



Construction of a Five-Super-Enhancer-Associated-Genes Prognostic Model for Osteosarcoma Patients

Zhanbo Ouyang¹, Guohua Li¹, Haihong Zhu¹, Jiaojiao Wang¹, Tingting Qi¹, Qiang Qu², Chao Tu³, Jian Qu^{1*} and Qiong Lu^{1*}

¹ Department of Pharmacy, The Second Xiangya Hospital, Central South University; Institute of Clinical Pharmacy, Central South University, Changsha, China, ² Department of Pharmacy, Xiangya Hospital, Central South University, Changsha, China, ³ Department of Orthopaedics, The Second Xiangya Hospital, Central South University, Changsha, China

OPEN ACCESS

Edited by:

Changjun Li,
Xiangya Hospital, Central South
University, China

Reviewed by:

Lanxiang Wu,
Chongqing Medical University, China
Xiaolong Tang,
Shenzhen University Health Science
Centre, China
Donna Fu,
University of California,
San Francisco, United States

*Correspondence:

Jian Qu
qujianstanley@csu.edu.cn
Qiong Lu
christy_luq@csu.edu.cn

Specialty section:

This article was submitted to
Molecular Medicine,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 25 August 2020

Accepted: 05 October 2020

Published: 30 October 2020

Citation:

Ouyang Z, Li G, Zhu H, Wang J,
Qi T, Qu Q, Tu C, Qu J and Lu Q
(2020) Construction of a Five-Super-
Enhancer-Associated-Genes
Prognostic Model for Osteosarcoma
Patients.
Front. Cell Dev. Biol. 8:598660.
doi: 10.3389/fcell.2020.598660

Osteosarcoma is a malignant tumor most commonly arising in children and adolescents and associated with poor prognosis. In recent years, some prognostic models have been constructed to assist clinicians in the treatment of osteosarcoma. However, the prognosis and treatment of patients with osteosarcoma remain unsatisfactory. Notably, super-enhancer (SE)-associated genes strongly promote the progression of osteosarcoma. In the present study, we constructed a novel effective prognostic model using SE-associated genes from osteosarcoma. Five SE-associated genes were initially screened through the least absolute shrinkage and selection operator (Lasso) penalized Cox regression, as well as univariate and multivariate Cox regression analyses. Meanwhile, a risk score model was constructed using the expression of these five genes. The excellent performance of the five-SE-associated-gene-based prognostic model was determined via time-dependent receiver operating characteristic (ROC) curves and Kaplan–Meier curves. Inferior outcome of overall survival (OS) was predicted in the high-risk group. A nomogram based on the polygenic risk score model was further established to validate the performance of the prognostic model. It showed that our prognostic model performed outstandingly in predicting 1-, 3-, and 5-year OS of patients with osteosarcoma. Meanwhile, these five genes also belonged to the hub genes associated with survival and necrosis of osteosarcoma according to the result of weighted gene co-expression network analysis based on the dataset of GSE39058. Therefore, we believe that the five-SE-associated-gene-based prognostic model established in this study can accurately predict the prognosis of patients with osteosarcoma and effectively assist clinicians in treating osteosarcoma in the future.

Keywords: osteosarcoma, super-enhancer, prognostic model, Lasso, weighted gene co-expression network analysis, overall survival

Abbreviations: AIC, Akaike information criterion; AUC, area under the ROC; C-index, concordance index; GEO, Gene Expression Omnibus; GS, gene significance; HR, hazard ratio; IGV, Integrative Genomics Viewer; K–M, Kaplan–Meier; Lasso, least absolute shrinkage and selection operator; MM, module membership; OS, overall survival; ROC, receiver operating characteristic; SE, super-enhancer; TCGA, The Cancer Genome Atlas; WGCNA, weighted gene co-expression network analysis.

INTRODUCTION

Osteosarcoma, the most common primary bone tumor, is most commonly arising in children and adolescents (Bielack et al., 2009). The survival rate of patients with osteosarcoma is markedly low due to the tendency of osteosarcoma for early metastasis and rapid progression (Kempf-Bielack et al., 2005). The overall 5-year survival rate of patients with non-metastatic osteosarcoma is approximately 60–70%, while the survival rate of osteosarcoma patients with metastatic symptoms is only 20–30% (Kempf-Bielack et al., 2005; Luetke et al., 2014; Meazza and Scanagatta, 2016). Moreover, high-grade patients comprise nearly 80–90% of those diagnosed with osteosarcoma; the treatment of these patients is challenging (Bielack et al., 2009). Furthermore, despite advancements in treatment in recent years, the therapeutic outcomes remain unsatisfactory. Although numerous traditional predictive factors, such as age, necrosis, recurrence, and clinical stage, are important contributors to clinical outcome, they are less effective in predicting the survival status due to the complex molecular mechanisms of osteosarcoma progression. Thus, it is urgent to investigate novel effective molecular biomarkers for the more precise prediction of the prognosis of patients with osteosarcoma.

Super-enhancer is a large cluster of *cis*-regulatory DNA elements densely bound by transcription factors and cofactors, playing vital roles in defining cell fate and identity (Hnisz et al., 2013). Histone marks (H3K27ac, H3K4me1, and H3K4me3), cofactor (p300), and mediators (CDK7 and BRD4) are common biomarkers used to define SE (Lovén et al., 2013; Hnisz et al., 2015; Ma et al., 2019); among them, H3K27ac is the most efficient biomarker for identifying SE (Creighton et al., 2010; Hnisz et al., 2013; Whyte et al., 2013). Therefore, we used H3K27ac, H3K4me1, and H3K4me3 as SE biomarkers in this study based on the ChIP-seq profiles data from Cistrome. Moreover, SE plays a critical role in the progression of osteosarcoma through activating the expression of its downstream target oncogenes, such as *MYC*, *LIF*, and *STAT3* (Chen et al., 2018; Lu et al., 2020; Zhang et al., 2020). Furthermore, SE inhibitors (e.g., JQ1, THZ1, and THZ2), which were developed for the treatment of osteosarcoma, could restore the expression levels of these genes (Lee et al., 2015; Chen et al., 2018; Zhang et al., 2020). Thus, screening and identifying hub oncogenes driven by SE will provide novel insights into the diagnosis, prognosis, and treatment of osteosarcoma.

The Lasso penalized Cox regression analysis, designed by Tibshirani (1997), is an accurate statistical method frequently used for the selection of variables in the Cox regression model. It can remarkably improve the accuracy of the model it generates compared with the stepwise regression method, as it involves fewer variables and produces more interpretable models (Tibshirani, 1997; Xiong et al., 2019). The reason for Lasso screening fewer variables is that it can regularize the impact of some variables by shrinking their coefficients to zero (Zhang et al., 2018). Nowadays, Lasso is widely used in constructing prognostic models for the prediction of

survival based on complicated high-throughput genomic data (Zhang et al., 2013; Lu et al., 2019; Wang et al., 2019; Wu et al., 2019). Based on this advantage, we applied the Lasso regression method to the construction of a prognostic model of osteosarcoma.

Weighted gene co-expression network analysis (WGCNA) is a system bioinformatics method involving molecular interaction mechanism analysis and construction of correlation gene networks (Zhang and Horvath, 2005). WGCNA has been frequently utilized in identifying probable oncogenes and in screening candidate targets for various cancers (Udyavar et al., 2013). Using WGCNA, Tian et al. (2018) showed that the insulin-like growth factor-binding-associated genes may play a critical role in osteosarcoma metastasis progression. Guan et al. (2020) defined four key genes (*ALOX5AP*, *HLA-DMB*, *HLA-DRA*, and *SPINT2*) as prognostic markers for osteosarcoma metastasis. We adapted WGCNA to identify the five screened genes from SE-associated genes that were crucial hub genes of osteosarcoma. Prognostic models based on some cancer-related genes (Long et al., 2018; Zuo et al., 2019), long non-coding RNAs (Liu et al., 2020; Yang et al., 2020; Zhang et al., 2020), or other biomarkers (Park et al., 2017; Elwood et al., 2018) for improving patient prognosis attract considerable attention. For instance, Zhu et al. (2020) constructed a seven-gene signature associated with osteosarcoma energy metabolism for predicting the prognosis of osteosarcoma through WGCNA and Lasso Cox regression analysis. Dong et al. (2019) constructed a risk score model based on eight genes for predicting the metastasis of osteosarcoma mainly through Lasso logistic regression analysis. Thus, we attempted to perform WGCNA for osteosarcoma and screen hub genes related to clinical traits of this disease. Nomograms are interactive tools developed in recent years, which have been widely used to predict the probability of survival (Wang et al., 2018; Yap et al., 2018; Pan et al., 2019; Xiao et al., 2019; Zhang et al., 2019). Similarly, we constructed a prognostic model and a visualized nomogram based on this model to improve the prognosis of patients with osteosarcoma.

In this study, Lasso penalized Cox regression analysis was initially performed using 349 selected SE-associated genes. A gene cluster containing five SE-associated genes, namely, *AMN1*, *LIMS1*, *SAMD4A*, *SPARC*, and *ZP3*, was screened. Subsequently, a risk score model based on these five genes was constructed and verified using training and validation datasets. Meanwhile, univariate and multivariate Cox regression analyses and ROC curve analysis were also used to assess the predictive ability of the model. We further screened module genes related to clinical traits of patients with osteosarcoma via WGCNA. The five SE-associated genes were also part of the module genes linked to survival and recurrence of osteosarcoma. Finally, an interactive nomogram was constructed based on the risk score model including the five-gene risk group and clinical traits. The five-SE-associated-gene-based prognostic model developed in this study will provide promising inspiration for the clinical therapy of patients with osteosarcoma.

MATERIALS AND METHODS

Data Collection and Expression Matrix Preparation

We downloaded the osteosarcoma dataset GSE39058 as training data from the National Center for Biotechnology Information GEO¹. The 20,819 genes in 42 samples of osteosarcoma were annotated according to the probe information from Illumina HumanHT-12 WG-DASL V4.0 R2 Expression BeadChip (GPL14951) (Kelly et al., 2013). We processed the raw data of the expression matrix by transformation of log₂- via the R package “limma.” The clinical information was organized into a matrix and is shown in **Table 1**. One sample (GSM954821) was excluded due to the lack of information on the survival status. The validation dataset Target-osteosarcoma (Target-OS) was downloaded from the publicly available database of TCGA². The 434 SE-associated genes from the osteosarcoma cell line U2OS were obtained from the website of the SE database (version 1.03³). Following the intersection of SE-associated genes and 20,819 genes from GSE39058, 349 SE-associated genes were used for the subsequent analysis. Chromatin immunoprecipitation (ChIP) signals of H3K27ac-seq, H3K4me1-seq, and H3K4me3-seq in U2OS were also downloaded from Cistrome⁴ and visualized using the IGV (Thorvaldsdottir et al., 2013).

Lasso Penalized Cox Proportional Hazard Regression Model

Lasso penalized Cox regression analysis was performed to screen important prognostic genes for predicting the OS of patients with osteosarcoma using R package “glmnet” (Friedman et al., 2010; Simon et al., 2011; Tibshirani et al., 2012). The shrinking penalty was identified based on the tuning parameter lambda (λ), which was identified based on 10-fold cross-validation. Two optimal λ values (λ_{\min} and λ_{1se}) were highlighted by the vertical lines with a minimizing mean-square error. To determine the optimal λ value, we performed the Wilcoxon test and ROC curve analysis for the comparison of λ_{\min} and λ_{1se} values. These two λ values

were used to construct a new fitted Lasso model, and two lists of genes were generated.

Cox Proportional Hazard Regression Model

Univariate and multivariate Cox proportional hazard regression analyses were performed to validate the associations between the expression levels of selected genes and OS of the patients via the R package “survival” and “survminer.” The regression coefficient (β -value) and HR were predicted. Harrell’s C-index, likelihood ratio test, and Wald test were also calculated, and the results are presented in the forest plot. Survival analysis of the single gene in the prognostic model was performed using the K–M survival curve and log-rank test via the R package “survival” and “survminer.” Time-dependent ROC curve analysis was calculated to evaluate the predictive capability of OS of the prognostic model via the R package “timeROC.”

Polygenic Risk Score Model for Prognostic Prediction

To further investigate the prognostic model based on the gene group, a polygenic risk score model was constructed used the following formula $RiskScore = \sum \beta_i * X_i$. X_i is the level of gene expression and β_i is the regression coefficient. The risk score model was defined as the linear combination of the expression levels of selected genes. The patients with osteosarcoma in the two datasets were divided into high- and low-risk groups according to the risk score. We verified the power of the polygenic risk score model through single-gene expression status analysis, event distribution analysis, K–M survival analysis, and time-dependent ROC curve analysis.

Predictive Nomogram for Prognostic Prediction

A nomogram based on independent prognostic factors of clinical traits and the polygenic risk score was constructed to predict the probability of 1-, 3-, and 5-year OS of patients with osteosarcoma. Subsequently, the discrimination of the nomogram was verified using the C-index obtained through a bootstrap method with 1,000 resamples. Calibration curves were plotted to compare the nomogram-predicted and ideal probabilities against the observed rates. Total points were calculated according to the parameters of prognostic factors using the R package “nomogramEx.” Each variable of the nomogram yields points, and their sum represents the total points that a patient receives. Finally, we developed the interactive nomogram for predicting the probability of OS of patients with osteosarcoma using the R package “replot.”

WGCNA

Weighted gene co-expression network analysis was performed to screen SE-associated hub genes linked to clinical traits using the dataset GSE39058. The analysis was performed as previously described (Ma et al., 2019). The optimal soft-thresholding value was estimated based on scale independence and mean connectivity analysis. Clinical traits-related genes were clustered into different modules. Moreover, the correlation between

¹<https://www.ncbi.nlm.nih.gov/geo/>

²<https://cancergenome.nih.gov/>

³<http://www.lipathway.net/sedb>

⁴<http://cistrome.org/db/#/>

TABLE 1 | The clinical information of 41 OS patients from dataset GSE39058.

Variable		Number
Gender	Male	21
	Female	20
Necrosis	Low	14
	High	27
Recurrence	No	21
	Yes	20
Death	No	29
	Yes	12

modules and clinical traits was highlighted in the heatmap. The relationship between the clinically significant MM and GS was displayed in the scatter plots.

Protein–Protein Interaction Analysis and Pathway Enrichment Analysis

Eigengenes within clinically significant modules and SE-associated genes were intersected using the Venn plot and produced a list of hub genes. An interaction network between hub genes and drugs was also constructed to identify some drugs related with these hub genes via “NetworkAnalyst 3.0” (Zhou et al., 2019). Subsequently, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was performed using these hub genes. The top 10 items of the enriched pathway were selected and shown in the dot plot.

RESULTS

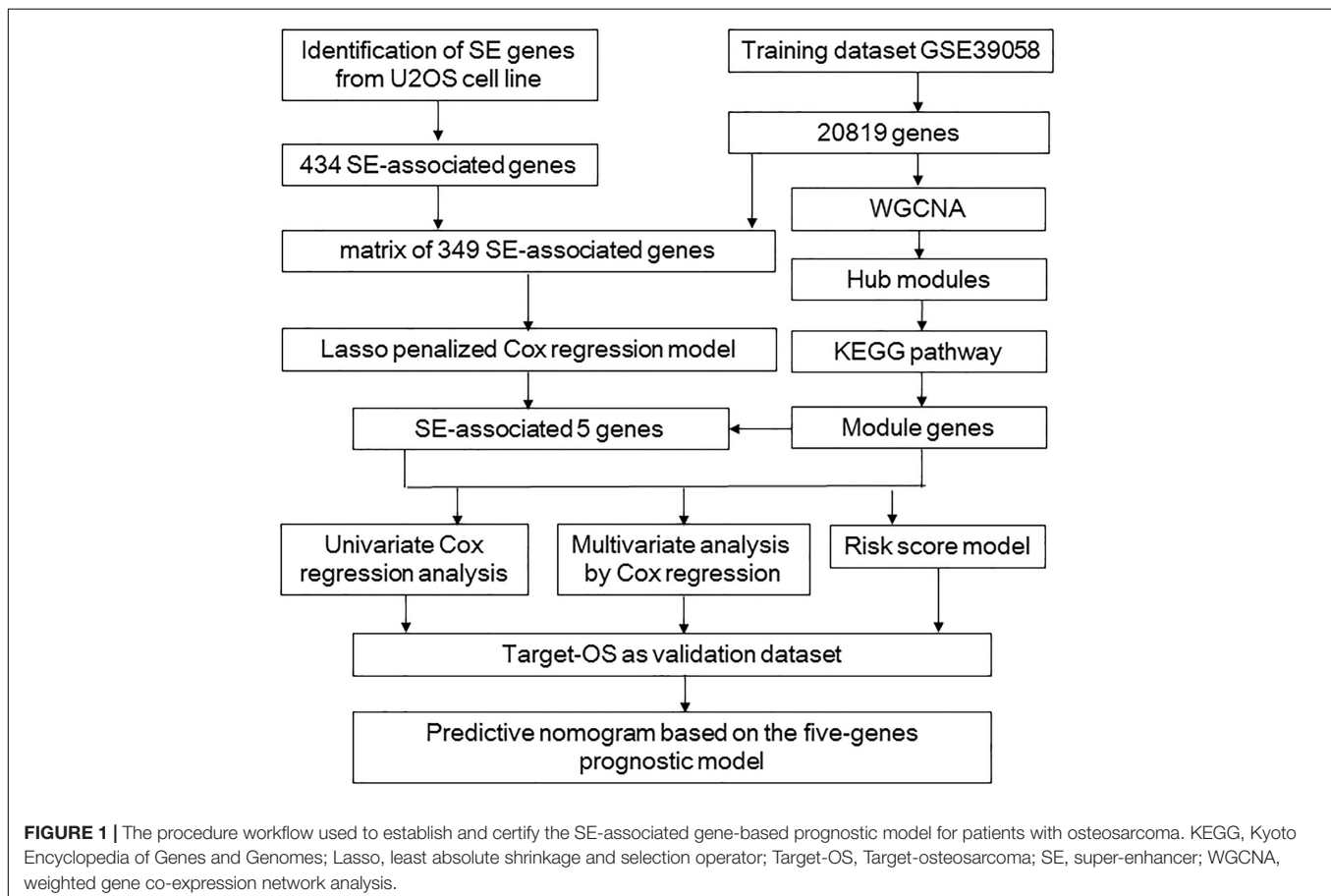
Establishment and Validation of the Lasso Penalized Cox Regression Model

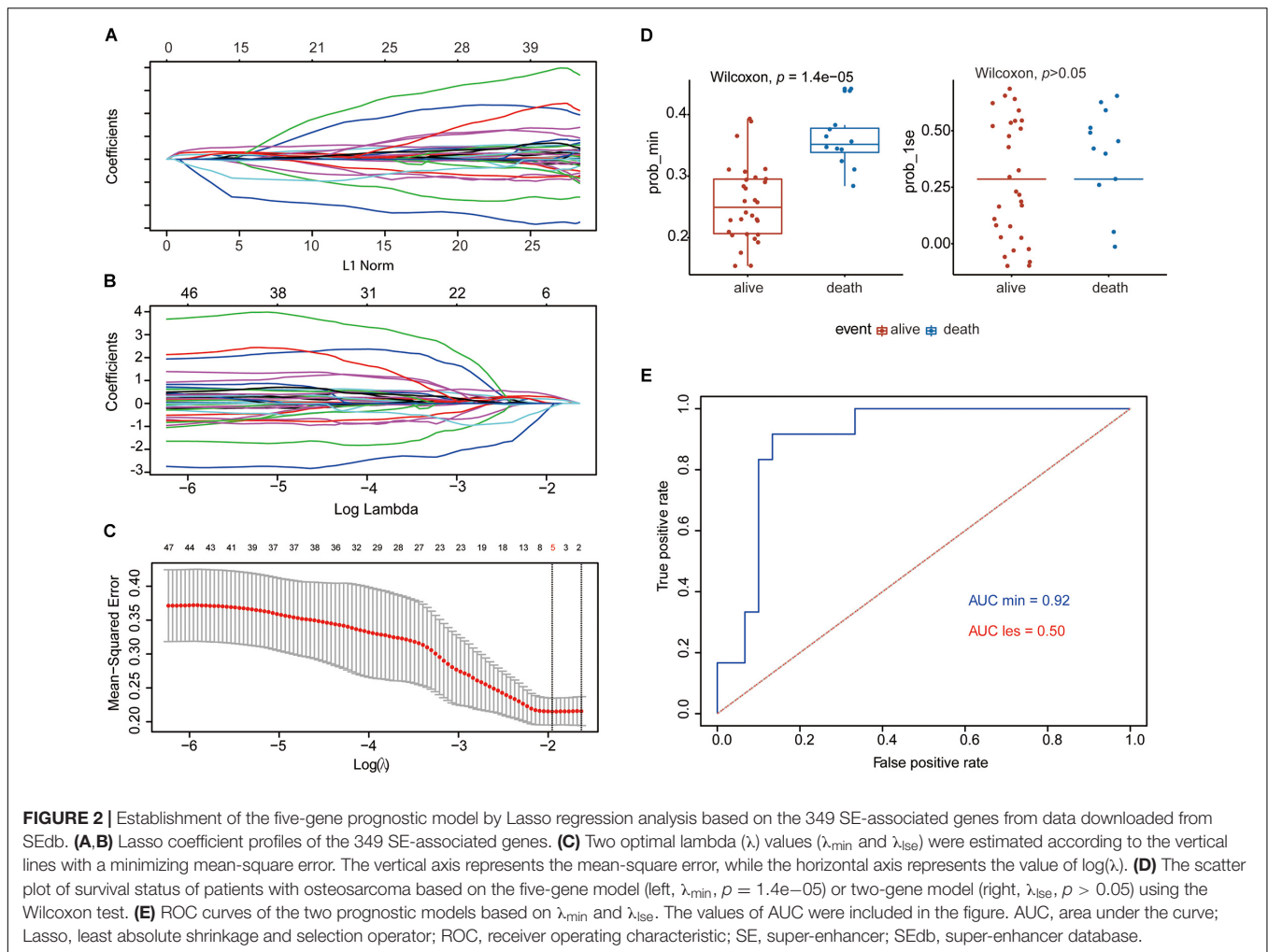
The flowchart shows our analysis procedures for the construction of the SE-associated gene-based prognostic model of osteosarcoma (Figure 1). The 349 SE-associated genes based on the dataset of GSE39058 were selected to perform the

Lasso penalized Cox regression analysis. Firstly, to identify the best-fit parameter of the λ parameter, two important optional λ values ($\log\lambda_{\min} = -1.91$ and $\log\lambda_{\text{lse}} = -1.48$) were calculated from the vertical lines with a minimizing mean-square error and further used to select two groups of genes (Figures 2A–C). Moreover, Lasso models were reconstructed according to the λ_{\min} and λ_{lse} , and survival probabilities were further estimated based on two gene lists (Figure 2D). As shown in Figure 2D, the use of the five gene-based prognostic model based on the λ_{\min} (Wilcoxon test, $p = 1.4e-05$) facilitated the obvious distinction of survival probabilities between samples obtained from living and expired patients. However, when a two-gene-based prognostic model was used based on the λ_{lse} , there was no statistically significant difference observed (Wilcoxon test, $p > 0.05$). Similarly, the outcome of ROC curve analysis showed that the area under the curve minimum (AUC_{\min}) was 0.92, implying that the five-gene-based Lasso model performed well in predicting the probability of OS of patients with osteosarcoma (Figure 2E).

Validation of Independent Prognostic Factors by the Cox Regression Model

For the validation of the Lasso penalized Cox regression model, we performed univariate and multivariate Cox regression analyses to determine whether these genes were independent





prognostic factors for the OS of patients with osteosarcoma. In the univariate Cox regression analysis, all log-rank p -values of these five genes were <0.05 (Figure 3A). Following multivariate Cox regression analysis, the global p -value (log-rank test) of the five-gene prognostic model was only 0.000171 (Figure 3B). The AIC was 59.74, and the C-index was 0.89, indicating that these five genes may be favorable prognostic factors for the OS of patients with osteosarcoma. Moreover, the outcome of K-M survival analysis was consistent with that of the univariate Cox regression analysis (Figure 3C). Furthermore, *AMN1* and *ZP3* may be protective factors in osteosarcoma (HR: 1.26e-36 and 1.13e-07, respectively), whereas *LIMS1*, *SAMD4A*, and *SPARC* appear to be harmful factors in this setting (HR: 699.2, 167, and 298.7, respectively). Thus, a Lasso penalized Cox regression model including five SE-associated genes may be used to predict the OS of patients with osteosarcoma.

Establishment and Validation of the Polygenic Risk Score Model

We integrated the expression data of the five genes and corresponding coefficient derived from the above multivariate

regression analysis to establish the risk score model. All patients in the training dataset of GSE39058 ($N = 41$) were divided into high-risk (risk score > 0) and low-risk groups (risk score < 0) (Figure 4A). As shown in Figure 4B, survival was more commonly observed in the low-risk group, whereas death was more frequent in the high-risk group. *LIMS1*, *SAMD4A*, and *SPARC* tended to upregulated, whereas *AMN1* and *ZP3* were downregulated in patients of the high-risk group (Figure 4C). The K-M OS analysis predicted an inferior outcome of OS in the high-risk group (log-rank test, $p = 0.0006$) (Figure 4D). Meanwhile, the AUCs of a time-dependent ROC curve calculated using the five-gene-based risk score model were >0.8 (Figure 4E), indicating that the forecast model had high sensitivity and specificity.

The Target-OS dataset was further used to verify the predictive values of the polygenic risk score. Following the multivariate Cox regression analysis, the global p -value (log-rank test) of the five-gene prognostic model for the Target-OS dataset was <0.05 (Supplementary Figure S1A). A total of 88 patients were classified into the low- and high-risk groups using the optimal cutoff value of the risk score (Supplementary Figures S1B-D). The K-M curves of two groups based on risk scores were

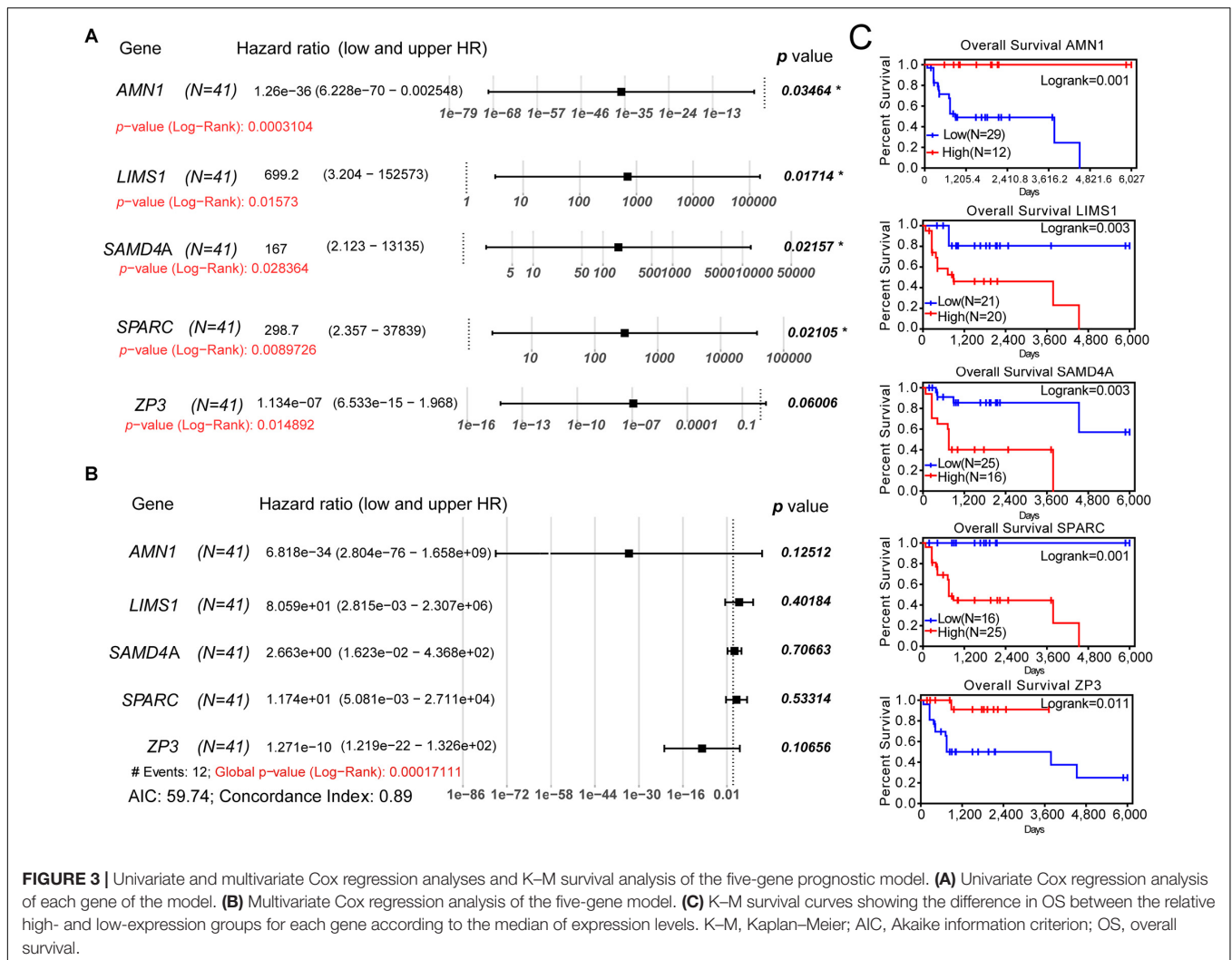


FIGURE 3 | Univariate and multivariate Cox regression analyses and K-M survival analysis of the five-gene prognostic model. **(A)** Univariate Cox regression analysis of each gene of the model. **(B)** Multivariate Cox regression analysis of the five-gene model. **(C)** K-M survival curves showing the difference in OS between the relative high- and low-expression groups for each gene according to the median of expression levels. K-M, Kaplan-Meier; AIC, Akaike information criterion; OS, overall survival.

significantly different (log-rank test, $p < 0.05$), and AUCs of time-dependent ROC curves also showed that the five-gene risk model could be a favorable approach to predicting the OS of patients with osteosarcoma (Supplementary Figures S1E,F).

Construction of the Predictive Nomogram Based on the Risk Score Model for Prognostic Prediction

To visualize the Cox regression model, a nomogram was constructed for predicting the 1-, 3-, and 5-year OS probability for patients with osteosarcoma in the GSE39058 dataset ($N = 41$). The predictors of the nomogram included age, sex, necrosis, recurrence, and risk group (Figure 5A). The calibration plot shown in Figure 5B illustrates the performance of the nomogram. The calibration plots were close to the ideal prediction gray line (45° line), indicating that our nomogram performed well in predicting survival probability (Figure 5B). To assess the predictive effect of the risk group as an indicator of OS, we applied the nomogram to a specific patient (GSM954850) in GSE39058 (Figures 5C,D). The patient GSM954850 expired at

750 days. According to the model containing the risk group, the predicted probability of death at 1,095 days was 0.815; this value was markedly higher than that obtained for the model lacking the risk group (0.659). This finding implies that the predictive model containing the risk group as a parameter is more accurate than that lacking the risk group.

Identification of SE-Related Hub Genes in Osteosarcoma Using WGCNA and a Protein-Protein Interaction Network

Weighted gene co-expression network analysis is another bioinformatics method for the analysis of clinical traits linked to the levels of gene expression. Cluster analysis was performed to display the heatmap of clinical traits (Figure 6A). As shown in Figure 6B, $\beta = 9$ was selected as the best soft-thresholding value via prediction of the scale independence and mean connectivity (Figure 6B) to construct the gene co-expression network. Subsequently, eigengenes were divided into different colored modules, producing 17 different modules. A module-trait relationship heatmap was developed according to Pearson's

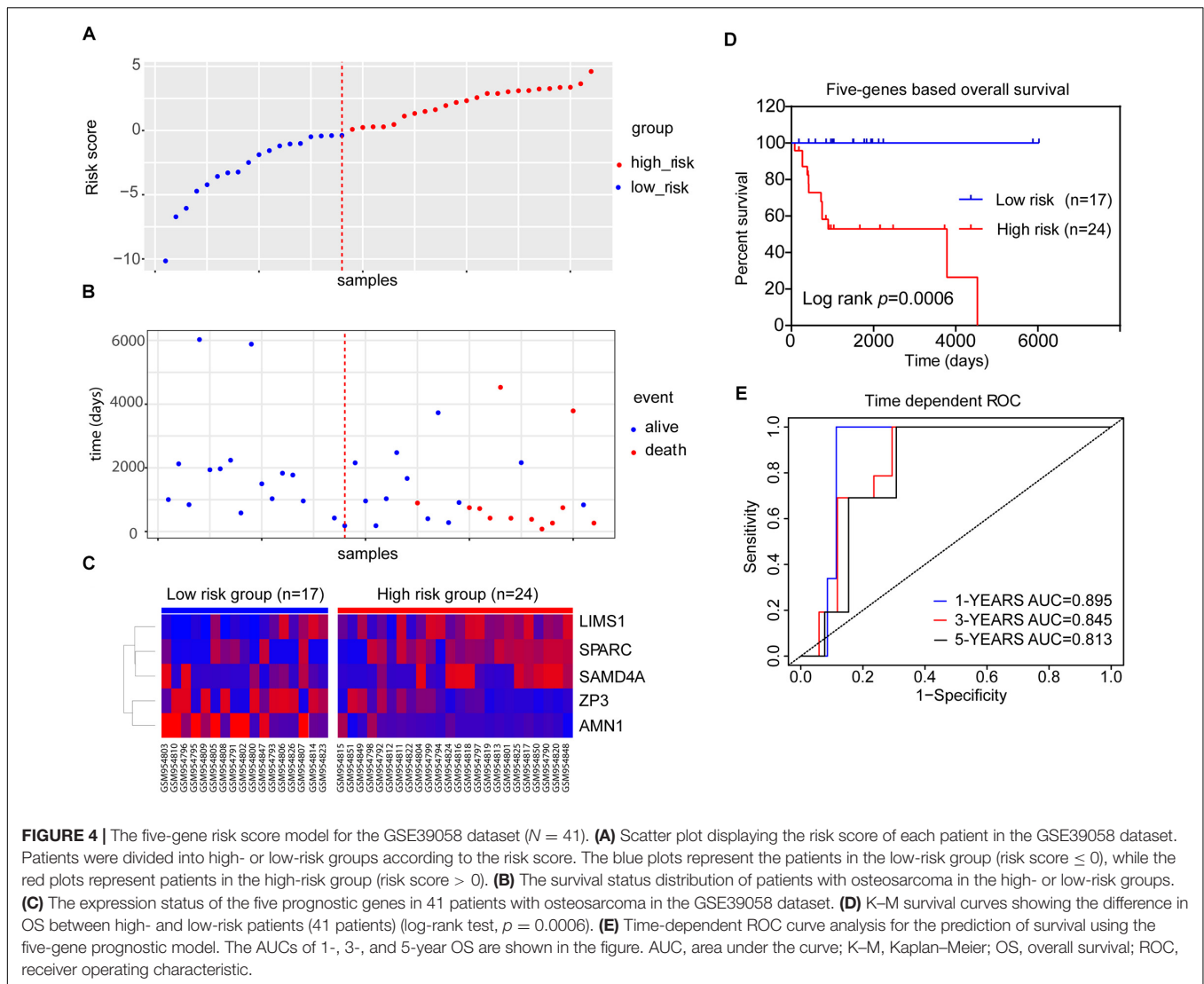


FIGURE 4 | The five-gene risk score model for the GSE39058 dataset ($N = 41$). **(A)** Scatter plot displaying the risk score of each patient in the GSE39058 dataset. Patients were divided into high- or low-risk groups according to the risk score. The blue plots represent the patients in the low-risk group (risk score ≤ 0), while the red plots represent patients in the high-risk group (risk score > 0). **(B)** The survival status distribution of patients with osteosarcoma in the GSE39058 dataset. **(C)** The expression status of the five prognostic genes in 41 patients with osteosarcoma in the GSE39058 dataset. **(D)** K–M survival curves showing the difference in OS between high- and low-risk patients (41 patients) (log-rank test, $p = 0.0006$). **(E)** Time-dependent ROC curve analysis for the prediction of survival using the five-gene prognostic model. The AUCs of 1-, 3-, and 5-year OS are shown in the figure. AUC, area under the curve; K–M, Kaplan–Meier; OS, overall survival; ROC, receiver operating characteristic.

correlation coefficient (**Figure 6C**). Three modules presented higher correlation with clinical traits: MEgray60 with recurrence ($r = 0.37, p = 0.01$); MEblue with death ($r = 0.33, p = 0.03$); and MEpink with survival time ($r = 0.47, p = 0.002$) (**Figure 6C**). A scatter plot also showed that the MM of these three modules was positively correlated with GS (**Figure 6D**). A Venn plot of SE-associated genes derived from the SE-associated gene matrix was drawn, and the genes within three modules were intersected (**Figure 6E**). There were 13, 70, and 16 interaction genes between module gray60, blue, pink, and SE-associated genes. The interaction genes were enriched in numerous cancer-related pathways, such as proteoglycans in cancer and transcriptional misregulation in cancer (**Figure 6F**). In addition, we constructed a gene–drug interaction network using the five genes and related drugs (**Supplementary Figure S2**). All five genes interacted with JQ1 (an established SE inhibitor). The network depicted that these five genes were probably regulated by SE to a certain extent. Finally, we displayed the signal tracks for the H3K27ac, H3K4me1, and H3K4me3 ChIP–seq profiles of the five genes

(**Figure 7**). We observed five predicted SEs near these five genes. These data suggest that the expression of the five genes is regulated by SE. In addition, the SE inhibitor JQ1 may regulate the expression pattern in U2OS cells.

DISCUSSION

Notably, there is a lack of effective target treatment regimens in clinical practice, except for the traditional method of surgical resection. Recently, prognostic models based on some disease-related genes or other biomarkers for improving the prognosis of osteosarcoma have attracted considerable attention. The Lasso regression Cox model, gene risk score model, and univariate and multivariate Cox regression models are widely used for the screening and mining of hub genes related to survival and other clinical characteristics. Particularly, the Lasso model can be applied to solve “curse of dimensionality” (small sample size combined with a very large number of genes) when large amounts

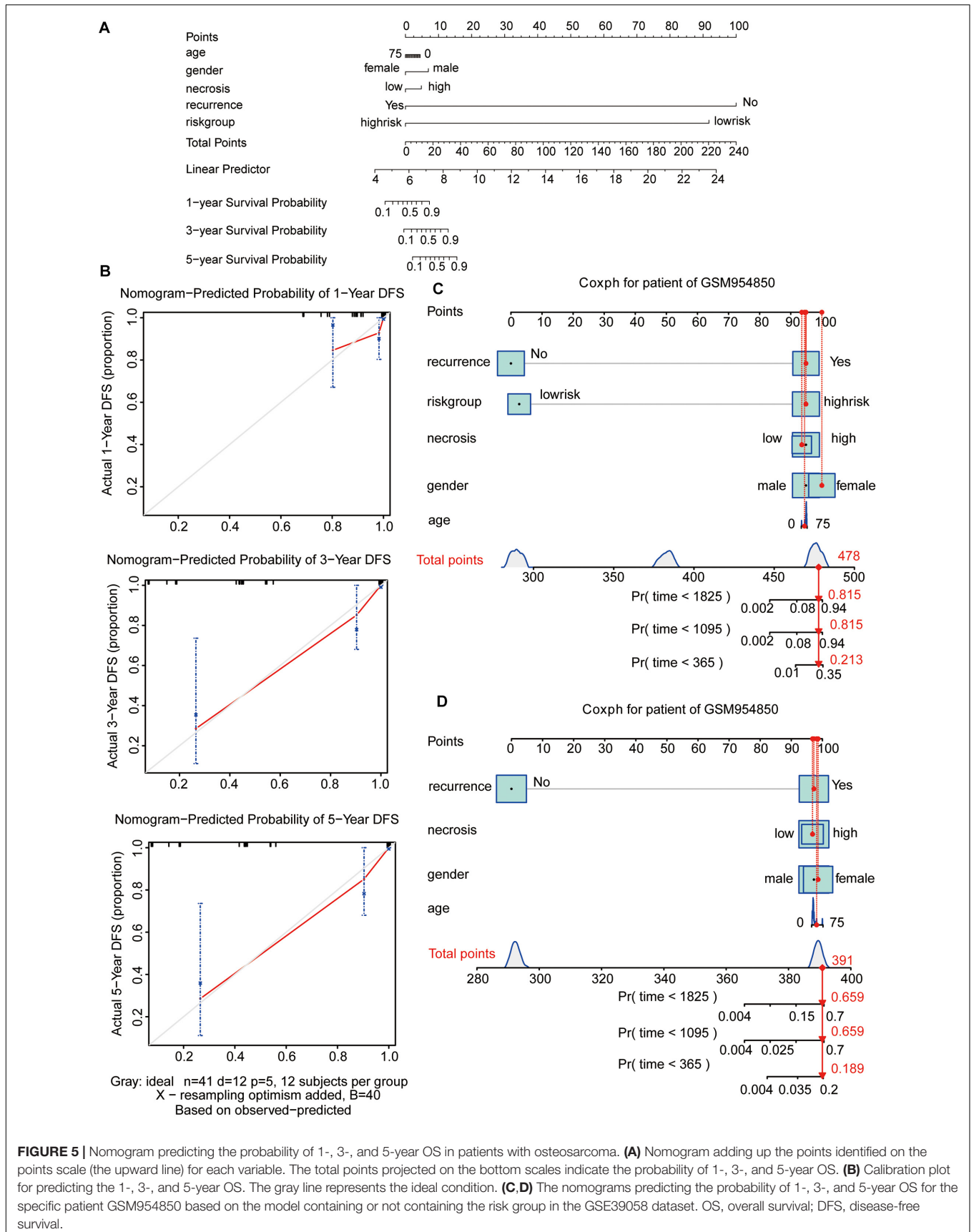
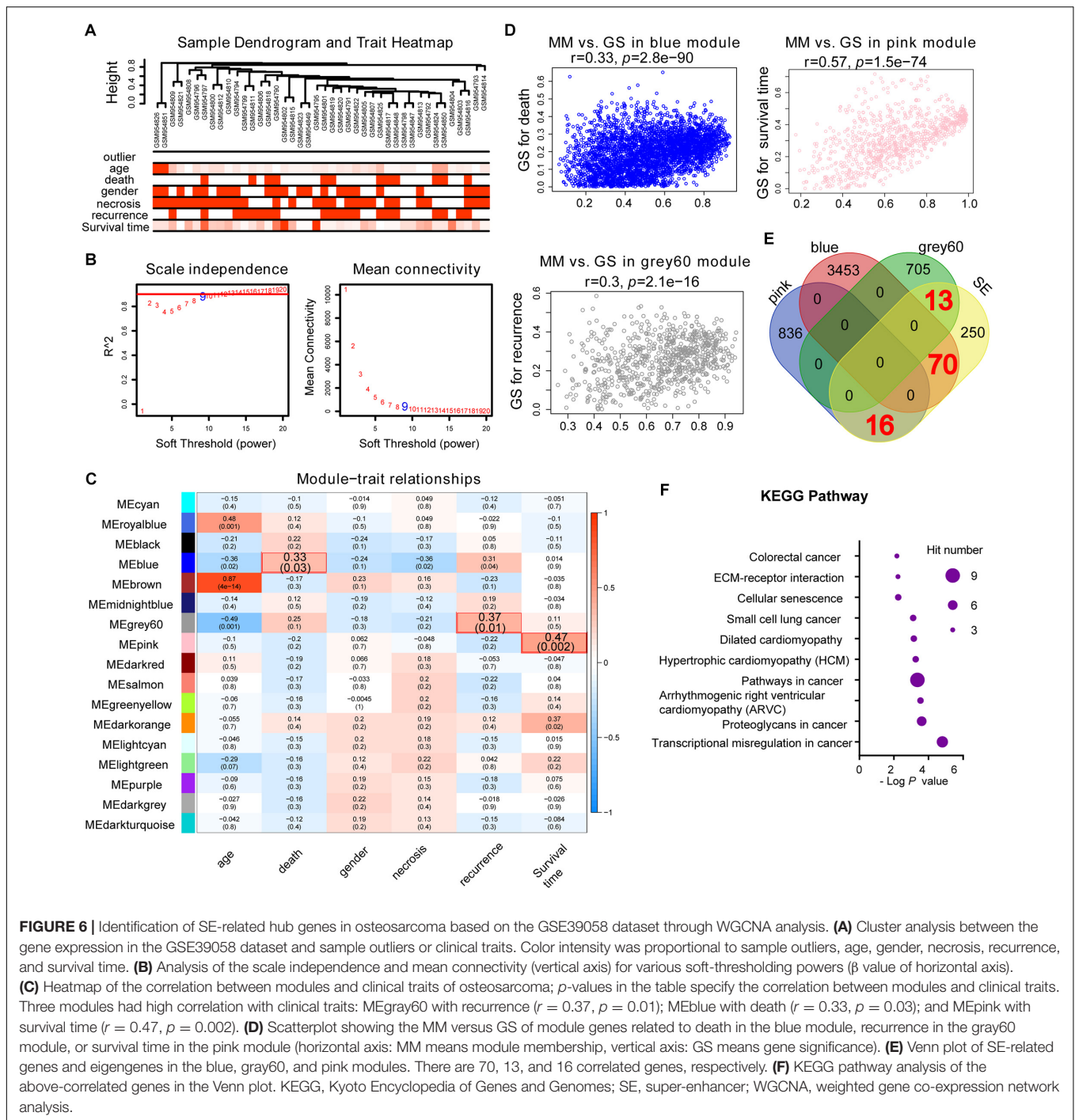


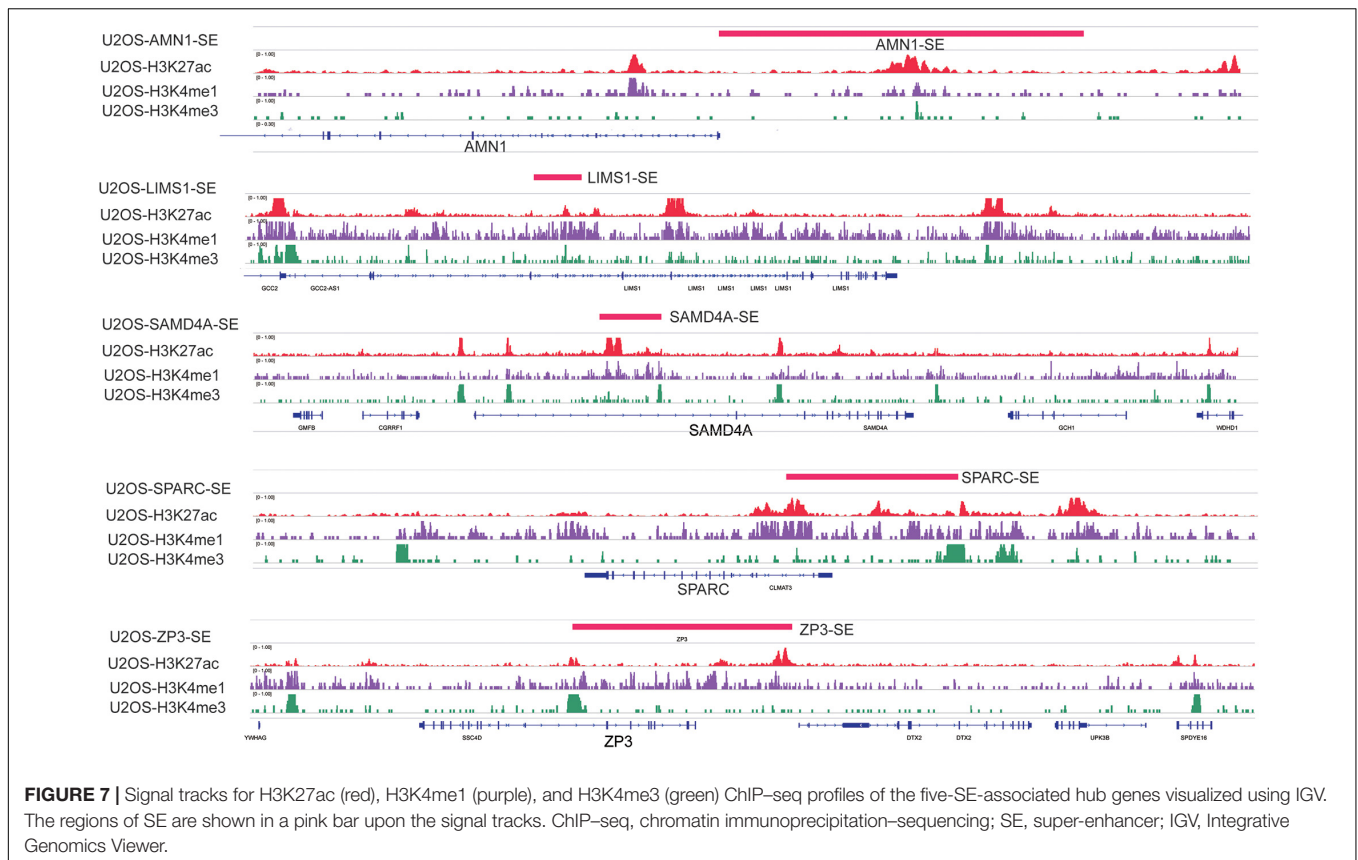
FIGURE 5 | Nomogram predicting the probability of 1-, 3-, and 5-year OS in patients with osteosarcoma. **(A)** Nomogram adding up the points identified on the points scale (the upward line) for each variable. The total points projected on the bottom scales indicate the probability of 1-, 3-, and 5-year OS. **(B)** Calibration plot for predicting the 1-, 3-, and 5-year OS. The gray line represents the ideal condition. **(C,D)** The nomograms predicting the probability of 1-, 3-, and 5-year OS for the specific patient GSM954850 based on the model containing or not containing the risk group in the GSE39058 dataset. OS, overall survival; DFS, disease-free survival.



of throughput biological data are processed. Recently, Lasso and WGCNA were used to construct a nomogram of radiomics of diagnostic computed tomography for predicting the survival of patients with high-grade osteosarcoma, which showed favorable performance (Wu et al., 2018). Another study also identified a seven-gene signature as a predictive biomarker for energy metabolism in osteosarcoma by Lasso and multivariate Cox regression (Zhu et al., 2020). Lasso coefficient profiles of the genes associated with the metastasis of osteosarcoma were established.

These profiles were used to construct a risk score model for the prediction of metastasis of osteosarcoma (Dong et al., 2019). These findings imply that prognostic models based on different clinical traits and gene expression levels may help to potentiate the personalized treatment of patients with osteosarcoma.

In the present study, we constructed a five-SE-associated-gene-based nomogram prognostic model to improve the prognosis of patients with osteosarcoma. We first constructed an SE-associated gene expression matrix based on the 349



genes derived from the intersection of SE-associated genes and 20,819 genes from the GSE39058 dataset. According to this expression matrix, Lasso penalized Cox regression analysis was performed to screen a gene cluster containing *AMN1*, *LIMS1*, *SAMD4A*, *SPARC*, and *ZP3*, which was used to construct the risk score model based on the expression levels of five genes. Considering the group risk score as an independent indicator for prognosis, the K–M curve and log-rank test showed that a high risk score was correlated to augmented death ratio. We assessed the capability of the five-gene prognostic model in the GSE39058 dataset and also in the validation dataset Target-OS via univariate and multivariate Cox regression analyses and ROC curve analysis. The AUCs of the prognostic model for predicting the 1-, 3-, and 5-year OS were 0.895, 0.845, and 0.813, respectively, for the GSE39058 dataset and 0.549, 0.721, and 0.724, respectively, for the Target-OS dataset. The AUCs and C-index also showed that the five-gene prognostic model had a great performance for the prediction of survival and could be a robust and efficient model of prognosis. An interactive nomogram based on the risk score model and other clinical traits was constructed and identified using the calibration plots. To determine the predictive effect of the risk group as an indicator of OS, we also applied the nomogram to a specific patient (GSM954850) in the GSE39058 dataset. The predictive model containing the risk group as a parameter is more accurate compared with the nomogram model lacking the risk group. Through this approach, the

prognosis of patients in the high-risk group may improve following the adjustment of individual treatment regimens for these patients.

The WGCNA of the 20,819 genes from the GSE39058 dataset revealed that the above five SE-associated genes also belonged to module genes. The KEGG analysis showed that the module genes associated with SE were enriched in numerous cancer-related pathways, such as proteoglycans in cancer and transcriptional misregulation in cancer. *LIMS1* functions as an oncogene to promote the survival of pancreatic cancer cells under oxygen–glucose deprivation conditions (Huang et al., 2019). The protein *LIMS1* (also known as *PINCH1*) encoded by *LIMS1* acts as an adaptor protein which contains five LIM domains or double zinc fingers (Hobert et al., 1999). *LIMS1* can be involved in integrin signaling through interacting with integrin-linked kinase mediated by its LIM domain. Besides, *LIMS1* can bind to integrin-linked kinase and parvin, which forms a protein complex, which is critical for the cell extracellular matrix adhesion (Ito et al., 2010). High expression of *LIMS1* is associated with poor prognosis in human laryngeal carcinomas (Tsinias et al., 2018). Inhibition of *LIMS1* through knockdown of *LIMS1* can inhibit cell proliferation of neuroblastoma cells (Saeki et al., 2018). Thus, *LIMS1* may be considered as a biomarker for osteosarcoma. Secreted protein acidic and rich in cysteine (*SPARC*) encodes a cysteine-rich acidic matrix-associated protein, which is essential for extracellular matrix synthesis and epithelial–mesenchymal transition that

promotes migration and invasion of many cancers (Sun et al., 2018; Jiang et al., 2019; López-Moncada et al., 2019). In particular, overexpression of *SPARC*, which was critical for the growth and maintenance of osteosarcoma, was found in 51 of 55 osteosarcoma tumor samples (Dalla-Torre et al., 2006). Overexpression of *SPARC* also mediated the suppressing effect of *TP53INP1* on the migration of pancreatic cancer cells and promoted the progression of malignant cancers (Seux et al., 2011). Therefore, it is convincing that *SPARC* can serve as a biomarker or treatment target for osteosarcoma. *SAMD4A*, encoded by the sterile alpha motif domain containing 4A (*SAMD4A*), was reported to involve in processes such as maternal RNA destabilization, translational repression, and early embryo development in the role of a posttranscriptional regulator (Smibert et al., 1996; Baez et al., 2011). Although studies about the function of *SAMD4A* in cancer are still rare, we believe that *SAMD4A* will become a novel biomarker for predicting the prognosis of osteosarcoma based on the findings of this study. Antagonist of the mitotic exit network 1 (*AMN1*) is essential for mitotic checkpoints and chromosome stability; the protein encoded by *AMN1* is involved in the progression of cell cycle (Wang et al., 2003; Xie et al., 2019). Although there are no relative studies about the role of *ZP3* in osteosarcoma, even if in cancer, combined with the findings of our study, we should also pay attention to *AMN1* and *ZP3* in osteosarcoma in a certain extent. In the present study, the gene–drug interaction network of the five genes showed that all genes interacted with JQ1 (an established SE inhibitor). Moreover, according to the ChIP–seq signal of H3K27ac, H3K4me1, and H3K4me3, we detected strong signals of SEs around these five genes. These data support that the expression of the five genes may be regulated by SE, while JQ1 as an SE inhibitor probably regulated the expression pattern in U2OS cells. Further molecular mechanism studies are warranted to recover the relationship between expression and SE in the future.

In conclusion, this was the first study that used the Lasso model to screen prognostic indicators from the profile of SE-associated genes. The five-gene-based interactive nomogram may be used as a clinical tool for improving the prognosis of patients with osteosarcoma, offering novel insights into personalized therapy in this setting.

REFERENCES

- Baez, M. V., Luchelli, L., Maschi, D., Habif, M., Pascual, M., Thomas, M. G., et al. (2011). Smaug1 mRNA-silencing foci respond to NMDA and modulate synapse formation. *J. Cell. Biol.* 195, 1141–1157. doi: 10.1083/jcb.201108159
- Bielack, S., Carrle, D., and Casali, P. G. (2009). Osteosarcoma: ESMO clinical recommendations for diagnosis, treatment and follow-up. *Ann. Oncol.* 20(Suppl. 4), 137–139. doi: 10.1093/annonc/mdp154
- Chen, D., Zhao, Z., Huang, Z., Chen, D.-C., Zhu, X.-X., Wang, Y.-Z., et al. (2018). Super enhancer inhibitors suppress MYC driven transcriptional amplification and tumor progression in osteosarcoma. *Bone Res.* 6:11.
- Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., et al. (2010). Histone H3K27ac separates active from poised enhancers and

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

JQ and QQ conceived and designed the theme of this study. ZO drafted the manuscript. TQ, GL, HZ, and JW provided some helpful advices in several statistical methods. QQ revised the manuscript. All authors involved in this study critically revised the manuscript and approved the final version with one accord.

FUNDING

This work was supported by grants awarded to QL by the Hunan Provincial Department of Finance Grant (Nos. 2019-93 and 2018-92) and to CT by the National Scientific foundation of China (No. 81902745).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2020.598660/full#supplementary-material>

Supplementary Figure 1 | Multivariate Cox regression analysis and risk score model analysis of the validation dataset Target-OS ($N = 88$). **(A)** Multivariate Cox regression analysis of dataset Target-OS based on the five-gene model. **(B)** Scatterplot displaying the risk score of each patient in the Target-OS dataset. The patients were divided into the high- or low-risk groups according to the risk score. The blue plots represent the patients in the low-risk group (risk score ≤ 0), while the red plots represent those in the high-risk group (risk score > 0). **(C)** Survival status distribution of the high- or low-risk groups in the Target-OS dataset. **(D)** The expression status of the five prognostic genes for patients in the Target-OS dataset. **(E)** K–M survival curves showing the difference in OS between high- and low-risk patients in the Target-OS dataset (log-rank test, $p = 0.0438$). **(F)** Time-dependent ROC curves analysis of survival prediction for patients in the Target-OS dataset using the five-gene prognostic model. The AUCs for 1-, 3-, and 5-year OS are shown in the figure. AUC, area under curve; K–M, Kaplan–Meier; ROC, receiver operating characteristic; Target-OS, Target-osteosarcoma, OS, overall survival; AIC, Akaike information criterion.

Supplementary Figure 2 | Gene–drug interaction network analysis of the five genes and drugs.

- predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.* 107, 21931–21936. doi: 10.1073/pnas.1016071107
- Dalla-Torre, C. A., Yoshimoto, M., Lee, C. H., Joshua, A. M., de Toledo, S. R., Petrilli, A. S., et al. (2006). Effects of THBS3, SPARC and SPP1 expression on biological behavior and survival in patients with osteosarcoma. *BMC Cancer* 6:237. doi: 10.1186/1471-2407-6-237
- Dong, S., Huo, H., Mao, Y., Li, X., and Dong, L. (2019). A risk score model for the prediction of osteosarcoma metastasis. *FEBS Open Biol.* 9, 519–526. doi: 10.1002/2211-5463.12592
- Elwood, J. M., Tawfiq, E., TinTin, S., Marshall, R. J., Phung, T. M., Campbell, I., et al. (2018). Development and validation of a new predictive model for breast cancer survival in New Zealand and comparison to the Nottingham prognostic index. *BMC Cancer* 18:897. doi: 10.1186/s12885-018-4791-x

- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33, 1–22.
- Guan, X., Guan, Z., and Song, C. (2020). Expression profile analysis identifies key genes as prognostic markers for metastasis of osteosarcoma. *Cancer Cell. Int.* 20:104.
- Hnisz, D., Abraham, B. J., Lee, T. I., Lau, A., Saint-André, V., Sigova, A. A., et al. (2013). Super-enhancers in the control of cell identity and disease. *Cell* 155, 934–947. doi: 10.1016/j.cell.2013.09.053
- Hnisz, D., Schuijers, J., Lin, C. Y., Weintraub, A. S., Abraham, B. J., Lee, T. I., et al. (2015). Convergence of developmental and oncogenic signaling pathways at transcriptional super-enhancers. *Mol. Cell.* 58, 362–370. doi: 10.1016/j.molcel.2015.02.014
- Hobert, O., Moerman, D. G., Clark, K. A., Beckerle, M. C., and Ruvkun, G. (1999). A conserved LIM protein that affects muscular adherens junction integrity and mechanosensory function in *Caenorhabditis elegans*. *J. Cell. Biol.* 144, 45–57. doi: 10.1083/jcb.144.1.45
- Huang, C., Li, Y., Li, Z., Xu, Y., Li, N., Ge, Y., et al. (2019). LIMS1 promotes pancreatic cancer cell survival under oxygen-glucose deprivation conditions by enhancing HIF1A protein translation. *Clin. Cancer Res.* 25, 4091–4103. doi: 10.1158/1078-0432.ccr-18-3533
- Ito, S., Takahara, Y., Hyodo, T., Hasegawa, H., Asano, E., Hamaguchi, M., et al. (2010). The roles of two distinct regions of PINCH-1 in the regulation of cell attachment and spreading. *Mol. Biol. Cell.* 21, 4120–4129. doi: 10.1091/mbc.E10-05-0459
- Jiang, X., Liu, F., Wang, Y., and Gao, J. (2019). Secreted protein acidic and rich in cysteine promotes epithelial-mesenchymal transition of hepatocellular carcinoma cells and acquisition of cancerstem cell phenotypes. *J. Gastroenterol. Hepatol.* 34, 1860–1868. doi: 10.1111/jgh.14692
- Kelly, A. D., Haibe-Kains, B., Janeway, K. A., Hill, K. E., Howe, E., Goldsmith, J., et al. (2013). MicroRNA paraffin-based studies in osteosarcoma reveal reproducible independent prognostic profiles at 14q32. *Genome Med.* 5:2. doi: 10.1186/gm406
- Kempf-Bielack, B., Bielack, S. S., Jurgens, H., Branscheid, D., Berdel, W. E., Exner, G. U., et al. (2005). Osteosarcoma relapse after combined modality therapy: an analysis of unselected patients in the cooperative osteosarcoma study group (COSS). *J. Clin. Oncol.* 23, 559–568. doi: 10.1200/jco.2005.04.063
- Lee, D. H., Qi, J., Bradner, J. E., Said, J. W., Doan, N. B., Forscher, C., et al. (2015). Synergistic effect of JQ1 and rapamycin for treatment of human osteosarcoma. *Int. J. Cancer* 136, 2055–2064. doi: 10.1002/ijc.29269
- Liu, Y., Gou, X., Wei, Z., Yu, H., Zhou, X., and Li, X. (2020). Bioinformatics profiling integrating a four immune-related long non-coding RNAs signature as a prognostic model for papillary renal cell carcinoma. *Aging* 12, 15359–15373. doi: 10.18632/aging.103580
- Long, J., Zhang, L., Wan, X., Lin, J., Bai, Y., Xu, W., et al. (2018). A four-gene-based prognostic model predicts overall survival in patients with hepatocellular carcinoma. *J. Cell. Mol. Med.* 22, 5928–5938. doi: 10.1111/jcmm.13863
- López-Moncada, F., Torres, M. J., Castellón, E. A., and Contreras, H. R. (2019). Secreted protein acidic and rich in cysteine (SPARC) induces epithelial-mesenchymal transition, enhancing migration and invasion, and is associated with high Gleason score in prostate cancer. *Asian J. Androl.* 21, 557–564. doi: 10.4103/aja.aja_23_19
- Lovén, J., Hoke, H. A., Lin, C. Y., Lau, A., Orlando, D. A., Vakoc, C. R., et al. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* 153, 320–334. doi: 10.1016/j.cell.2013.03.036
- Lu, B., He, Y., He, J., Wang, L., Liu, Z., Yang, J., et al. (2020). Epigenetic profiling identifies LIF as a super-enhancer-controlled regulator of stem cell-like properties in osteosarcoma. *Mol. Cancer Res.* 18, 57–67. doi: 10.1158/1541-7786.mcr-19-0470
- Lu, L., Wu, Y., Feng, M., Xue, X., and Fan, Y. (2019). A novel seven-miRNA prognostic model to predict overall survival in head and neck squamous cell carcinoma patients. *Mol. Med. Rep.* 20, 4340–4348. doi: 10.3892/mmr.2019.10665
- Luetke, A., Meyers, P. A., Lewis, I., and Juergens, H. (2014). Osteosarcoma treatment - where do we stand? A state of the art review. *Cancer Treat. Rev.* 40, 523–532. doi: 10.1016/j.ctrv.2013.11.006
- Ma, H., Qu, J., Luo, J., Qi, T., Tan, H., Jiang, Z., et al. (2019). Super-enhancer-associated hub genes in chronic myeloid leukemia identified using weighted gene co-expression network analysis. *Cancer Manag. Res.* 11, 10705–10718. doi: 10.2147/CMAR.S214614
- Meazza, C., and Scanagatta, P. (2016). Metastatic osteosarcoma: a challenging multidisciplinary treatment. *Expert Rev. Anticancer Ther.* 16, 543–556. doi: 10.1586/14737140.2016.1168697
- Pan, X., Yang, W., Chen, Y., Tong, L., Li, C., and Li, H. (2019). Nomogram for predicting the overall survival of patients with inflammatory breast cancer: a SEER-based study. *Breast* 47, 56–61. doi: 10.1016/j.breast.2019.05.015
- Park, H. S., Park, J. S., Chun, Y. J., Roh, Y. H., Moon, J., Chon, H. J., et al. (2017). Prognostic factors and scoring model for survival in metastatic biliary tract cancer. *Cancer Res. Treat.* 49, 1127–1139. doi: 10.4143/crt.2016.538
- Saeki, N., Saito, A., Sugaya, Y., Amemiya, M., Ono, H., Komatsuzaki, R., et al. (2018). Chromatin immunoprecipitation and DNA sequencing identified a LIMS1/ILK pathway regulated by LMO1 in neuroblastoma. *Cancer Genomics Proteomics* 15, 165–174. doi: 10.21873/cgp.20074
- Seux, M., Peugeot, S., Montero, M. P., Siret, C., Rigot, V., Clerc, P., et al. (2011). TP53INP1 decreases pancreatic cancer cell migration by regulating SPARC expression. *Oncogene* 30, 3049–3061. doi: 10.1038/onc.2011.25
- Simon, N., Friedman, J., Hastie, T., and Tibshirani, R. (2011). Regularization paths for Cox's proportional hazards model via coordinate descent. *J. Stat. Softw.* 39, 1–13. doi: 10.18637/jss.v039.i05
- Smibert, C. A., Wilson, J. E., Kerr, K., and Macdonald, P. M. (1996). smaug protein represses translation of unlocalized nanos mRNA in the *Drosophila embryo*. *Genes Dev.* 10, 2600–2609. doi: 10.1101/gad.10.20.2600
- Sun, W., Feng, J., Yi, Q., Xu, X., Chen, Y., and Tang, L. (2018). SPARC acts as a mediator of TGF- β 1 in promoting epithelial-to-mesenchymal transition in A549 and H1299 lung cancer cells. *Biofactors* 44, 453–464. doi: 10.1002/biof.1442
- Thorvaldsdottir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* 14, 178–192. doi: 10.1093/bib/bbs017
- Tian, H., Guan, D., and Li, J. (2018). Identifying osteosarcoma metastasis associated genes by weighted gene co-expression network analysis (WGCNA). *Medicine* 97:e10781. doi: 10.1097/md.00000000000010781
- Tibshirani, R. (1997). The lasso method for variable selection in the Cox model. *Stat. Med.* 16, 385–395. doi: 10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3
- Tibshirani, R., Bien, J., Friedman, J., Hastie, T., Simon, N., Taylor, J., et al. (2012). Strong rules for discarding predictors in lasso-type problems. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 74, 245–266. doi: 10.1111/j.1467-9868.2011.01004.x
- Tsinias, G., Nikou, S., Papadas, T., Pitsos, P., Papadaki, H., and Bravou, V. (2018). High PINCH1 expression in human laryngeal carcinoma associates with poor prognosis. *Anal. Cell. Pathol.* 2018:2989635. doi: 10.1155/2018/2989635
- Udyavar, A. R., Hoeksema, M. D., Clark, J. E., Zou, Y., Tang, Z., Li, Z., et al. (2013). Co-expression network analysis identifies spleen tyrosine kinase (SYK) as a candidate oncogenic driver in a subset of small-cell lung cancer. *BMC Syst. Biol.* 7(Suppl. 5):S1. doi: 10.1186/1752-0509-7-S5-S1
- Wang, L., Shi, J., Huang, Y., Liu, S., Zhang, J., Ding, H., et al. (2019). A six-gene prognostic model predicts overall survival in bladder cancer patients. *Cancer Cell. Int.* 19:229. doi: 10.1186/s12935-019-0950-7
- Wang, S., Yang, L., Ci, B., Maclean, M., Gerber, D. E., Xiao, G., et al. (2018). Development and validation of a nomogram prognostic model for SCLC patients. *J. Thorac. Oncol.* 13, 1338–1348. doi: 10.1016/j.jtho.2018.05.037
- Wang, Y., Shirogane, T., Liu, D., Harper, J. W., and Elledge, S. J. (2003). Exit from exit: resetting the cell cycle through Amn1 inhibition of G protein signaling. *Cell* 112, 697–709. doi: 10.1016/s0092-8674(03)00121-1
- Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., et al. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153, 307–319. doi: 10.1016/j.cell.2013.03.035
- Wu, M., Li, X., Zhang, T., Liu, Z., and Zhao, Y. (2019). Identification of a nine-gene signature and establishment of a prognostic nomogram predicting overall survival of pancreatic cancer. *Front. Oncol.* 9:996. doi: 10.3389/fonc.2019.00996
- Wu, Y., Xu, L., Yang, P., Lin, N., Huang, X., Pan, W., et al. (2018). Survival prediction in high-grade osteosarcoma using radiomics of diagnostic computed tomography. *EBioMedicine* 34, 27–34. doi: 10.1016/j.ebiom.2018.07.006
- Xiao, Z., Shi, Z., Hu, L., Gao, Y., Zhao, J., Liu, Y., et al. (2019). A new nomogram from the SEER database for predicting the prognosis of gallbladder cancer

- patients after surgery. *Ann. Transl. Med.* 7, 738. doi: 10.21037/atm.2019.11.112
- Xie, B., Becker, E., Stuparevic, I., Wery, M., Szachnowski, U., Morillon, A., et al. (2019). The anti-cancer drug 5-fluorouracil affects cell cycle regulators and potential regulatory long non-coding RNAs in yeast. *RNA Biol.* 16, 727–741. doi: 10.1080/15476286.2019.1581596
- Xiong, Y., Ling, Q. H., Han, F., and Liu, Q. H. (2019). An efficient gene selection method for microarray data based on LASSO and BPSO. *BMC Bioinformatics* 20(Suppl. 22):715. doi: 10.1186/s12859-019-3228-0
- Yang, S., Wang, Y., Ren, J., Zhou, X., Cai, K., Guo, L., et al. (2020). Identification of diagnostic and prognostic lncRNA biomarkers in oral squamous carcinoma by integrated analysis and machine learning. *Cancer Biomark.* [Epub ahead of print]. doi: 10.3233/cbm-191215
- Yap, W. K., Shih, M. C., Kuo, C., Pai, P. C., Chou, W. C., Chang, K. P., et al. (2018). Development and validation of a nomogram for assessing survival in patients with metastatic lung cancer referred for radiotherapy for bone metastases. *JAMA Netw. Open* 1:e183242. doi: 10.1001/jamanetworkopen.2018.3242
- Zhang, B., and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* 4:17.
- Zhang, D., Zeng, S., and Hu, X. (2020). Identification of a three-long noncoding RNA prognostic model involved competitive endogenous RNA in kidney renal clear cell carcinoma. *Cancer Cell. Int.* 20:319. doi: 10.1186/s12935-020-01423-4
- Zhang, F., Tapera, T. M., and Gou, J. (2018). Application of a new dietary pattern analysis method in nutritional epidemiology. *BMC Med. Res. Methodol.* 18:119. doi: 10.1186/s12874-018-0585-8
- Zhang, J., Fan, J., Yin, R., Geng, L., Zhu, M., Shen, W., et al. (2019). A nomogram to predict overall survival of patients with early stage non-small cell lung cancer. *J. Thorac. Dis.* 11, 5407–5416. doi: 10.21037/jtd.2019.11.53
- Zhang, J., Liu, W., Zou, C., Zhao, Z., Lai, Y., Shi, Z., et al. (2020). Targeting super-enhancer-associated oncogenes in osteosarcoma with THZ2, a covalent CDK7 inhibitor. *Clin. Cancer Res.* 26, 2681–2692. doi: 10.1158/1078-0432.CCR-19-1418
- Zhang, J. X., Song, W., Chen, Z. H., Wei, J. H., Liao, Y. J., Lei, J., et al. (2013). Prognostic and predictive value of a microRNA signature in stage II colon cancer: a microRNA expression analysis. *Lancet Oncol.* 14, 1295–1306. doi: 10.1016/s1470-2045(13)70491-1
- Zhou, G., Soufan, O., Ewald, J., Hancock, R. E. W., Basu, N., and Xia, J. (2019). NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res.* 47, W234–W241. doi: 10.1093/nar/gkz240
- Zhu, N., Hou, J., Ma, G., Guo, S., Zhao, C., and Chen, B. (2020). Co-expression network analysis identifies a gene signature as a predictive biomarker for energy metabolism in osteosarcoma. *Cancer Cell. Int.* 20:259. doi: 10.1186/s12935-020-01352-2
- Zuo, S., Wei, M., Zhang, H., Chen, A., Wu, J., Wei, J., et al. (2019). A robust six-gene prognostic signature for prediction of both disease-free and overall survival in non-small cell lung cancer. *J. Transl. Med.* 17:152. doi: 10.1186/s12967-019-1899-y

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Ouyang, Li, Zhu, Wang, Qi, Qu, Tu, Qu and Lu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.