

A Five-Immune-Related lncRNA Signature Predicts Survival of Patients With Osteosarcoma

Jie Huang (✉ 360480883@qq.com)

Department of Orthopedics, The Second Affiliated Hospital of Guangxi Medical University

<https://orcid.org/0000-0003-2728-3402>

Hongyi Lai

Department of Traditional Chinese Medicine, The Second Affiliated Hospital of Guangxi Medical University

Wentao Qin

Department of Bone and Joint Surgery, The First Affiliated Hospital of Guangxi Medical University

Zhandong Bo

Department of Bone and Joint Surgery, The First Affiliated Hospital of Guangxi Medical University

Zhen Tan

Department of Orthopedics, The Second Hospital of Guangxi Medical University

Zhenjie Wu

Department of Bone and Soft Tissue Surgery, Guangxi Medical University Cancer Hospital

Research Article

Keywords: Osteosarcoma, Immunity, LncRNA, Prognosis, Cox prognostic model

Posted Date: September 17th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-800834/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Osteosarcoma (OS) is the most common primary solid malignant bone tumor, and its metastasis is a prominent cause of high mortality in patients.

Methods: A risk signature was constructed based on re-annotating the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) data matrix, of the lncRNAs related to OS prognosis and immunity. From the OS transcription data, which is downloaded from the TARGET, the 1126 lncRNAs those harbour co-expressions with immunity genes were selected by Pearson correlation test and later divided into the training set ($n=44$) and validation set ($n=41$) with the caret package of R. With the training set we build the model related to Osteosarcoma prognosis by the univariate and multivariate Cox, and the Lasso regression analysis, and in combination with the clinical factors we conducted the multivariate Cox regression analysis to build the 1-year, 3-year and 5-year survival rate nomograms. Afterwards, we validated the ROC and the calibration curve of the subjects with the validation set and the whole dataset. Lastly, we performed functional enrichment analysis with the GSEA, GO and KEGG to figure out the biological functions of the prognosis genes.

Results: The training set was performed in univariate and multivariate Cox regression analysis, identifying 25 lncRNAs correlated with prognosis. Eleven lncRNAs were selected by the least absolute shrinkage and selection operator (LASSO) regression for multivariate cox analysis and Kaplan-Meier (KM) survival analysis. Finally, lncRNAs (RP11-69E11.4, SNHG6, MIR210HG, RP11-750H9.5 and CTD-2341M24.1) risk signature was constructed, and the validation set and the whole dataset were used to evaluate the prediction stability and accuracy of the signature. The survival times of high- and low-risk groups were significantly different in the training set, validation set and the whole dataset. Further, function enrichment and gene set enrichment analysis revealed that the lncRNAs in the signature may affect the proliferation, migration, chemotaxis and combination of Osteosarcoma-related immune cells, and involve in every pathways of OS metabolism.

Conclusion: The five lncRNAs survival risk signature could potentially predict the prognosis of OS patients, additionally, may provide novel insights for future clinical diagnosis and treatment of OS.

Background

Osteosarcoma (OS) derives from primitive bone-forming mesenchymal cells and is the most common primary bone malignancy [1]. Characterized by strong invasiveness and metastasis, it's affecting children and adolescents worldwide [2]. Current treatment of OS includes surgical resection of all gross disease in conjunction with systemic chemotherapy to control micro-metastatic disease [3]. Despite this, OS is highly susceptible to recurrence and metastasis, resulting in a 5-year survival rate of less than 25% for patients [3, 4], indicating that OS metastasis is associated poor long-term prognosis. Among the novel treatments currently available for OS, immunotherapy may be one of the most effective ones [5, 6]. Osteosarcoma is with a distinctive tumor microenvironment driven by micromolecular features that are

importantly linked to the immune response [6]. It has been shown that immune infiltration in the tumor microenvironment affects the prognosis of OS patients [7], and the immune gene CXCL12 plays an important role of mediation in the invasiveness and metastasis of OS [8], and immunity factors play a very key role in the development of OS. Long noncoding RNA (lncRNA) is an RNA greater than 200 bp in length and does not have the ability to encode proteins because of lack of open reading frame [9]. However, it is an important regulator factor in the human body and is involved in many cellular processes such as chromatin remodeling, transcriptional regulation and gene expression, which promotes or inhibits the invasiveness and metastasis of many malignant tumors including OS [10]. In recent years, lncRNAs have been found to play important regulatory roles in various processes of tumor immune response [11], and their overexpression leads to the growth and proliferation of OS cells and promotes the development of OS. Previous studies have shown that lncRNA TUG1, lncRNA SNHG4, and lncRNA TTN-AS1 can promote the invasiveness and metastasis of OS and can predict the prognosis of OS patients [12–14]. As the research goes deeper, more and more prognostic genes related to the survival of OS patients have been discovered. However, there are currently few studies on immune-related lncRNAs in the prognosis of OS. This study was conducted to explore the prognostic relationship of immune-related lncRNAs in OS and to construct a stable nomogram by obtaining Osteosarcoma-related transcriptome information from the TARGET database, in expectation to offer referencing and assistances to the clinicians.

Methods

Gene expression profiles and data preparation

UCSC Xena(<https://xena.ucsc.edu/>) is an institution that integrates The Cancer Genome Atlas (TCGA), International Cancer Genome Consortium (ICGC) and the TARGET database (Therapeutically Applicable Research To Generate Effective Treatments), among others. TARGET database was downloaded in September 2020 using the terms “osteosarcoma”. Transcriptomic data in FPKM (Fragments per Kilobase Million) format for 88 OS patients and clinical information data for 150 cases were contained in the gene expression data and clinical information files of OS patients, and finally, by screening, 85 cases containing complete survival data and 45 cases containing complete clinical data were available, including 18 female patients and 27 male patients, 34 cases without metastasis and 11 cases with metastasis, 40 patients that had surgical amputation and 5 patients didn't.

Analysis methods

Data processing and acquisition of immune-related lncRNAs

Immune-related genes were directly downloaded from Gene Lists in the IMMPORt database (<https://www.immport.org/>), with a total of 2498 genes. We converted ensembl gene ID to gene symbol ID according to the human genome annotation file (ENSEMBL, <http://asia.ensembl.org/index.html>), and separately extracted the expression of lncRNAs and protein-codable mRNAs (Protein coding). Further, the genes that were duplicated in the data were averaged using the limma R package, and the expression of immune-related genes were identified from the resulting expression matrix of mRNAs using Strawberry

Perl software. Subsequently, The $P < 0.001$ and $|cor| > 0.4$ were used as the cut-off criteria, lncRNAs and immune-related genes were performed co-expression analysis using R package.

For the stability of the model, the data of 85 patients with immune-related lncRNAs were randomly divided into a training set ($n=44$) and a validation set ($n=41$) using the R package. The training set was used to establish a prognostic model, and the validation set was used to evaluate the stability of the model.

Construction and evaluation of prognosis risk score model

Univariate and multivariate Cox regression analysis of the training set was performed using R package Survival to screen for lncRNAs with independent prognosis. Lasso regression with R package glmnet was used to select appropriate variables to simplify the final signature and avoid over-fitting of the model. The minimum error λ was selected by 10-fold cross-validation to obtain minimum partial likelihood deviance [15]. Further, Kaplan-Meier (KM) survival analysis combined with the log-rank test was used to determine the relationship between the selected lncRNAs and the overall survival rate of OS patients, and high- and low-expression was divided by the median of lncRNAs expression. The risk score of the training set was calculated from the obtained independent prognostic lncRNAs by multiplying the coefficient of regression (Coef) of the multivariate Cox regression analysis by the sum of the expression of each independent prognostic lncRNAs for that patient. Additionally, the risk scores of the 85 patients with OS were calculated using the risk signature to divide patients into high- and low-risk groups, taking the median of risk scores as the cut-off. The prognosis differences between the two patients with OS groups, was evaluated by KM survival analysis. Finally, a nomogram that was capable of predicting the survival of OS patients based on the risk scores for the overall data set combined with OS clinical information, including age, sex, metastasis and amputation, was formulated in the univariate and multivariate Cox regression analysis.

Validation of the risk score model

High- or low-risk of the validation set and the overall data set was determined by the median of risk of the training set, and the time-dependent receiver operating characteristic (ROC) curve was used to judge the accuracy of the 1-year, 3-year and 5-year overall survival of the OS patients. The area under the ROC curve (AUC) greater than 0.7, indicates that the model was of good accuracy. The accuracy of the nomogram was judged by the calibration curve [16].

Functional enrichment analysis

The mRNAs associated with the lncRNAs in the prognostic model were extracted, on which gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment analysis were conducted to understand the molecular biological function of the lncRNAs in the prognostic model.

Gene set enrichment analysis (GSEA)

The overall data set was divided into high- and low-risk groups by median of the risk score of the model, and then, the high- and low-risk groups were analyzed by weighted enrichment method of the GSEA 4.0.3 (<http://www.broadinstitute.org/gsea>). The gene set was selected from MSIGDB (<http://www.broadinstitute.org/gsea/msigdb>) , where genes with $P < 0.05$, absolute normalized enrichment score (NES) ≥ 1 , and corrected P (FDR) < 0.05 were used as significantly enriched gene sets.

Results

Identification of immune-related lncRNAs correlated with OS prognosis

A total of 1126 immune-related lncRNAs and their expression were obtained through co-expression analysis with immune genes. Based on the training set, a total of 25 lncRNAs correlated with OS prognosis at $P < 0.01$ were detected in univariate Cox regression analysis with the R package Survival (Table 1).

Table 1 Univariate cox regression analysis for 25 immune-related lncRNAs were correlated with Osteosarcoma

id	HR	Lower 95% CI	Upper 95% CI	P value
ZFAS1	2.832085	1.378306	5.819248	0.00460853
RP11-69E11.4	2.01796	1.287032	3.163994	0.00221606
RP11-304F15.3	2.056553	1.203551	3.51411	0.00834548
EPB41L4A-AS1	2.919649	1.305154	6.531299	0.00910008
RP11-535M15.1	2.235202	1.246264	4.008885	0.00696358
SNHG7	2.843028	1.41978	5.692998	0.00318482
RP3-460G2.2	0.44072	0.244987	0.792835	0.0062419
GAS5	2.700873	1.667261	4.375268	0.0000542
OLMALINC	3.011264	1.318521	6.877184	0.0088915
ELFN1-AS1	2.148336	1.40772	3.278599	0.00039184
UNC5B-AS1	2.272839	1.403979	3.679398	0.00083625
RP5-894A10.2	4.0175	1.454702	11.09527	0.00729447
SNHG6	2.827236	1.511355	5.288806	0.00114413
MIR210HG	2.141462	1.25106	3.665579	0.00549075
RP11-77403.3	0.255121	0.09958	0.653615	0.00442894
RP11-21L23.3	2.629617	1.320111	5.238108	0.00596264
RP11-750H9.5	0.302797	0.137799	0.665361	0.00293673
RP11-136I14.5	1.618776	1.135072	2.308607	0.00782557
CTD-2341M24.1	0.229614	0.075576	0.697613	0.00945731
RP11-45P15.4	0.268225	0.098707	0.728871	0.00987941
RP1-239B22.5	2.383866	1.249241	4.549015	0.00841514
RP11-81A22.5	1.439271	1.106446	1.872212	0.00665094
PARD6G-AS1	2.333596	1.31905	4.128477	0.00359924
RP5-894A10.6	3.627491	1.576219	8.348263	0.002446
CTD-2227E11.1	2.839231	1.422439	5.667188	0.00308429

Abbreviations: TARGET, Therapeutically Applicable Research to Generate Effective Treatment; HR, hazard ratio;

CI, confidence interval

Construction and evaluation of the risk score signature

For the 25 prognosis-related lncRNAs, LASSO Cox regression analysis was used to select appropriate lncRNAs to construct the model [17]. Eleven lncRNAs were obtained by 10-fold cross-validation, in which the partial likelihood deviance was at its minimum (Fig. 1). Subsequently, multivariate Cox regression analysis and KM survival analysis were conducted to construct the risk signature. A combination of five lncRNAs (RP11-69E11.4, SNHG6, MIR210HG, RP11-750H9.5 and CTD-2341M24.1) (Table 2) was selected as the most suitable predictor for prognosis. A risk score model was constructed to predict the prognosis of patients with OS as follows: RP11-69E11.4 × 1.219 + MIR210HG × 0.905 + MIR210HG × 1.000 - RP11-750H9.5 × 1.479 - CTD-2341M24.1 × 1.763.

Table 2

Multivariate cox regression analysis for RP11-69E11.4, SNHG6, MIR210HG, RP11-750H9.5, and CTD-2341M24.1

id	Coefficient	HR	Lower 95% CI	Upper 95% CI	P value
RP11-69E11.4	1.219242	3.384621	1.834678	6.243962	0.000095
SNHG6	0.904814	2.471473	1.102726	5.53916	0.027988
MIR210HG	0.999502	2.716928	1.409203	5.238208	0.002844
RP11-750H9.5	-1.4786	0.227956	0.080779	0.643281	0.005215
CTD-2341M24.1	-1.76314	0.171506	0.043114	0.68225	0.012325

Abbreviations: coef, coefficient

Taking the median of risk scores as the cut-off, the OS patients were classified into high-and low-risk groups. Across the dataset, the KM survival analysis also showed a significant correlation of the 5 lncRNAs with survival and prognosis, that high-risk group patients had significantly poorer prognosis compared to low-risk (P< 0.01) (Fig. 2).

The high-risk and low-risk groups was compared in survival analysis curves, indicating that the overall survival period of the patients in the low-risk group was much higher than that of the high-risk group, and the risk score and number of deaths in the high-risk group were also much higher than that of the low-risk group (Figs. 3.1, 3.2 and 3.3).

The time-dependent AUCs for the risk signature of training set were 0.925 at 1-year, 0.929 at 3-year, and 0.961 at 5-year (Fig. 4a), of validation set were 0.921 at 1-year, 0.898 at 3-year, and 0.878 at 5-year (Fig. 4b), of overall data set were 0.913 at 1-year, 0.881 at 3-year, and 0.882 at 5-year (Fig. 4c). All of sets were greater than 0.87, indicating good prediction accuracy of the model.

Furthermore, the validation set and the overall data set were well consistent with the training set that suggested by the calibration curves for 1-year, 3-year and 5-year survival rates (Fig. 5).

Construction and verification of the nomogram

Incorporating clinical factors of the OS dataset into the model of the overall dataset, the results of the univariate and multivariate Cox regression analysis showed that our model also played a significant role in clinical predictability (Table 2). Therefore, a nomogram integrating the risk score and four clinical factors was established to predict OS prognosis (Fig. 6). Each variable has its own score, and the 1-year, 3-year and 5-year predicted survival can be obtained by summing the scores of the different variables on the nomogram.

GO and KEGG enrichment analysis

In GO term, the results of immune genes co-expressed with modeled lncRNAs, showed that the biological process (BP) of co-expressed immune-related genes was enriched to the leukocyte migration and proliferation, cell chemotaxis and monocyte proliferation. The cellular component (CC) was enriched to the MHC protein complex, the outer side of the plasma membrane, the MHC class II protein complex and so on. The molecular function (MF) was enriched to cytokine binding, receptor-ligand activity and cytokine receptor activity (Fig. 7). All above of them were associated with the risk signature, which affect the proliferation, migration, chemotaxis and combination of Osteosarcoma-related immune cells. In addition, co-expressed immune-related genes were involved in KEGG pathways, such as B-cell receptor signaling, NF- κ B signaling pathway, MAPK signaling pathway, chemokine signaling pathway, JAK-STAT signaling pathway and some other cancer-related pathways (Fig. 8). The lncRNAs play a key role in the development of OS and contribute to it was indicated in above enrichment results.

Gene set enrichment analysis

The GSEA software provides a method for analyzing large gene expression data, interpreting the gene expression data, and aggregating and visualizing genes with common biological functions [18]. With the data of high- and low-risk groups were identified through GSEA, it was found that the high-risk group in the model significantly clustered into the immune-related gene set, while it gradually decreased when it came to the low-risk group, as shown in Fig. 9. Combining with the former overall survival analysis suggested that the survival of the low-risk group was significantly higher than that of the high-risk group, indicating that lncRNAs may play a crucial role in the prognosis of OS.

Discussion

OS is a malignant tumor, a highly liable to recurrence and metastasis, poses a significant health threat to adolescents and children. Although OS can be treated at present, once metastasis occurs, the mortality rate will increase dramatically. The survival rate of OS patients with metastasis has been shown to be lower than patients with OS without metastasis [19, 20]. It has long been found that lncRNAs are significantly linked to metastasis in a variety of cancer cells [21-23], so finding prognosis related lncRNAs is of great value for improving the survival of OS patients. In recent years, lncRNAs have increasingly been found to play a key role in the prognosis of tumors and have become the focus of research in the field of oncology. Immunity has long been used as a measure to treat OS, however, immune-related

lncRNAs are rare in the study of OS, so it is non-trivial to find a prognostic model of immune-related lncRNAs that can predict the survival of OS patients.

In our study, we re-annotated the OS gene chips from the TARGET database, among which 25 survival-associated lncRNAs were detected in univariate Cox regression analysis. Using LASSO regression, 11 lncRNAs were selected to construct the risk signature. Further, multivariate Cox regression analysis and KM survival analysis were performed to construct and optimize the prognosis risk signature, five immune-related lncRNAs were found, that are closely associated with OS prognosis, including RP11-69E11.4, SNHG6, MIR210HG, RP11-750H9.5 and CTD-2341M24.1. On these immune-related lncRNAs, a prognostic risk scoring model was constructed. The overall survival analysis of the model and the survival analysis of modeling-related lncRNAs were significant, verifying the relationship between these lncRNAs and the prognosis of OS. Meanwhile, the analysis of the model combined with clinical factors showed that the model we constructed made better predictions than clinical factors such as age, sex, metastasis and amputation. In addition, in GSEA analysis, the model was strongly enriched in the molecular set of immunity, which suggesting a close link with immune response.

Among the modeling lncRNAs in this study, several have been reported to be associated with various tumors. RP11-69E11.4 was found to serve as a novel clinical marker for predicting recurrent OS [24], and was also found in this study to be a clinical indicator for predicting the prognosis of OS. In liver cancer, glioma and colorectal cancer [25-27], lncRNA SNHG6 is significantly differentially expressed and involved in cancer initiation and progression, and is a prognostic marker for a variety of cancers. LncRNA MIR210HG, like lncRNA SNHG6, is a novel long non-coding RNA found to be oncogenic in a variety of cancers [28, 29]. MIR210HG was found to promote invasion and metastasis of OS by sponging miR-503 [30], which was also shown to influence the prognosis of OS. Both RP11-750H9.5 and CTD-2341M24.1, however, have not been reported in OS and are of potential value. The above results indicate that the five lncRNAs in this study harbour some prognostic value in OS, but they have not been specifically investigated and need to be further explored.

To figure out the molecular biological functions of the modeled lncRNAs, GO and KEGG functional enrichment and annotation were conducted on co-expressed immune-related mRNAs using clusterProfiler R package. The results indicated that these genes were mainly enriched in monocytes, leukocyte migration and proliferation, MHC, MHC class II protein complexes, all of which were related to immune response. In addition, the NF- κ B signaling pathway, MAPK signaling pathway, chemokine signaling pathway, and JAK-STAT signaling pathway were associated with these lncRNAs. NF- κ B, a protein complex present in almost all cells, is involved in a variety of in vivo responses, especially in the immune response, and NF- κ B false expression is to a great extent associated with the development of cancer and its metastasis [31, 32]. The MAPK pathway is one of the common intersecting pathways of signaling pathways such as inflammation, apoptosis, stress, through which signals are transmitted to cells and are involved in cell proliferation, carcinogenesis and metastasis [33]. The different types of chemokines in the chemokine signaling pathway can modify the expression of signals from other pathways and play an indirect or direct role in the occurrence of diseases, including the JAK-STAT signaling pathway. The JAK-

STAT signaling pathway is a cytokine-stimulated signal transduction pathway in which JAK, by phosphorylating STAT to form a STAT dimer, initiates transcription into the nucleus. However, JAK can also activate SHP2, an upstream molecule of MAPK, thereby inducing activation of the MAPK pathway. The above analysis suggests that the five lncRNAs may be involved in the development of OS by influencing these pathways.

There are number of limitations to our study. First, the results have not been validated by in vivo, in vitro, and clinical studies. Thus, further studies are needed to confirm the correlation between lncRNAs and OS prognosis in clinical samples. Second, some important lncRNAs may have been missed in our multiple screening processes. Further, heterogeneity between datasets was unavoidable.

Conclusion

In conclusion, in this study, by the method of re-annotating the TARGET database, we obtained five lncRNAs that are associated with OS prognosis and immunity, on which we constructed lncRNA molecular markers that are associated with OS prognosis. Modeling was performed and then further validated its stability and accuracy, found that the model was capable of making good predictions to the prognosis of OS patients and can be used as a potential marker combination for predicting the prognosis of OS. Further, lncRNAs may provide novel insights for future clinical diagnosis and treatment of OS.

Declarations

Ethics approval and consent to participate

The study had been approved by the ethical committee of the participating hospitals.

Consent for publication

Consent for publication was obtained from every individual whose data are included in this manuscript.

Availability of data and materials

UCSC Xena(<https://xena.ucsc.edu/>)

The IMMPORt database (<https://www.immport.org/>)

ENSEMBL (<http://asia.ensembl.org/index.html>)

The GSEA 4.0.3 (<http://www.broadinstitute.org/gsea>)

MSIGDB (<http://www.broadinstitute.org/gsea/msigdb>)

Competing interests

The authors declare that they have no competing interests.

Funding

The research projects were supported by Guangxi Natural Science Fundation (NO:2020GXNSFBA159016)

Authors' contributions

HJ conceptualized and designed the study, drafted the initial manuscript, and reviewed and revised the manuscript. LHY, QWT, and WZJ collected data, carried out the initial analyses, constructed the risk assessment model of osteosarcoma prognosis, and reviewed and revised the manuscript. TZ, and BZD coordinated and supervised data analysis, and critically reviewed the manuscript for important intellectual content. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Acknowledgements

Not applicable.

Authors' information

Jie Huang,¹ Hongyi Lai,² Wentao Qin,³ Zhandong Bo,⁴ Zhen Tan,⁵ Zhenjie Wu,⁶

¹ Department of Orthopedics, The Second Affiliated Hospital of Guangxi Medical University, Nanning, China, 530007, Email: 360480883@qq.com

² Department of Traditional Chinese Medicine, The Second Affiliated Hospital of Guangxi Medical University, Nanning, China, 530007

³ Department of Bone and Joint Surgery, The First Affiliated Hospital of Guangxi Medical University, Nanning, China, 530021

⁴ Department of Bone and Joint Surgery, The First Affiliated Hospital of Guangxi Medical University, Nanning, China, 530021

⁵ Department of Orthopedics, The Second Affiliated Hospital of Guangxi Medical University, Nanning, China, 530007, Email: tanzhen99@126.com

⁶ Department of Bone and Soft Tissue Surgery, Guangxi Medical University Cancer Hospital, Nanning, China, 530021

References

1. Ottaviani G, Jaffe N. The epidemiology of osteosarcoma. *Cancer Treat Res.* 2009;152:3–13.
2. Angulo P, et al., *Natural compounds targeting major cell signaling pathways: a novel paradigm for osteosarcoma therapy*. *Journal of Hematology & Oncology*, 2017. 10(1).
3. Harrison DJ, et al. Current and future therapeutic approaches for osteosarcoma. *Expert Rev Anticancer Ther.* 2017;18(1):39–50.
4. Kansara M, et al. Translational biology of osteosarcoma. *Nat Rev Cancer.* 2014;14(11):722–35.
5. Lettieri CK, et al. Progress and opportunities for immune therapeutics in osteosarcoma. *Immunotherapy.* 2016;8(10):1233–44.
6. Nathenson MJ, Conley AP, Sausville E. Immunotherapy: A New (and Old) Approach to Treatment of Soft Tissue and Bone Sarcomas. *Oncologist.* 2017;23(1):71–83.
7. Koirala P, et al., *Immune infiltration and PD-L1 expression in the tumor microenvironment are prognostic in osteosarcoma*. *Scientific Reports*, 2016. 6(1).
8. Li B, et al. Epigenetic Regulation of CXCL12 Plays a Critical Role in Mediating Tumor Progression and the Immune Response In Osteosarcoma. *Can Res.* 2018;78(14):3938–53.
9. Jandura A, Krause HM. The New RNA World: Growing Evidence for Long Noncoding RNA Functionality. *Trends Genet.* 2017;33(10):665–76.
10. Smolle M, Pichler M. *The Role of Long Non-Coding RNAs in Osteosarcoma*. *Non-Coding RNA*, 2018. 4(1).
11. Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet.* 2016;17(1):47–62.
12. Fu D, et al., *LncRNA TTN-AS1 regulates osteosarcoma cell apoptosis and drug resistance via the miR-134-5p/MBTD1 axis*. *Aging (Albany NY)*, 2019. 11(19).
13. Sheng K, Li Y. LncRNA TUG1 promotes the development of osteosarcoma through RUNX2. *Experimental and Therapeutic Medicine*; 2019.
14. Xu R, et al., *LncRNA SNHG4 promotes tumour growth by sponging miR-224-3p and predicts poor survival and recurrence in human osteosarcoma*. *Cell Proliferation*, 2018. 51(6).
15. Mao, et al., *A seven-lncRNA signature predicts overall survival in esophageal squamous cell carcinoma*. *Scientific reports*, 2018.
16. Harrell FEJ, et al. Evaluating the Yield of Medical Tests. *Jama.* 1982;247(18):2543–6.
17. Vidaurre D, Bielza C, Larrañaga P. Lazy lasso for local regression. *Computational Stats.* 2012;27(3):531–50.
18. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102(43):15545–50.
19. Bacci G, et al., *Osteogenic sarcoma of the extremity with detectable lung metastases at presentation*. *Cancer*, 1997.
20. Rasalkar DD, et al. Pulmonary metastases in children with osteosarcoma: characteristics and impact on patient survival. *Pediatr Radiol.* 2011;41(2):227–36.

21. Dhamija S, Diederichs S. From junk to master regulators of invasion: lncRNA functions in migration, EMT and metastasis. *Int J Cancer*. 2016;139(2):269–80.
22. Kim J, et al. Long noncoding RNA MALAT1 suppresses breast cancer metastasis. *Nat Genet*. 2018;50(12):1705–15.
23. Wang X, et al. The lncRNA TP73-AS1 promotes ovarian cancer cell proliferation and metastasis via modulation of MMP2 and MMP9. *J Cell Biochem*. 2018;119(9):7790–9.
24. Ying T, et al., *The lncRNAs RP1-261G23.7, RP11-69E11.4 and SATB2-AS1 are a novel clinical signature for predicting recurrent osteosarcoma*. Bioscience Reports, 2020. 40(1).
25. Birgani MT, et al. Long Non-Coding RNA SNHG6 as a Potential Biomarker for Hepatocellular Carcinoma. *Pathology Oncology Research*. 2017;24(2):329–37.
26. Cai G, et al., *LncRNA SNHG6 acts as a prognostic factor to regulate cell proliferation in glioma through targeting p21*. *Biomedicine & Pharmacotherapy*, 2018. 102: p. 452–457.
27. Li Z, et al. SNHG6 Promotes Tumor Growth via Repression of P21 in Colorectal Cancer. *Cell Physiol Biochem*. 2018;49(2):463–78.
28. Wang AH, et al., *MIR210HG promotes cell proliferation and invasion by regulating miR-503-5p/TRAF4 axis in cervical cancer*. *Aging (Albany NY)*, 2020. 12(4).
29. Wang Y, et al. MIR210HG predicts poor prognosis and functions as an oncogenic lncRNA in hepatocellular carcinoma. *Biomed Pharmacother*. 2019;111:1297–301.
30. Li J, et al. Long Noncoding RNA miR210HG Sponges miR-503 to Facilitate Osteosarcoma Cell Invasion and Metastasis. *DNA Cell Biol*. 2017;36(12):1117–25.
31. Liu B, et al. A Cytoplasmic NF-κB Interacting Long Noncoding RNA Blocks IκB Phosphorylation and Suppresses Breast Cancer Metastasis. *Cancer Cell*. 2015;27(3):370–81.
32. Patel M, et al. NF-κB pathways in the development and progression of colorectal cancer. *Translational Research*. 2018;197:43–56.
33. Masliah-Planchon J, Garinet S, Pasman E. *RAS-MAPK pathway epigenetic activation in cancer: miRNAs in action*. *Oncotarget*, 2016. 7(25).

Figures

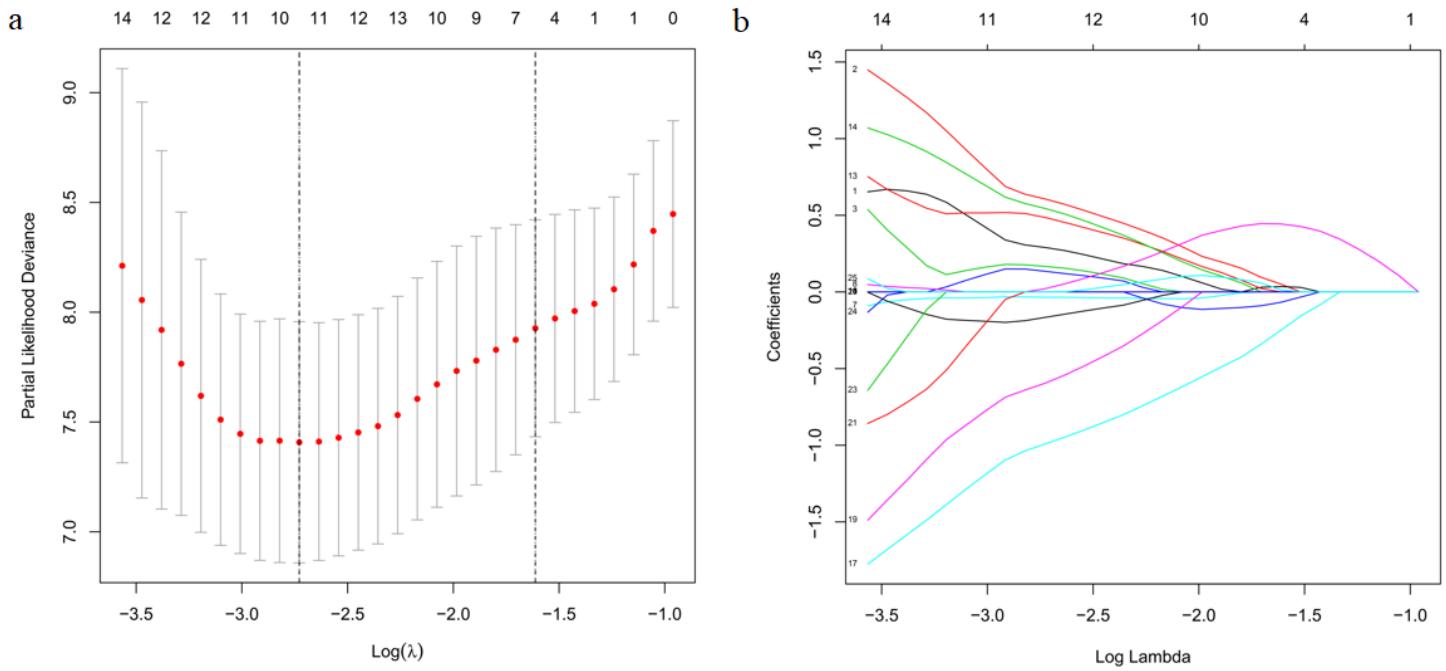


Figure 1

LASSO Cox regression analysis. (a) Partial likelihood deviance was plotted versus $\log(\lambda)$. The vertical dotted line indicates the λ value with the minimum error and the largest λ value in which deviance was within one SE of the minimum. (b) LASSO coefficient profiles of the genes associated with survival of OS. LASSO, least absolute shrinkage and selection operator

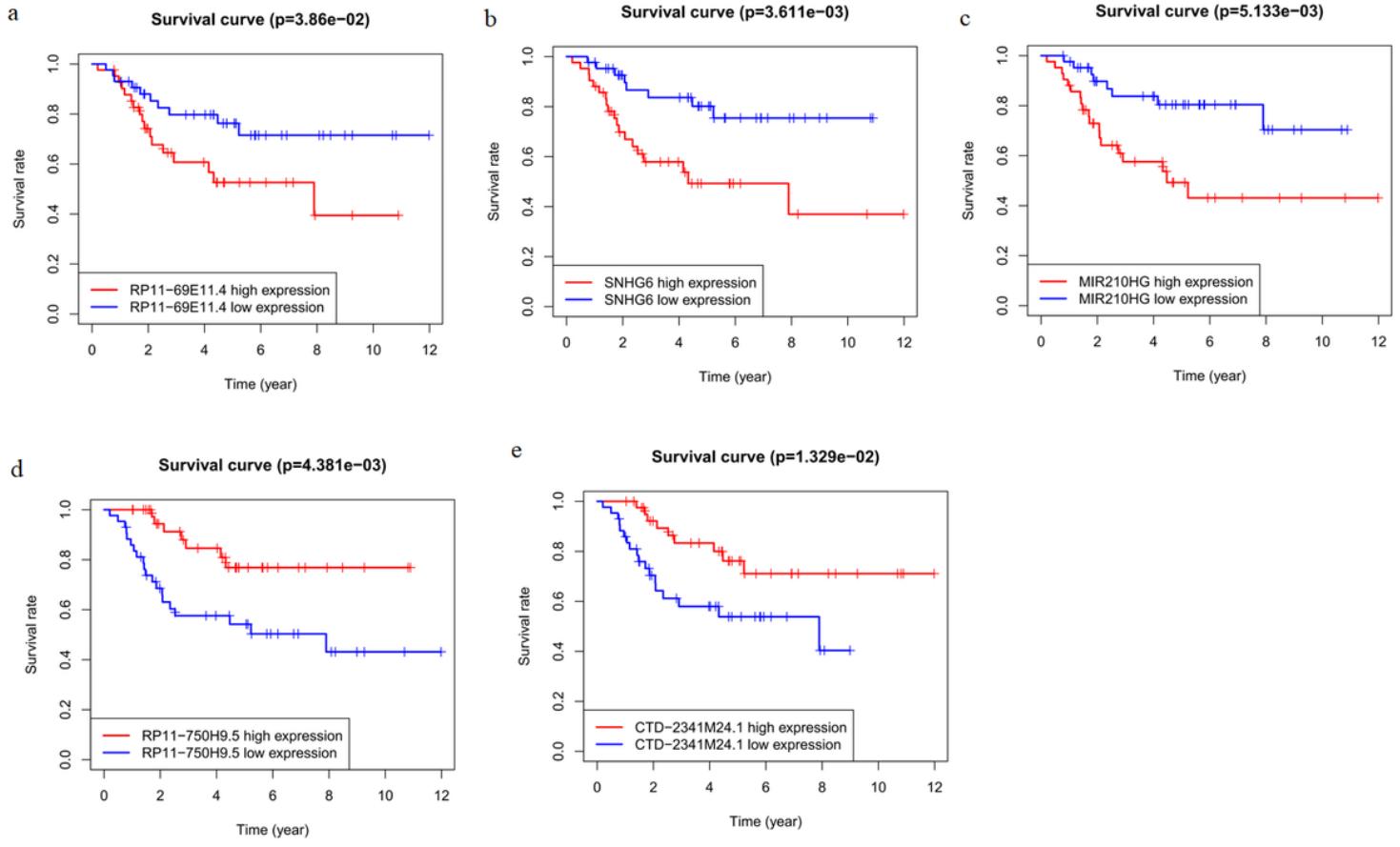


Figure 2

(a) Kaplan-Meier survival curve for high- and low-expression groups of RP11-69E11.4, and (b) SNHG6, (c) MIR210HG, (d) RP11-750H9.5, (e) CTD-2341M24.1

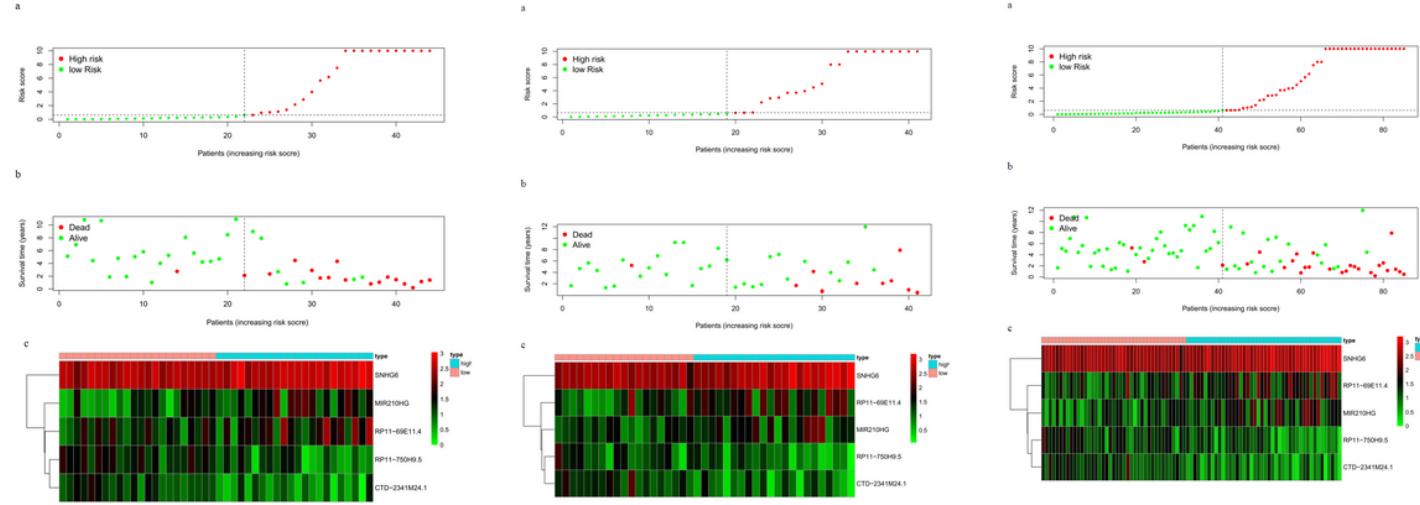


Figure 3

1 Risk scores and survival status of patients in high- and low-risk groups in the training set of the TARGET dataset. (a) Risk scores of patients in the high (red)- and low (green)-risk groups. (b) Survival

status of patients in these two groups. Green and red represent alive and dead patients, respectively. (c) Expression of SNHG6, MIR210HG, RP11-69E11.4, RP11-750H9.5 and CTD-2341M24.1 in each patient. 2 Risk scores and survival status of patients in high- and low-risk groups in the validation set of the TARGET dataset. (a) Risk scores of patients in the high (red)- and low (green)-risk groups. (b) Survival status of patients in these two groups. Green and red represent alive and dead patients, respectively. (c) Expression of SNHG6, RP11-69E11.4, MIR210HG, RP11-750H9.5 and CTD-2341M24.1 in each patient. 3 Risk scores and survival status of patients in high- and low-risk groups in the overall data set of the TARGET. (a) Risk scores of patients in the high (red)- and low (green)-risk groups. (b) Survival status of patients in these two groups. Green and red represent alive and dead patients, respectively. (c) Expression of SNHG6, RP11-69E11.4, MIR210HG, RP11-750H9.5 and CTD-2341M24.1 in each patient.

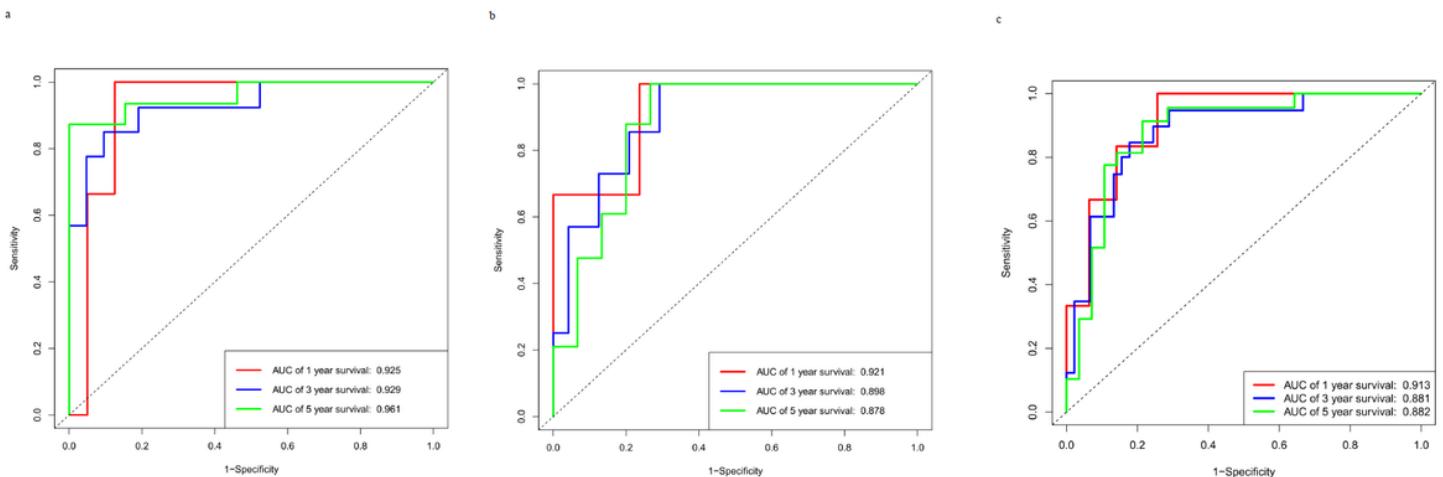


Figure 4

(a) Time-dependent ROC curves evaluating the prediction efficiency of the risk score in the training set of the TARGET database. (b) Time-dependent ROC curves evaluating the prediction efficiency of the risk score in the validation set of the TARGET database. (c) Time-dependent ROC curves evaluating the prediction efficiency of the risk score in the overall data set of the TARGET database. AUC, Area under the ROC curve; ROC, receiver operating characteristic

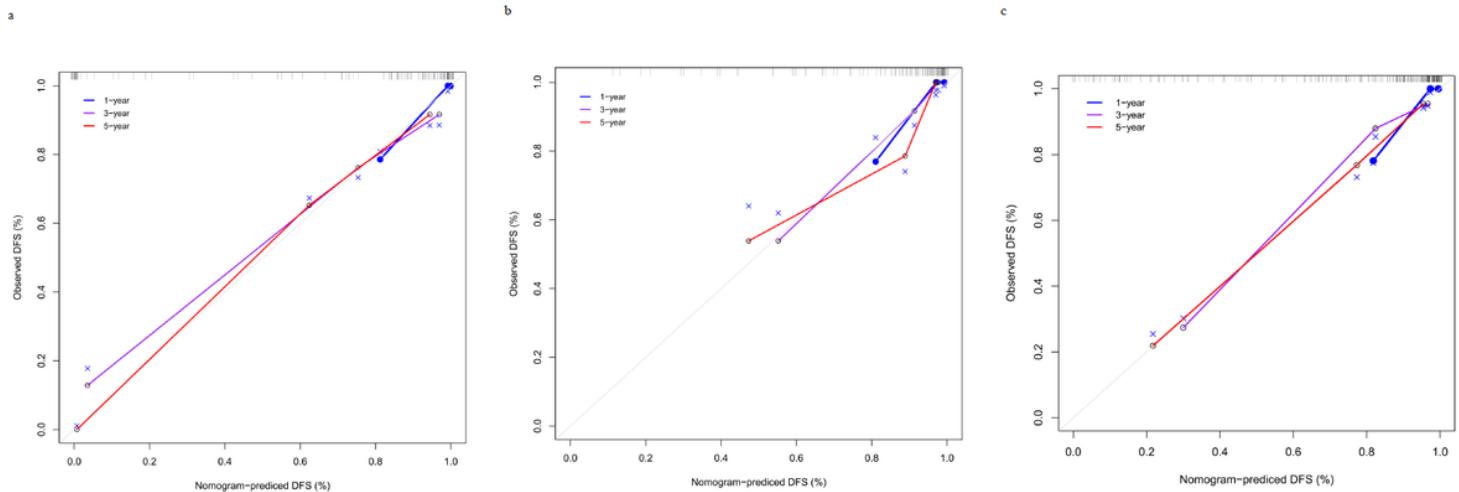


Figure 5

(a) Calibration curve of the nomogram in the training set of the TARGET database. (b) Calibration curve of the nomogram in the validation set of the TARGET database. (c) Calibration curve of the nomogram in the overall data set of the TARGET database.

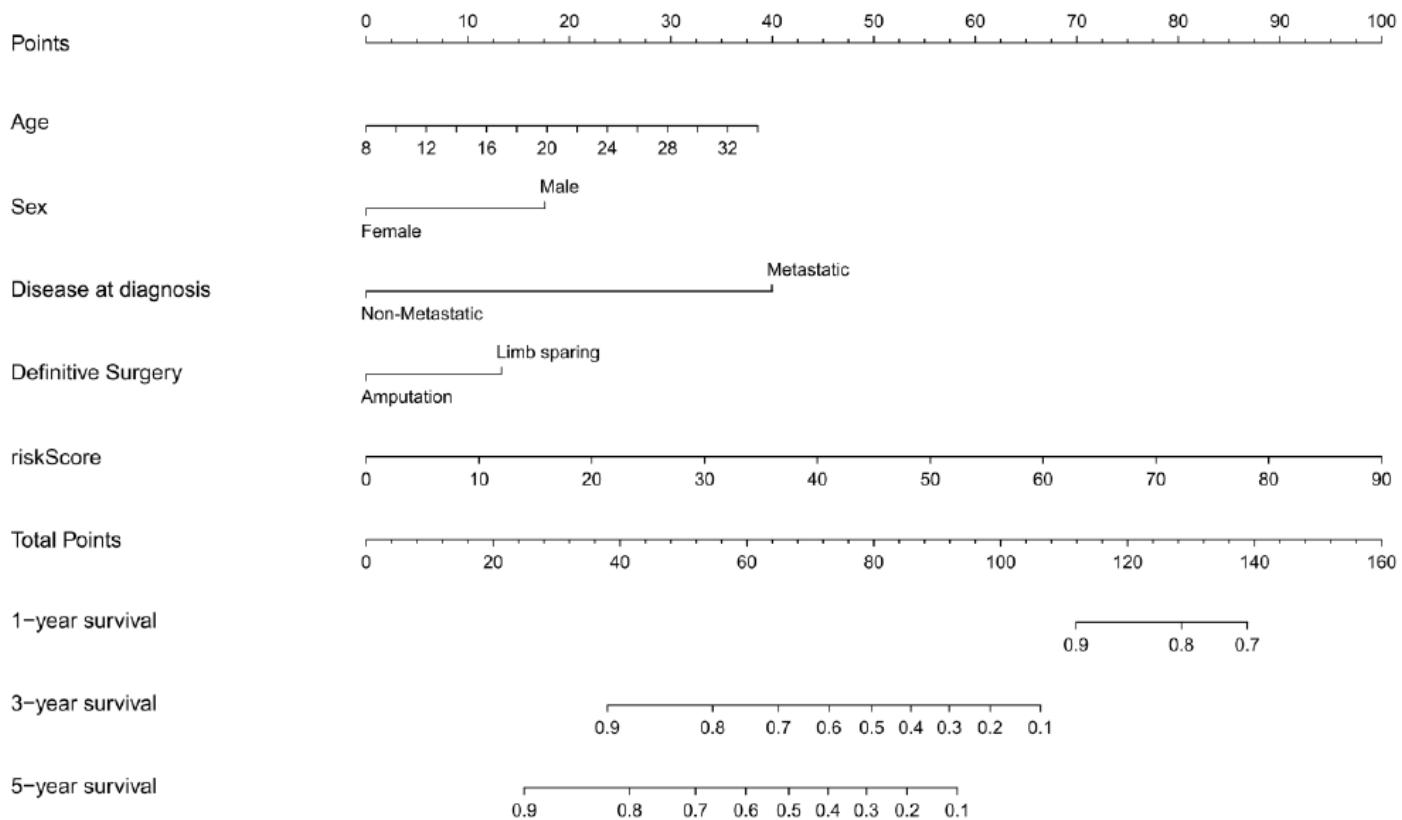


Figure 6

Nomogram to predict survival of patients with OS. OS, osteosarcoma

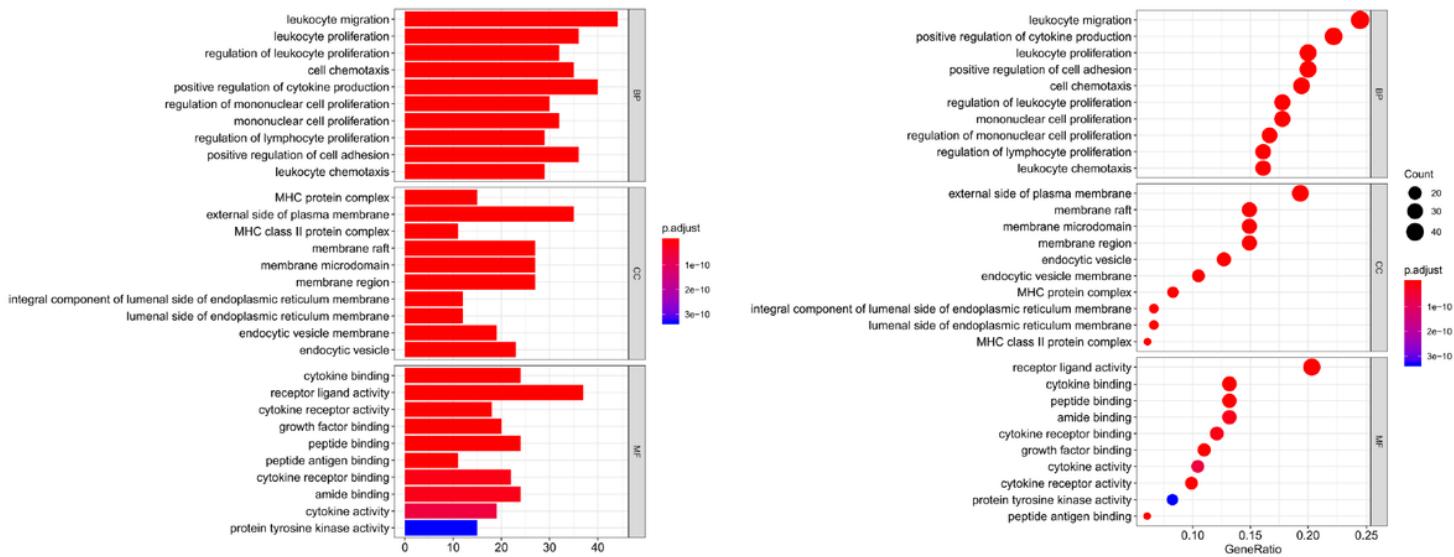


Figure 7

Functional enrichment analysis depicted the biological pathways and processes associated with correlated lncRNAs. The results of GO biological process enrichment. GO, gene ontology

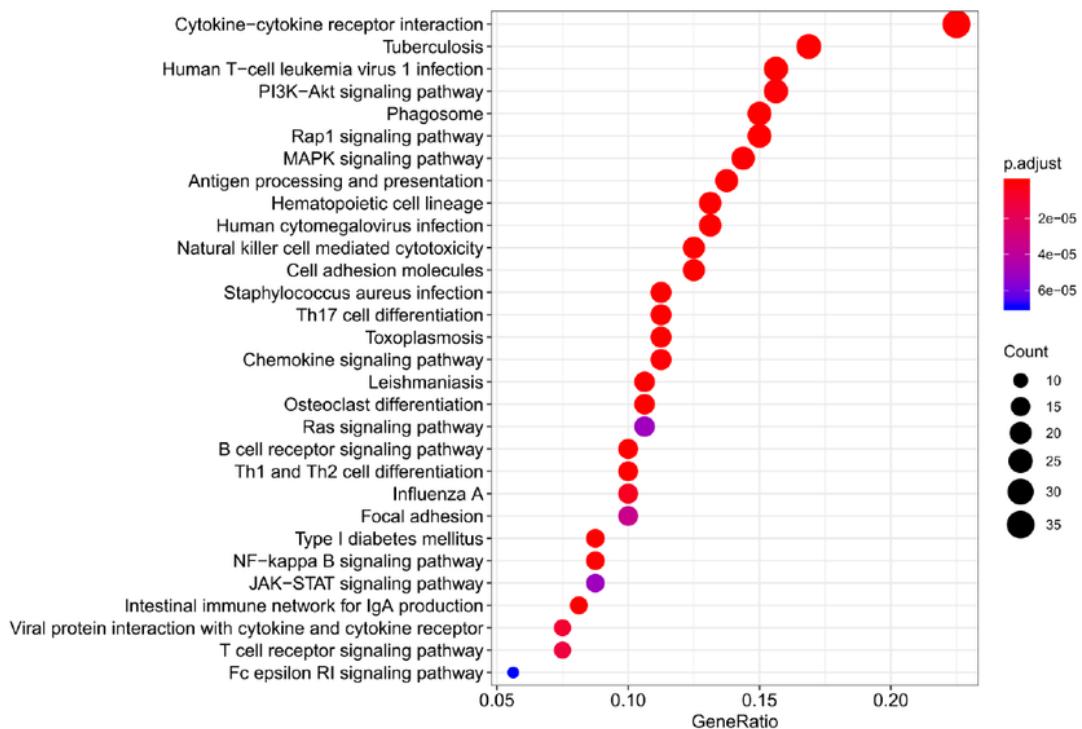
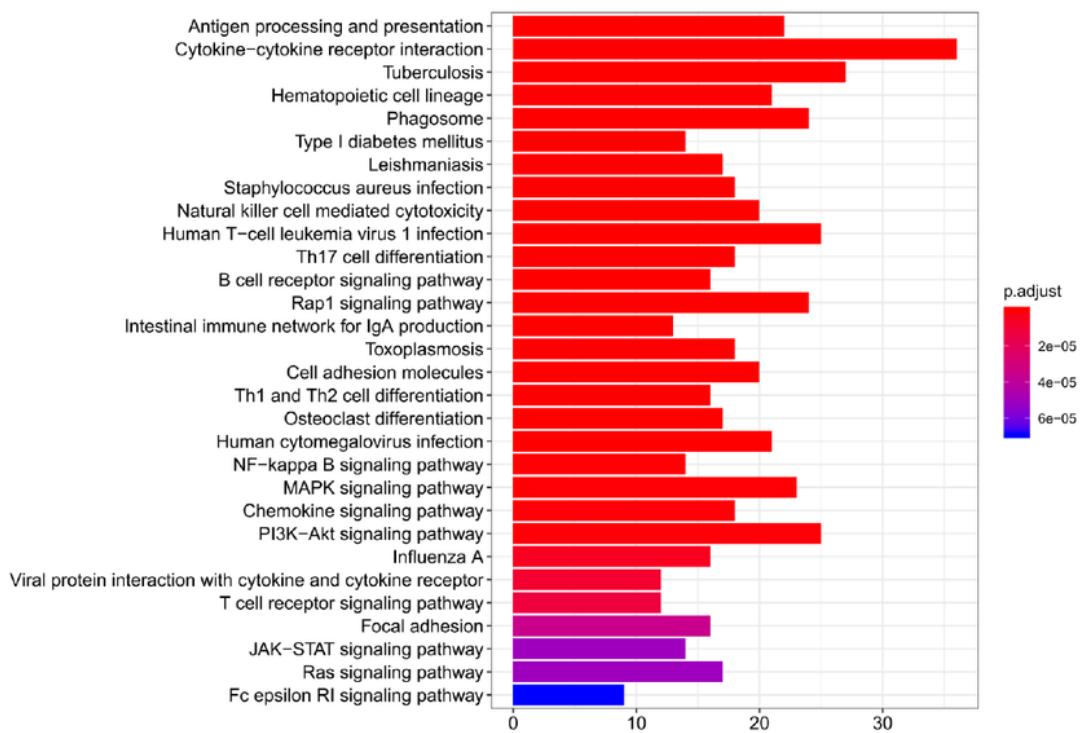


Figure 8

Functional enrichment analysis depicted the biological pathways and processes associated with correlated lncRNAs. The results of KEGG signaling pathways analysis. KEGG, Kyoto encyclopedia of genes and genomes

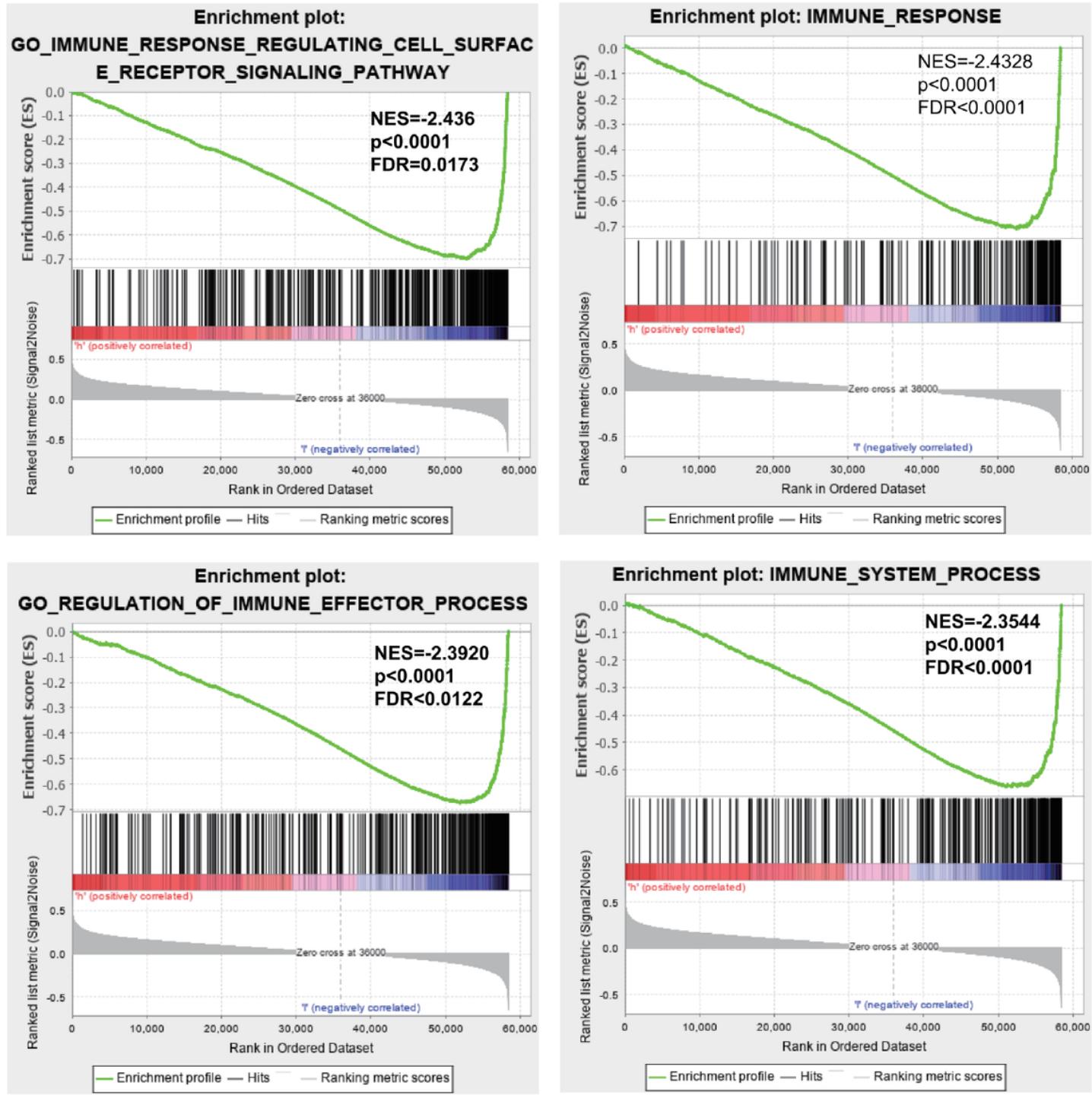


Figure 9

GSEA for risk signature-based on the TARGET database. GSEA, gene set enrichment analysis