scientific reports

Check for updates

OPEN

An immune-related eleven-RNA signature-drived risk score model for prognosis of osteosarcoma metastasis

Jia-Song Teng[1] & Yang Wang[2✉]

This study aimed to determine an immune-related RNA signature as a prognostic marker, in this study, we developed a risk score model for predicting the prognosis of osteosarcoma metastasis. We first downloaded the clinical information and expression data of osteosarcoma samples from the UCSC Xena and GEO databases, of which the former was the training set and the latter was the validation set. Immune infiltration was assessed using the ssGSEA and ESTIMATE algorithms, and the osteosarcoma samples were divided into the Immunity_L and Immunity_H groups. Then, eleven RNAs were identified as the optimal prognostic RNA signatures using LASSO Cox regression analysis for establishing a risk score (RS) model. Kaplan–Meier approach indicated the high-risk group exhibited a shorter survival. Furthermore, we analyzed the tumor metastasis, age, and RS model status were determined to be independent clinical prognostic factors using Cox regression analysis. Decision curve analysis (DCA) indicated that the prognostic factor + RS model had the best net benefit. Finally, nine tumor-infiltrating immune cells (TIICs) showed significant differences in abundance between high- and low-risk groups via CIBERSORT deconvolution algorithm. In conclusion, the immune-related eleven-RNA signature be could served as a potential prognostic biomarker for osteosarcoma metastasis.

Keywords Osteosarcoma, Metastasis, Risk score model, Prognosis

Osteosarcoma is the most common malignant primary bone tumor in young adolescents and children[1]. Early metastasis and poor prognosis are the main features of osteosarcoma. Approximately 20% of patients exhibit metastasis at diagnosis [2]. The current standard management of newly diagnosed osteosarcoma involves in surgery, neoadjuvant chemotherapy, and postoperative adjuvant chemotherapy, which has raised the 5-year survival of localized osteosarcoma to approximately 70%[3,4]. However, metastasis remains the main challenge in osteosarcoma therapy as a considerable cause of death in patients with osteosarcoma [5]. The outlook of metastatic osteosarcoma seems to be grim as the 5-year survival drops to as little as 20% and has stagnated over the past few decades, underscoring the critical need for novel therapeutic strategies[6]. Metastasis is closely related to OS prognosis [7]. Thus, the development of valuable prognostic biomarkers is imperative to improve the survival of osteosarcoma patients, especially in the case of metastatic osteosarcoma.

Immunotherapy is becoming an appealing application in osteosarcoma therapy[8]. The tumor microenvironment (TME) shows communication between tumors and immune cells and plays a strong role in the development of osteosarcoma [9,10]. TIICs in the TME are reported to be relevant for the prognosis of osteosarcoma[11]. Thus, a comprehensive analysis of immune infiltration in osteosarcoma may contribute to identifying potential prognostic biomarkers for osteosarcoma. In this study, the immune infiltration of osteosarcoma metastasis was evaluated using several algorithms such as ESTIMATE and CIBERSORT. An immune-related RNA signature was identified as a prognostic marker through LASSO Cox regression analysis and used to build a risk score (RS) model to assess the prognosis of osteosarcoma metastasis. This study might enhance the knowledge of the clinical prognostic outcomes of osteosarcoma metastasis and provide new clues to the potential immune-related therapeutic targets for osteosarcoma metastasis.

[1]The Second Affiliated Hospital of Nanjing University of Chinese Medicine, No. 23 Nanhu Road, Jianye District, Nanjing 210017, Jiangsu, China. [2]Department of Orthopedics, China-Japan Union Hospital of Jilin University, No. 126 Xiantai Street, Changchun 130033, Jilin, China. ✉email: wangyang0859@163.com

nature portfolio

Methods

Data collection and preprocessing

RNA-Seq expression data of GDC TCGA Sarcoma and phenotype data were downloaded from the UCSC Xena database, produced by the Illumina HiSeq 2000 RNA Sequencing platform. In total, 176 samples with expression data, metastatic diagnosis information, and survival data were obtained, which served as the training set for further analysis.

The GSE39055 dataset was downloaded from the NCBI Gene Expression Omnibus, and the platform used was the Illumina HumanHT-12 WG-DASL V4.0 R2 expression beadchip. This dataset was public on Jan 22, 2013[12]. A total of 36 samples with expression and survival data were obtained, which served as the validation set in the following analysis.

Using the annotation file of each platform, both sets were re-annotated to determine the expression levels of the lncRNAs and mRNAs.

Immune infiltration analysis and grouping

ssGSEA was employed to perform immune infiltration analysis of osteosarcoma samples, which was implemented using the GSVA package in R3.6.1. The stromal score, immune score, estimate score, and tumor purity of the osteosarcoma samples were calculated using the ESTIMATE algorithm. According to the results, osteosarcoma samples were split into two immunity groups: Immunity_H and Immunity_L.

Identification of differential expression RNAs (DERs)

According to metastatic diagnosis information, the samples were divided into metastatic and non-metastatic groups. DERs (including DElncRNAs and DEmRNAs) were screened between Immunity_H vs. Immunity_L groups as well as metastasis versus non-metastasis groups by the cutoff of $|\log_2 FC| > 0.263$ and false discovery rate (FDR) < 0.05, using the limma package in R3.6.1. Two DER sets from different comparisons between groups were intersected to obtain the common DERs. Functional enrichment analysis including gene ontology (GO) biology process (BP) and kyoto encyclopedia of genes and genomes (KEGG) [13] was applied to the common DERs using DAVID. The threshold was set at $p < 0.05$.

Identification of prognostic feature DERs and development of RS model for prognosis

In the training set, the survival package in R3.6.1 was utilized to perform univariate Cox regression analysis on the common DERs to identify significant prognostic DERs. Multivariate Cox regression analysis was performed to determine the significant independent prognostic DERs. The threshold was set at log-rank $p < 0.05$. LASSO Cox regression analysis was performed to select prognostic feature RNAs using the penalized package in R3.6.1. Based on the optimal prognostic RNA signature, an RS model was constructed using the following formula:

$$RS = \sum \beta_{RNAs} \times Exp_{RNAs}$$

where β_{RNAs} represents the regression coefficient of the RNA in the LASSO model, while Exp_{RNAs} denotes the expression value of the RNA. Based on this formula, the RS value of each sample in both sets was computed. Samples were split into high- and low-risk groups with the median RS as the cutoff criterion. The two groups were then analyzed for survival using the Kaplan–Meier method, implemented using the survival package in R3.6.1. The RS model was validated using GSE39055.

Identification of independent clinical prognostic factors

In the training set, univariate and multivariate Cox regression analyses were applied to screen for independent clinical prognostic factors using the survival package in R3.6.1. The threshold was a log-rank $p < 0.05$. The results were visualized using the forest plot package in R3.6.1. Decision curve analysis (DCA) was then applied to the single prognostic factor, RS, and prognostic factors + RS models, which were implemented using the rmda package in R3.6.1, to observe the net benefit of each model, thereby analyzing the effect of different factors on prognosis.

Correlational analysis between prognostic feature RNAs and TIIC types

In the training set, the abundance of 22 TIIC types was assessed using the CIBERSORT deconvolution algorithm. A t-test was used to select TIIC types with significant differences in abundance between the high- and low-risk groups. The correlation between the expression of prognostic feature RNAs and TIIC types was measured using Pearson's correlation coefficient, which was calculated using the cor function in R3.6.1.

Results

Immune infiltration assessment and grouping

The flow diagram was shown in Fig. S1. After re-annotation, the expression data of 1398 lncRNAs and 14,631 mRNAs in the training set were obtained. Using ssGSEA, the samples were divided into two clusters (Fig. 1A). Cluster 1 included 69 samples, whereas Cluster 2 had 107 samples. Cluster 1 exhibited more favorable overall survival (Fig. 1B). As depicted in Fig. 2, the stromal, immune, and estimated scores were higher in cluster 2 than in cluster 1, while tumor purity was markedly lower. According to the assessment scores, Cluster 1 was defined as the Immunity_L group, whereas Cluster 2 was defined as the Immunity_H group.
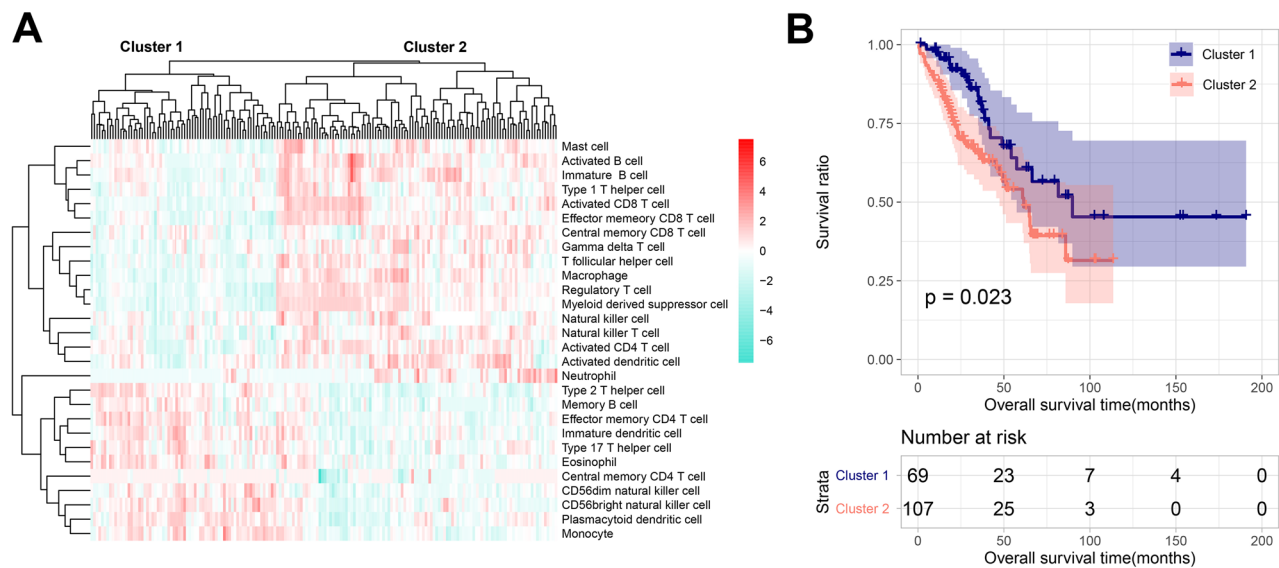
Figure 1. The hierarchical clustering results of immune infiltration of osteosarcoma samples based on ssGSEA (**A**) and Kaplan–Meier curves of survival of different clusters (**B**).
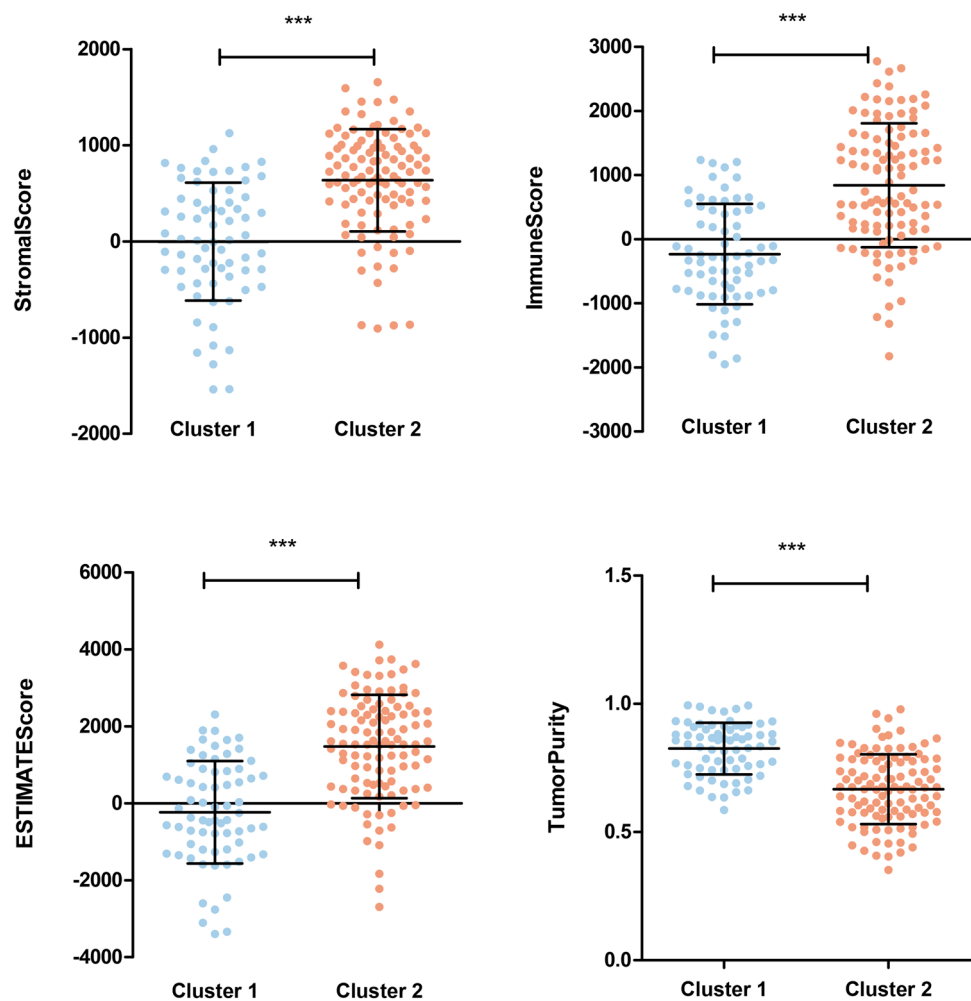

Figure 2. The distribution of stromal score, immune score, estimate score, and tumor purity of different clusters.

DERs and functional enrichment

A total of 542 DERs were obtained between the metastatic and non-metastatic groups (Fig. S2A). Meanwhile, 3624 DERs were obtained between the Immunity_H and Immunity_L groups Fig. S2B). A total of 268 common DERs, including 13 DElncRNAs and 255 DEmRNAs, were found between the two DER sets (Fig. S2C). Functional annotation showed that DEmRNAs were significantly enriched in 14 GO biological processes (Fig. S3A) and 12 KEGG pathways (Figure S3B), such as the immune response and B cell receptor signaling pathway.

RS prognostic model

In the training set, 86 DERs were identified to be significantly related to prognosis via univariate Cox regression analysis, including seven DElncRNAs and 79 DEmRNAs. Through multivariate Cox regression analysis, 60 significant independent prognostic DERs were screened, consisting of 6 DElncRNAs and 54 DEmRNAs. Furthermore, 11 DERs were identified as prognostic feature RNAs (Fig. S4A,B), comprising two DElncRNAs (LINC00324 and LINC00672) and nine DEmRNAs (MAP3K8, SLC2A6, ABCB1, C1orf127, RIPK3, GGT1, RAB7B, BDKRB1, and APCDD1L).

Based on the optimal prognostic RNA signature, the RS model was built as follows:

$$\begin{aligned} RS = &(-0.037703623) * Exp_{MAP3K8} + (-0.026616209) * Exp_{SLC2A6} \\ &+ (-0.023773466) * Exp_{ABCB1} + (-0.013457687) * Exp_{C1orf127} \\ &+ (-0.012065782) * Exp_{LINC00672} + (-0.009766063) * Exp_{LINC00324} \\ &+ (-0.006871635) * Exp_{RIPK3} + (-0.005448974) * Exp_{GGT1} \\ &+ (-0.002570943) * Exp_{RAB7B} + (-0.001708685) * Exp_{BDKRB1} \\ &+ (0.005993841) * Exp_{APCDD1L} \end{aligned}$$

Based on the median RS, the samples were split into high- and low-risk groups in both training dataset (Fig. S5A) and validation dataset (Fig. S5B). The Kaplan–Meier curve revealed that the RS was markedly related to prognosis, and the high-risk group showed a shorter survival (Fig. 3A). Also, the above results were validated using GSE39055 dataset (Fig. S5B, Fig. 3B), suggesting that a high RS was notably relevant to poor prognosis.

Independent clinical prognostic factors

Three independent clinical prognostic factors were identified: including age, tumor metastasis, and RS model status (Fig. 4A). DCA showed that, with respect to the net benefit, the prognostic factors + RS model was superior to other models in most ranges of the threshold probability, indicating that this prediction model had the best clinical utility (Fig. 4B).
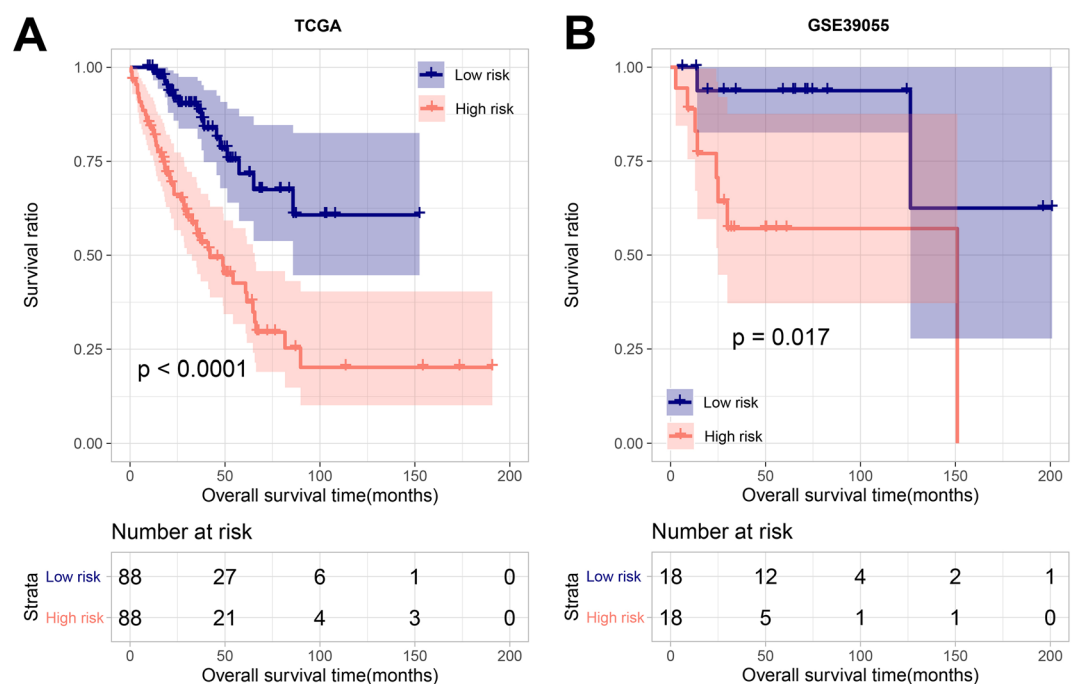

Figure 3. Kaplan–Meier curves of survival of high and low risk groups in the training set (**A**) and the validation set (**B**).
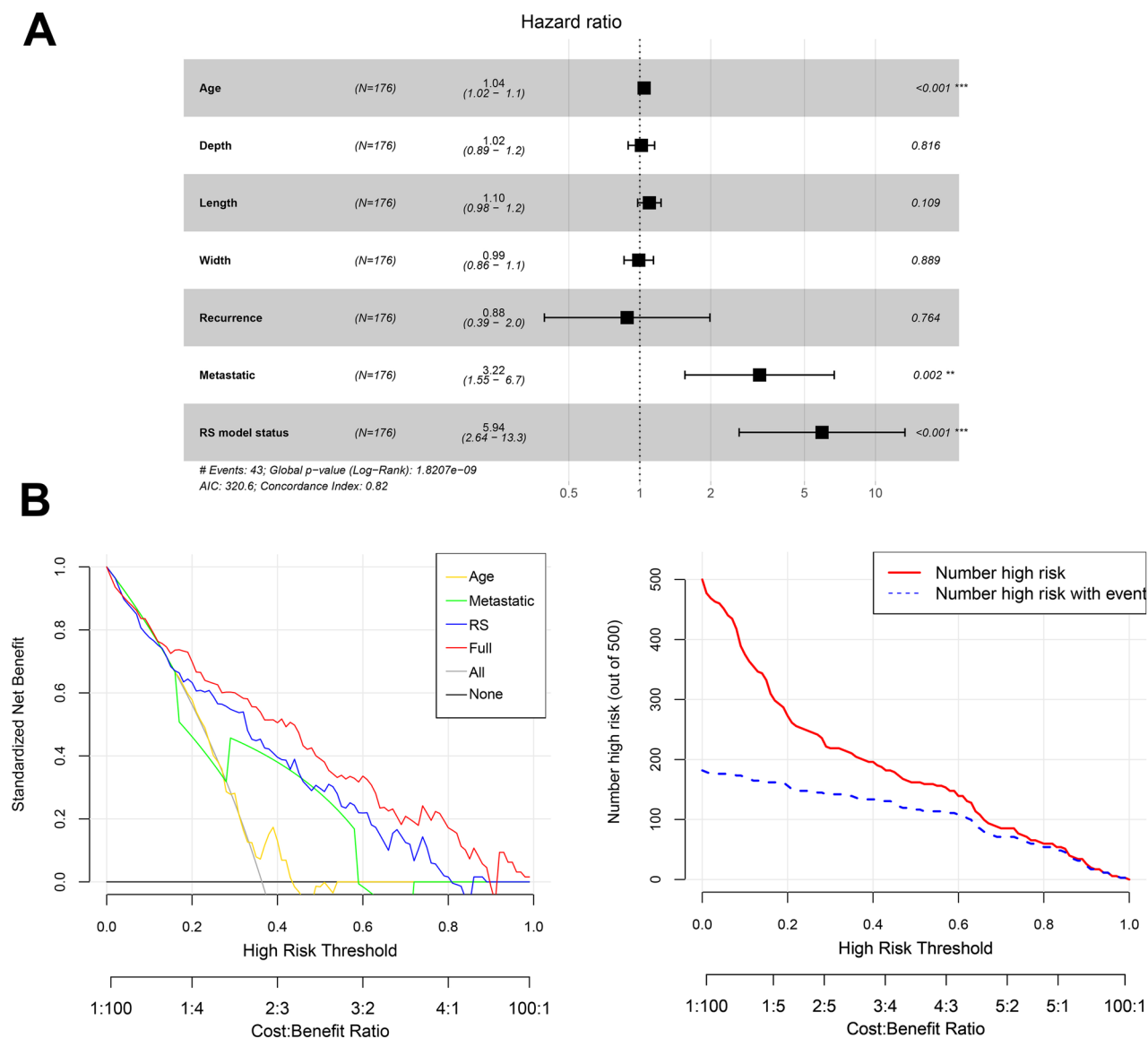
A

B

Figure 4. Identification of independent clinical prognostic factors (**A**) and decision curve analysis of different models (**B**).

Correlational between prognostic feature RNAs and TIIC types

Using the CIBERSORT deconvolution algorithm, 9 of 22 TIIC types exhibited significant differences in abundance between the high- and low-risk groups, including naïve B cell, CD 8 + T cells, γ δ T cells, monocytes, resting NK cells, M0 macrophages, M1 macrophages, M2 macrophages, and resting myeloid dendritic cells (Fig. 5). The correlation between 11 prognostic feature RNAs and 9 TIIC types was assessed by calculating the Pearson correlation coefficient (Fig. 6). For instance, MAP3K8 is associated with eight TIIC types, except for resting myeloid dendritic cells. MAP3K8 is the only gene related to monocytes. The expression of MAP3K8 was relevant to that of LINC00324, C1orf127, GGT1, BDKRB1, and APCDD1L.

Discussion

Osteosarcoma is a high-grade malignant bone tumor with a low overall survival (OS) rate. As one of the key determinants of the lethality of osteosarcoma, metastasis usually leads to a poor prognosis, with only about 20% becoming long-term survivors compared to 70 that of patients with localized disease. Treatment modalities and clinical consequences for osteosarcoma have remained unchanged for decades, prompting the interest in novel strategies for osteosarcoma management and therapy. Immunotherapy has gained attention for osteosarcoma [14,15]. The exploration of valuable immune-related prognostic biomarkers is beneficial for predicting high-risk patients and improving immunotherapy efficacy to improve the long-term survival outcomes of patients with osteosarcoma metastasis.

In the present study, 11 immune-related independent prognostic RNAs were used to generate an RS model to predict the prognosis of osteosarcoma samples. The AUC values of the RS model for assessing 1-, 3-, and 5-year survival were 0.958, 0.927, and 0.920 in the training set, indicating that the RNA signature exhibited a
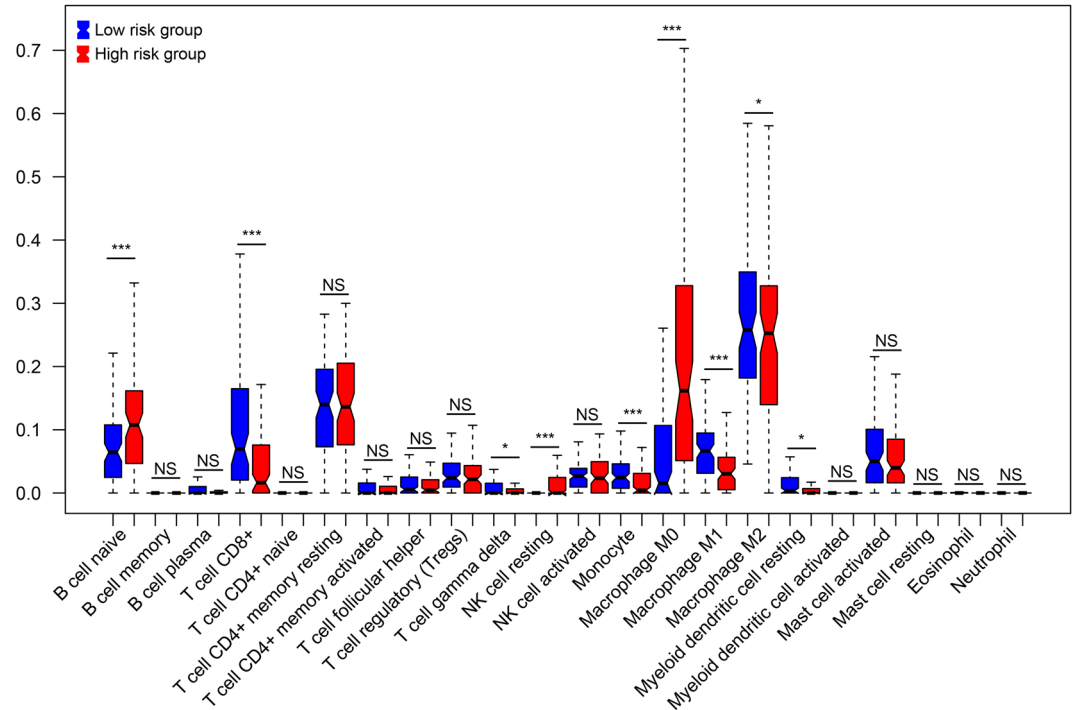
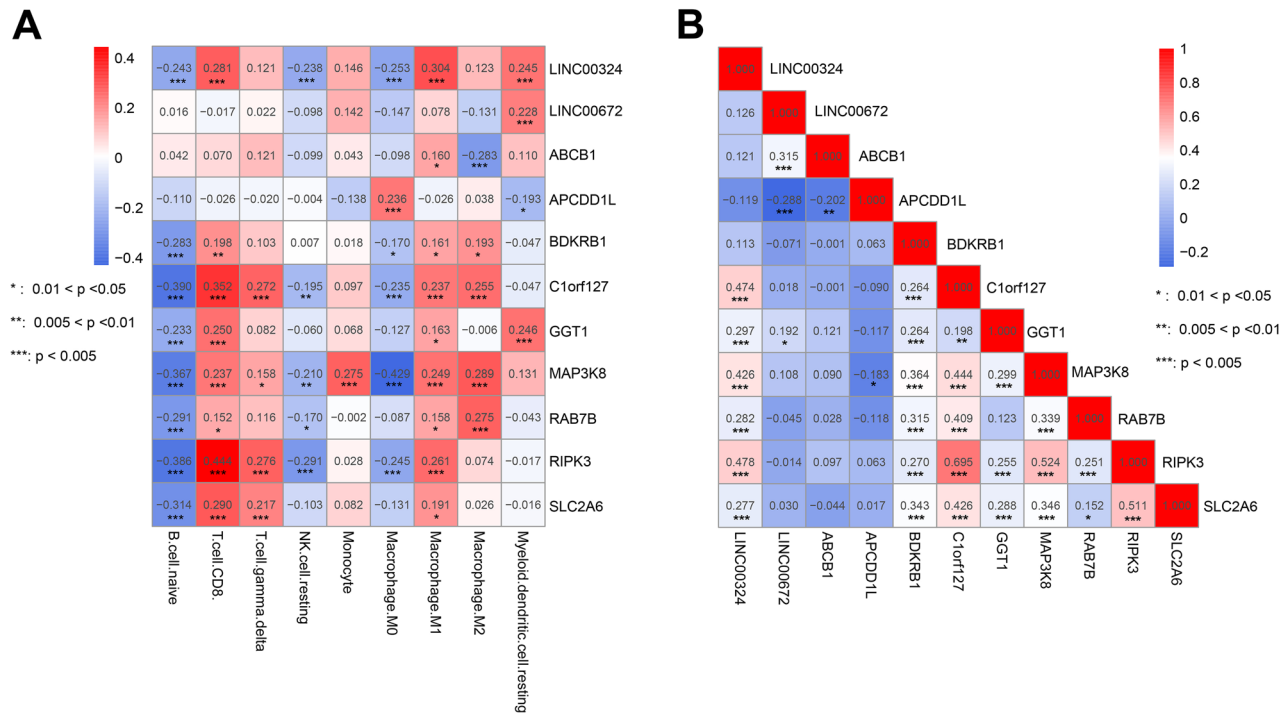**Figure 5.** The abundances of 22 TIIC types between high and low risk groups.

Figure 6. The correlation analysis. (**A**) The correlation matrix of 11 prognostic feature DERs and 9 TIIC types. (**B**) The correlation matrix of all the prognostic feature DERs.

good capability for predicting survival in osteosarcoma samples, which was also verified in the validation set. Cases were divided into high- and low-risk groups based on the median RS. RS was significantly associated with prognosis. Patients in the high-risk group showed poor survival outcomes. More than 60% of the samples in the high-risk group died within 5 years, whereas in the low-risk group, about 25% died. Therefore, appropriate monitoring and intervention should be considered for high-risk groups, for instance, more continuous follow-up visits and active treatments.

The RS model status, tumor metastasis, and age were served as independent prognostic factors for osteosarcoma survival in the current study. It is well admitted that metastasis is one of the most disadvantages among known prognostic factors at diagnosis [16]. The lung is the main site of osteosarcoma metastasis, accounting for approximately 80% of all metastatic cases. Osteosarcoma also metastasizes to bones and lymph nodes. As a multistep complex process that involves invasion, intravasation, dissemination, extravasation, and colonization, metastasis is difficult to control[17]. Despite the introduction of multimodal chemotherapy and surgery, the dismal long-term survival of osteosarcoma metastasis remains for decades [18]. Bimodal age distribution is a characteristic of osteosarcoma, the first peak of which is documented in young children and adolescents, and the second peak is observed in geriatric patients[19]. The first peak is attributed to intense linear bone growth, whereas the second peak is associated with increased bone resorption by osteoclasts[18]. Recently, a meta-analysis demonstrated that older osteosarcoma patients fare worse than younger[20]. However, other studies have shown that age is not a significant independent prognostic factor, which blurs the prognostic role of age in osteosarcoma[21]. Some are optimistic that with better medical technology, age will cease to be a vital factor in prognosis[22].

The immune infiltration landscape of the osteosarcoma samples was observed using the CIBERSORT algorithm, and nine TIICs exhibited significant differences in abundance between the high- and low-risk groups. Macrophages are reported to be the primary members of the immune environment in osteosarcoma[23], which is in line with our results. Monocytes, as precursors of macrophages, can enter tumor tissues and polarize into M1 or M2 macrophages in the TME[24]. In the current study, the high-risk group that showed poor survival had increased levels of M0 macrophages and markedly decreased levels of monocytes and M1 and M2 macrophages, implying that phenotypic changes and polarization levels of macrophages may be associated with the osteosarcoma survival. Zhang et al.[25] a similar view that the polarization of M0 to M1 or M2 macrophages is relevant to the ameliorative outcome of osteosarcoma. Moreover, M2 macrophages are related to osteosarcoma metastasis and poor patient prognosis, whereas a phenotype change from M2 to M1 macrophages causes the regression of lung metastasis of osteosarcoma [26,27]. Some strategies targeting macrophages have displayed a bright future in osteosarcoma therapy, such as blocking the polarization of M1 to M2 macrophages or inhibiting M2 macrophages directly, activating macrophages, and enhancing non-macrophage recruitment[28]. Further basic experiments and clinical investigations focusing on immunotherapy are warranted to improve the survival of osteosarcoma patients.

Bioinformatics plays a crucial role in the discovery of new potential biomarkers for cancer by integrating omics data from various sources. Omics data, including genomics, transcriptomics, proteomics, and metabolomics, provide a comprehensive view of the molecular mechanisms underlying cancer development and progression.

Integration of omics data allows researchers to uncover complex interactions between different molecular layers, leading to a more comprehensive understanding of the underlying biological processes in cancer. In this study, the bioinformatics was performed on the omics data to identify the novel biomarkers for osteosarcoma. This holistic approach facilitates the identification of novel biomarkers that can be used for early detection, prognosis, and prediction of treatment response in cancer patients.

Conclusion

In conclusion, an immune-related eleven-RNA signature-derived risk score model performed well in assessing the prognosis of osteosarcoma metastasis, and these immune-related RNAs may be promising prognostic biomarkers and targets for osteosarcoma metastasis.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

References

1. Savage, S. A. & Mirabello, L. Using epidemiology and genomics to understand osteosarcoma etiology. *Sarcoma* **2011**, 1–13 (2011).
2. Luetke, A., Meyers, P. A., Lewis, I. & Juergens, H. Osteosarcoma treatment–where do we stand? A state of the art review. *Cancer Treat. Rev.* **40**, 523–532 (2014).
3. Mirabello, L., Troisi, R. J. & Savage, S. A. Osteosarcoma incidence and survival rates from 1973 to 2004: data from the Surveillance, Epidemiology, and End Results Program. *Cancer* **115**, 1531–1543 (2009).
4. Isakoff, M. S., Bielack, S. S., Meltzer, P. & Gorlick, R. Osteosarcoma: Current treatment and a collaborative pathway to success. *J. Clin. Oncol.* **33**, 3029 (2015).
5. Yang, C. *et al.* Bone microenvironment and osteosarcoma metastasis. *Int. J. Mol. Sci.* **21**, 6985 (2020).
6. Gianferante, D. M., Mirabello, L. & Savage, S. A. Germline and somatic genetics of osteosarcoma—connecting aetiology, biology and therapy. *Nat. Rev. Endocrinol.* **13**, 480–491 (2017).
7. Sasaki, R., Osaki, M. & Okada, F. MicroRNA-based diagnosis and treatment of metastatic human osteosarcoma. *Cancers* **11**, 553 (2019).
8. Lettieri, C. K. *et al.* Progress and opportunities for immune therapeutics in osteosarcoma. *Immunotherapy* **8**, 1233–1244 (2016).
9. Gomez-Brouchet, A. *et al.* CD163-positive tumor-associated macrophages and CD8-positive cytotoxic lymphocytes are powerful diagnostic markers for the therapeutic stratification of osteosarcoma patients: An immunohistochemical analysis of the biopsies fromthe French OS2006 phase 3 trial. *Oncoimmunology* **6**, e1331193 (2017).
10. Wang, C. *et al.* Macrophage migration inhibitory factor promotes osteosarcoma growth and lung metastasis through activating the RAS/MAPK pathway. *Cancer Lett.* **403**, 271–279 (2017).
11. Liu, P., Xiao, Q., Zhou, B., Dai, Z. & Kang, Y. Prognostic significance of programmed death ligand 1 expression and tumor-infiltrating lymphocytes in axial osteosarcoma. *World Neurosurg.* **129**, e240–e254 (2019).

12. Kelly, A. D. *et al.* MicroRNA paraffin-based studies in osteosarcoma reveal reproducible independent prognostic profiles at 14q32. *Genome Med.* **5**, 2. https://doi.org/10.1186/gm406 (2013).
13. Kanehisa Laboratories. *KEGG Database*, www.kegg.jp/kegg/kegg1.html (1995).
14. Mori, K., Rédini, F., Gouin, F., Cherrier, B. & Heymann, D. Osteosarcoma: Current status of immunotherapy and future trends. *Oncol. Rep.* **15**, 693–700 (2006).
15. Gill, J. & Gorlick, R. Advancing therapy for osteosarcoma. *Nat. Rev. Clin. Oncol.* **18**, 609–624 (2021).
16. Smeland, S. *et al.* Survival and prognosis with osteosarcoma: Outcomes in more than 2000 patients in the EURAMOS-1 (European and American Osteosarcoma Study) cohort. *Eur. J. Cancer* **109**, 36–50 (2019).
17. Yang, Z. *et al.* Long noncoding RNAs in the progression, metastasis, and prognosis of osteosarcoma. *Cell Death Disease* **7**, e2389–e2389 (2016).
18. Czarnecka, A. M. *et al.* Molecular biology of osteosarcoma. *Cancers* **12**, 2130 (2020).
19. Sadykova, L. R. *et al.* Epidemiology and risk factors of osteosarcoma. *Cancer Investig.* **38**, 259–269 (2020).
20. Xin, S. & Wei, G. Prognostic factors in osteosarcoma: A study level meta-analysis and systematic review of current practice. *J. Bone Oncol.* **21**, 100281 (2020).
21. Harting, M. T. *et al.* Age as a prognostic factor for patients with osteosarcoma: An analysis of 438 patients. *J. Cancer Res. Clin. Oncol.* **136**, 561–570 (2010).
22. Bakhshi, S. & Radhakrishnan, V. Prognostic markers in osteosarcoma. *Expert Rev. Anticancer Therapy* **10**, 271–287 (2010).
23. Heymann, M.-F., Lézot, F. & Heymann, D. The contribution of immune infiltrates and the local microenvironment in the pathogenesis of osteosarcoma. *Cell. Immunol.* **343**, 103711 (2019).
24. Lee, H.-W., Choi, H.-J., Ha, S.-J., Lee, K.-T. & Kwon, Y.-G. Recruitment of monocytes/macrophages in different tumor microenvironments. *Biochim. Biophys. Acta Rev. Cancer* **1835**, 170–179 (2013).
25. Zhang, C. *et al.* Profiles of immune cell infiltration and immune-related genes in the tumor microenvironment of osteosarcoma. *Aging* **12**, 3486 (2020).
26. Cersosimo, F. *et al.* Tumor-associated macrophages in osteosarcoma: From mechanisms to therapy. *Int. J. Mol. Sci.* **21**, 5207 (2020).
27. Dhupkar, P., Gordon, N., Stewart, J. & Kleinerman, E. S. Anti-PD-1 therapy redirects macrophages from an M2 to an M1 phenotype inducing regression of OS lung metastases. *Cancer Med.* **7**, 2654–2664 (2018).
28. Wang, Z. *et al.* Innate immune cells: a potential and promising cell population for treating osteosarcoma. *Front. Immunol.* **10**, 1114 (2019).

Author contributions

J.-S.T.: Conceptualization, Data curation, writing—original draft, writing—review, and editing. Y.W.: Writing—review and editing, Methodology, Validation. All authors have read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-54292-6.

Correspondence and requests for materials should be addressed to Y.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.