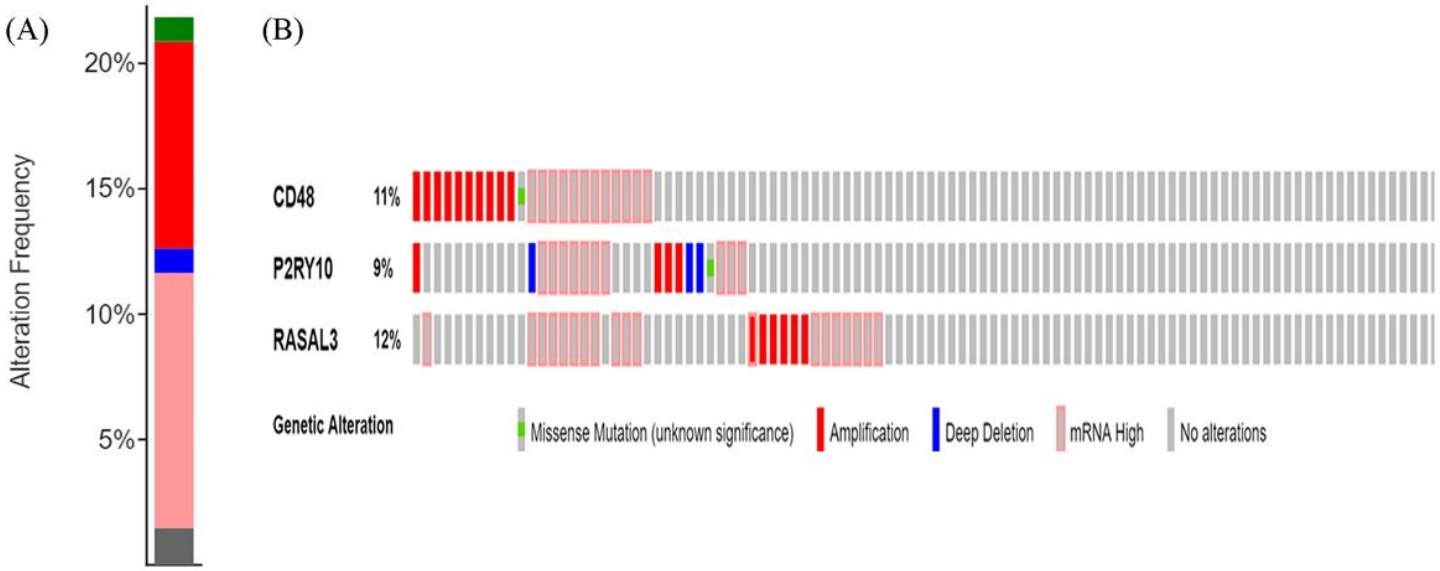

Figure 13

Genetic alterations. (A) Alteration frequency of each of the three hub genes in sarcoma tumors; (B) Visual summary of genetic alternations in the three hub genes.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryList2.xls
- SFigure2.jpg
- SFigure3.jpg
- SFigure4.jpg
- SFigure5.jpg
- SFigure1.jpg
- SupplementaryList1.xls