

Identification of a Programmed Cell Death-related Signature for Prediction of Osteosarcoma

Ankai Xu (✉ 21718246@zju.edu.cn)

Zhejiang University School of Medicine Second Affiliated Hospital <https://orcid.org/0000-0003-4971-6168>

Jinti Lin

Zhejiang University

Wei Yu

Zhejiang University

Jiakang Jin

Zhejiang University

Bing Liu

Zhejiang University

Huimin Tao

Zhejiang University

Primary research

Keywords: osteosarcoma, programmed cell death, MUC1, TCF7L2, TGFB2, TRIAP1, CBS

Posted Date: March 24th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-326918/v1>

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

[Read Full License](#)

Abstract

Background: This study aims to perform bioinformatics analysis of programmed cell death-related genes (PCDGs) in osteosarcoma, and to construct a multi-gene signature for predicting the prognosis.

Methods: The functional enrichment analysis was applied for prognostic PCDGs, and PPI network as well as drug-gene interactions were constructed. In order to set up the prognosis evaluation system, we established a prognosis model by integrating PCDGs. The Least Absolute Shrinkage and Selection Operator (LASSO) logistic regression analysis as well as the multivariate Cox proportional hazard regression analysis were used to construct the five-genes signature (MUC1, TCF7L2, TGFB2, TRIAP1, CBS) for prediction in Gene Expression Omnibus (GEO) cohort. According to the median risk score, survival analysis was conducted to evaluate the prognostic value of the risk score in GEO cohort. Next, by combining other clinic-pathological independent prognostic factor, stage at diagnosis, a nomogram was established to predict individual survival probability.

Result: GO analysis showed that the 15 prognostic PCDGs were mainly enriched in apoptotic signaling pathway, regulation of secretion and p53 signaling pathway. KEGG analysis showed that aforesaid genes were mainly related to PI3K-Akt signaling pathway, diverse neoplasms signaling pathway and hepatitis B. Drug-gene interactions displayed available drugs for influencing osteosarcoma via programmed cell death, such as adalimumab, tacrolimus and sirolimus. The risk score was constructed based on 5 genes and patients were significantly divided into high-risk and low-risk groups according to overall survival. In multivariate Cox regression analysis, risk score was still an independent prognostic factor (HR = 2.526, 95% CI = 1.597–3.994, $P < 0.001$). Cumulative curve showed that low-risk score patients were significantly had better prognosis than that of patients with high-risk score ($P < 0.001$). The external independent TARGET cohort proved the validity of risk score model and the nomograph.

Conclusion: From the facet of programmed cell death, we provided a multi-gene signature

and the nomograph for the prognostic predictors of osteosarcoma patients, and available drugs displayed may provide promising treatment strategies.

Background

Osteosarcoma, a type of malignant tumor, whose survival rate has not been obviously improved for decades. As previous researches exhibited, osteosarcoma often showed resistance to the methods of immunity, such as inhibitor for PD-1/PD-L1 [1, 2]. Thus, brand-new and efficient therapeutics are urgently needed. In the past few decades, people have learned increasingly about the death of cell [3]. And programmed cell death (PCD) as one of the breakthroughs has huge potential in application to the treatments for cancers.

PCD, also named regulated cell death, is a biologically controlled process. There are many kinds of PCD, including apoptosis, ferroptosis, autophagy-dependent cell death, pyroptosis, entotic cell death and so on

[3]. Among them, apoptosis, ferroptosis and autophagy have the most intensively investigated and these processes are suitable for bioinformatics analysis. Apoptosis, widely participating in early development and tumor progression, is different from necrosis in many facets, such as morphology, availability of caspases and biochemical features [4]. Although Ford et al. discovered that apoptotic tumor cells may stimulate the proliferation of tumors, triggering apoptosis is still one of the important approaches for killing cancer cells [5]. Ferroptosis, is a special form of PCD, dependent on iron and excessive peroxidation of polyunsaturated fatty acids. Additionally, targeting the sensibility of ferroptosis in tumors, many kinds of drugs like altretamine, sorafenib, and silica nanoparticles had already shown the therapeutic potential [6]. Autophagy, widely exists in eukaryotic cells, disintegrating cellular organelles and macromolecules, may enhance the tolerance of tumor to treatments [7, 8].

Increasingly, compelling evidence has shown the significance of PCD in the therapeutics for osteosarcoma. Via inducing cell apoptosis, many drugs play roles in antineoplastic activity in osteosarcoma, such as apatinib, As2S₂, artocarpin and metformin [9–12]. Further, Liansheng and co-workers observed that, inhibited *TGFB111* which is up-regulated in exosome from osteosarcoma can promote apoptosis and inhibit proliferation [13]. As previous researches exhibited, autophagy process in osteosarcoma may have diverse effects owing to the stage of autophagy [14]. In general, inhibition of autophagy often has a negative impact on osteosarcoma [15–17]. Even though there are few researches about ferroptosis in osteosarcoma, researchers still discovered relational drugs like EF24 and PEITC [18, 19].

In the present research, we explored prognostic PCD-related genes (PCDGs) in osteosarcoma. Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) and protein-protein interactions (PPI) network map analyses were next carried out and drug-gene interactions was displayed. In order to construct a prognostic model of osteosarcoma integrating PCDGs, the least absolute shrinkage and selection operator regression (LASSO) method as well as multivariate Cox regression with stepwise method analyses were used to further screen the genes for constructing a risk score model [20]. Based on the median risk score, survival analysis was conducted to evaluate the prognostic value of the model. Then, the risk score and tumor metastasis stage at diagnosis were recognized independent prognostic factors. Further, a nomogram was established to predict individual survival probability. Finally, an external independent TARGET cohort proved the validity of the risk score model as well as the nomograph.

Materials And Methods

Data Sources and Data Preprocessing

We firstly paid attention to all the open samples from patients with overall survival (OS) rate for osteosarcoma. Total three datasets (GSE21257, GSE16091, GSE39055) in GEO (<https://www.ncbi.nlm.nih.gov/gds>) were collected to make GEO merged cohort, and the “ComBat” algorithm was applied to reduce the likelihood of batch effects from non-biological technical biases [21].

And then samples from TARGET (<https://ocg.cancer.gov/programs/target/data-matrix>) were also gathered, and FPKM values were transformed into transcripts per kilobase million (TPM) values and then $\log_2(\text{TPM} + 1)$ values, which are more comparable between samples and more similar to data from microarrays [22, 23]. All the gene matrices were standardized by quantiles normalization. Finally, 121 patients in GEO merged cohort and 95 patients in TARGET cohort were selected.

Identification and enrichment analysis of PCDGs

GO and KEGG pathway analyses were realized via “clusterProfiler” package, and $p.\text{adjust} < 0.05$ was believed to be statistically significant [24]. PPI network was acquired from the website portal ([25], <https://string-db.org/>). Drug-gene interactions were displayed via the drug-gene interaction database ([26], www.dgidb.org). Cytoscape software (version 3.8.2) was also used.

Construction of the prognostic gene signatures

First, respectively, Kaplan-Meier survival analysis and univariate Cox hazard analysis were applied on the entire GEO merged cohort to screen PCDGs significantly associated with OS rate. Then, the LASSO method was implemented to reduce dimensionality and using the R package “glmnet”. Next, the multivariate Cox regression analysis with stepwise method was performed to build the risk score model [27]. As weights, beta (β) coefficients of the prognostic genes in the multivariate logistic analysis were used to calculate the risk score.

$$\text{riskscore} = \sum_{i=1}^N (\text{Exp}_i \times \beta_i),$$

where EXP stands for the expression level of each gene. The specificity and sensitivity of the prognosis prediction based on the risk score model, were depicted by a receiver operating characteristic (ROC) curve and quantified by the area under the ROC curve (AUC) using “timeROC” package [28].

In order to learn whether PCDGs related risk index can be used as an independent predictor, univariate as well as multivariate Cox regression analyses were conducted. Risk score, gender, age, tumor site and metastasis stage at diagnosis were used as covariates in both GEO merged cohort and TARGET cohort.

The construction of nomogram

Independent predictors, metastasis stage at diagnosis and risk score were used to construct the nomogram together using the “rms” package in R. Then, calibration curves were drawn to assess the consistency between actual and predicted survival.

Results

Identification and enrichment analysis of PCDGs

The GEO merged cohort consisted of 121 patients with OS rate. The clinical characteristics of the patients in both GEO and TARGET cohort were collected. The flow chart of this study was illuminated in Supplement Fig. 1. We searched for PCDGs in the Harmonizome database [29]. A total of 836 PCDGs were selected (Supplement table 1). To further get prognostic PCDGs for osteosarcoma, Kaplan-Meier and univariate Cox hazard analyses were applied to GEO merged cohort ($p < 0.05$ was considered significant). Results showed 15 PCDGs were significantly correlated with OS rate (Fig. 1A). To further understand PCDGs in osteosarcoma, GO, KEGG and PPI network were applied to GEO merged cohort (Fig. 1B-1D). Via GO analysis, among PCDGs, prognostic genes for osteosarcoma mainly participate in apoptotic signaling pathway, regulation of secretion and p53 signaling pathway, suggesting correlative treatments for osteosarcoma may be advantageous. Via KEGG analysis, the aforesaid genes were of universal applicability and they were significantly correlated with other tumors like hepatocellular carcinoma, colorectal cancer, prostate cancer, gastric cancer and pancreatic cancer. Via KEGG analysis, the prognostic PCDGs also were mainly enriched in PI3K-Akt signaling pathway, PD-L1 expression and PD-1 checkpoint pathway, TGF-beta signaling pathway. Next, with drug-gene interactions, interrelated available drugs were exhibited (Fig. 2A). In facet of PCDGs of osteosarcoma, some drugs or chemical compounds may exert influence, such as adalimumab, aspirin, cysteine, fasudil, pravastatin and repaglinide. And some of them were already used to the cure of cancer or other diseases.

Construction of the prognostic gene signatures

15 prognostic PCDGs entered the LASSO regression analysis. Nine genes (CREB3L1, DNAJC10, LAMP1, MUC1, TCF7L2, TGFB2, TRIAP1, CBS, KEAP1) were identified and subsequently were exerted multivariate Cox regression analysis with the stepwise method (Fig. 2B-2C). Finally, 5 genes (MUC1, TCF7L2, TGFB2, TRIAP1 and CBS) related to the prognosis of osteosarcoma were obtained (Fig. 2D). The coefficients of each gene are shown in Table 1. The median of risk score was used to define the groups.

In GEO merged cohort, the distribution of risk scores in patients and the relationship between risk scores and survival time were displayed (Fig. 3A-3B). By heatmap, gene expression profiles in high-risk and low-risk were exhibited (Fig. 3C). The genes with $HR > 1$ (MUC1, TCF7L2, TRIAP1 and CBS) were considered dangerous, while the gene with $HR < 1$ (TGFB2) protective. Additionally, the high-risk group has the higher expression dangerous genes with the lower protective one. With the Kaplan-Meier cumulative curve in GEO merged cohort, the validity of the risk model was exhibited (Fig. 4A). By the time-dependent ROC curves, the AUC for 1-year, 2-year, 3-year, and 5-year OS were gathered (Fig. 4B).

The risk model based on PCDGs as an independent prognostic factor

In GEO merged cohort, both univariate and multivariate analysis showed that the risk score an independent prognostic indicator ($HR = 2.764$, $95\%CI = 1.750-4.364$, $P < 0.001$; $HR = 2.526$, $95\%CI = 1.597-3.994$, $P < 0.001$) (Fig. 4C-4D). Further, we validated this model in the external independent TARGET cohort with the same way (Fig. 5A-5B). The risk score was still proved an independent prognostic factor the same as tumor stage at diagnosis (Fig. 5C-5D).

The construction of nomogram

A nomogram is not only a highly legible illustration of a mathematical model, but also a convenient tool for predicting the outcome of individual patients in the clinical environment. The above analysis showed that the risk score as well as tumor stage at diagnosis were independent prognostic indicators. And in this study, the nomogram consisting of above significant risk factors was established, with C-index of 0.798 in GEO merged cohort and 0.700 in TARGET cohort (Fig. 6A). The calibration curves showed non-significant deviations between predicted and actual probability in both GEO merged cohort and TARGET cohort (Fig. 6B-6C).

Discussion

Programmed cell death is crucial for the normal growth and development of an organism [30]. As a special kind of cell, cancer cell, is also controlled by programmed cell death via its specific patterns. Additionally, the imbalance of programmed cell death in a cell may be also cancerogenic [30]. Among them, apoptosis, ferroptosis and autophagy have been the most extensively studied. Further, programmed cell death is relational to the cancer metastasis as well as prognosis [31–33].

In the previous studies, there are several gene expression-based signatures for the prognosis of osteosarcoma [34–40]. Moreover, gene signatures based on cell death have been reported in diverse cancers, such as lung cancer, hepatocellular carcinoma and gastric cancer [32, 33, 41]. Although an earlier study indicated the importance of apoptosis, little attention was paid to PCDGs in prognosis prediction of osteosarcoma [36]. In this study, we explored and analyzed the prognostic PCDGs in GEO merged osteosarcoma dataset. Via drug-gene interactions, relational available drugs were excavated. LASSO regression as well as multivariate Cox regression with stepwise method analyses were exerted to develop five prognostic markers for GEO merged cohort. In the GEO merged cohort, risk scores significantly stratified patient outcomes. Further, both univariate and multivariate analysis showed that the risk score an independent prognostic indicator. Subsequently, the nomogram consisting of significant risk factors was established. More importantly, in the external independent TARGET cohort, the prognostic efficacy of the five-gene signature as well as the nomogram were verified.

In previous researches, it has been shown that the five genes contained in the signature are all related to cancer. MUC1, a membrane tethered mucin glycoprotein, as a high-risk related gene in our study, was proved a kind of oncogene by most human carcinomas [42, 43]. In the facet of apoptosis, MUC1 often acted as an inhibitory of apoptosis in cancer cells while inducing apoptosis of activated human T cells in the study by Gimmi et al. [44–47]. As an inhibitor of apoptosis, MUC1 conferred the resistance of a variety of drugs as well as radiation treatment on cancer cells [48–51]. Further, many studies indicated the relationship between the high expression of MUC1 and poor prognosis with enhanced metastasis [52, 53]. A possible underlying mechanism by Zhao and his colleagues is that, MUC1 could prolong the survival of circulating tumor cells [54]. TCF7L2, transcription factor 7-like 2, is a critical member of the Wnt/beta-catenin canonical signaling pathway and a moderator of apoptosis [55, 56]. In a variety of

tumors, such as colorectal cancer, acute myeloid leukemia and clear cell renal cell carcinoma, this factor was proved related to the metastasis and poor prognosis [57–60]. Further, in beta-estradiol treatment process of osteosarcoma, TCF7L2 was observed reduce the expression with other members involved in Wnt/beta-catenin canonical signaling pathway [61]. TRIAP1 (TP53 regulated inhibitor of apoptosis 1) is a notable inhibitor of apoptosis controlled by TP53 and an available target by epirubicin in osteosarcoma [62]. CBS, cystathionine β -synthase, was associated with homocysteine metabolism, remethylation metabolism, the generation of H₂S, drug resistance, oxidative stress and ferroptosis [63–65]. In many neoplasms, CBS was proven differentially expressed to normal tissue and harmful to prognosis. As for osteosarcoma, HU Yang and his group revealed the difference of gene expression of CBS between tumor and normal samples, but did not explore the impact of CBS and its mechanism [66]. TGFB2 encodes a secreted ligand of the transforming growth factor-beta (TGF- β 2) superfamily of proteins. TGF- β 2 is capable of suppressing interleukin-2-dependent growth of T lymphocytes, and TGF- β 2 can be either a proto-oncogene or a tumor suppressor [67]. According to early studies, the function of TGF- β 2 was controversial and heterogeneous. On the one hand, TGF- β 2 promoted the progress of malignant tumors, such as glioblastoma and breast cancer [68, 69]. On the other hand, TGF- β 2 may play a tumor inhibition in pancreatic ductal adenocarcinoma as well as squamous cell carcinomas [70, 71]. Further, TGF- β 2 introduced the dormancy of malignant disseminated tumor cells in the bone marrow and TGF- β 2 introduced the restriction of metastasis in lungs [72–74].

For osteosarcoma population, our 5-gene risk score model based on PCDGs was shown to be a reliable predictor of prognosis. However, the present study still has some limitations. Firstly, the race and district of patient were heterogeneous because of the retrospective multiple-platform study. Secondly, integrated and uniform clinical characteristic was incomplete for osteosarcoma cohort, needing preferable clinical follow-up and data collection.

Conclusion

In conclusion, our study established a novel 5-gene signature and nomogram to forecast the prognosis of osteosarcoma patients, which may contribute to the clinical decision-making of individual therapy.

Abbreviations

OS: overall survival; **GEO:** Gene Expression Omnibus; **TARGET:** Therapeutically Applicable Research to Generate Effective Treatments; **LASSO:** least absolute shrinkage and selection operator; **ROC:** receiver operating characteristic; **AUC:** the area under the receiver operating characteristic curve; **PCD:** programmed cell death; **PCDGs:** programmed cell death -related genes; **GO:** Gene Ontology; **KEGG:** Kyoto Encyclopedia of Genes and Genomes; **PPI:** protein-protein interactions; **TPM:** transcripts per kilobase million; **TGF- β 2:** transforming growth factor-beta.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

Publicly available datasets were analyzed in this study. The data are accessible at the following repositories: <https://www.ncbi.nlm.nih.gov/gds>, <https://ocg.cancer.gov/programs/target/data-matrix>.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by grant from the National Natural Science Foundation of China (NO.81872181, NO. 82002836 and NO. 81702662), Natural Science Foundation of Zhejiang Province (NO.LY20H160025 and LY18H160015) and the Science and Technology Planning Project of Zhejiang Province (NO. 2019KY079).

Author's Contributions

Ankai Xu and Jinti Lin contributed to the planning of the study. Ankai Xu and Jiakang Jin carried out the study and collected important background information. Ankai Xu and Wei Yu drafted the manuscript and revised the manuscript. Bing Liu and Huimin Tao are in charge of the whole research conduction and paper writing. All authors read and approved the final manuscript.

Acknowledgements

A tool by Welsh, E.A. et al in Excel was used for us to prevent gene symbols' conversion errors [75]. And via a convenient website portal, we generated Venn diagrams in our study (<http://bioinformatics.psb.ugent.be/webtools/Venn/>).

References

1. Duerinck J, Du Four S, Vandervorst F, D'Haene N, Le Mercier M, Michotte A, Van Binst AM, Everaert H, Salmon I, Bouttens F *et al*: **Randomized phase II study of axitinib versus physicians best alternative choice of therapy in patients with recurrent glioblastoma**. *J Neurooncol* 2016, **128**(1):147-155.
2. Tawbi HA, Burgess M, Bolejack V, Van Tine BA, Schuetze SM, Hu J, D'Angelo S, Attia S, Riedel RF, Priebat DA *et al*: **Pembrolizumab in advanced soft-tissue sarcoma and bone sarcoma (SARC028): a**

- multicentre, two-cohort, single-arm, open-label, phase 2 trial. *LANCET ONCOL* 2017, **18**(11):1493-1501.
3. Tang D, Kang R, Berghe TV, Vandenabeele P, Kroemer G: **The molecular machinery of regulated cell death.** *CELL RES* 2019, **29**(5):347-364.
 4. Elmore S: **Apoptosis: a review of programmed cell death.** *TOXICOL PATHOL* 2007, **35**(4):495-516.
 5. Ford CA, Petrova S, Pound JD, Voss JJ, Melville L, Paterson M, Farnworth SL, Gallimore AM, Cuff S, Wheadon H *et al*: **Oncogenic properties of apoptotic tumor cells in aggressive B cell lymphoma.** *CURR BIOL* 2015, **25**(5):577-588.
 6. Hassannia B, Vandenabeele P, Vanden BT: **Targeting Ferroptosis to Iron Out Cancer.** *CANCER CELL* 2019, **35**(6):830-849.
 7. Yang S, Wang X, Contino G, Liesa M, Sahin E, Ying H, Bause A, Li Y, Stommel JM, Dell'Antonio G *et al*: **Pancreatic cancers require autophagy for tumor growth.** *Genes Dev* 2011, **25**(7):717-729.
 8. Qiu J, Sun M, Wang Y, Chen B: **Identification and validation of an individualized autophagy-clinical prognostic index in gastric cancer patients.** *CANCER CELL INT* 2020, **20**:178.
 9. Liu K, Ren T, Huang Y, Sun K, Bao X, Wang S, Zheng B, Guo W: **Apatinib promotes autophagy and apoptosis through VEGFR2/STAT3/BCL-2 signaling in osteosarcoma.** *CELL DEATH DIS* 2017, **8**(8):e3015.
 10. Wang G, Zhang T, Sun W, Wang H, Yin F, Wang Z, Zuo D, Sun M, Zhou Z, Lin B *et al*: **Arsenic sulfide induces apoptosis and autophagy through the activation of ROS/JNK and suppression of Akt/mTOR signaling pathways in osteosarcoma.** *Free Radic Biol Med* 2017, **106**:24-37.
 11. Lee CW, Chi MC, Chang TM, Liu JF: **Artocarpin induces cell apoptosis in human osteosarcoma cells through endoplasmic reticulum stress and reactive oxygen species.** *J CELL PHYSIOL* 2019, **234**(8):13157-13168.
 12. Li B, Zhou P, Xu K, Chen T, Jiao J, Wei H, Yang X, Xu W, Wan W, Xiao J: **Metformin induces cell cycle arrest, apoptosis and autophagy through ROS/JNK signaling pathway in human osteosarcoma.** *INT J BIOL SCI* 2020, **16**(1):74-84.
 13. Sha L, Ma D, Chen C: **Exosome-mediated Hic-5 regulates proliferation and apoptosis of osteosarcoma via Wnt/beta-catenin signal pathway.** *Aging (Albany NY)* 2020, **12**(23):23598-23608.
 14. He G, Ma Y, Zhu Y, Yong L, Liu X, Wang P, Liang C, Yang C, Zhao Z, Hai B *et al*: **Cross Talk Between Autophagy and Apoptosis Contributes to ZnO Nanoparticle-Induced Human Osteosarcoma Cell Death.** *ADV HEALTHC MATER* 2018, **7**(17):e1800332.
 15. Yu W, Wang Y, Zhu J, Jin L, Liu B, Xia K, Wang J, Gao J, Liang C, Tao H: **Autophagy inhibitor enhance ZnPc/BSA nanoparticle induced photodynamic therapy by suppressing PD-L1 expression in osteosarcoma immunotherapy.** *BIOMATERIALS* 2019, **192**:128-139.
 16. Akin D, Wang SK, Habibzadegah-Tari P, Law B, Ostrov D, Li M, Yin XM, Kim JS, Horenstein N, Dunn WJ: **A novel ATG4B antagonist inhibits autophagy and has a negative impact on osteosarcoma tumors.** *AUTOPHAGY* 2014, **10**(11):2021-2035.

17. Ren T, Zheng B, Huang Y, Wang S, Bao X, Liu K, Guo W: **Osteosarcoma cell intrinsic PD-L2 signals promote invasion and metastasis via the RhoA-ROCK-LIMK2 and autophagy pathways.** *CELL DEATH DIS* 2019, **10**(4):261.
18. Lv HH, Zhen CX, Liu JY, Shang P: **PEITC triggers multiple forms of cell death by GSH-iron-ROS regulation in K7M2 murine osteosarcoma cells.** *ACTA PHARMACOL SIN* 2020, **41**(8):1119-1132.
19. Lin H, Chen X, Zhang C, Yang T, Deng Z, Song Y, Huang L, Li F, Li Q, Lin S *et al*: **EF24 induces ferroptosis in osteosarcoma cells through HMOX1.** *BIOMED PHARMACOTHER* 2021, **136**:111202.
20. Tibshirani R: **The lasso method for variable selection in the Cox model.** *STAT MED* 1997, **16**(4):385-395.
21. Johnson WE, Li C, Rabinovic A: **Adjusting batch effects in microarray expression data using empirical Bayes methods.** *BIostatistics* 2007, **8**(1):118-127.
22. Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, Bin J, Liao Y, Rao J, Liao W: **Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures.** *CANCER IMMUNOL RES* 2019, **7**(5):737-750.
23. Wagner GP, Kin K, Lynch VJ: **Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples.** *Theory Biosci* 2012, **131**(4):281-285.
24. Yu G, Wang LG, Han Y, He QY: **clusterProfiler: an R package for comparing biological themes among gene clusters.** *OMICS* 2012, **16**(5):284-287.
25. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P *et al*: **The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible.** *NUCLEIC ACIDS RES* 2017, **45**(D1):D362-D368.
26. Cotto KC, Wagner AH, Feng YY, Kiwala S, Coffman AC, Spies G, Wollam A, Spies NC, Griffith OL, Griffith M: **DGIdb 3.0: a redesign and expansion of the drug-gene interaction database.** *NUCLEIC ACIDS RES* 2018, **46**(D1):D1068-D1073.
27. Zhang Z: **Variable selection with stepwise and best subset approaches.** *Ann Transl Med* 2016, **4**(7):136.
28. Blanche P, Dartigues JF, Jacqmin-Gadda H: **Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks.** *STAT MED* 2013, **32**(30):5381-5397.
29. Rouillard AD, Gunderson GW, Fernandez NF, Wang Z, Monteiro CD, McDermott MG, Ma'Ayan A: **The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins.** *Database (Oxford)* 2016, **2016**.
30. Degterev A, Yuan J: **Expansion and evolution of cell death programmes.** *Nat Rev Mol Cell Biol* 2008, **9**(5):378-390.
31. Su Z, Yang Z, Xu Y, Chen Y, Yu Q: **Apoptosis, autophagy, necroptosis, and cancer metastasis.** *MOL CANCER* 2015, **14**:48.

32. Liang JY, Wang DS, Lin HC, Chen XX, Yang H, Zheng Y, Li YH: **A Novel Ferroptosis-related Gene Signature for Overall Survival Prediction in Patients with Hepatocellular Carcinoma.** *INT J BIOL SCI* 2020, **16**(13):2430-2441.
33. Ahluwalia P, Ahluwalia M, Mondal AK, Sahajpal N, Kota V, Rojiani MV, Rojiani AM, Kolhe R: **Immunogenomic Gene Signature of Cell-Death Associated Genes with Prognostic Implications in Lung Cancer.** *Cancers (Basel)* 2021, **13**(1).
34. Zhu N, Hou J, Ma G, Guo S, Zhao C, Chen B: **Co-expression network analysis identifies a gene signature as a predictive biomarker for energy metabolism in osteosarcoma.** *CANCER CELL INT* 2020, **20**:259.
35. Zhang H, Guo L, Zhang Z, Sun Y, Kang H, Song C, Liu H, Lei Z, Wang J, Mi B *et al.* **Co-Expression Network Analysis Identified Gene Signatures in Osteosarcoma as a Predictive Tool for Lung Metastasis and Survival.** *J CANCER* 2019, **10**(16):3706-3716.
36. Dong S, Huo H, Mao Y, Li X, Dong L: **A risk score model for the prediction of osteosarcoma metastasis.** *FEBS OPEN BIO* 2019, **9**(3):519-526.
37. Liu F, Xing L, Zhang X, Zhang X: **A Four-Pseudogene Classifier Identified by Machine Learning Serves as a Novel Prognostic Marker for Survival of Osteosarcoma.** *Genes (Basel)* 2019, **10**(6).
38. Luo X, Tang J, Xuan H, Liu J, Li X: **Identification and Validation of a Potent Multi-miRNA Signature for Prediction of Prognosis of Osteosarcoma Patients.** *Med Sci Monit* 2020, **26**:e919272.
39. Wen C, Wang H, Wang H, Mo H, Zhong W, Tang J, Lu Y, Zhou W, Tan A, Liu Y *et al.* **A three-gene signature based on tumour microenvironment predicts overall survival of osteosarcoma in adolescents and young adults.** *Aging (Albany NY)* 2020, **13**(1):619-645.
40. Wu G, Zhang M: **A novel risk score model based on eight genes and a nomogram for predicting overall survival of patients with osteosarcoma.** *BMC CANCER* 2020, **20**(1):456.
41. Qiu J, Sun M, Wang Y, Chen B: **Identification and validation of an individualized autophagy-clinical prognostic index in gastric cancer patients.** *CANCER CELL INT* 2020, **20**:178.
42. Raina D, Ahmad R, Kumar S, Ren J, Yoshida K, Kharbanda S, Kufe D: **MUC1 oncoprotein blocks nuclear targeting of c-Abl in the apoptotic response to DNA damage.** *EMBO J* 2006, **25**(16):3774-3783.
43. Yin L, Kharbanda S, Kufe D: **Mucin 1 oncoprotein blocks hypoxia-inducible factor 1alpha activation in a survival response to hypoxia.** *J BIOL CHEM* 2007, **282**(1):257-266.
44. Singh PK, Hollingsworth MA: **Cell surface-associated mucins in signal transduction.** *TRENDS CELL BIOL* 2006, **16**(9):467-476.
45. Kawano T, Ito M, Raina D, Wu Z, Rosenblatt J, Avigan D, Stone R, Kufe D: **MUC1 oncoprotein regulates Bcr-Abl stability and pathogenesis in chronic myelogenous leukemia cells.** *CANCER RES* 2007, **67**(24):11576-11584.
46. Agata N, Ahmad R, Kawano T, Raina D, Kharbanda S, Kufe D: **MUC1 oncoprotein blocks death receptor-mediated apoptosis by inhibiting recruitment of caspase-8.** *CANCER RES* 2008, **68**(15):6136-6144.

47. Gimmi CD, Morrison BW, Mainprice BA, Gribben JG, Boussiotis VA, Freeman GJ, Park SY, Watanabe M, Gong J, Hayes DF *et al*: **Breast cancer-associated antigen, DF3/MUC1, induces apoptosis of activated human T cells.** *NAT MED* 1996, **2**(12):1367-1370.
48. Ren J, Agata N, Chen D, Li Y, Yu WH, Huang L, Raina D, Chen W, Kharbanda S, Kufe D: **Human MUC1 carcinoma-associated protein confers resistance to genotoxic anticancer agents.** *CANCER CELL* 2004, **5**(2):163-175.
49. Siragusa M, Zerilli M, Iovino F, Francipane MG, Lombardo Y, Ricci-Vitiani L, Di Gesu G, Todaro M, De Maria R, Stassi G: **MUC1 oncoprotein promotes refractoriness to chemotherapy in thyroid cancer cells.** *CANCER RES* 2007, **67**(11):5522-5530.
50. Reynolds IS, Fichtner M, McNamara DA, Kay EW, Prehn J, Burke JP: **Mucin glycoproteins block apoptosis; promote invasion, proliferation, and migration; and cause chemoresistance through diverse pathways in epithelial cancers.** *Cancer Metastasis Rev* 2019, **38**(1-2):237-257.
51. Zhang J, Wang L, Jiang J, Qiao Z: **Elevation of microRNA-512-5p inhibits MUC1 to reduce radioresistance in cervical cancer.** *CELL CYCLE* 2020, **19**(6):652-665.
52. Zhao P, Meng M, Xu B, Dong A, Ni G, Lu L: **Decreased expression of MUC1 induces apoptosis and inhibits migration in pancreatic cancer PANC-1 cells via regulation of Slug pathway.** *CANCER BIOMARK* 2017, **20**(4):469-476.
53. Hattrup CL, Gendler SJ: **MUC1 alters oncogenic events and transcription in human breast cancer cells.** *BREAST CANCER RES* 2006, **8**(4):R37.
54. Zhao Q, Barclay M, Hilkens J, Guo X, Barrow H, Rhodes JM, Yu LG: **Interaction between circulating galectin-3 and cancer-associated MUC1 enhances tumour cell homotypic aggregation and prevents anoikis.** *MOL CANCER* 2010, **9**:154.
55. Agalliu I, Suuriniemi M, Prokunina-Olsson L, Johanneson B, Collins FS, Stanford JL, Ostrander EA: **Evaluation of a variant in the transcription factor 7-like 2 (TCF7L2) gene and prostate cancer risk in a population-based study.** *PROSTATE* 2008, **68**(7):740-747.
56. Pooja T, Karunagaran D: **Emodin suppresses Wnt signaling in human colorectal cancer cells SW480 and SW620.** *EUR J PHARMACOL* 2014, **742**:55-64.
57. Chen HJ, Wei Z, Sun J, Bhattacharya A, Savage DJ, Serda R, Mackeyev Y, Curley SA, Bu P, Wang L *et al*: **A recellularized human colon model identifies cancer driver genes.** *NAT BIOTECHNOL* 2016, **34**(8):845-851.
58. Zhang B, Jia WH, Matsuda K, Kweon SS, Matsuo K, Xiang YB, Shin A, Jee SH, Kim DH, Cai Q *et al*: **Large-scale genetic study in East Asians identifies six new loci associated with colorectal cancer risk.** *NAT GENET* 2014, **46**(6):533-542.
59. Saenz DT, Fiskus W, Mill CP, Perera D, Manshoury T, Lara BH, Karkhanis V, Sharma S, Horrigan SK, Bose P *et al*: **Mechanistic basis and efficacy of targeting the beta-catenin-TCF7L2-JMJD6-c-Myc axis to overcome resistance to BET inhibitors.** *BLOOD* 2020, **135**(15):1255-1269.
60. Kojima T, Shimazui T, Horie R, Hinotsu S, Oikawa T, Kawai K, Suzuki H, Meno K, Akaza H, Uchida K: **FOXO1 and TCF7L2 genes involved in metastasis and poor prognosis in clear cell renal cell**

- carcinoma.** *Genes Chromosomes Cancer* 2010, **49**(4):379-389.
61. Chen B, Liu Z, Zhang J, Wang H, Yu B: **RNA sequencing identifies gene expression profile changes associated with beta-estradiol treatment in U2OS osteosarcoma cells.** *Onco Targets Ther* 2017, **10**:3421-3427.
62. Yu L, Meng M, Bao Y, Zhang C, Gao B, Sa R, Luo W: **miR-1301/TRIAP1 Axis Participates in Epirubicin-Mediated Anti-Proliferation and Pro-Apoptosis in Osteosarcoma.** *YONSEI MED J* 2019, **60**(9):832-841.
63. Yamamoto T, Takano N, Ishiwata K, Ohmura M, Nagahata Y, Matsuura T, Kamata A, Sakamoto K, Nakanishi T, Kubo A *et al*: **Reduced methylation of PFKFB3 in cancer cells shunts glucose towards the pentose phosphate pathway.** *NAT COMMUN* 2014, **5**:3480.
64. Wang M, Mao C, Ouyang L, Liu Y, Lai W, Liu N, Shi Y, Chen L, Xiao D, Yu F *et al*: **Long noncoding RNA LINC00336 inhibits ferroptosis in lung cancer by functioning as a competing endogenous RNA.** *CELL DEATH DIFFER* 2019, **26**(11):2329-2343.
65. Phillips CM, Zatarain JR, Nicholls ME, Porter C, Widen SG, Thanki K, Johnson P, Jawad MU, Moyer MP, Randall JW *et al*: **Upregulation of Cystathionine-beta-Synthase in Colonic Epithelia Reprograms Metabolism and Promotes Carcinogenesis.** *CANCER RES* 2017, **77**(21):5741-5754.
66. HU Yang QWLK: **Expression of endogenous hydrogen sulfide and its synthases in human osteosarcoma cells.** *Orthopaedics* 2016, **7**(02):116-119.
67. Kubickova L, Sedlarikova L, Hajek R, Sevcikova S: **TGF-beta - an excellent servant but a bad master.** *J TRANSL MED* 2012, **10**:183.
68. Rodon L, Gonzalez-Junca A, Inda MM, Sala-Hojman A, Martinez-Saez E, Seoane J: **Active CREB1 promotes a malignant TGFbeta2 autocrine loop in glioblastoma.** *CANCER DISCOV* 2014, **4**(10):1230-1241.
69. Beisner J, Buck MB, Fritz P, Dippon J, Schwab M, Brauch H, Zugmaier G, Pfizenmaier K, Knabbe C: **A novel functional polymorphism in the transforming growth factor-beta2 gene promoter and tumor progression in breast cancer.** *CANCER RES* 2006, **66**(15):7554-7561.
70. Abraham CG, Ludwig MP, Andrysik Z, Pandey A, Joshi M, Galbraith MD, Sullivan KD, Espinosa JM: **DeltaNp63alpha Suppresses TGFB2 Expression and RHOA Activity to Drive Cell Proliferation in Squamous Cell Carcinomas.** *CELL REP* 2018, **24**(12):3224-3236.
71. Zhang S, Chung WC, Xu K: **Lunatic Fringe is a potent tumor suppressor in Kras-initiated pancreatic cancer.** *ONCOGENE* 2016, **35**(19):2485-2495.
72. Bragado P, Estrada Y, Parikh F, Krause S, Capobianco C, Farina HG, Schewe DM, Aguirre-Ghiso JA: **TGF-beta2 dictates disseminated tumour cell fate in target organs through TGF-beta-RIII and p38alpha/beta signalling.** *NAT CELL BIOL* 2013, **15**(11):1351-1361.
73. Derynck R, Turley SJ, Akhurst RJ: **TGFbeta biology in cancer progression and immunotherapy.** *NAT REV CLIN ONCOL* 2021, **18**(1):9-34.
74. Yumoto K, Eber MR, Wang J, Cackowski FC, Decker AM, Lee E, Nobre AR, Aguirre-Ghiso JA, Jung Y, Taichman RS: **Axl is required for TGF-beta2-induced dormancy of prostate cancer cells in the bone marrow.** *Sci Rep* 2016, **6**:36520.

Tables

Due to technical limitations, table 1 is only available as a download in the Supplemental Files section.

Figures

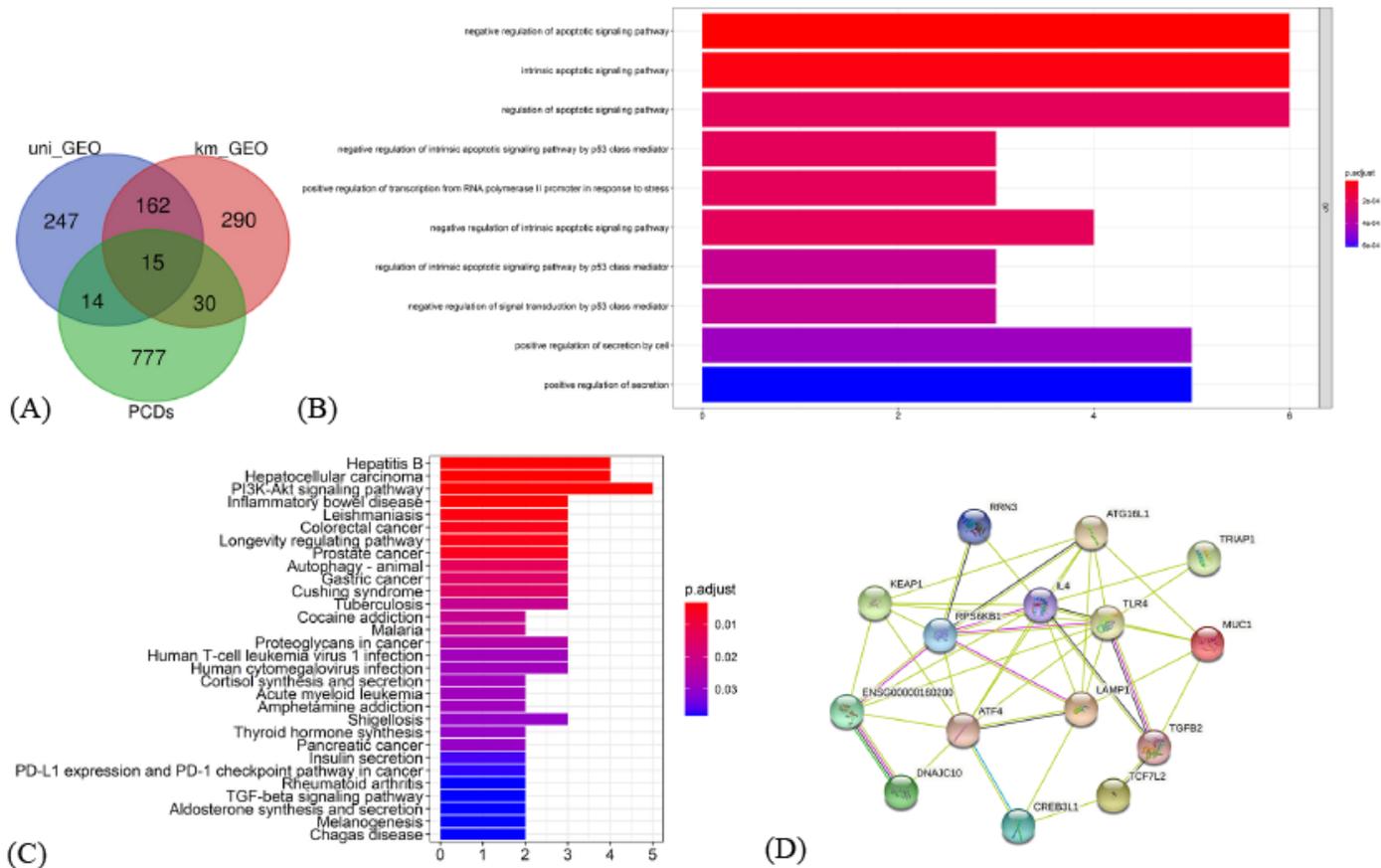


Figure 1

We searched for PCDGs in the Harmonizome database [29]. A total of 836 PCDGs were selected (Supplement table 1). To further get prognostic PCDGs for osteosarcoma, Kaplan-Meier and univariate Cox hazard analyses were applied to GEO merged cohort ($p < 0.05$ was considered significant). Results showed 15 PCDGs were significantly correlated with OS rate (Fig. 1A). To further understand PCDGs in osteosarcoma, GO, KEGG and PPI network were applied to GEO merged cohort (Fig. 1B-1D).

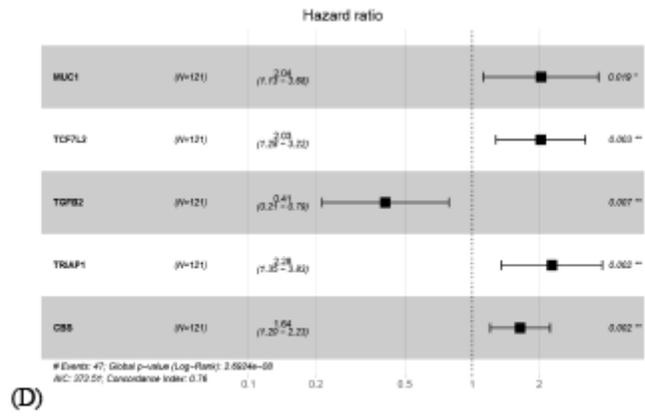
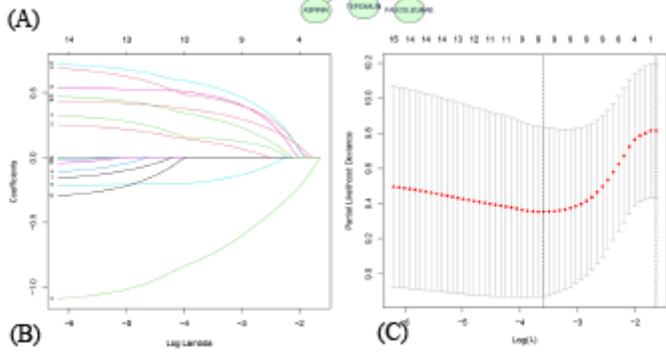
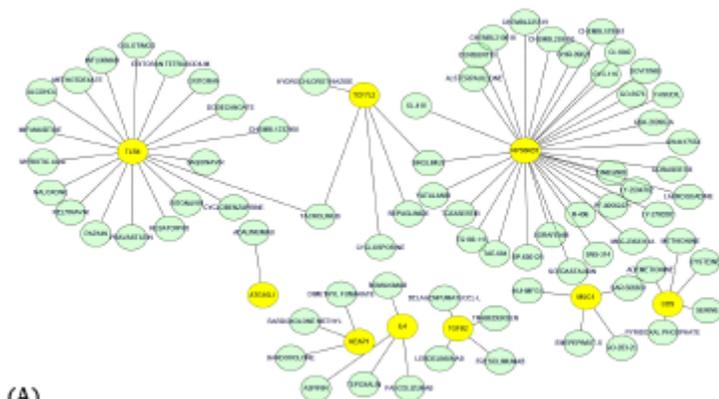


Figure 2

Next, with drug-gene interactions, interrelated available drugs were exhibited (Fig. 2A). 15 prognostic PCDGs entered the LASSO regression analysis. Nine genes (CREB3L1, DNAJC10, LAMP1, MUC1, TCF7L2, TGFB2, TRIAP1, CBS, KEAP1) were identified and subsequently were exerted multivariate Cox regression analysis with the stepwise method (Fig. 2B-2C). Finally, 5 genes (MUC1, TCF7L2, TGFB2, TRIAP1 and CBS) related to the prognosis of osteosarcoma were obtained (Fig. 2D). The coefficients of each gene are shown in Table 1. The median of risk score was used to define the groups.

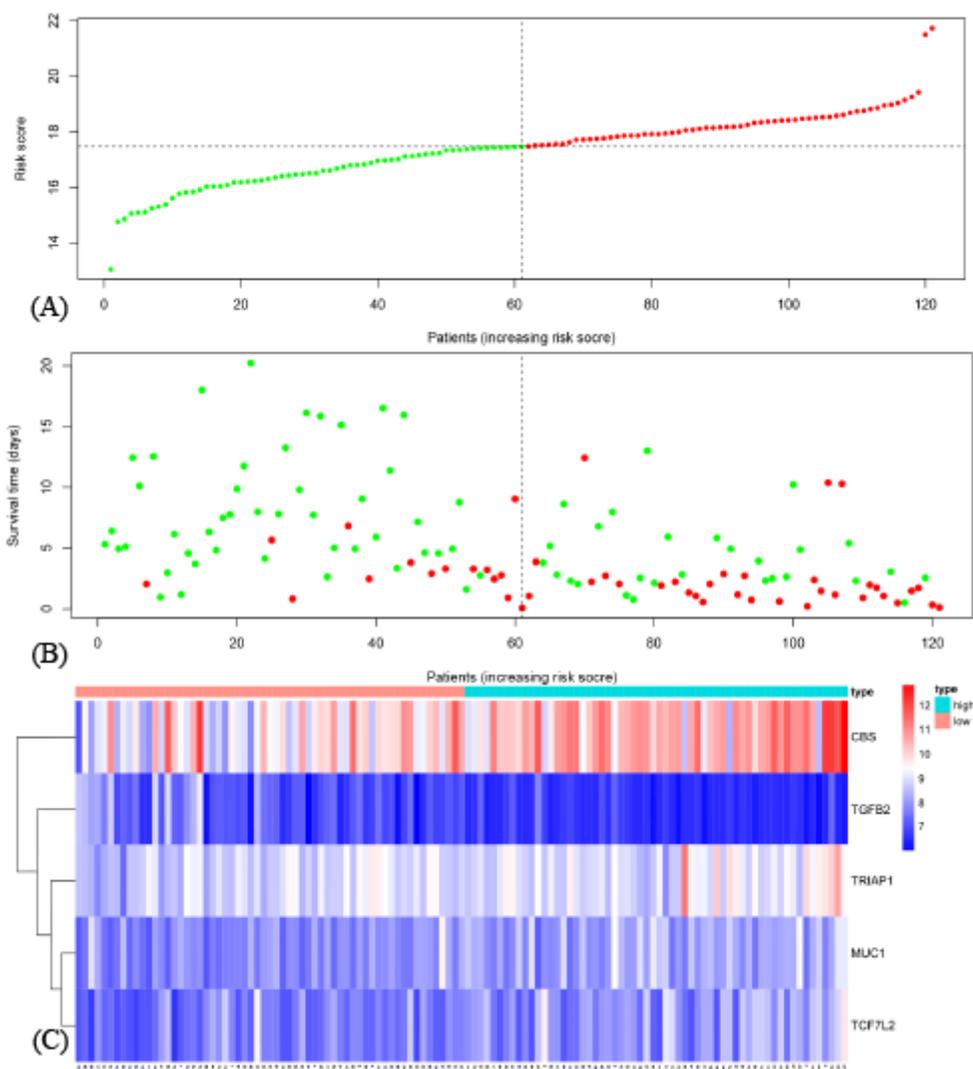


Figure 3

In GEO merged cohort, the distribution of risk scores in patients and the relationship between risk scores and survival time were displayed (Fig. 3A-3B). By heatmap, gene expression profiles in high-risk and low-risk were exhibited (Fig. 3C).

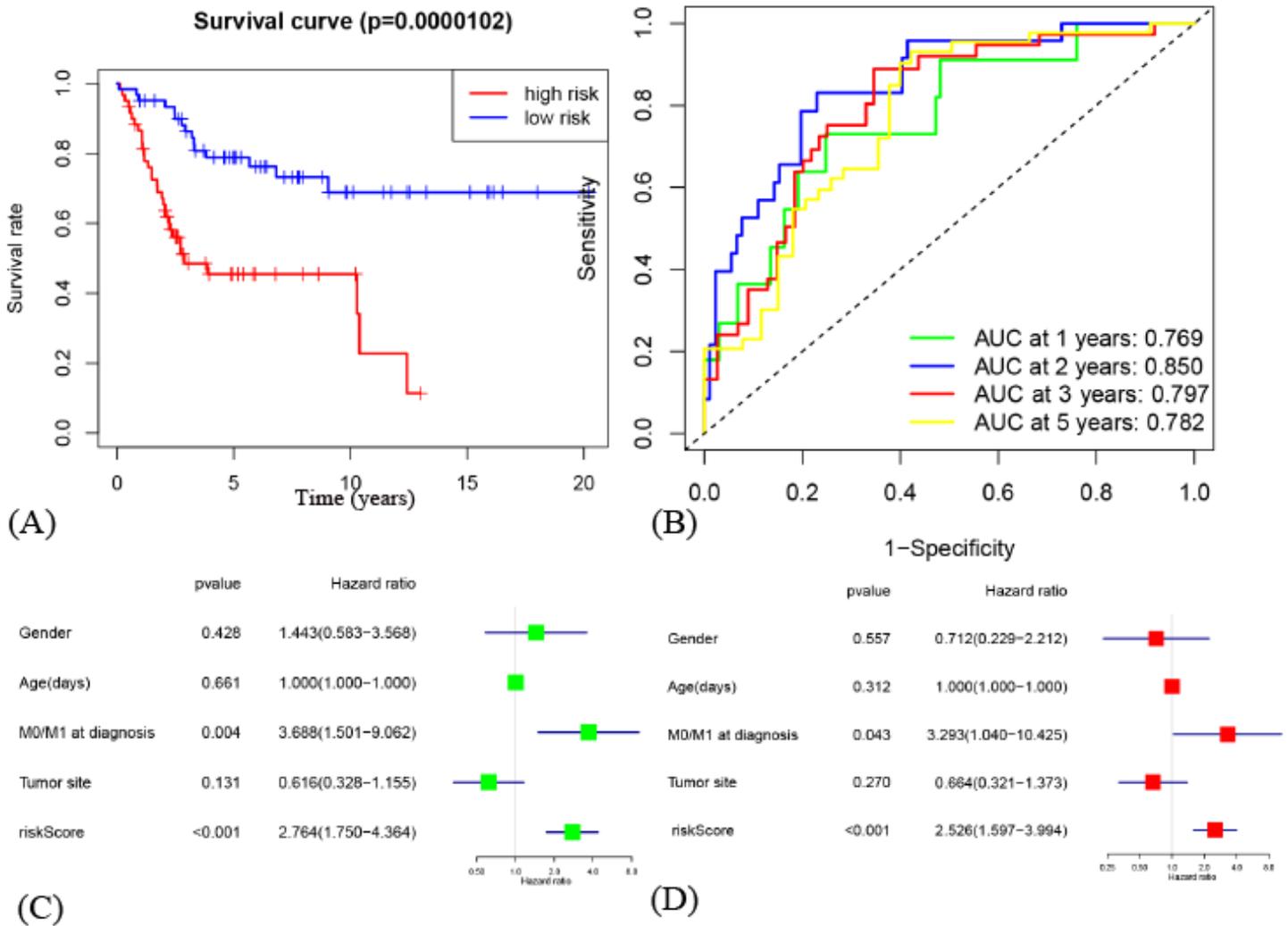


Figure 4

With the Kaplan-Meier cumulative curve in GEO merged cohort, the validity of the risk model was exhibited (Fig. 4A). By the time-dependent ROC curves, the AUC for 1-year, 2-year, 3-year, and 5-year OS were gathered (Fig. 4B). In GEO merged cohort, both univariate and multivariate analysis showed that the risk score an independent prognostic indicator (HR = 2.764, 95%CI = 1.750–4.364, P < 0.001; HR = 2.526, 95%CI = 1.597–3.994, P < 0.001) (Fig. 4C-4D).

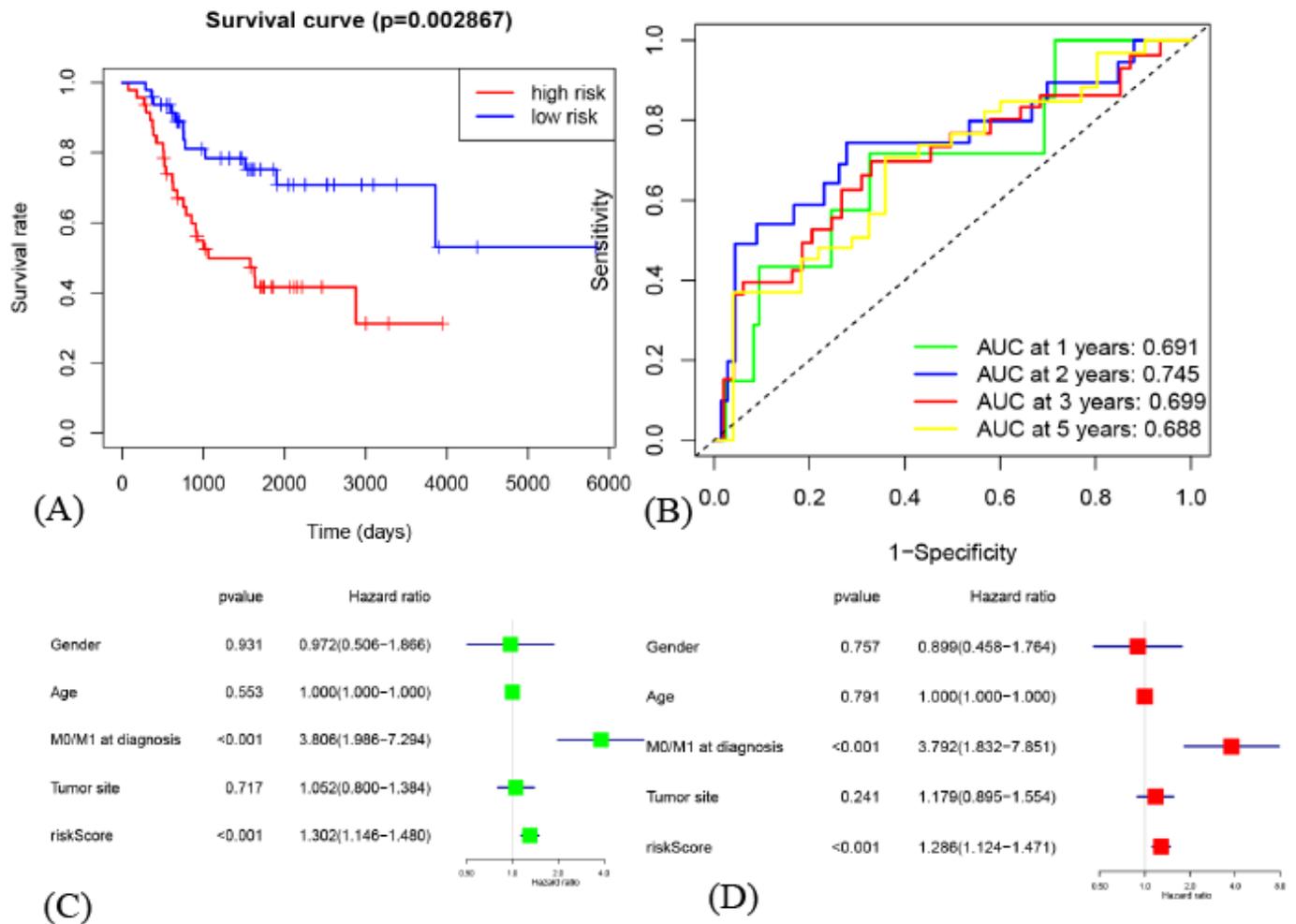


Figure 5

Further, we validated this model in the external independent TARGET cohort with the same way (Fig. 5A-5B). The risk score was still proved an independent prognostic factor the same as tumor stage at diagnosis (Fig. 5C-5D).

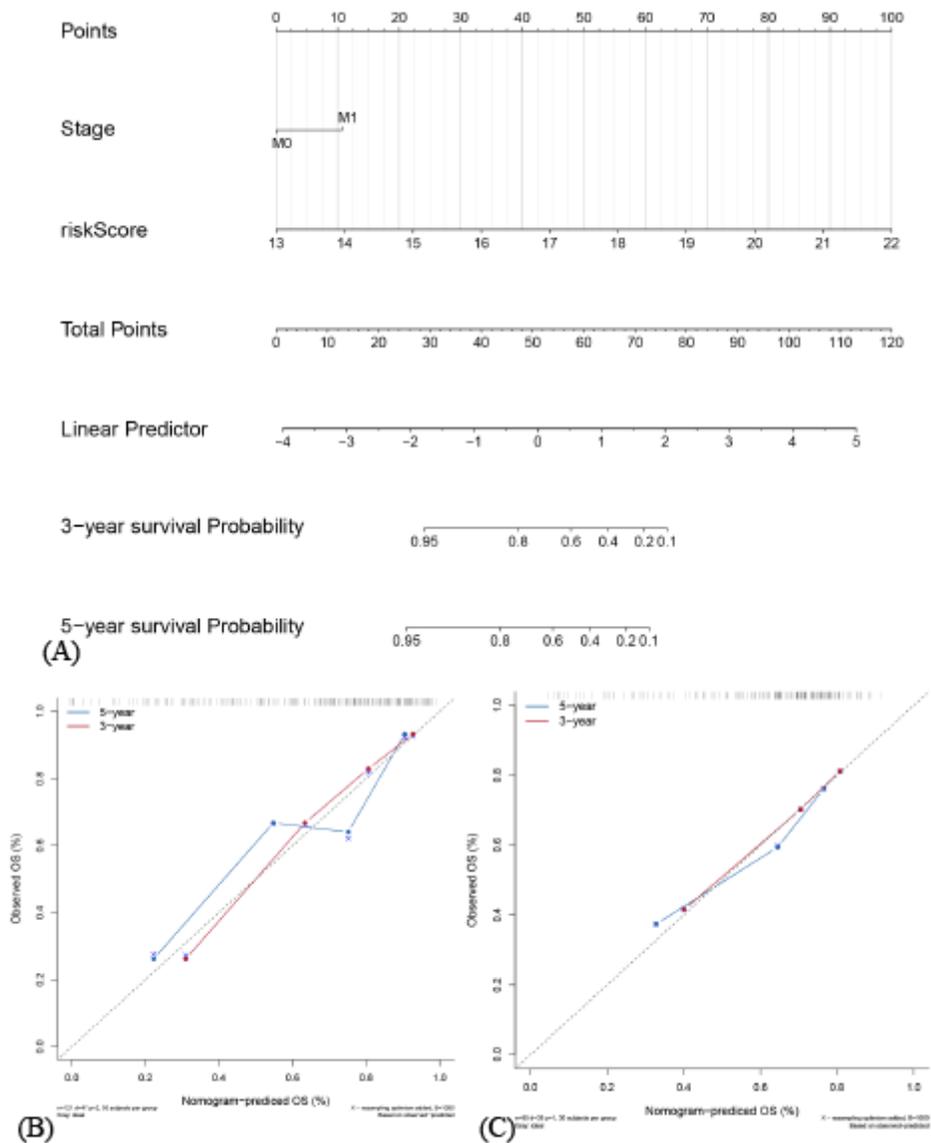


Figure 6

And in this study, the nomogram consisting of above significant risk factors was established, with C-index of 0.798 in GEO merged cohort and 0.700 in TARGET cohort (Fig. 6A). The calibration curves showed non-significant deviations between predicted and actual probability in both GEO merged cohort and TARGET cohort (Fig. 6B-6C).

Supplementary Files

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

- [Table1.xlsx](#)
- [SupplementFigureS1.png](#)
- [SupplementTableS1.xlsx](#)