Journal of Bone Oncology 26 (2021) 100331

Contents lists available at ScienceDirect

Journal of Bone Oncology

journal homepage: www.elsevier.com/locate/jbo

Identifying biomolecules and constructing a prognostic risk prediction model for recurrence in osteosarcoma

Minglei Zhang [a], Yang Liu [b], Daliang Kong [a,*]

[a] Departments of Orthopaedics, China-Japan Union Hospital of Jilin University, No.126, Xiantai Street, Changchun, Jilin 130033, China
[b] Department of Radiological, The Second Clinical Hospital of Jilin University, NO.218, Ziqiang Street, Nanguan District, Changchun, Jilin 130000, China

ARTICLE INFO

Article history:
Received 21 July 2020
Revised 30 September 2020
Accepted 2 October 2020
Available online 21 October 2020

Keywords:
Osteosarcoma
Recurrence
LncRNA
mRNA
Prognostic signature

ABSTRACT

Introduction: Osteosarcoma is a high-morbidity bone cancer with an unsatisfactory prognosis. The aim of this study is to develop novel potential prognostic biomarkers and construct a prognostic risk prediction model for recurrence in osteosarcoma.
Methods: By analyzing microarray data, univariate and multivariate Cox regression analyses were performed to screen prognostic RNA signatures and to build a prognostic model. The RNA signatures were validated using Kaplan–Meier curves. Then, we developed and validated a nomogram combining age, recurrence, metastatic, and Prognostic score (PS) models to predict the individual's overall survival at the 3- and 5-year points. Pathway enrichment of RNA was conducted based on the significant co-expressed RNAs.
Results: A total of 319 mRNAs and 14 lncRNAs were identified in the microarray data. One lncRNA (LINC00957) and six mRNAs (METL1, CA9, B3GALT4, ALDH1A1, LAMB3, and ITGB4) were identified as RNA signatures and showed good performances in survival prediction for both the training and validation cohorts. Cox regression analysis showed that the seven RNA signatures could independently predict overall survival. Furthermore, age, recurrence, metastatic, and PS models were identified as independent prognostic factors via univariate and multivariate Cox analyses ($P < 0.05$) and included in the prognostic nomogram. The C-index values for the 3- and 5-year overall survival predictions of the nomogram were 0.809 and 0.740, respectively.
Conclusions: The current study provides the novel potential of seven RNA candidates as prognostic biomarkers. Nomograms were constructed to provide accurate and individualized survival prediction for recurrence in osteosarcoma patients.

1. Introduction

Osteosarcoma is the most common primary non-hematological bone cancer, affecting adolescents and children more than adults, histologically characterized by the production of osteoid in malignant cells [1]. Recurrence and metastasis are principal pathological problems in the malignant progression of osteosarcoma, and the long-term survival rate of such patients is around 20%, which evidently hampers the effectiveness of osteosarcoma clinical treatments and brings unfavorable outcomes to osteosarcoma patients [2]. The survival rate of these patients has improved as a result of comprehensive management in the form of intensive chemotherapy and surgery [3]. However, the progress has dwindled despite modern therapy over the past three decades, the prognosis remains poor for most patients with recurrent osteosarcoma [4–7]. Therefore, accurate prognosis and efficient therapy are urgently needed to improve the treatment of recurrence in osteosarcoma.

As a complicated disease, osteosarcoma results from interactions between genetic and other factors. Numerous studies have demonstrated that a variety of factors may contribute to the development of osteosarcoma, including age, gender, ethnicity, or physical, chemical, and biological agents [8–10]. In addition, some studies showed that genetic factors may play a more important role in the pathogenesis of osteosarcoma [11–13]. Recently, several studies have investigated the potential role of lncRNAs and mRNAs as diagnostic or prognostic targets in osteosarcoma [14–16]. Decreased expression levels of lncRNA maternally expressed 3 (MEG3) have been reported to be an independent predictor of a shorter overall survival period in patients with osteosarcoma,

* Corresponding author.
E-mail addresses: zhangml1997@126.com (M. Zhang), 1303334113@qq.com (Y. Liu), kongdl666@163.com (D. Kong).

https://doi.org/10.1016/j.jbo.2020.100331
2212-1374/

suggesting that MEG3 may be a useful prognostic biomarker and may exhibit an important pro-oncogenic effect for the prognosis of osteosarcoma [17–19]. Hongzeng Wu et al. were the first to report that high expression of mRNA TfR1 and VEGF was significantly correlated to poor overall survival, and both TfR1 and VEGF were the independent prognostic indicators of osteosarcoma patients [3,20]. However, because the risk of relapse may differ among patients at the same disease stage in the osteosarcoma population, the roles of lncRNA and mRNA are still unclear, and the association of lncRNAs and mRNAs in the prognosis of osteosarcoma patients remains elusive. Consequently, individual and early prediction of recurrence while planning osteosarcoma management is crucial as it could result in better-adapted treatments and survival rates.

This study aimed to develop and validate a survival prediction nomogram for an individualized prediction of survival in patients with recurring osteosarcoma. Note that in this study, patients with osteosarcoma recurrence only represented local recurrence, not including metastasis. The lncRNA and mRNA expression profiles and matched clinical information in samples of patients with osteosarcoma were integrated to identify the prognostic biomarkers associated with the overall survival of patients with osteosarcoma and to establish an RNA prognostic risk model that can effectively predict clinical survival. The significant prognostic power of an RNA prognostic risk model was assessed and validated. Subsequently, an effective prognostic nomogram that incorporates both RNA signatures and clinical risk factors to predict 3- and 5-year overall survival rates and cancer-specific survival rates for patients with osteosarcoma was constructed.

2. Materials & methods

2.1. Patient information and data preprocessing

The RNA-seq dataset (including lncRNA and mRNA) of the 265 osteosarcoma samples and the clinical survival data was downloaded from TCGA (https://tcga-data.nci.nih.gov/) and obtained by Illumina HiSeq 2000 RNA Sequencing. After retaining the osteosarcoma samples with recurrence and clinical survival prognosis information, 169 osteosarcoma samples were included in this study. Simultaneously, the microarray data GSE39055 [21] was downloaded from GEO (https://www.ncbi.nlm.nih.gov/) with "osteosarcoma, recurrence, Homo sapiens" as the search terms [22]. A total of 37 osteosarcoma samples (including 18 recurrence and 19 non-recurrence) were downloaded, and all of the samples had clinical survival prognosis information. The Illumina HumanHT-12 WG-DASL V4.0 R2 Expression BeadChip detection platform was used to obtain the dataset.

2.2. Differential expression analysis

The lncRNA and mRNA in the TCGA datasets were re-annotated using the Ref Seq ID information supported by the detection platform of the HUGO Gene Nomenclature Committee (http://www.genenames.org/) [23], which includes 4526 lncRNAs and 19,205 mRNAs. The limma package (version 3.34.7, https://bioconductor.org/packages/release/bioc/html/limma.html) [24] in R was used to perform screening differentially expressed RNAs [DERs, including differentially expressed lncRNA (DE-lncRNAs) and differentially expressed mRNA (DE-mRNAs)] between recurrence and non-recurrence osteosarcoma samples. A significance analysis of microarrays with false discovery rate (FDR) < 0.05 and $|\log_2$ fold change (FC)| > 0.263 were chosen as the cutoff criteria to define the DERs. Hierarchical clustering analyses of DERs were performed using the pheatmap package (version 1.0.8, https://cran.r-project.org/web/packages/pheatmap/index.html) in R[25] and were presented in two-way hierarchical clustering heatmaps based on centered Pearson correlation [26]. $P < 0.05$ was considered statistically significant.

2.3. Construction prognostic score model

The R package "survival" (http://bioconductor.org/packages/survivalr/) [27] was used to identify the independent prognosis-related DERs for the 169 osteosarcoma samples by univariate and multivariate COX regression analyses. Log-rank *P*-value < 0.05 was set as the significant correlation threshold. On the basis of these prognosis-related DERs, a Cox Proportional Hazards (Cox-PH) model was applied to select the optimal panel of prognostic DERs. The optimal lambda was determined after running 1,000 stimulations through a cross-validation likelihood using the R package "penalized" (version 0.9-50, http://bioconductor.org/packages/penalized/) [28]. Afterwards, the prognostic score model was built based on the expression levels and independent prognostic coefficients of the independent prognosis-associated DERs. The prognostic risk score (PS) of the osteosarcoma samples were calculated using the formula below [29]:

$$\mathrm{PS} = \sum \beta_{\mathrm{DERs}} \times \mathrm{Exp}_{\mathrm{DERs}}$$

β_{DERs} indicates the independent prognostic coefficients of independent prognosis-related DERs, and $\mathrm{Exp}_{\mathrm{DERs}}$ denotes the expression levels of independent prognosis-related DERs in the training dataset.

The median of the PS value of osteosarcoma samples in the training dataset were calculated, then the samples were divided into high-risk and low-risk groups using the median PS as the cutoff point. The overall survival differences between high-risk groups and low-risk groups were determined using Kaplan–Meier survival curves in the R package "survival" (version 2.41-1, http://bioconductor.org/packages/survivalr/) [27]. Then, P-values and hazard ratio (HR) with a 95% confidence interval were generated by log-rank tests. Further to this, time-dependent receiver operating characteristic (ROC) curves were employed to measure predictive performance, and the GSE39055 dataset from the GEO database was used to validate the prognostic model.

2.4. Construction of an osteosarcoma-specific prognostic model

Univariate and multivariate Cox regression analyses were used to screen the independent predictive value of the DERs prognostic model in 169 osteosarcoma patients with clinical information from the TCGA, including age (>20 years), gender, tumor multifocal, tumor recurrence, tumor metastasis, radiotherapy, and tumor necrosis. Statistically significant correlation was performed with a log-rank *P*-value < 0.05 and was set as the cutoff. Kaplan–Meier analysis was performed, and Kaplan–Meier curves were plotted for independent predictive clinical characteristic models using the *P*-value < 0.05 as a cutoff value.

We included each independent predictive factor selected by the multivariate Cox regression analysis to generate a nomogram using the "rms" package (version 5.1-2, https://cran.r-project.org/web/packages/rms/index.html) in R3.4.1 [30,31]. The calibration and discrimination of the independent predictive model were included as validation steps. The ROC curve analyses with 3- and 5-years as the defining points were performed using the R package "pROC" (version 1.14.0, https://cran.r-project.org/web/packages/pROC/index.html) [32], which has been used to evaluate prognostic performance for survival prediction. The area under the ROC curve (AUC) for evaluating discriminatory ability was calculated, and the values ranged from 0.5 to 1, with those closer to 1 indicat-

ing better efficiency. Furthermore, the prognostic capacity of the nomogram was assessed by calculating the Harrell's concordance index (C-index) in the R3.4.1 "survcomp" package (http://www.bioconductor.org/packages/release/bioc/html/survcomp.html) [33]. The value of the C-index ranged from 0.5 to 1.0 (C-index = 0.5 indicated random chance, and C-index = 1.0 indicated perfect predictive accuracy). Generally, C-index > 0.7 indicated an acceptable discrimination accuracy for prognosis.

2.5. Construction of lncRNA and mRNA co-expression network and function enrichment analysis

The co-expression analysis of the DERs significantly associated with prognosis was conducted based on the Pearson correlation coefficient of the Cor function in the R language (https://stat.ethz.ch/R-manual/R-devel/library/stats/html/cor.test.html) [34]. Their expression levels were measured on a network visualization display via Cytoscape (version 3.6.1, http://www.cytoscape.org/) [35]. Then, KEGG pathway enrichment analysis was performed on DERs in the lncRNA–mRNA network using Gene Set Enrichment Analysis (http://software.broadinstitute.org/gsea/index.jsp) in R [36]. P < 0.05 was considered to screen KEGG pathways that were significantly enriched in the relevant DERs.

3. Results

3.1. Data preprocessing and DERs screening

After data preprocessing, a total of 10,700 mRNAs and 1029 lncRNAs were detected (Table S1). A total of 333 DERs were obtained among them, including 319 mRNAs (120 up-regulated and 199 down-regulated) and 14 lncRNAs (two up-regulated and 12 down-regulated) in osteosarcoma samples with recurrence (n = 28) compared with non-recurrence (n = 141) when p < 0.05 and $|\log_2 FC| > 0.263$ was the cutoff criteria (Table S2). We identified all of the DERs that were shown in the volcano map, according to the value of $|\log_2 FC|$, and displayed the DERs on a volcano map (Fig. 1A). The expression values of differentially expressed lncRNAs and mRNAs were two-way hierarchically clustered, and the color contrast of the heatmap indicated that there was a significant difference in the expression levels between the non-recurrence and recurrence osteosarcoma samples (Fig. 1B).

3.2. Construction of DERs signature prognostic model

A total of 87 mRNAs and four lncRNAs that were significantly associated with independent predictive biomarkers by univariate Cox proportional hazards regression analyses (P < 0.05; Table S3) were identified. Subsequently, multivariate Cox regression analysis was performed using these DERs to further screen independent predictive biomarkers, including 25 mRNAs and two lncRNAs (P < 0.05; Table S4). A total of seven optimal panels of prognostic signature DERs (one lncRNA and six mRNAs) were selected using a Cox-PH model (Table 1). Then, a predictive model was constructed based on the coefficient of these seven DERs. As a result, LINC00957, METTL1, and CA9 had positive coefficients and HR > 1, which indicated that higher expression levels of these DERs were associated with a shorter overall survival time, and B3GALT4, ALDH1A1, LAMB3, and ITGB4 may be protective factors because of negative coefficients and HR < 1. Higher expression levels of B3GALT4, ALDH1A1, LAMB3, and ITGB4 predicted better overall survival.

The K-M survival curve results showed that the survival ratio was significantly different between the high-risk and low-risk groups in the training ($p = 1.911e{-}09$) (Fig. 2A) and validation ($p = 7.818e{-}03$) (Fig. 2B), demonstrating that OS patients with high-risk scores had significantly poorer overall survival times compared with patients with lower risk scores. The AUCs of these seven DER signatures were 0.900 and 0.789 for the training and validation datasets, respectively, and the PS scores and grouping are shown in Table S5.

3.3. Construction of an Osteosarcoma-Specific prognostic model

Univariate and multivariate Cox regression analyses were performed to screen the independent prognostic clinical factors, and the clinical information of the samples is shown in Table S6. A total of four independent prognostic factors were significantly screened in the training set, including age, recurrence, metastatic, and PS models (Table 2). Furthermore, the K-M survival analysis indicated that for patients ≤60 years old, a significant longer overall survival time was observed in the patients >60 years old group (p = 1.509E−02) (Fig. 3A), but for osteosarcoma patients with developed recurrence (p = 4.61E−05) (Fig. 3B) and metastasis (p = 1.48E−04) (Fig. 3C), the overall survival time of patients in the recurrence and metastasis group were significantly shorter

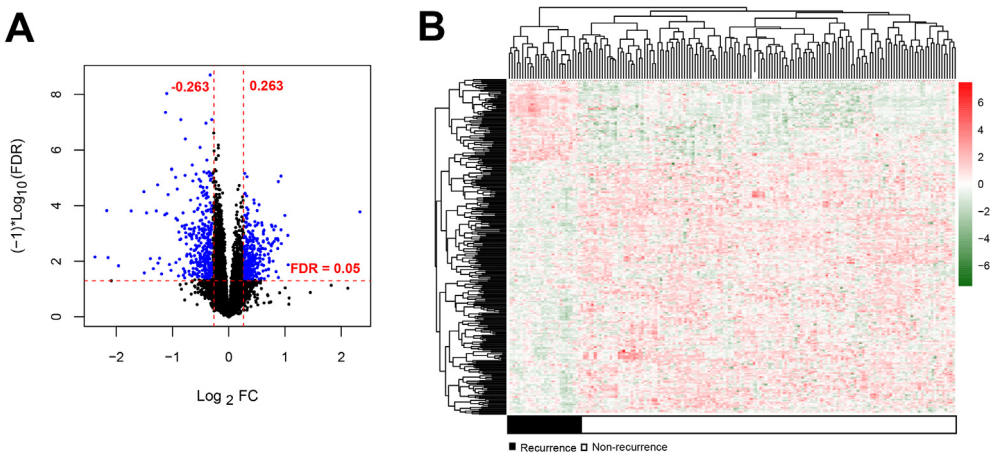

Fig. 1. A: The volcano plot for DERs related to recurrence. The x-axis is the $\log_2$ fold change (FC), and the y-axis is $-\log_{10}$ false discovery rate (FDR). Blue dots indicate significant differentially expressed genes (DERs), the red horizontal dashed line indicates FDR < 0.05, and the two red vertical lines indicate $|\log_2 FC| > 0.263$. B: Two-way hierarchically clustered heatmap for TCGA using the DERs: red indicates up-regulated DERs, green indicates down-regulated DERs, black indicates recurrence osteosarcoma samples, and white indicate non-recurrence osteosarcoma samples. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Information of 7 differentially expressed genes prognostic signature.

Symbol	Type	Multi-variate Cox regression analysis			LASSO coefficient
		HR	95%CI	P value	
Risky RNAs					
LINC00957	lncRNA	1.121	1.072–1.629	4.55E-02	0.0153
METTL1	mRNA	1.099	1.004–1.304	2.75E-02	0.0826
CA9	mRNA	1.174	1.017–1.356	2.88E-02	0.1327
Protective RNAs					
B3GALT4	mRNA	0.557	0.381–0.814	2.51E-03	−0.3057
ALDH1A1	mRNA	0.735	0.619–0.872	4.33E-04	−0.2538
LAMB3	mRNA	0.843	0.678–0.948	1.24E-02	−0.1132
ITGB4	mRNA	0.808	0.586–0.957	1.96E-02	−0.0162

LASSO: The least absolute shrinkage and selection operator; HR: Hazard risk; CI: Confidence interval

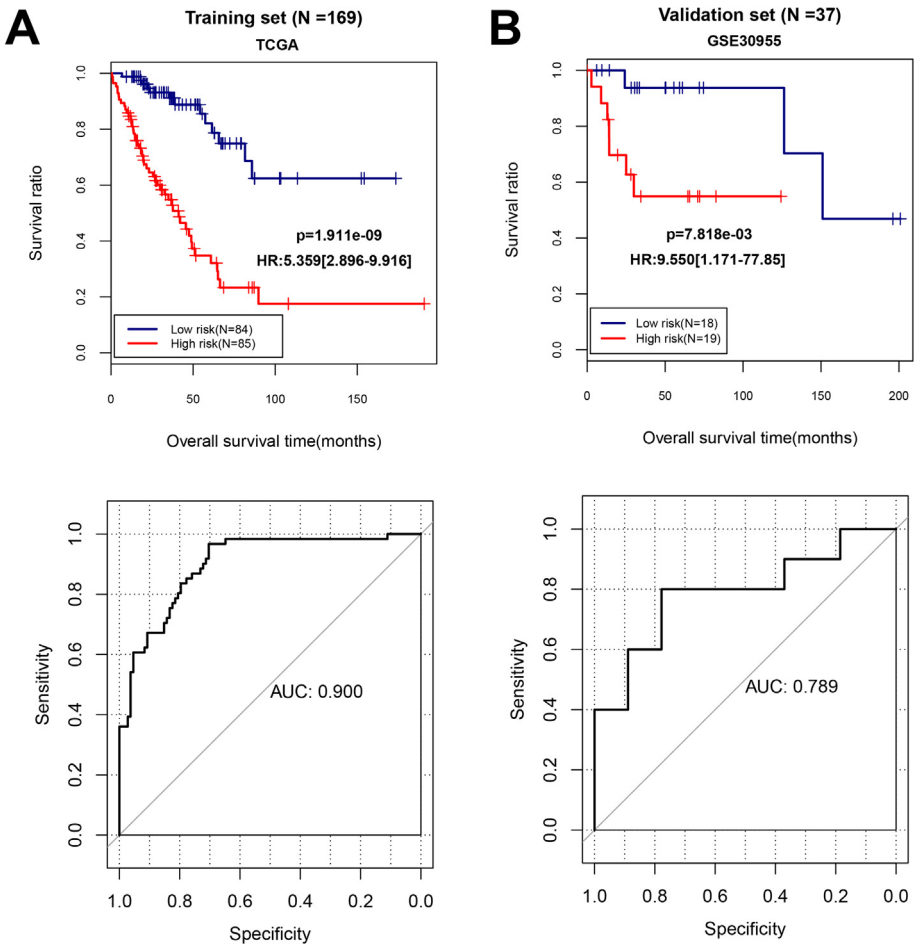

Fig. 2. Kaplan–Meier (K–M) survival curves classified osteosarcoma patients into high-risk and low-risk groups using the seven RNAs signature in the training and test datasets. P-values were calculated by log-rank test. Blue and red curves are for low-risk and high-risk groups, respectively. The survival ratio differences between the two groups were determined using the two-sided log-rank test. A (Above): Training (TCGA, N = 169). A (Below): Receiver operator characteristic (ROC) curve of the prediction results based on the PS prognosis model. B (Above): Validation (GSE30955, N = 37). B (Below): ROC curve of the prediction results based on the PS prognosis model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

compared to that of patients in the non-recurrence and non-metastasis group.

To construct a more sensitive predictive tool in clinical practice, through a stepwise Cox proportional hazard analysis, risk score, age, recurrence, metastatic, and PS models were selected to establish a nomogram model (Fig. 4A). According to the calibration plot, the prediction of the 3- or 5-year survival probability of patients with overall survival provided by the nomogram was consistent with the actual observation (Fig. 4B). The C-index values for the 3- and 5-year overall survival predictions of the nomogram were 0.809 and 0.740, respectively, which further indicated favorable discrimination performance. We compared the predictive power of the nomogram models (Fig. 5), and the different model parameters (AUC and C-index) by ROC curve analyses are shown in Table 3. The multi-RNA-based model (AUC = 0.900; C-index = 0.799) and multi-RNAs combined stage model (AUC = 0.939; C-index = 0.0.829) were shown the best predictive power. Due to the ages of the samples used in the study are all>20 years old, this prognostic model is not suitable for the pediatric patients with recurrence osteosarcoma.

Table 2
Univariate and multivariate analyses of cancer-specific survival in the training cohort.

Clinical characteristics	TCGA (N = 169)	Uni-variables cox			Multi-variables cox		
		HR	95%CI	P	HR	95%CI	P
Age(years,mean ± sd)	61.40 ± 15.24	1.019	1.001–1.037	4.24E−02	1.025	1.005–1.045	1.23 E−02
Gender(Male/Female)	69/100	1.155	0.693–1.925	5.80 E−01	–	–	–
Tumor multifocal(Yes/No/-)	33/127/9	1.614	0.896–2.910	1.08 E−01	–	–	–
Tumor recurrence(Yes/No)	28/141	2.692	1.581–4.585	1.48 E−04	1.855	1.061–3.241	4.21 E−02
Tumor metastatic(Yes/No)	56/113	2.754	1.657–4.578	4.61 E−05	2.679	1.517–4.732	6.86 E−04
Radiotherapy(Yes/No)	61/108	0.817	0.483–1.381	4.50 E−01	–	–	–
Tumor necrosis(No/Slight/Moderate/Severe/-)	59/34/59/9/8	1.193	0.927–1.535	1.69 E−01	–	–	–
mRNA status based model(High/ Low)	84/85	6.454	3.437–12.12	6.26 E−11	4.509	2.387–8.514	3.44 E−06
Dead(Death/Alive/-)	61/108	–	–	–	–	–	–
Overall survival time(months,mean ± sd)	40.32 ± 32.59	–	–	–	–	–	–

TCGA: The Cancer Genome Atlas database; HR: Hazard risk; CI: Confidence interval.

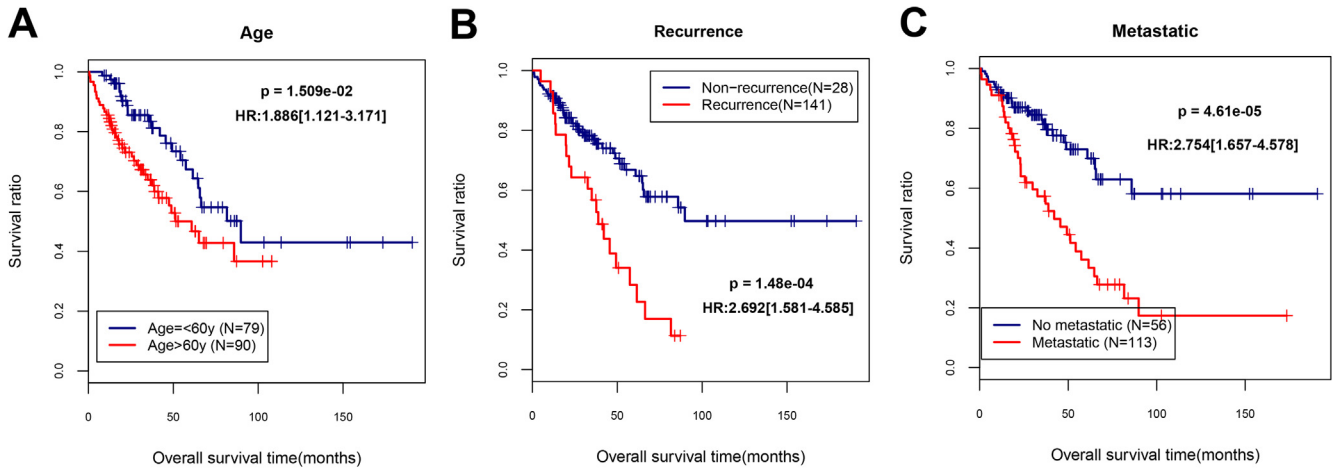

Fig. 3. A. Age prognosis-related Kaplan–Meier (K–M) survival curves; red indicates >60 years old of osteosarcoma samples, and blue indicates ≤60 years old of osteosarcoma samples. B. Recurrence factor prognosis-related K-M curve; red indicates recurrence in osteosarcoma samples, and blue indicates non-recurrence in osteosarcoma samples. C. Metastatic factor prognosis-related K-M; red indicates metastatic osteosarcoma samples, and blue indicates non-metastatic osteosarcoma samples. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.4. Co-expression network construction and function enrichment analyses

We established the genes co-expression network by using screened independent predictive biomarkers. The network contained 56 lncRNA–mRNA pairs (Fig. 6A). All of the co-expression relationship pairs are shown in Table S7. Among them, the mRNAs (B3GALT4, ALDH1A1, LAMB3, and ITGB4) have a co-expression relationship with LINC00957, suggesting that it might regulate the relapse of osteosarcoma. Furthermore, one KEGG pathway (pathways in cancer) (p = 0.0166) was obviously positively correlated with the signature lncRNA LINC00957 (Fig. 6B).

4. Discussion

Recurrence in osteosarcoma has become a serious health burden worldwide and results in poor survival for patients [37]. Therefore, it is crucial to identify the functional genes in order to develop novel therapeutic strategies for the treatment of patients with recurrent osteosarcoma. Emerging evidence has shown that lncRNAs and mRNAs play pivotal roles in the initiation and development of osteosarcoma, and the clinical significance of lncRNAs and mRNAs has also attracted increasing attention for their diagnostic and prognostic value in recurring osteosarcoma [38].

In this research, our results robustly supported that seven DER signatures, including one lncRNA (LINC00957) and six mRNAs (METTL1, CA9, B3GALT4, ALDH1A1, LAMB3, and ITGB4), could be promising assets in predicting the prognosis of recurrent osteosarcoma patients. Furthermore, we developed and validated a nomogram, combining age, recurrence, metastatic, and PS models. The C-index values for 3- and 5-year overall survival predictions of the nomogram were 0.809 and 0.740, respectively, which further indicated favorable discrimination performance. The multi-RNA-based model (AUC = 0.900; C-index = 0.799) and multi-RNAs combined stage model (AUC = 0.939; C-index = 0.0.829) were shown to have the best predictive power. Finally, the genes co-expression network was established, including 56 lncRNA–mRNA pairs. Among them, the mRNAs (B3GALT4, ALDH1A1, LAMB3, and ITGB4) have a co-expression relationship with LINC00957, suggesting that it might regulate the relapse of osteosarcoma.

The long non-coding RNA00957 (LINC00957) was demonstrated to be involved in tumor progression, and Zhang et al. [39]. found that the expression levels of LINC00957 were significantly associated with advanced TNM stage, poor chemotherapy outcomes, and worse prognosis. The β-1, 3-galactosyltransferase-4 (B3GALT4) gene belongs to the β-1, 3-galactosyltransferase (β3GalT) gene family, which plays an essential role in the o-glycosylation process. The surface of cancer cells expresses glycoproteins, which are rich

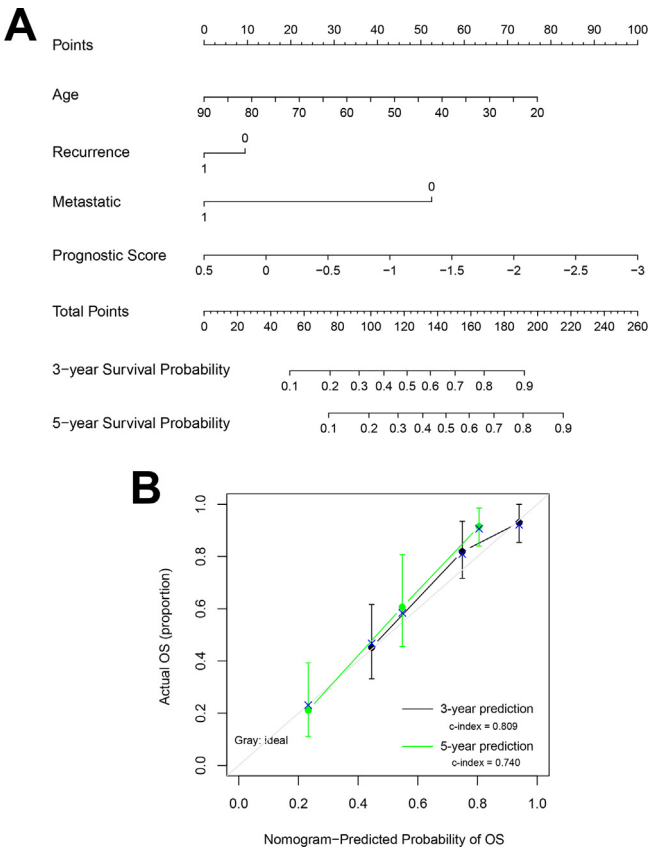

Fig. 4. Nomogram predicting overall survival for patients with osteosarcoma. A. For each patient, three lines are drawn upward to determine the points received from the three predictors in the nomogram. The sum of these points is located on the "Total Points" axis. Then, a line is drawn downward to determine the possibility of 3- and 5-year overall survival of osteosarcoma. B. The calibration plot for the internal validation of the nomogram. The y-axis represents actual survival, and the x-axis represents nomogram-predicted survival.

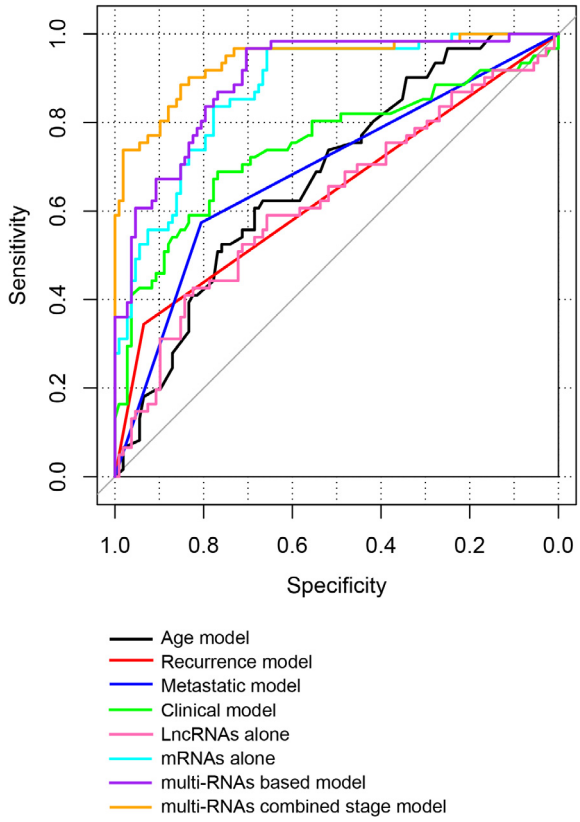

Fig. 5. ROC analysis of the sensitivity and specificity for survival prediction by differently factors prognostic model.

Table 3
The different model parameters information by Receiver Operating Characteristic (ROC) curves analyses.

	AUROC	C-index	P value
Age model	0.679	0.590	3.378 E−02
Recurrence model	0.640	0.571	1.585 E−02
Metastatic model	0.690	0.613	1.250 E−03
Clinical model	0.747	0.675	5.139 E−06
LncRNAs alone	0.625	0.612	3.364 E−03
mRNAs alone	0.875	0.796	0
multi-RNAs based model	0.900	0.799	0
multi-RNAs combined stage model	0.939	0.829	0

AUROC: Area under Receiver Operating Characteristic curve; C-index: index of concordance.

in o-glycosylation domains. Therefore, the gene family may be closely related to the tumor. Furthermore, Seko et al. had confirmed that B3GALT4 could be used as a novel biomarker for the diagnosis of gynecological cancers. In addition, Ting Zhang revealed that B3GALT4 is a novel prognostic biomarker for CRC and highlighted the significance of B3GALT4 as a promising therapeutic target for CRC.[40] Overexpression of ALDH1A1 and ALDH3A1 has been linked to resistance to oxazaphosphorines, such as CPA, in a variety of human cancers, presumably by directing the metabolism of 4-HI to the inactive metabolite. ALDH1A1 can mediate epithelial-to-mesenchymal transition, an important phenomenon associated with tumor invasion and metastasis [41]. Wan-Ting Liu et al. indicated that ALDH1A1 plays an important role in tumor invasion, metastasis, and prognosis [42]. Furthermore, Pooja Hingorani et al. concluded that the ALDH3A1 overexpression might be an active agent in resistant and relapsed osteosarcoma in patients [43]. Laminin subunit beta-3 (LAMB3) is a major component of the basement membrane zone. Liu L. et al. investigated that high LAMB3 expression was significantly associated with positive lymph node metastasis and poor prognosis in patients with head and neck squamous cell carcinoma and supported that LAMB3 silencing could induce the sensitivity of anti-cancer drugs [44].

The integrin β4 (ITGB4), forming a dimer with integrin α6 (ITGA6), during carcinoma progression, the dimer can cause cell deformation and promote tumor cell metastasis behavior [45,46]. ITGB4 may be positively associated with poor prognosis as it is aberrantly expressed in breast, colorectal, and lung cancers [47]. The seven prognostic RNAs have all been reported to be associated with human cancer, suggesting the reliability of the methods that we used in this study.

Our study provided novel evidence that higher expression levels of B3GALT4, ALDH1A1, LAMB3, and ITGB4 predicted better overall survival rates, which might be potential predictors of recurrent osteosarcoma prognosis. Further studies are needed to validate these results and investigate the molecular characteristics. Our research results constitute an improvement in the prediction accuracy of the model by combining separate signature RNAs and clinical factors, and this is sufficiently effective as an independent component of the nomogram. Therefore, our multi-RNA-based model and multi-RNAs combined stage model prediction nomogram for survival prediction of recurrent osteosarcoma might allow physicians to provide a more appropriate treatment strategy for each patient. Furthermore, the nomogram can enable a personalized survival prediction for each patient.

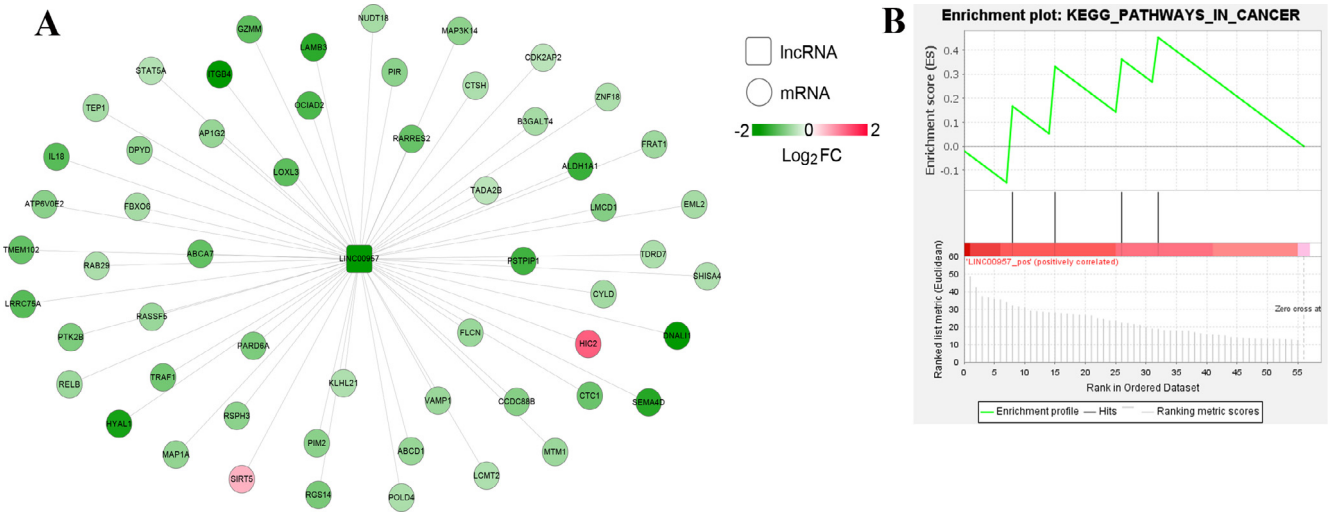

Fig. 6. Co-expression network. The correlation between signature lncRNAs and genes with a significant prognosis was determined by Pearson correlation analysis. A. The square indicates lncRNAs, and circles indicate mRNAs. The changes in color from green to red indicate significant down-regulation to up-regulation. B. KEGG signaling pathway significantly positively related to the signature lncRNA. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Conclusions

In conclusion, our study highlighted the prognostic value of the one lncRNA and seven mRNAs, which might be powerful biomarkers for recurrence in osteosarcoma survival, and suggested practical applications in prognostic predictions and targeted therapy of recurrent osteosarcoma. The predictive nomogram showed robust performance in predicting osteosarcoma prognosis. Therefore, our model might provide an effective and reliable guide to prognosis assessment and treatment decision-making in the clinic.

6. Ethics in publishing

Not applicable.

7. Authors' Contributions

Daliang Kong and Minglei Zhang were responsible for the conception and design of the research, and drafting the manuscript. Yang Liu performed the data acquisition. Minglei Zhang performed the data analysis and interpretation. Yang Liu participated in the design of the study and performed the statistical analysis. All authors have read and approved the manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jbo.2020.100331.

References

[1] M.E. Anderson, Update on survival in osteosarcoma, Orthop. Clin. North Am. 47 (2016) 283–292.
[2] H. Sheen, W. Kim, B.H. Byun, C.B. Kong, W.S. Song, W.H. Cho, I. Lim, S.M. Lim, S. K. Woo, Metastasis risk prediction model in osteosarcoma using metabolic imaging phenotypes: A multivariable radiomics model, PLoS ONE 14 (11) (2019) e0225242.
[3] X. Shi, X. Guan, MicroRNA-99b predicts clinical outcome of osteosarcoma and suppresses tumor cell proliferation, migration and invasion, Diagn. Pathol. 14 (1) (2019).
[4] Z. Guan, PET CT in the diagnosis and prognosis of osteosarcoma, Front Biosci 23 (11) (2018) 2157–2165.
[5] M.M. Ben-Dayan, T.J. Ow, T.J. Belbin, J. Wetzler, R.V. Smith, G. Childs, B. Diergaarde, D.N. Hayes, J.R. Grandis, M.B. Prystowsky, N.F. Schlecht, Nonpromoter methylation of the CDKN2A gene with active transcription is associated with improved locoregional control in laryngeal squamous cell carcinoma, Cancer Med. 6 (2) (2017) 397–407.
[6] N. Wong, S.S. Khwaja, C.M. Baker, H.A. Gay, W.L. Thorstad, M.D. Daly, J.S. Lewis, X.J.C.M. Wang, Prognostic microRNA signatures derived from The Cancer Genome Atlas for head and neck squamous cell carcinomas, Cancer Med. 5 (7) (2016) 1619–1628.
[7] G. Zhang, E. Fan, Q. Zhong, G. Feng, Y.u. Shuai, M. Wu, Q. Chen, X. Gou, Identification and potential mechanisms of a 4-lncRNA signature that predicts prognosis in patients with laryngeal cancer, Hum Genomics 13 (1) (2019).
[8] R. Nandra, M. Parry, J. Forsberg, R. Grimer, Can a Bayesian Belief Network Be Used to Estimate 1-year Survival in Patients With Bone Sarcomas?, Clin Orthop. Relat. Res. 475 (6) (2017) 1681–1689.
[9] J.W. Martin, J.A. Squire, M. Zielenska, The Genetics of Osteosarcoma, Sarcoma 2012 (2012) 1–11.
[10] G. Ottaviani, N. Jaffe, The etiology of osteosarcoma, Cancer Treat Res. 152 (2010) 15–32.
[11] A.L. Sarver, V. Thayanithy, M.C. Scott, MicroRNAs at the human 14q32 locus have prognostic significance in osteosarcoma, Orphanet. J. Rare Dis. 8 (1) (2013) 7.
[12] W.U. Xing, Z.R. Chen, Apoptosis-related gene expression and its clinical significance of human osteosarcoma, Zhonghua Zhong Liu Za Zhi 26 (11) (2004) 678.
[13] H. Feng, P. Guo, J. Wang, J. Xu, C. Xie, F. Gao, Practice: Expression of Leptin and Sirtuin-1 is associated with poor prognosis in patients with osteosarcoma, Pathology – Res. Practice (2016).
[14] S. Hu, J. Chang, Y. Li, W. Wang, E.J. Oncotarget, Long non-coding RNAs for osteosarcoma in the mouse: a meta-analysis, Oncotarget 8 (59) (2015).
[15] Z. Li, P. Dou, T. Liu, S. He, JCP, Biochemistry BIJoECP, Pharmacology: Application of Long Noncoding RNAs in Osteosarcoma: Biomarkers and Therapeutic Targets.1407-1419.
[16] M. Cong, J. Li, R. Jing, Z.J. Li, Long non-coding RNA tumor suppressor candidate 7 functions as a tumor suppressor and inhibits proliferation in osteosarcoma, Tumor Biol. 37 (7) (2016).
[17] B. Shen, N. Zhou, T. Hu, W. Zhao, D. Wu, S. Wang, LncRNA MEG3 negatively modified osteosarcoma development through regulation of miR-361-5p and FoxM1, J. Cell. Physiol. 234 (8) (2019) 13464–13480.
[18] Y. Wang, D. Kong, Knockdown of lncRNA MEG3 inhibits viability, migration, and invasion and promotes apoptosis by sponging miR-127 in osteosarcoma cell, J. Cell. Biochem. 119 (1) (2018) 669–679.

[19] Z.Z. Tian, X.J. Guo, Y.M. Zhao, Y. Fang, Decreased expression of long non-coding RNA MEG3 acts as a potential predictor biomarker in progression and poor prognosis of osteosarcoma, Int. J. Clin. Exp. Pathol. 8 (11) (2014) 15138–15142.
[20] S. Zhang, L. Ding, X. Li, H. Fan, Identification of biomarkers associated with the recurrence of osteosarcoma using ceRNA regulatory network analysis, Int. J. Mol. Med. 43 (4) (2019) 1723–1733.
[21] A.D. Kelly, B. Haibe-Kains, et al., MicroRNA paraffin-based studies in osteosarcoma reveal reproducible independent prognostic profiles at 14q32, Genome Med. 5 (1) (2013) 2.
[22] R. Edgar, Gene expression omnibus: NCBI gene expression and hybridization array data repository, Nucleic Acids Res. 30 (1) (2002) 207–210.
[23] M.W. Wright, A short guide to long non-coding RNA gene nomenclature, Hum Genomics 8 (1) (2014) 7.
[24] M.E. Ritchie, P. Belinda, W. Di, Y. Hu, C.W. Law, S. Wei, G.K. Smyth, Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 7 (2015) 7.
[25] L. Wang, C. Cao, Q. Ma, Q. Zeng, H. Wang, RNA-seq analyses of multiple meristems of soybean: novel and alternative transcripts, evolutionary and functional implications, BMC Plant Biol. 14 (1) (2014).
[26] M.B. Eisen, P.T. Spellman, P.O. Brown, D. Botstein, Cluster analysis and display of genome-wide expression patterns, Proc. Natl. Acad. Sci. 95 (25) (1998) 14863–14868.
[27] P. Wang, Y. Wang, B. Hang, X. Zou, J.H. Mao, A novel gene expression-based prognostic scoring system to predict survival in gastric cancer, Oncotarget 7 (34) (2016) 55343–55351.
[28] J.J. Goeman, L1 penalized estimation in the cox proportional hazards model, Biom. J. 52 (1) (2010).
[29] H. Yang, H.C. Lin, H. Liu, D. Gan, W. Jin, C. Cui, Y. Yan, Y. Qian, C. Han, Z. Wang, A 6 lncRNA-based risk score system for predicting the recurrence of colon adenocarcinoma patients, Front. Oncol. 10 (2020) 81.
[30] W.I. Anderson, D.H. Schlafer, K.R. Vesely, Thyroid follicular carcinoma with pulmonary metastases in a beaver (Castor canadensis), J. Wildl. Dis. 25 (4) (1989) 599–600.
[31] K.H. Eng, E. Schiller, K. Morrel, On representing the prognostic value of continuous gene expression biomarkers with the restricted mean survival curve, Oncotarget 6 (34) (2015).
[32] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, M. Müller, pROC: an open-source package for R and S+ to analyze and compare ROC curves, BMC Bioinformatics 12 (2011) 177, 70.
[33] S. Shan, W. Chen, J.D. Jia, Transcriptome analysis revealed a highly connected gene module associated with cirrhosis to hepatocellular carcinoma development, Front. Genet. 10 (2019) 305.
[34] L.E.J.R. Eberly, Correlation and Simple Linear Regression 227 (3) (2003) 617–622.
[35] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (11) (2003) 2498–2504.
[36] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, PNAS 102 (43) (2005) 15545–15550.
[37] Z L, P D, T L, physiology HSJC, biochemistry : international journal of experimental cellular physiology b, pharmacology: Application of Long Noncoding RNAs in Osteosarcoma: Biomarkers and Therapeutic Targets. 2017, 42(4):1407-1419.
[38] K. Xu, W. Xiong, S. Zhao, B. Wang, MicroRNA-106b serves as a prognostic biomarker and is associated with cell proliferation, migration, and invasion in osteosarcoma, Oncol. Lett. 18 (3) (2019) 3342–3348.
[39] L.H. Zhang, L.H. Li, P.F. Zhang, Y.F. Cai, D. Hua, LINC00957 acted as prognostic marker was associated with fluorouracil resistance in human colorectal cancer, Front. Oncol. 9 (2019) 776.
[40] T. Zhang, F. Wang, J.Y. Wu, Z.C. Qiu, Y. Wang, F. Liu, X.S. Ge, X.W. Qi, Y. Mao, D. Hua, Clinical correlation of B7–H3 and B3GALT4 with the prognosis of colorectal cancer, World J. Gastroenterol. 24 (31) (2018) 3538–3546.
[41] N.E. Sladek, R. Kollander, L. Sreerama, D.T. Kiang, Cellular levels of aldehyde dehydrogenases (ALDH1A1 and ALDH3A1) as predictors of therapeutic responses to cyclophosphamide-based chemotherapy of breast cancer: a retrospective study. Rational individualization of oxazaphosphorine-based cancer chemotherapeutic regimens, Cancer Chemother. Pharmacol. 49 (4) (2002) 309–321.
[42] W.T. Liu, W.B. Liu, M. Gao, Y.Y. Zhang, K.S. Gu, Expression of ALDH1A1 and CD133 is associated with the prognosis and effect of different chemotherapeutic regimens in gastric cancer, Oncol. Lett. 18 (5) (2019) 4573–4582.
[43] P. Hingorani, W. Zhang, S. Piperdi, L. Pressman, J. Lin, R. Gorlick, E.A. Kolb, Preclinical activity of palifosfamide lysine (ZIO-201) in pediatric sarcomas including oxazaphosphorine-resistant osteosarcoma, Cancer Chemother. Pharmacol. 64 (4) (2009) 733–740.
[44] L. Liu, S.N. Jung, C. Oh, K. Lee, H.R. Won, J.W. Chang, J.M. Kim, B.S. Koo, LAMB3 is associated with disease progression and cisplatin cytotoxic sensitivity in head and neck squamous cell carcinoma, Eur. J. Surgical Oncol. 45 (3) (2019) 359–365.
[45] I. Rabinovitz, A.M. Mercurio, The integrin alpha6beta4 functions in carcinoma cell migration on laminin-1 by mediating the formation and stabilization of actin-containing motility structures, J. Cell Biology 139 (7) (1997) 1873–1884.
[46] R.L. Stewart, K.L. O'Connor, Clinical significance of the integrin alpha6beta4 in human malignancies, Laboratory Investigation 95 (9) (2015) 976–986.
[47] M. Li, X. Jiang, G. Wang, C. Zhai, Y. Liu, H. Li, Y. Zhang, W. Yu, Z. Zhao, ITGB4 is a novel prognostic factor in colon cancer, J. Cancer 10 (21) (2019) 5223–5233.