Research Article

The Value of Immune-Related Genes Signature in Osteosarcoma Based on Weighted Gene Co-expression Network Analysis

Xin Wang,1 Li Gan,2 Ju Ye,3 Mengjie Tang4,4 and Wei Liu5

1Department of Bone and Soft Tissue, Hunan Cancer Hospital, The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha 410013, China
2Department of Anesthesiology, Hunan Cancer Hospital, The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha 410013, China
3Department of Pharmacy, Zunyi Medical University, Zunyi 563000, China
4Department of Pathology, Hunan Cancer Hospital, The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha 410013, China
5Department of Pharmacy, The Third Xiangya Hospital, Central South University, Changsha 410013, China

Correspondence should be addressed to Mengjie Tang; tangmengjie@hnca.org.cn and Wei Liu; liuweixy3@csu.edu.cn

Received 16 March 2021; Accepted 25 April 2021; Published 15 May 2021

Academic Editor: Jialiang Liang

Copyright © 2021 Xin Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Osteosarcoma (OS) is a serious malignant tumor that is more common in adolescents or children under 20 years of age. This study is aimed at obtaining immune-related genes (IRGs) associated with the progression and prognosis of OS.

Method. Expression profiling data and clinical data for OS were downloaded from the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) database. ESTIMATE calculates immune scores and stromal scores of samples and performs the prognostic analysis. Weighted gene coexpression network analysis (WGCNA) was used to find modules correlated with immune and stromal scores. Cox regression analysis and least absolute shrinkage and selection operator (LASSO) analysis were used to explore IRGs associated with OS prognosis and construct and validate a hazard score model. Finally, we verified the expression and function of EVI2B in OS. Results. WGCNA selected twenty-eight IRGs, 10 of which were associated with OS prognosis, and LASSO further obtained three key prognostic genes. An EVI2B model was constructed, and according to the risk score model, patients in the high-risk group had a worse prognosis than those in the low-risk group, and the prognosis was statistically significant in the high- and low-risk groups. Receiver operating characteristic (ROC) curves were used to assess the prognostic model’s accuracy and externally validate the independent GSE21257 cohort. The results of immunohistochemical staining and qPCR showed that EVI2B was a tumor suppressor gene. The differential genes in the high- and low-risk groups were analyzed by enrichment analysis of GO and KEGG, indicating that the EVI2B model is associated with immune response. Conclusion. In this study, IRG EVI2B is closely related to OS’s prognosis and can be used as a potential biomarker for prognosis and treatment of OS.

1. Introduction

Osteosarcoma is a malignant tumor that occurs more commonly in adolescents or children under 20 years of age [1]. Conventional therapies for osteosarcoma include surgery, adjuvant chemotherapy, and neoadjuvant chemotherapy [2–4]. For patients with osteosarcoma, the use of standard multiagent chemotherapy in combination with surgical resection produces a long-term survival of approximately 70% for the localized disease at diagnosis [5, 6] and 20–30% for metastatic disease at diagnosis or recurrence [7]. Although advances have been made in surgical techniques, targeted therapy and tumor immunity, and complications such as infection and poor survival due to limb salvage surgery, there are still many issues to be addressed in the treatment of OS; therefore, there is an urgent need to develop new prediction methods to improve OS patients’ survival [5, 8].
OS-related gene expression data were retrieved from the TARGRT (https://ocg.cancer.gov/programs/targ-

The prognosis of immune score and stromal score was analyzed using the R package survival in OS.

Gene co-expression module was identified in OS using WGCNA.

Univariate cox analysis and Multivariate cox analysis

1 gene model

Model evaluation and validation

Comparison with experimental validation

Figure 1: Flowchart of the present study.
A large body of evidence suggests that OS has an immune system with multiple therapeutic targets, including receptor T cell therapy [9], HER2-specific immunity [10–12], and adjuvant immune therapy with mifamurtide [13]. Immune checkpoint inhibitors may develop immune tolerance to tumor immunosuppressants, thereby increasing endogenous antitumor activity, and could improve therapy-induced tumor immunogenic chemotherapy, radiotherapy, and targeted therapy [13–17]. To date, the impact of IRGs on OS prognosis is not well clear, so prognostic
Figure 3: Continued.
Scale independence

Mean connectivity

Cluster dendrogram

Module membership vs. gene significance
\( \text{cor} = 0.024, p = 0.76 \)

Figure 3: Continued.
biomarkers with IRG are clinically essential for OS prognosis identification.

A standard analysis method relates the measurement of these genomic covariates to patients’ survival time, which often censored survival data [18]. A popular strategy is to use these covariates to fit a Cox regression model to the censored survival data and then predict new cancer prognosis based on this fitted model [18–20]. Compared with Cox risk regression analysis, data overfitting can be solved entirely by the LASSO analysis [21]. LASSO variants are a popular strategy to provide variable selection in regression analysis and have extended to the Cox regression model [18, 22, 23]. Regression with LASSO penalty is a commonly used method to provide variable selection in regression analysis and has been extended to the Cox regression model [18, 22, 23]. Systems biology algorithms for WGCNA have been used to assess the association between gene sets and clinical features by constructing scale-free gene coexpression networks [24–27].

This study focused on revealing IRGs involved in OS prognosis. The WGCNA analysis and PPI network analysis were first used by constructing a Cox proportional hazards (PH) prognostic model for LASSO [18, 19, 22]. Univariate and multivariate Cox regression analyses were performed for IRGs in PPI, and the hub genes identified could be used for prognostic risk assessment in OS patients. Finally, we validated the prediction model’s performance and accuracy by ROC analysis. Furthermore, the prognostic significance of this model was validated on an independent GEO dataset. To reveal the role of differential genes in these training groups, functional and pathway enrichment analysis was performed using differential genes. The results showed that these differential genes were mainly associated with immune response and inflammatory response. We also performed immunohistochemical staining and qPCR experimental validation, and the results were consistent with the bioinformatics results. In this study, IRG is involved in the immune process to affect the prognosis of OS patients, thereby intervening in patients’ immune response to improve the poor prognosis and enhance the quality of life of cancer patients.

2. Materials and Methods

2.1. Data Collection and Processing. Gene expression data (FPKM, fragments per kilobase million) and clinical data for OS obtained from the TARGET database (https://ocg.cancer.gov/programs/target). Log₂ (FPKM+1) conversion for the expression value of the TARGET database. Clinico-pathological data of the corresponding patients retrieved
from the database, including gender, ethnicity, age, tumor location, time to recurrence, and survival information. In this study, data from the GEO database were used (http://www.ncbi.nlm.nih.gov/geo/) of an independent dataset for external validation, downloaded, and collected a high-throughput gene expression dataset, numbered GSE21257 (series matrix file). The Ensemble IDs of the mRNAs in the TARGET database were extracted from the GENCODE project. The ID conversion for the GSE21257 data set is provided by GPL10295 (Illumina human-6 v2.0 expression beadchip).

### Table 1: 28 key genes selected in the pink module.

<table>
<thead>
<tr>
<th>Probes</th>
<th>Module color</th>
<th>GS. immune score</th>
<th>p.GS. immune score</th>
<th>MM pink</th>
<th>p.MM pink</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTPRJ</td>
<td>Pink</td>
<td>0.601266559</td>
<td>7.37E-10</td>
<td>8.004039444</td>
<td>1.38E-20</td>
</tr>
<tr>
<td>SCARF1</td>
<td>Pink</td>
<td>0.527434914</td>
<td>1.53E-07</td>
<td>8.05025206</td>
<td>5.65E-21</td>
</tr>
<tr>
<td>CD300C</td>
<td>Pink</td>
<td>0.540904961</td>
<td>6.35E-08</td>
<td>8.16077901</td>
<td>6.05E-22</td>
</tr>
<tr>
<td>ANPEP</td>
<td>Pink</td>
<td>0.548497065</td>
<td>3.80E-08</td>
<td>8.16287387</td>
<td>5.79E-22</td>
</tr>
<tr>
<td>TBXAS1</td>
<td>Pink</td>
<td>0.763212372</td>
<td>8.47E-18</td>
<td>8.21317927</td>
<td>1.99E-22</td>
</tr>
<tr>
<td>RNASE6</td>
<td>Pink</td>
<td>0.724597637</td>
<td>2.13E-15</td>
<td>8.34634834</td>
<td>9.93E-24</td>
</tr>
<tr>
<td>HSD3B7</td>
<td>Pink</td>
<td>0.5189629</td>
<td>2.61E-07</td>
<td>8.35214408</td>
<td>8.67E-24</td>
</tr>
<tr>
<td>EVI2B</td>
<td>Pink</td>
<td>0.668107494</td>
<td>1.56E-12</td>
<td>8.35865608</td>
<td>7.84E-24</td>
</tr>
<tr>
<td>NCKAP1L</td>
<td>Pink</td>
<td>0.771339655</td>
<td>2.31E-18</td>
<td>8.40965442</td>
<td>2.17E-24</td>
</tr>
<tr>
<td>PTK2B</td>
<td>Pink</td>
<td>0.624883331</td>
<td>9.88E-11</td>
<td>8.49623447</td>
<td>2.43E-25</td>
</tr>
<tr>
<td>SPI1</td>
<td>Pink</td>
<td>0.673046116</td>
<td>9.26E-13</td>
<td>8.49796638</td>
<td>2.32E-25</td>
</tr>
<tr>
<td>GRN</td>
<td>Pink</td>
<td>0.581838777</td>
<td>3.41E-09</td>
<td>8.50541409</td>
<td>1.91E-25</td>
</tr>
<tr>
<td>SQOR</td>
<td>Pink</td>
<td>0.656087588</td>
<td>5.28E-12</td>
<td>8.63819471</td>
<td>4.90E-27</td>
</tr>
<tr>
<td>GLRX</td>
<td>Pink</td>
<td>0.62368047</td>
<td>1.10E-10</td>
<td>8.64140646</td>
<td>4.46E-27</td>
</tr>
<tr>
<td>APOBR</td>
<td>Pink</td>
<td>0.600016784</td>
<td>8.15E-10</td>
<td>8.71667349</td>
<td>4.66E-28</td>
</tr>
<tr>
<td>GNA15</td>
<td>Pink</td>
<td>0.653638843</td>
<td>6.73E-12</td>
<td>8.81020828</td>
<td>2.28E-29</td>
</tr>
<tr>
<td>LILRB4</td>
<td>Pink</td>
<td>0.713838334</td>
<td>8.46E-15</td>
<td>8.88280775</td>
<td>1.25E-29</td>
</tr>
<tr>
<td>LAPT5M</td>
<td>Pink</td>
<td>0.630129195</td>
<td>6.18E-11</td>
<td>8.88368281</td>
<td>9.39E-30</td>
</tr>
<tr>
<td>CCR1</td>
<td>Pink</td>
<td>0.626964574</td>
<td>8.21E-11</td>
<td>8.94293907</td>
<td>1.99E-31</td>
</tr>
<tr>
<td>FERM3</td>
<td>Pink</td>
<td>0.728808967</td>
<td>1.22E-15</td>
<td>8.98203844</td>
<td>4.36E-32</td>
</tr>
<tr>
<td>CYTH4</td>
<td>Pink</td>
<td>0.698812331</td>
<td>5.23E-14</td>
<td>9.05139124</td>
<td>2.52E-33</td>
</tr>
<tr>
<td>LCP1</td>
<td>Pink</td>
<td>0.670851401</td>
<td>1.17E-12</td>
<td>9.10716239</td>
<td>2.16E-34</td>
</tr>
<tr>
<td>CD4</td>
<td>Pink</td>
<td>0.646606036</td>
<td>1.33E-11</td>
<td>9.1904082</td>
<td>4.02E-36</td>
</tr>
<tr>
<td>CSF1R</td>
<td>Pink</td>
<td>0.639120086</td>
<td>2.71E-11</td>
<td>9.203441</td>
<td>2.07E-36</td>
</tr>
<tr>
<td>PSTPIP1</td>
<td>Pink</td>
<td>0.507580891</td>
<td>5.23E-07</td>
<td>9.28772774</td>
<td>2.19E-38</td>
</tr>
<tr>
<td>LAT2</td>
<td>Pink</td>
<td>0.580893451</td>
<td>3.67E-09</td>
<td>9.30001029</td>
<td>1.05E-38</td>
</tr>
<tr>
<td>SELPLG</td>
<td>Pink</td>
<td>0.626714866</td>
<td>8.39E-11</td>
<td>9.35560532</td>
<td>3.49E-40</td>
</tr>
<tr>
<td>RAC2</td>
<td>Pink</td>
<td>0.517122434</td>
<td>2.92E-07</td>
<td>0.94981488</td>
<td>1.14E-44</td>
</tr>
</tbody>
</table>

### Table 2: 10 prognostic genes for overall survival of OS patients assessed by univariate regression analysis.

<table>
<thead>
<tr>
<th>Gene</th>
<th>HR</th>
<th>z</th>
<th>p value</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>EVI2B</td>
<td>0.627242</td>
<td>-2.6835</td>
<td>0.007286</td>
<td>0.446155</td>
<td>0.881827</td>
</tr>
<tr>
<td>GRN</td>
<td>0.588554</td>
<td>-2.57038</td>
<td>0.010159</td>
<td>0.392866</td>
<td>0.881716</td>
</tr>
<tr>
<td>NCKAP1L</td>
<td>0.570277</td>
<td>-2.51025</td>
<td>0.012065</td>
<td>0.367825</td>
<td>0.884159</td>
</tr>
<tr>
<td>SELPLG</td>
<td>0.653439</td>
<td>-2.50146</td>
<td>0.012368</td>
<td>0.46818</td>
<td>0.912005</td>
</tr>
<tr>
<td>LILRB4</td>
<td>0.625401</td>
<td>-2.22715</td>
<td>0.025937</td>
<td>0.413782</td>
<td>0.945248</td>
</tr>
<tr>
<td>FERM3</td>
<td>0.698225</td>
<td>-2.22338</td>
<td>0.02619</td>
<td>0.508714</td>
<td>0.958334</td>
</tr>
<tr>
<td>SQOR</td>
<td>0.601923</td>
<td>-2.18736</td>
<td>0.028716</td>
<td>0.381945</td>
<td>0.948596</td>
</tr>
<tr>
<td>CYTH4</td>
<td>0.676584</td>
<td>-2.13248</td>
<td>0.032968</td>
<td>0.472465</td>
<td>0.968888</td>
</tr>
<tr>
<td>GNA15</td>
<td>0.686282</td>
<td>-2.06624</td>
<td>0.038806</td>
<td>0.480192</td>
<td>0.980824</td>
</tr>
<tr>
<td>RNASE6</td>
<td>0.752659</td>
<td>-1.99037</td>
<td>0.04655</td>
<td>0.568959</td>
<td>0.995669</td>
</tr>
</tbody>
</table>

2.2 Correlation between Prognosis and Stromal/Immune Score. The ESTIMATE algorithm was applied to the normalized expression stromal [28] to estimate each osteosarcoma sample’s stromal and immune score. The overall survival needs the primary prognostic endpoint, the immune score, and the stromal score of each sample calculated with the R package ESTIMATE; it analyzed overall prognosis survival using the R package survival.

2.3 Construction of Coexpression Network. TARGET’s data set used as the training set, the WGCNA program package [29, 30] in R software used to remove outlier samples and find the modules related to immune score and stromal score toolkit of heat map drawn to analyze the strength of interacting. Module-trait correlations were estimated using correlations between module signature genes and traits (immune...
The soft-threshold power of $\beta$ calculates using a scale-free topology criterion. A weighted adjacency stromal generates, and the soft-threshold power $\beta$ can emphasize weak and robust correlations between penalized genes \cite{24, 31}. Modules of RNA were obtained using the dynamic tree cutting method. It is classifying genes with similar expression profiles by modules; the average linkage was performed to perform hierarchical clustering by TOM-based dissimilarity \cite{24}. Finally, the difference of module eigengenes (MES) of the module dendrogram was calculated, and some modules were merged. The pink module with the highest correlation coefficient was selected as the next research object. The conditions for the screening of IRGs in the pink module were that the correlation between genes and the pink module was more significant than 0.8, the correlation coefficient between genes and the immune score was greater than 0.5, and a total of 28 IRGs were selected.

2.4. PPI Network Analysis. Using search tools for reciprocal gene/protein retrieval (https://string-db.org/) database...
2.5 Construction of Risk Assessment Model. Then, we performed univariate Cox regression analysis of 28 IRGs followed by screening for IRGs with significant prognostic differences ($p < 0.05$); IRGs were further analyzed by LASSO to reduce genes. Finally, a multivariate Cox prognostic model was built by selecting the IRGs closely related to survival. The sample was divided into two groups based on the median risk score. Besides, we validated the model in the training set. External validation was performed in an independent GEO dataset, GSE21257.

Figure 5: The results of evaluating the accuracy of the EVI2B model. (a) Survival analysis between high- and low-risk groups in GSE21257. (b) The result for the ROC curve of the GSE21257 cohort.
Figure 6: Continued.
2.6. Establishment and Validation of the Prognostic Risk Scoring System. To generate a risk scoring system for genes, we performed multivariate Cox proportional hazards regression. First, we used the survival package of R to obtain the regression coefficients for each gene. The coefficient of each selected gene (parameter coefficient R) represents the estimated logarithm of the hazard ratio (HR, parameter exp(coefficient R)) in R. Then, a risk score formula was established for all patients. ROC curve (AUC) predictions over time calculated for 1-, 3-, and 5-year survival using the survival ROC package in R. The optimal cut-off point was selected as the maximum sensitivity and specificity. According to the optimal cut-off point, patients were divided into high- and low-risk groups; the survival difference between the two groups were assessed with the R package survival analysis.

2.7. GO and KEGG Enrichment Analysis. Differential genes between TARGET high- and low-risk groups (R for differential analysis wrapped as limma package, FDR < 0.05, |Log FC| >1) were subjected to enrichment analysis. GO analysis is a standard method to define genes and their RNA or protein products to identify high-throughput transcriptome or genomic data [34]. KEGG is a collection of databases dealing with genomes, diseases, biological pathways, drugs, and chemicals [35]. We used the R package ClusterProfiler [36] to enrich differential genes in the training set’s high- and low-risk groups.

2.8. Sample Collection. Osteosarcoma samples were collected from Hunan Cancer Hospital in 2020. OS tissues and adjacent normal tissues from 20 patients were collected, immediately placed in liquid nitrogen, and stored at −80°C. None of the OS patients received antineoplastic therapy. Both the patients and their families in this study were fully informed, and informed consent was obtained from the participants. The Hunan Cancer Hospital ethics committee approved the study.

2.9. Immunohistochemical Staining and Analysis. Immunohistochemistry was performed in 20 OS tissues and adjacent normal tissues. Paraffin-embedded sections were stained to determine the expression level of EVI2B protein. Sections were incubated overnight at 4°C with anti-EVI2B 1 antibody.
Figure 7: The relative EVI2B expression in OS tissues and adjacent normal tissues. (a) The immunohistochemical detection results showed the expression level of EVI2B in osteosarcoma tissues and adjacent normal tissues. (b) RT-qPCR analysis of relative EVI2B expression in OS tissues and adjacent normal tissues.
(Invitrogen) at a 1:100 dilution. After washing with phosphate-buffered saline (PBS), the slides were incubated with a goat anti-mouse IgG secondary antibody conjugated to fluorescein isothiocyanate (ZSDB-BIO, Beijing, China) for 30 mins. They were washed with PBS and then incubated with an antifade reagent (Invitrogen, Carlsbad, USA). Finally, staining was observed using an Olympus CX41 fluorescence microscope (Olympus, Tokyo, Japan). Positive staining intensity in OS and adjacent normal tissues were analyzed by integrated optical density (IOD) using Image-Pro Plus software (version 6.0; Media Cybernetics, United States). All images were taken using the same microscope and camera set. Image-Pro Plus software was used to calculate the mean IOD (μm²) per stained area (IOD/area) for positive staining. T-test was used to analyze the results. A p value < 0.05 was considered statistically significant.

2.10. Real-Time Quantitative PCR. Total RNA was isolated from tissues from the above patients using Trizol reagent (TaKaRa, Japan) according to the manufacturer’s instructions. One microgram of RNA was reverse transcribed into cDNA using the Revert Aid First Strand cDNA Synthesis Kit (Thermo, USA). Quantitative RT-PCR was then performed with Pro Taq HS Premix Probe qPCR Kit (Accurate, Hunan, China). The amplification program consisted of one cycle of pre-denaturation at 95°C for 5 mins, 37 cycles of denaturation at 95°C for 30 s, annealing at 60°C for 30 s, and extension at 72°C for 10 mins. The GAPDH gene was used as an endogenous control gene for normalizing the expression of target genes. Each sample was analyzed in triplicate. The sequences of primers were used for RT-qPCR and annealing temperature (Table S1).

2.11. Statistical Analyses. All statistical analyses were performed using the R software 3.5.0. Statistical significance set at a probability value of p < 0.05. Univariate, LASSO and multivariate Cox regression analyses were applied to predict the overall survival of OS patients. The effect of ROC and calibration curve on the prediction accuracy of the prognostic model was compared. Enrichment analysis of differential genes in the high- and low-risk groups in the training set was performed with the R package cluster profile.

3. Results

3.1. Relationship between Immune Score and Stromal Score for Prognosis in OS. The flowchart of this study design is shown in Figure 1. Overall survival analysis did with R package survival analysis of the immune score’s prognosis and stromal score. The immune score and stromal score were calculated based on the ESTIMATE algorithm can promote the quantification of immune and stromal components in tumors; in this algorithm, the immune and stromal score is calculated by analyzing the specific gene expression characteristics of the immune and stromal cells to predict the infiltration of nontumor cells. Patients with high immune scores had better overall survival than those with low immune scores (Figure 2(a)), log-rank p value = 0.001. Patients with a high stromal score had a more favorable prognosis than those with a low stromal score (Figure 2(b)), log-rank p value = 0.008. The above results indicated that the immune score and stromal score significantly correlated with OS patients’ prognosis. Patients with high immune scores and the stromal score had a good prognosis.

3.2. IRGs of OS Screened with WGCNA and PPI Network Diagram. The sample cluster tree had an outlying sample, and the red line was the cut-off value to filter data (Figure 3(a)). All samples were in clusters after removing one outlying sample based on immune score and stromal score (Figure 3(b)). Sample dendrogram and trait heatmap were drawn according to the immune score and stromal score (Figure 3(b)). The above results were obtained by further analysis of the modules associated with immune score and stromal score by WGCNA. In this study, R² and connectivity were highest when the power β set at 15 (Figures 3(c) and 3(d)).

Thus, β identifies distinct gene coexpression modules in OSs. The cluster dendrogram of the selected genes was clustered with an adjacency stromal, and we obtained a module visualization of the genes (Figure 3(e)). We also received the correlation between modules and traits (Figure 3(h)). The pink module has the highest correlation coefficient with an immune score, with a correlation coefficient of 0.58, p < 0.05, which was statistically significant, indicating that the immune score significantly correlated with clinical traits—the above results were based on the WGCNA analysis performed by the immune system stromal scoring. Therefore, we chose this module as the next research object. We screened the pink module’s IRGs with a correlation between genes and the pink module greater than 0.8 and a correlation coefficient between genes and immune score greater than 0.5 (Figure 3(f)). With conditions, we screened a total of 28 IRGs (PTPRJ, SCARFI, CD300C, ANPEP, TBXAS1, RNASE6, HSDB3, EV12B, NCKAP1L, PTK2B, SPI1, GNA15, LILRB4, LAPT4M, CCR1, FERMT3, CYTH4, LCP1, CD4, CSF1R, PSTPIP1, LAT2, SELPLG, and RAC2) shown in Table 1.

A PPI regulatory network of 28 IRGs was obtained based on the STRING online database (https://string-db.org/). We found that the direct interaction between hub genes was through the PPI network analysis (Figure 3(g)). Nineteen genes (PTPRJ, CD300C, RNASE6, EV12B, NCKAP1L, PTK2B, SPI1, GNA15, LILRB4, LAPT4M, CCR1, FERMT3, CYTH4, LCP1, CD4, CSF1R, PSTPIP1, SELPLG, and RAC2) were correlated well with immune score.

3.3. Recognize IRG Prognostic Gene and Survival Analysis and ROC Analysis. The 28 IRGs were subjected to univariate Cox analysis, resulting in 10 IRGs with prognostic significance (Table 2) (p value < 0.05). These ten genes, EV12B, GRN, NCKAP1L, SELPLG, LILRB4, FERMT3, SQOR, CYTH4, GNA15, and RNASE6, were subjected to Lasso regression analysis further to screen prognostic genes (Figures 4(a) and 4(b)) after 1,000 stimuli run through the crossvalidation possibility, the optimal lambda determined, and three genes, EV12B, GRN, and NCKAP1L, were screened. Next, multivariate Cox analysis was performed to construct a prognostic
model based on EVI2B, GRN, and NCKAP1L, resulting in a model containing one IRG EVI2B. EVI2B is highly expressed in normal bone tissue controls and lowly expressed in osteosarcoma (Table 2), which can be regarded as a tumor suppressor gene consistent with the results of a colorectal cancer study by Yuan et al. [37].

By dividing the sample into high- and low-risk groups according to the model’s median risk score, we found a significant difference in survival between the high- and low-risk groups (log-rank p value = 0.016) (Figure 4(c)). ROC analysis showed AUC values of 0.676, 0.697, and 0.708 for ROC curves at 1, 3, and 5 years, respectively (Figure 4(d)). All results showed that the prediction effect of the IRG EVI2B model was moderately accurate.

3.4. Evaluating the Accuracy of the IRG EVI2B Model. The time-dependent ROC curve analysis evaluated the accuracy of the OS prediction model constructed by EVI2B. The IRG EVI2B model evaluates in an independent GEO dataset GSE21257. By dividing GSE21257 samples into high- and low-risk groups according to the TARGET dataset samples’ median risk score in the model, we found a significant difference in survival between the high- and low-risk groups (log-rank p value = 0.003) (Figure 5(a)). ROC analysis showed that the AUC values for 1, 3, and 5 years of the ROC curve were 0.663, 0.690, and 0.658, respectively (Figure 5(b)). The survival rate of the low-risk group of the GSE21257 sample in the validation set based on the IRG EVI2B model was higher than that of the high-risk group, and the prolonged survival time was consistent with the results in the training set. The AUC values for the validation set GSE21257 samples were like the AUC results in the training set. The result fully validates the IRG EVI2B prognostic model’s accuracy and illustrates that the IRG EVI2B is feasible as a prognostic indicator in OS patients.

3.5. GO and KEGG Enrichment Analysis. Next, to further elucidate the molecular functions and signaling pathways in which differential genes between the TARGET high- and low-risk groups are involved, we performed functional enrichment analysis of all DEGs. GO results: these genes are involved not only in T cell activation, regulation of leukocyte differentiation, regulation of lymphocyte activation, and several other biological processes but also in cytokine binding, cytokine receptor complex activity, MHC protein binding, and molecular functions in other tumors. Some of the encoded proteins are important components of the plasma membrane’s external side, secretory granule membrane, and vacuolar membrane (Figure 6(a)). We further analyzed these genes’ signaling pathways, and KEGG results showed that these genes play a regulatory role in important immune-related pathways (such as the B cell receptor signaling pathway and chemokine signaling pathway) (Figure 6(b)). These results suggest that the OS prognostic model is significantly associated with immune response and inflammatory response, affecting cancer patients’ prognosis by participating in immune processes, indicating the correctness of our analysis results.

3.6. IRG EVI2B Expression Is Downregulated in OS Tissues. Finally, immunostaining analysis of IRG EVI2B showed that EVI2B expression was higher in adjacent normal tissues (Figure 7(a)). EVI2B-positive staining was significantly different between the control and OS, and positive staining was low and statistically significant in OS (Figure 7(a)). We examined IRG EVI2B expression in clinical samples from OS patients by qPCR. The qPCR results showed that IRG EVI2B was highly expressed at adjacent normal bone tissue levels (Figure 7(b)). The above results showed that IRG EVI2B was highly expressed in adjacent normal bone tissue (control) and lowly expressed in osteosarcoma, and Yuan et al. included colorectal cancer study [37].

4. Discussion

In the past decades, traditional cancer treatments, including surgical treatment and chemotherapy/radiotherapy, have incredibly prolonged osteosarcoma patients’ survival time [38, 39]. Although the 5-year overall survival rate is high [9, 40, 41], the survival rate of patients with metastatic disease (most commonly in the lung parenchyma and distal bone) is meager, 19-30% [39, 42]. Also, chemotherapy/radiotherapy acquired resistance to new tumor drugs and severe side effects of OS treatment’s drug resistance characteristics after extensive surgical resection of OS treatment reach the platform [39, 43].

The use of targeted therapy [44–46] also benefits OS patients; for example, therapies targeting the unfolded protein response (UPR) pathway may be a feasible approach for treating osteosarcoma. The components targeting the UPR may prove beneficial for patients refractory to conventional chemotherapy [47]. In addition to being a marker of metastatic disease, therapeutic targeting of cathepsin D (CTSD) may be a promising novel approach for osteosarcoma treatment; it may produce a good response in metastatic disease [47]. However, targeted therapy also has some limitations and cannot be suitable for all types of osteosarcoma. In recent years, immunotherapy [48] has dramatically developed in cancer. Immune checkpoint-based therapy has been shown to encouraging the antitumor effect by restoring immune response in the tumor microenvironment [49, 50]. At the basement of immune checkpoint inhibitors, ipilimumab (monoclonal antibody) anticytotoxic T lymphocyte antigen four antibodies (mAb) (CTLA4) and monoclonal antibody death protein 1 (PD1) or PD1 ligand (PDL1) against programmed cells show that immune checkpoints may be immune-tolerant to the tumor [13–17, 51]. A study has shown that for osteosarcoma, the immune checkpoint inhibitor PD-L1 is negatively correlated with prognosis, and PD-1 is negatively correlated with overall survival [13–17, 51]. To date, although immunotherapy [10, 11, 13, 51–53] has used in OS, the impact on OS prognosis is not well understood, so there is an urgent need for a new biomarker that can use to predict the prognosis of OS patients, and our study is to affect the prognosis of cancer patients by participating in the immune process.

At present, the use of EVI2B [54] is still relatively rare. The EVI2B is located in the intron of the neurofibromatosis type 1 (NF1) gene and transcribed in the opposite direction
to the NF1 gene [55–58]. Still, their expression is not related, indicating that they are independently regulated [59]. EVI2B is a transmembrane protein [60], while NF1 is a tumor suppressor gene [61]. EVI2B is expressed in many different cell types, including myeloid cells [62]. The EVI2B is also localized within EVI2, a common viral integration site found in retroviral induced myeloid tumors [54]. It has postulated that viral integration on EVI2 alters EVI2B expression and that this altered expression predisposes mice to the development of bone marrow tumors [54].

The successful construction of the IRG EVI2B prognostic model and the evaluation and validation of the accuracy of the IRG EVI2B prognostic model with the ROC curve method and in the GEO dataset GSE21257 give us hope to use the IRG EVI2B gene as a new biomarker and therapeutic target for predicting the prognosis of osteosarcoma patients. In the same study, EVI2B was found to correlate with the prognosis of patients significantly negatively with colon cancer \( (p < 0.05) \), and EVI2B may be involved in the prognosis of colon cancer patients [37, 63]. EVI2B can similarly be used as a prognosis for osteosarcoma compared to the study by Yang et al. [64].

EVI2B is expressed in various tumors [62] and breast cancer mutations [65]. EVI2B is involved in the differentiation of melanocytes and keratinocytes [66, 67]. In fibroblast-like cells derived from neurofibromas, increased EVI2B mRNA levels were found [66]. EVI2B gene, their related studies indicate that the EVI2B gene is a direct target of CCAAT/enhancer-binding protein (C/EBP). The product of this gene, the transmembrane glycoprotein EVI2B (CD361) [68], shows to be abundantly expressed on the surface of primary hematopoietic cells, reaching the highest expression level in mature granulocytes [69]. They use shRNA-mediated downregulation of EVI2B in human and murine cell lines and primary hematopoietic stem and progenitor cells (HSPC), demonstrating impaired myeloid lineage development and altered progenitor function in EVI2B-depleted cells. Its study suggests that EVI2B is an essential modulator of HSPC and bone marrow differentiation and that low levels of EVI2B may contribute to the differentiation inhibitory profile of acute myelocytic leukemia (AML) [69].

It has shown that EVI2B appears highly expressed in patients with 11q deletion in chronic lymphocytic leukemia (CLL). It speculated that the EVI2B gene might be associated with poor prognosis in patients with CLL 11q deletion [70]. It also studied that EVI2B was transiently expressed on cellular adjuvants and used to directly immunize BALB/c mice to obtain a higher specific anti-EVI2B Ab response in immunized mice than stimulated by other cellular adjuvants, which used to accelerate the development of Ab drugs [71]. Our study also showed that EVI2B is associated with the prognosis of osteosarcoma. These results suggest that it is feasible for us to use EVI2B as a marker and potential target for the diagnosis and treatment of osteosarcoma, but further studies should be on its possible action mechanism.

5. Conclusions

In conclusion, we selected IRG EVI2B by WGCNA analysis and constructed an IRG-based prognostic model, which was validated and experimentally analyzed to obtain that IRG EVI2B can significantly predict the prognosis of osteosarcoma patients.

Data Availability

The publicly available datasets were analyzed in this study; these can be found in the TARGET and the NCBI Gene Expression Omnibus (GSE21257).

Ethical Approval

The Hunan Cancer Hospital ethics committee approved the study.

Conflicts 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.

Authors’ Contributions

Li Gan and Xin Wang contributed equally and are co-first authors to this work.

Acknowledgments

This study was supported by the Scientific Research Project of Hunan Health Commission (Nos. 202102041763 and 20200985), Changsha Municipal Natural Science Foundation (No. kq2014267), and Hunan Cancer Hospital Climb Plan (No. 2020QH001).

Supplementary Materials

Table S1: the sequences of primers used for RT-qPCR and annealing temperature. (Supplementary Materials)

References


