Immune-Related Long Non-coding RNA Signature and Clinical Nomogram to Evaluate Survival of Patients Suffering Esophageal Squamous Cell Carcinoma

Ting Zhu*, Zhifeng Ma, Haiyong Wang, Desheng Wei, Bin Wang, Chu Zhang, Linhai Fu, Zhupeng Li and Guangmao Yu*

Department of Thoracic Surgery, Shaoxing People's Hospital, Shaoxing, China

Esophageal squamous cell carcinoma (ESCC) turns out to be one of the most prevalent cancer types, leading to a relatively high mortality among worldwide sufferers. In this study, gene microarray data of ESCC patients were obtained from the GEO database, with the samples involved divided into a training set and a validation set. Based on the immune-related differential long non-coding RNAs (lncRNAs) we identified, a prognostic eight-lncRNA-based risk signature was constructed following regression analyses. Then, the predictive capacity of the model was evaluated in the training set and validation set using survival curves and receiver operation characteristic curves. In addition, univariate and multivariate regression analyses based on clinical information and the model-based risk score also demonstrated the ability of the risk score in independently determining the prognosis of patients. Besides, based on the CIBERSORT tool, the abundance of immune infiltrates in tumor samples was scored, and a significant difference was presented between the high- and low-risk groups. Correlation analysis with immune checkpoints (PD1, PDL1, and CTLA4) indicated that the eight-lncRNA signature–based risk score was negatively correlated with PD1 expression, suggesting that the eight-lncRNA signature may have an effect in immunotherapy for ESCC. Finally, GO annotation was performed for the differential mRNAs that were co-expressed with the eight lncRNAs, and it was uncovered that they were remarkably enriched in immune-related biological functions. These results suggested that the eight-lncRNA signature–based risk model could be employed as an independent biomarker for ESCC prognosis and might play a part in evaluating the response of ESCC to immunotherapy with immune checkpoint blockade.

Keywords: risk model, nomogram, immunity, long non-coding RNA, esophageal squamous cell carcinoma

INTRODUCTION

Esophageal cancer is the eighth most prevalent cancer worldwide and the sixth in cancer mortality (Siegel et al., 2017). There are two main histological types: esophageal adenocarcinoma (EAC) and esophageal squamous cell carcinoma (ESCC), which differ in etiology, pathogenesis, and biological characteristics. Despite the rapid increase in the incidence of EAC in western countries,

ESCC remains dominant in East Asia (Torre et al., 2015), with a poor 5-year overall survival (OS) rate and a high incidence of recurrence and metastasis (Aquino et al., 2013). Although the TNM staging system has been extensively employed as a factor indicating prognosis, due to the heterogeneity of ESCC, there still exist differences in the survival of patients at the same clinical stage. Hence, a comprehensive study of key molecular mechanisms related to the prognosis of ESCC is urgently needed.

Long non-coding RNAs (lncRNAs) as RNA transcripts are more than 200 nucleotides in length and are short of the ability to encode proteins (Gupta et al., 2010). As reported, lncRNAs are significant players in the regulation of a variety of biological processes, and they participate in gene expression regulation via chromatin modification and transcriptional and post-transcriptional processing (Rinn and Chang, 2012). In addition to regulating biological processes, recent studies reveal that lncRNAs have the potential to act as prognostic biomarkers, and multiple lncRNAs are identified and validated for prognostic use in many cancers (including gastric cancer, colorectal cancer, and renal clear cell carcinoma) (Hu et al., 2014; Zhu et al., 2016; Shi et al., 2017). Nevertheless, there have been few studies regarding the function of lncRNAs on ESCC prognosis, mainly attributed to the scarcity of relevant comprehensive and systematic analysis (Sun et al., 2014). Currently, ESCC gene expression data and related prognostic information are available in public databases, including the GEO database. Therefore, lncRNA and mRNA data of ESCC were downloaded from GEO for analysis here.

An increasing number of studies have displayed the importance of the immune microenvironment for the development of digestive tract cancers, including ESCC (Ahtiainen et al., 2019; Masugi et al., 2019; Xue et al., 2019), which can provide reliable potential biomarkers for cancer diagnosis and prognosis. The immune microenvironment, a promoter or a suppressor for the growth and progression of tumors, can efficiently target tumors through drugs and show an association with survival of cancer patients (Tamborero et al., 2018). Despite the immune microenvironment having recently been studied in pan-cancers or certain tumors (Tamborero et al., 2018; Thorsson et al., 2018), no study has provided a comprehensive immune spectrum specifically for cancers of the digestive system. In this study, lncRNAs associated with immune pathways in ESCC were identified through the ImmLnc website, and this method was previously verified in an independent data set (Li et al., 2020). CIBERSORT, a new algorithm employed to count immune cell subsets and provide the possibility of identifying immune biomarkers for diagnosis and prognosis (Newman et al., 2015), was carried out in this study as well.

Here, a prognostic model based on immune-related lncRNAs that could be applied to predict the prognosis of ESCC was determined through the GEO training set and ImmLnc, and a nomogram that could be applied in clinical practice was constructed. The prognostic value of the model and the nomogram was then verified in another independent data set. In addition, the distribution of immune infiltrates in groups with high/low risk scores was also revealed, and the possible role of the lncRNA signature we identified in immunotherapy was explored. In all, the results of this research will help to achieve an accurate prognosis for ESCC patients and may have the potential to predict patients' response to immunotherapy with immune checkpoint blockade (ICB).

MATERIALS AND METHODS

Data Downloading and Preprocessing

Esophageal squamous cell carcinoma gene expression (normal: 119, tumor: 119) and clinical information (**Supplementary Table 1**) of microarray GSE53624 (GPL18109) were downloaded from the GEO database as the training set, and 8,975 lncRNAs (**Supplementary Table 2**) and 19,361 mRNAs (**Supplementary Table 3**) were obtained by data annotation. Similarly, ESCC gene expression data (normal: 60, tumor: 60) and clinical information (**Supplementary Table 4**) of GSE53622 (GPL18109) were downloaded from the GEO database as the validation set. In all, 8,975 lncRNAs (**Supplementary Table 5**) and 19,361 mRNAs (**Supplementary Table 6**) were gained by annotation.

Screening of Immune-Related Differentially Expressed lncRNAs

Expression levels of lncRNAs and mRNAs in tumor and normal groups were analyzed by the *limma* package, respectively, with the same parameters ($| \text{logFC}| > 1$, FDR < 0.05). The data set of Lnc_Pathways_Sig.txt was downloaded from ImmLnc[1], and a total of 3,271 immune-related lncRNAs of ESCC were obtained (**Supplementary Table 7**). These 3,271 lncRNAs were intersected with differentially expressed lncRNAs (DElncRNAs) to obtain immune-related DElncRNAs.

Construction of Prognostic Risk Model

The *survival* package was employed to conduct a univariate regression analysis of immune-related DElncRNAs ($p < 0.05$), and lncRNAs related to prognosis were screened. Multivariate regression analysis was then conducted with the *survminer* package on ESCC survival-related genes to obtain a risk gene signature, and the risk score based on the signature was formulated as below:

$$\text{Risk score} = \sum_{i=1}^{n} (Coef_i \times x_i)$$

where, $Coef_i$ represents the synergetic coefficient and x_i represents the relative expression of genes standardized by Z-score.

Validation of the Prognostic Risk Model

Both the training and the validation sets were taken to testify to the validity of the model. Each patient was conferred a risk score based on the model. The median risk score of all samples in each set was used as the critical value to form high-risk and low-risk groups, and the *survminer* package was employed for

---

[1]http://bio-bigdata.hrbmu.edu.cn/ImmLnc/index.jsp

survival analysis. Receiver operation characteristic (ROC) curves were drawn with the *timeROC* package for the model, and area under the curve (AUC) values corresponding to 1, 2, and 3 years were calculated, respectively. In addition, the model-based risk score was taken as a single characteristic factor, and clinical information was combined for further regression analyses to assess the independence of the risk model.

Construction and Evaluation of Nomogram

Considering both clinical information and the risk score (high/low), a nomogram was then drawn with the *rms* package to predict 1-, 2-, and 3-year mortality. The predictive performance of the nomogram was identified using calibration curves with the *foreign* package. In addition, the *timeROC* package was used to draw ROC curves and calculate corresponding AUC values in combination with clinical information and risk scores. Then, the same operation was conducted in the validation set GSE53622 to evaluate the predictive ability of the nomogram.

Correlation Between Risk Score and Immune Infiltration

In order to explore the abundance of immune infiltrates in the high- and low-risk groups, the CIBERSORT algorithm was used to score the infiltration abundance of each immune cell in samples to evaluate the proportion of 22 types of immune cells in each sample. A *P* value less than 0.05 was defined as the standard for successful deconvolution of samples, and the samples with CIBERSORT $p < 0.05$ were analyzed in the next step.

Correlation Between lncRNA Signature and Immune Checkpoints

It has been confirmed that PD1, PDL1, and CTLA4 can be used as immune checkpoints for ESCC (Jiao et al., 2019; Baba et al., 2020). Here, Pearson correlation analysis was performed for the signature-based risk score and immune checkpoint expression to search for the potential function of the lncRNA signature in immunotherapy.

GO Enrichment Analysis

Pearson correlation analysis was employed to assess the correlation of the signature lncRNAs with differentially expressed mRNAs (DEmRNAs). The mRNAs with $|\text{cor}| > 0.5$ were selected as the co-expressed mRNAs. The *clusterProfiler* package was used for GO biological annotation of the co-expressed mRNA gene set (q value < 0.05) to explore the biological function regulated by the lncRNA signature in ESCC.

RESULTS

Screening of Immune-Related DElncRNAs

The differential analysis results uncovered that a total of 972 DElncRNAs were obtained, including 404 upregulated and 568 downregulated genes (**Figure 1A**), while a total of

3,026 DEmRNAs were obtained, with 1,283 upregulated and 1,743 downregulated (**Figure 1B**). In order to obtain immune-related DElncRNAs, 3,271 immune-related lncRNAs in ESCC downloaded from the ImmLnc website were intersected with the DElncRNAs from GEO. Finally, a total of 111 lncRNAs were obtained (**Figure 1C**). It was considered that the above 111 lncRNAs were immune-related DElncRNAs in ESCC.

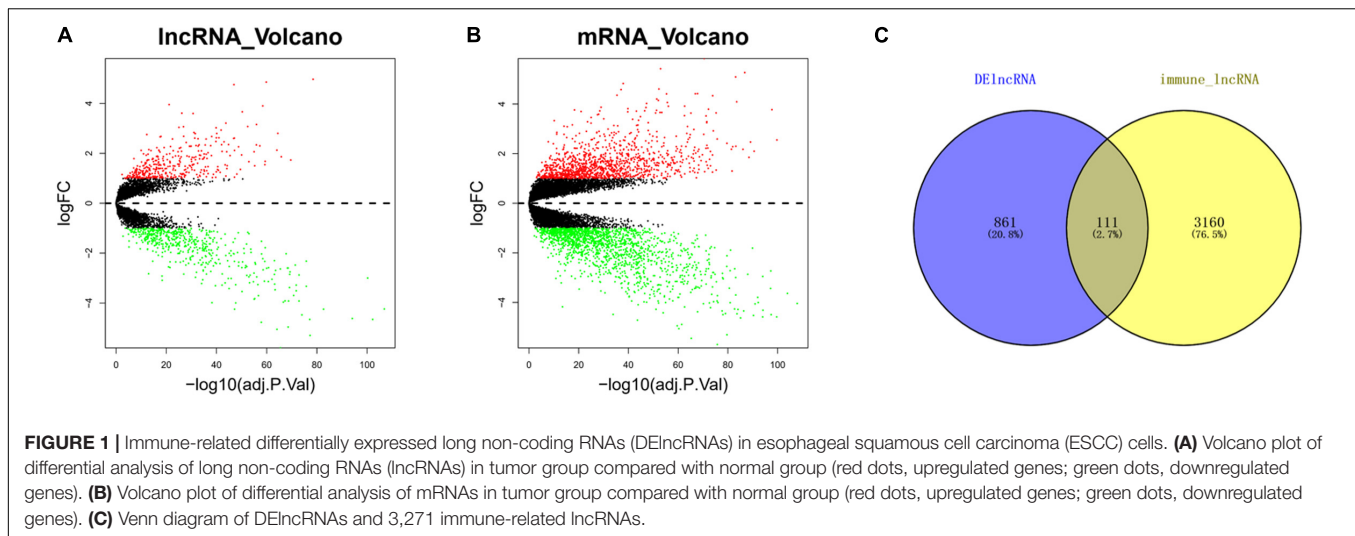## Construction and Validation of Risk Model

For the purpose of identifying prognostic lncRNAs of ESCC patients, the 111 lncRNAs identified before were analyzed by univariate analysis, and finally, 14 lncRNAs markedly related to prognosis of ESCC were obtained (**Supplementary Table 8**). A series of multivariate regression models were constructed for these 14 lncRNAs, and eventually, an eight-lncRNA signature–based model was identified (SMC5-AS1, MAMDC2-AS1, LINC01828, CASC8, AC112721.1, LINC00626, MIR100HG, and LINC02159) (**Figure 2A**). The formula of the model-based risk score was: risk score = 0.241 × SMC5-AS1+0.362 × MAMDC2-AS1-0.361 × LINC01828-0.318 × CASC8 +0.139 × AC112721.1 +0.110 × LINC00626-0.194 × MIR100HG-0.192 × LINC02159. The score distribution diagrams (**Figures 2B,C**), survival status diagrams (**Figures 2D,E**) and survival curves (**Figures 2F,G**) of patients with high or low risk scores in the training set and validation set all indicated that patients with a high risk score had a prominently lower survival rate than those with a low risk score. The ROC curves revealed that the AUC values in the training set for survival in 1, 2, and 3 years were 0.82, 0.8, and 0.82, respectively (**Figure 2H**), while in the validation set, the values were 0.74, 0.69, and 0.66, respectively (**Figure 2I**). In view of these data, it was proved that the eight-lncRNA signature based on the training set possessed a certain power in predicting the prognosis of ESCC patients.

The Risk Model Has the Ability to Independently Evaluate the Prognosis of LUAD Patients

To assess the independence of the model in predicting prognosis, the risk score was subjected to a univariate regression analysis, with clinical information of the training set taken into account as well. As analyzed, age, stage, and the risk score were of crucial significance in determining the prognosis (**Figure 3A**). Further multivariate analysis results suggested that tobacco, stage, and the risk score were of vital significance for prognosis prediction (**Figure 3B**). These results displayed that the risk model had the ability to assess prognosis independently of other clinical factors.

Construction and Evaluation of Nomogram

In order to help clinicians better determine the prognosis of ESCC patients, a nomogram was drawn using the eight-lncRNA signature–based risk score plus clinical information in the training set. As revealed, tobacco, stage, and the risk score had considerable significance in the prognosis of patients. In the training set, the proportion of patients with a total score of 442

FIGURE 1 | Immune-related differentially expressed long non-coding RNAs (DElncRNAs) in esophageal squamous cell carcinoma (ESCC) cells. **(A)** Volcano plot of differential analysis of long non-coding RNAs (lncRNAs) in tumor group compared with normal group (red dots, upregulated genes; green dots, downregulated genes). **(B)** Volcano plot of differential analysis of mRNAs in tumor group compared with normal group (red dots, upregulated genes; green dots, downregulated genes). **(C)** Venn diagram of DElncRNAs and 3,271 immune-related lncRNAs.

was the largest, and the corresponding risk of death at 1, 2, and 3 years was 0.289, 0.584, and 0.745, respectively (**Figure 4A**). The prognostic capability of the nomogram was further evaluated by ROC analysis visualized by calibration curves. The degree of fit of the calibration curves corresponding to 1, 2, and 3 years in the training set and validation set was good (**Figures 4B–G**). ROC curves were drawn by combining clinical information and risk scores. The AUC values corresponding to 1, 2, and 3 years in the training set were 0.79, 0.83, and 0.8, respectively (**Figure 4H**), while those in the validation set were 0.8, 0.78, and 0.73, respectively (**Figure 4I**). These results demonstrated that the nomogram had a good ability to predict the prognosis of patients.

Correlation Between Risk Score and Immunity

CIBERSORT was implemented to score the abundance of 22 kinds of immune infiltrates in tumor samples, and Pearson correlation analysis was then conducted among the immune cells (**Figures 5A–C**). With p value < 0.05 as the threshold, 59 cases having a low risk score and 57 cases having a high risk score were selected. As analyzed, remarkable differences in the abundance of immune infiltrates were noted in patients of the high- and low-risk groups, including resting memory CD4 T cells, NK cells activated, M0 macrophages, plasma cells, follicular helper T cells, and resting NK cells (**Figure 5D**). Then, correlation analysis was conducted between the eight-lncRNA signature–based risk score and expression levels of PD1, PDL1, and CTLA4 immune checkpoints, uncovering that there was a negative linkage between the risk score and PD1 expression (Cor = −0.33) (**Figures 5E–G**), which indicated that the eight-lncRNA signature may play a part in assessing patients' response to ICB immunotherapy.

The Eight-lncRNA Signature Regulates Immune-Related Biological Functions

GO annotation analysis was employed to further clarify the biological functions that may be affected by the eight-lncRNA

signature. Since lncRNAs could not be directly analyzed, 305 mRNAs were firstly obtained by co-expression analysis of the eight lncRNAs. Then, GO annotation of these 305 mRNAs was carried out and uncovered a notable enrichment in the biological functions involved in immunity, such as the extracellular matrix, endoplasmic reticulum, and basement membrane, etc (**Figure 6**). Since these mRNAs were markedly correlated with the eight-lncRNA signature we identified, it was inferred that the eight-lncRNA signature may participate in regulating the immune-related biological functions mentioned above.

DISCUSSION

For most cancers involving ESCC, TNM staging remains the primary reference that helps to guide therapeutic strategy and is identified to be a prognostic indicator. Nevertheless, due to the heterogeneity at both the molecular and the genetic levels, clinical outcomes and prognoses of tumors vary even when patients are diagnosed at the same stage and receive similar treatment (Navin et al., 2011; Gerlinger et al., 2012). At present, as high-throughput technologies, such as chip and RNA sequencing, develop continuously, gene expression profiling has turned out to be a potent tool that helps to discover molecular biomarkers associated with the phenotype or prognosis of esophageal cancer (Zhan et al., 2016). Multigene signatures that are identified for evaluating activity of cancer-associated genes, like Oncotype DX for breast cancer or ColoPrint chips for colon cancer, have been in the market, and the described signatures can be employed to aid in the treatment and prognostic management of tumors (Birnbaum et al., 2017).

Accumulating literature shows the abnormal expression of certain lncRNAs, which may be one of the main factors affecting the malignant progression of tumors. At present, most studies concentrate on lncRNA signatures, which can be potential markers in prognosis prediction of colorectal cancer, glioma, and pancreatic cancer in an independent manner (Zhang et al., 2012, 2013; Hu et al., 2014;
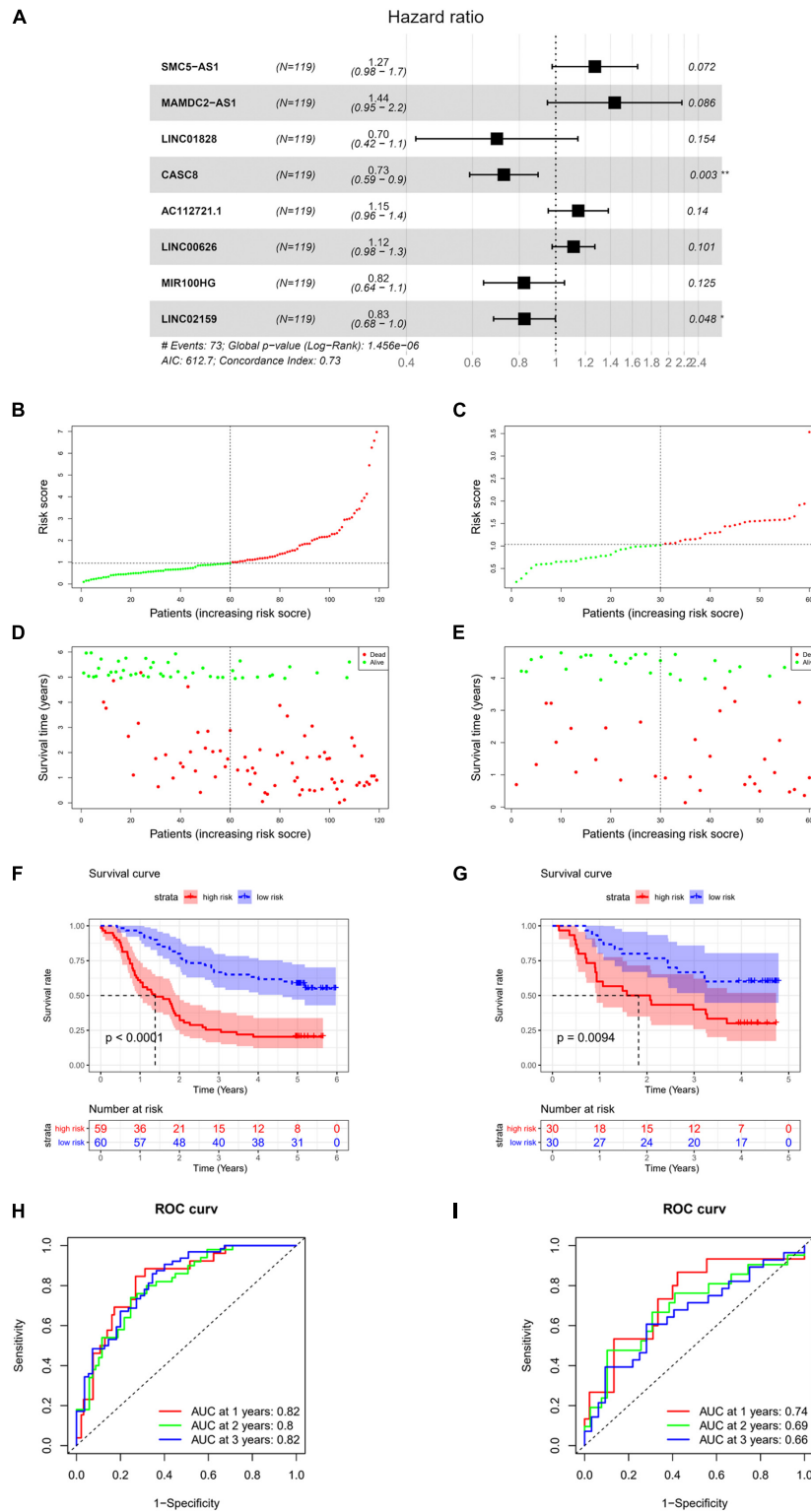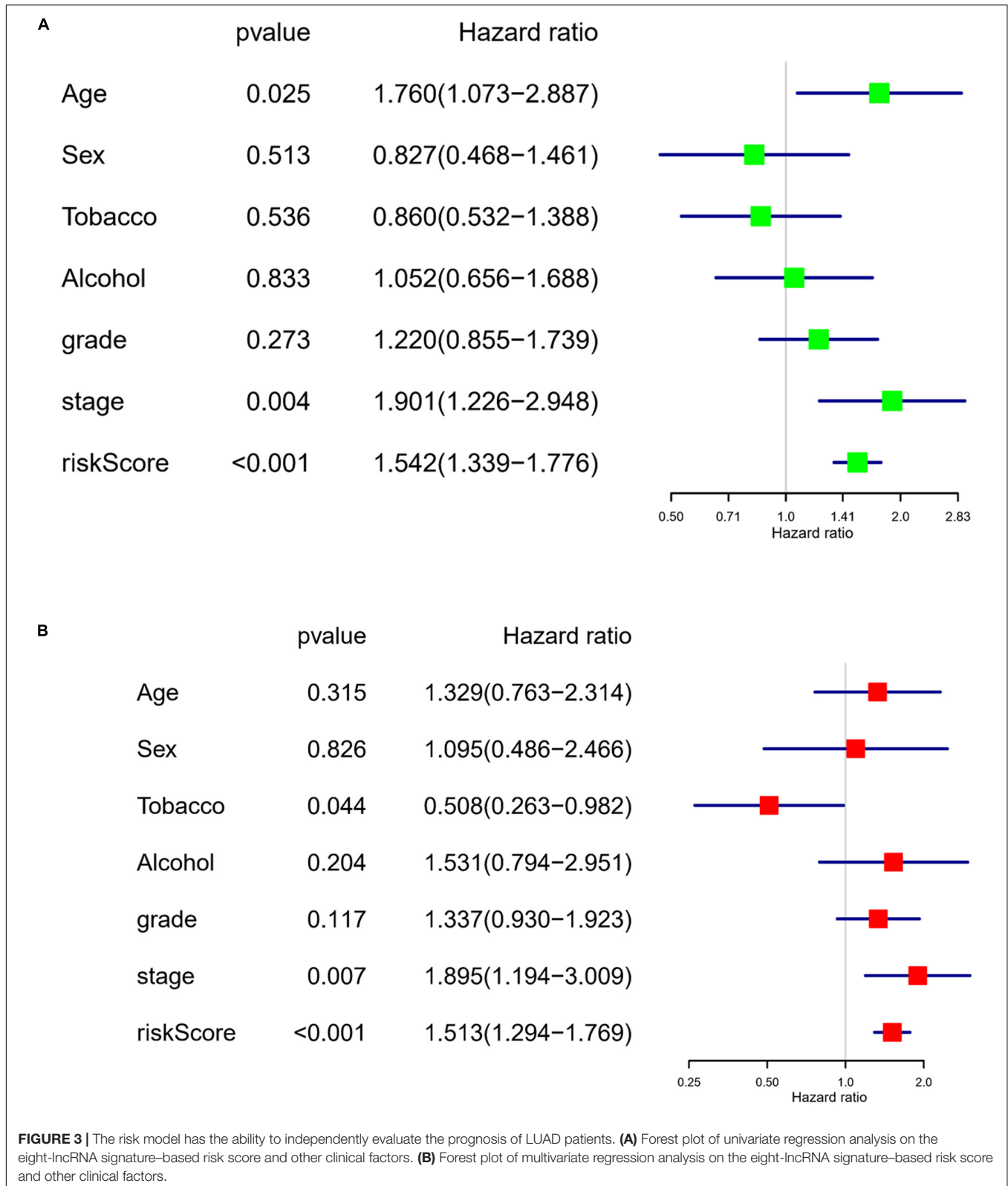
FIGURE 2 | Construction and validation of risk model. **(A)** Forest plot of multivariate regression analysis on eight-lncRNA signature. **(B)** Score distribution diagram of samples with high or low risk in GSE53624 data set (red: high risk; green: low risk). **(C)** Score distribution diagram of samples with high or low risk in GSE53622 data set (red, high risk; green, low risk). **(D)** Survival status diagram of samples in GSE53624 data set; red dots represent death, and green dots represent survival. **(E)** Survival status diagram of samples in GSE53622 data set; red dots represent death, and green dots represent survival. **(F)** Survival curves for samples with high or low risk in GSE53624 data set. **(G)** Survival curves for samples with high or low risk in GSE53622 data set. **(H)** Receiver operation characteristic (ROC) curves of GSE53624 data set. **(I)** ROC curves of GSE53622 data set. * indicates $p < 0.05$, while ** indicates $p < 0.01$.

FIGURE 3 | The risk model has the ability to independently evaluate the prognosis of LUAD patients. (A) Forest plot of univariate regression analysis on the eight-lncRNA signature–based risk score and other clinical factors. (B) Forest plot of multivariate regression analysis on the eight-lncRNA signature–based risk score and other clinical factors.

Birnbaum et al., 2017). Despite the fact that multiple lncRNAs, such as HOTAIR (Lv et al., 2013), CCAT2 (Zhang et al., 2015), and MALAT1, display potential prognostic value in ESCC (Cao et al., 2015; Deng et al., 2016), the role of immune-related lncRNA signatures in prognosis has not been clarified in literature.
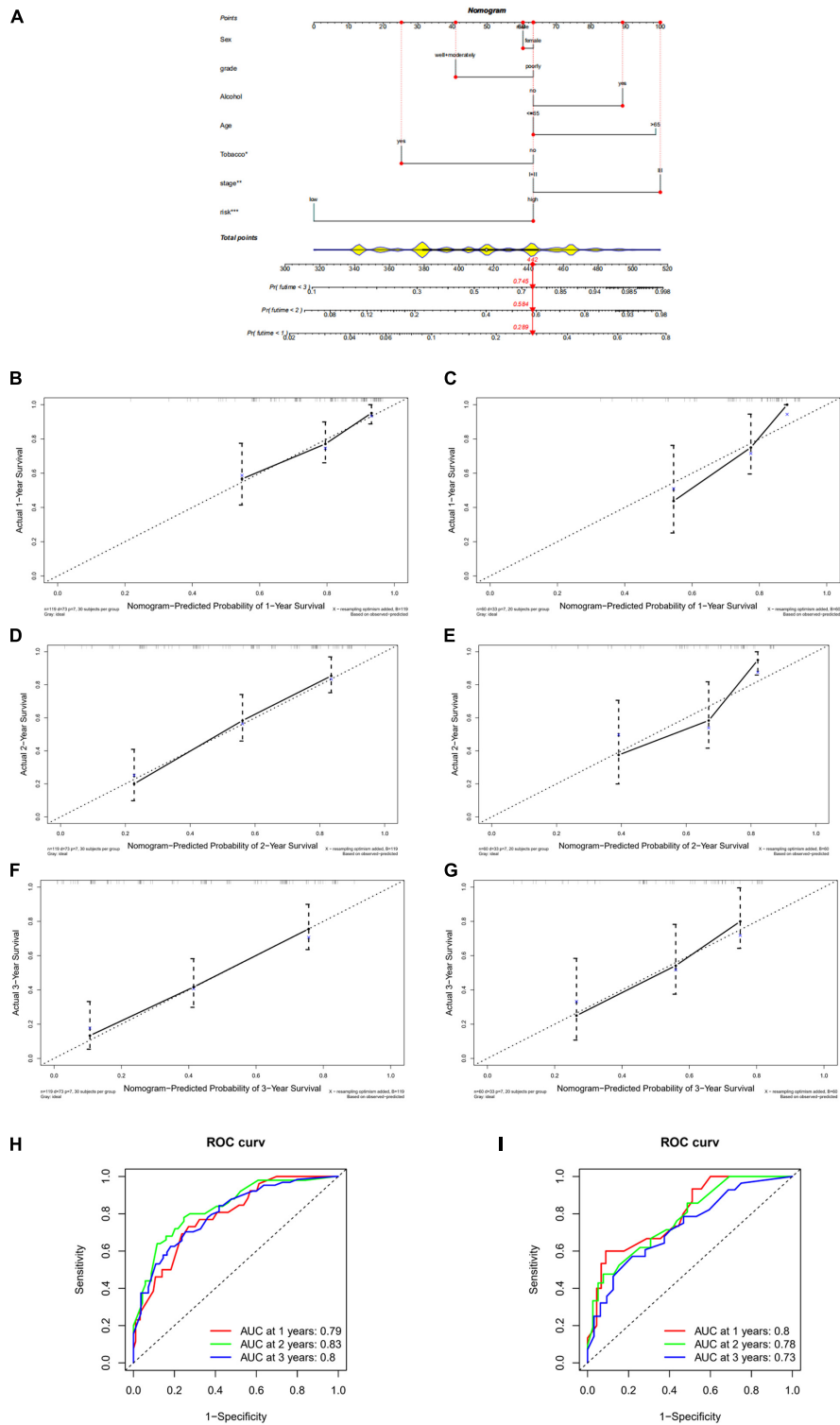
FIGURE 4 | Construction and evaluation of nomogram. **(A)** Nomogram based on eight-lncRNA signature–based risk score and clinical information in GSE53624 data set. **(B)** Calibration curve of nomogram-predicted probability of 1-year survival in GSE53624 data set. **(C)** Calibration curve of nomogram-predicted probability of 1-year survival in GSE53622 data set. **(D)** Calibration curve of nomogram-predicted probability of 2-year survival in GSE53624 data set. **(E)** Calibration curve of nomogram-predicted probability of 2-year survival in GSE53622 data set. **(F)** Calibration curve of nomogram-predicted probability of 3-year survival in GSE53624 data set. **(G)** Calibration curve of nomogram-predicted probability of 3-year survival in GSE53622 data set. **(H)** ROC curves based on clinical information combined with risk score in GSE53624 data set. **(I)** ROC curves based on clinical information combined with risk score in GSE53622 data set. *, **, and *** indicate $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.
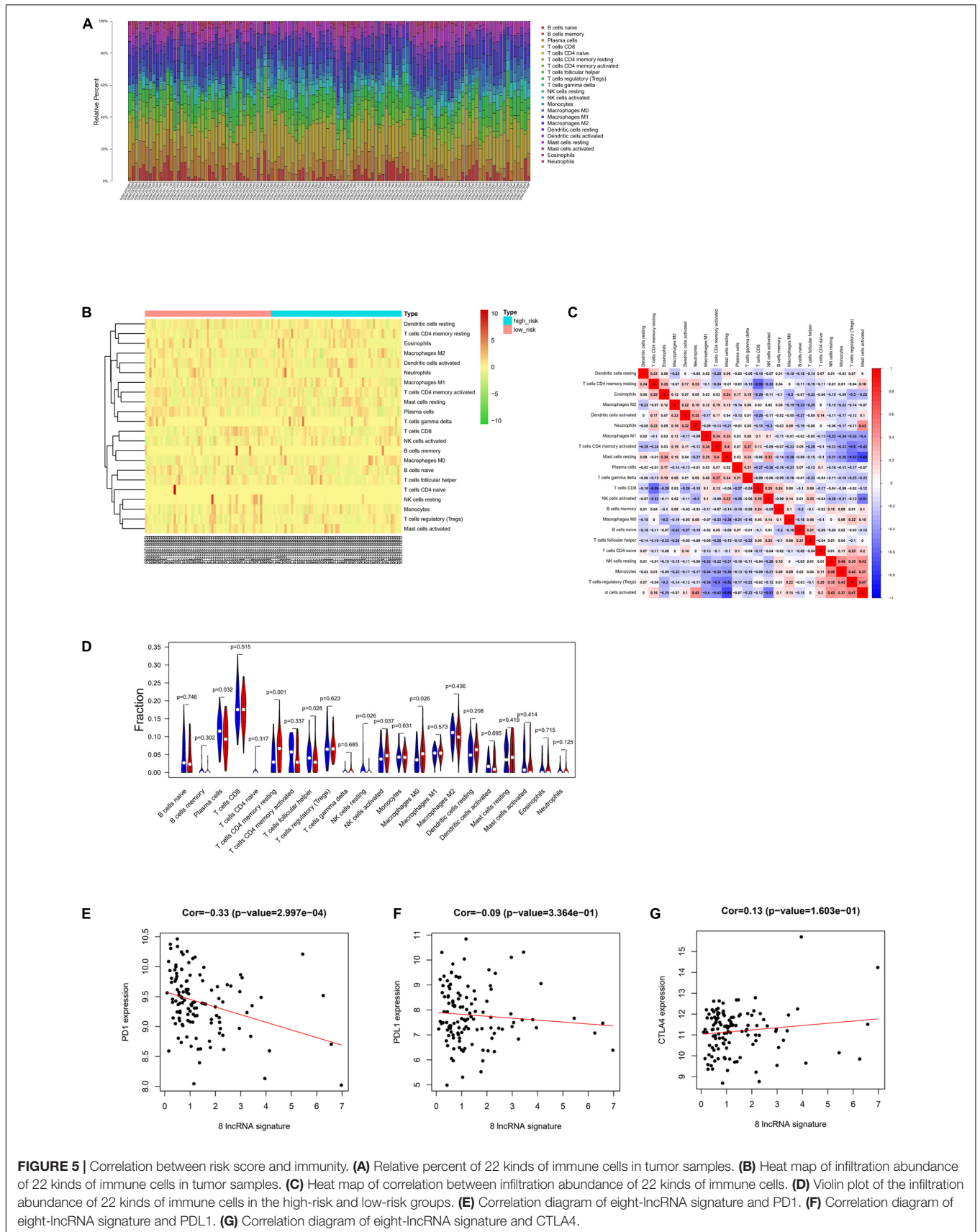
FIGURE 5 | Correlation between risk score and immunity. **(A)** Relative percent of 22 kinds of immune cells in tumor samples. **(B)** Heat map of infiltration abundance of 22 kinds of immune cells in tumor samples. **(C)** Heat map of correlation between infiltration abundance of 22 kinds of immune cells. **(D)** Violin plot of the infiltration abundance of 22 kinds of immune cells in the high-risk and low-risk groups. **(E)** Correlation diagram of eight-lncRNA signature and PD1. **(F)** Correlation diagram of eight-lncRNA signature and PDL1. **(G)** Correlation diagram of eight-lncRNA signature and CTLA4.

FIGURE 6 | GO annotation analysis of mRNAs co-expressed with the eight signature lncRNAs.
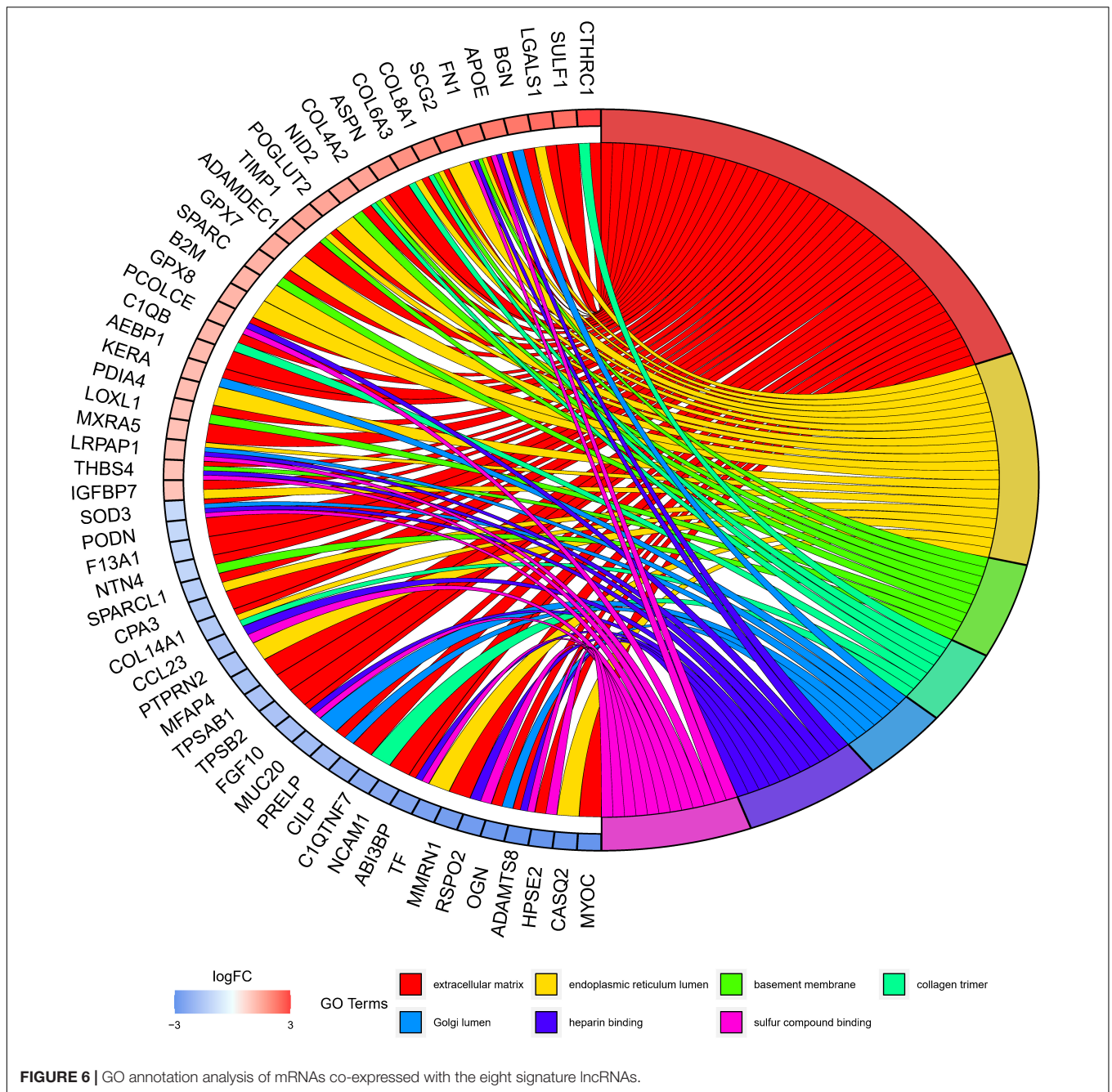
In recent years, the study of the tumor immune microenvironment has taken a leading position in cancer research (Fridman et al., 2012; Bhatia et al., 2019). In this study, an eight-lncRNA signature–based risk model was constructed. At present, there is no literature that reports on the SMC5-AS1, LINC00626, LINC01828, and LINC02159 genes. However, in this study, it was uncovered that SMC5-AS1 and LINC00626 were high-risk genes in ESCC, while LINC01828 and LINC02159 were low-risk genes. A study on MAMDC2-AS1 indicated that MAMDC2-AS1 can promote the progression of HSV-1 disease (Wang et al., 2020). Hu et al. (2017) revealed the role of

CASC8 as a tumor suppressor in bladder cancer and revealed that CASC8 can be a promising biomarker for cancer diagnosis. Through the TCGA database, Wang et al. (2019) analyzed the correlation between AC112721.1 and the OS time of patients with bladder urothelial carcinoma, finding that AC112721.1 expression is negatively associated with the survival of patients. Li et al. (2019) observed that the upregulation or downregulation of MIR100HG depends on different tumor types, and elevated MIR100HG expression can be an independent indicator for poor OS of patients suffering from gastric cancer. In this paper, the predictive accuracy of the model was verified in both the

training and the validation sets. A nomogram is a tool that can assist clinicians to better design the treatment strategy of patients according to medical condition and help to realize personalized treatment of patients (Long et al., 2018). Here, a nomogram was also established based on the risk score of the eight-lncRNA signature together with clinical information of ESCC patients, and the predictive power of the nomogram was evaluated as well.

Recently, there has been increasing evidence that some lncRNAs play a regulatory role in tumor immuno-response, including antigen release and immune cell infiltration (Carpenter and Fitzgerald, 2018; Denaro et al., 2019). In this study, CIBERSORT was implemented to score the abundance of immune infiltrates in tumor samples. It was found that the immune-related lncRNA signature was correlated to immune cell infiltration (M0 macrophages, T cells, and NK cells, etc.) in ESCC, indicating the vital role of the signature in immune infiltration in ESCC. Clara Di Vito (Di Vito et al., 2019) mentioned the dual role of NK cells regarding cancer progression or boosting the onset of immuno-suppressant tumor microenvironments, which might be the possible reason for the upregulation of NK cells activated in the high-risk group in the results of this paper. Literature showed that plasma cells are antibody producers and can promote immune response (Ribatti, 2017), which is consistent with the downward trend of plasma cells in the infiltrating abundance of the high-risk group in this paper. Immunotherapy offers a promising new treatment option for ESCC patients (Mimura et al., 2018). However, only some patients respond to ICB treatment, and the relevant literature and clinical trials both indicated that neither immune checkpoint gene expression nor mutational load can reliably predict the response of ESCC patients to ICB treatment (Fukuoka et al., 2019; Song et al., 2020). Therefore, identification of biomarkers to predict a patient's response to ICB immunotherapy is critical. In addition, some lncRNAs were uncovered to be correlated with immune responses and can predict a patient's response to immunotherapy (Xu et al., 2018; Vishnubalaji et al., 2020). In this study, Pearson correlation analysis was conducted between the eight-lncRNA signature–based risk score and three immune checkpoints. The results indicated that the signature-based risk score was remarkably negatively correlated with PD1 checkpoint. Therefore, it was speculated that the eight-lncRNA signature might have the potential to predict the response of ESCC patients to immunotherapy.

Finally, in order to further understand the biological processes of ESCC that might be regulated by the eight-lncRNA signature, mRNA co-expressed with the eight signature lncRNAs was analyzed by GO annotation. It turned out that the co-expressed mRNAs were remarkably enriched in biological functions, such as the extracellular matrix, endoplasmic reticulum lumen, and basement membrane, etc. Moreover, it can be seen in the literature that the dysregulation of the extracellular matrix can promote tumor immune escape (Pickup et al., 2014). Endoplasmic reticulum lumen stress can induce differentiation of immune cells (So, 2018). Therefore, combined with the results of this study, it is believed that the eight-lncRNA signature may participate in regulating immune-associated biological functions of ESCC.

In conclusion, in this study, DElncRNAs in ESCC were obtained from the GEO database, and DElncRNAs related to ESCC immunity were further identified after a consultation on the ImmLnc website. Eventually, an eight-lncRNA signature–based risk model and a nomogram that could be used to determine the prognosis of ESCC patients were established. However, it is still challenging to apply the eight-lncRNA signature in clinical treatment as a therapeutic target. In the future, a series of cell function experiments and clinical trials are required to advance the clinical application of the eight-lncRNA signature we constructed in treatment so as to improve the survival rate of ESCC patients.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

ZL and LF contributed to the study design. CZ and BW conducted the literature search. DW acquired the data. TZ wrote the article. HW performed data analysis and drafted the article. ZM revised the article. GY gave the final approval of the version to be submitted. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.641960/full#supplementary-material

Supplementary Table 1 | Clinical information of samples from GSE53624.

Supplementary Table 2 | LncRNAs obtained by GSE53624 annotation.

Supplementary Table 3 | mRNAs obtained by GSE53624 annotation.

Supplementary Table 4 | Clinical information of samples from GSE53622.

Supplementary Table 5 | LncRNAs obtained by GSE53622 annotation.

Supplementary Table 6 | mRNAs obtained by GSE53622 annotation.

Supplementary Table 7 | LncRNAs associated with ESCC immunity downloaded from ImmLnc website.

Supplementary Table 8 | 14 lncRNAs significantly related to prognosis of ESCC obtained by univariate analysis.

REFERENCES

Ahtiainen, M., Wirta, E. V., Kuopio, T., Seppälä, T., Rantala, J., Mecklin, J. P., et al. (2019). Combined prognostic value of CD274 (PD-L1)/PDCDI (PD-1) expression and immune cell infiltration in colorectal cancer as per mismatch repair status. *Mod. Pathol.* 32, 866–883. doi: 10.1038/s41379-019-0219-7

Aquino, J. L., Said, M. M., Pereira, D. A., Cecchino, G. N., and Leandro-Merhi, V. A. (2013). Complications of the rescue esophagectomy in advanced esophageal cancer. *Arq. Bras. Cir. Dig.* 26, 173–178. doi: 10.1590/s0102-67202013000300004

Baba, Y., Nomoto, D., Okadome, K., Ishimoto, T., Iwatsuki, M., Miyamoto, Y., et al. (2020). Tumor immune microenvironment and immune checkpoint inhibitors in esophageal squamous cell carcinoma. *Cancer Sci.* 111, 3132–3141. doi: 10.1111/cas.14541

Bhatia, S., Oweida, A., Lennon, S., Darragh, L. B., Milner, D., Phan, A. V., et al. (2019). Inhibition of EphB4-Ephrin-B2 signaling reprograms the tumor immune microenvironment in head and neck cancers. *Cancer Res.* 79, 2722–2735. doi: 10.1158/0008-5472.CAN-18-3257

Birnbaum, D. J., Finetti, P., Lopresti, A., Gilabert, M., Poizat, F., Raoul, J. L., et al. (2017). A 25-gene classifier predicts overall survival in resectable pancreatic cancer. *BMC Med.* 15:170. doi: 10.1186/s12916-017-0936-z

Cao, X., Zhao, R., Chen, Q., Zhao, Y., Zhang, B., Zhang, Y., et al. (2015). MALAT1 might be a predictive marker of poor prognosis in patients who underwent radical resection of middle thoracic esophageal squamous cell carcinoma. *Cancer Biomark.* 15, 717–723. doi: 10.3233/CBM-150513

Carpenter, S., and Fitzgerald, K. A. (2018). Cytokines and long noncoding RNAs. *Cold Spring Harb. Perspect. Biol.* 10:a028589. doi: 10.1101/cshperspect.a028589

Denaro, N., Merlano, M. C., and Lo Nigro, C. (2019). Long noncoding RNAs as regulators of cancer immunity. *Mol. Oncol.* 13, 61–73. doi: 10.1002/1878-0261.12413

Deng, H. Y., Wang, Y. C., Ni, P. Z., Lin, Y. D., and Chen, L. Q. (2016). Long noncoding RNAs are novel potential prognostic biomarkers for esophageal squamous cell carcinoma: an overview. *J. Thorac. Dis.* 8, E653–E659. doi: 10.21037/jtd.2016.07.01

Di Vito, C., Mikulak, J., Zaghi, E., Pesce, S., Marcenaro, E., and Mavilio, D. (2019). NK cells to cure cancer. *Semin. Immunol.* 41:101272. doi: 10.1016/j.smim.2019.03.004

Fridman, W. H., Pages, F., Sautes-Fridman, C., and Galon, J. (2012). The immune contexture in human tumours: impact on clinical outcome. *Nat. Rev. Cancer* 12, 298–306. doi: 10.1038/nrc3245

Fukuoka, E., Yamashita, K., Tanaka, T., Sawada, R., Sugita, Y., Arimoto, A., et al. (2019). Neoadjuvant chemotherapy increases PD-L1 expression and CD8(+) tumor-infiltrating lymphocytes in esophageal squamous cell carcinoma. *Anticancer Res.* 39, 4539–4548. doi: 10.21873/anticanres.13631

Gerlinger, M., Rowan, A. J., Horswell, S., Math, M., Larkin, J., Endesfelder, D., et al. (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* 366, 883–892. doi: 10.1056/NEJMoa1113205

Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., et al. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464, 1071–1076. doi: 10.1038/nature08975

Hu, R., Zhong, P., Xiong, L., and Duan, L. (2017). Long noncoding RNA cancer susceptibility candidate 8 suppresses the proliferation of bladder cancer cells via regulating glycolysis. *DNA Cell Biol.* 36, 767–774. doi: 10.1089/dna.2017.3785

Hu, Y., Chen, H. Y., Yu, C. Y., Xu, J., Wang, J. L., Qian, J., et al. (2014). A long non-coding RNA signature to improve prognosis prediction of colorectal cancer. *Oncotarget* 5, 2230–2242. doi: 10.18632/oncotarget.1895

Jiao, R., Luo, H., Xu, W., and Ge, H. (2019). Immune checkpoint inhibitors in esophageal squamous cell carcinoma: progress and opportunities. *Onco Targets Ther.* 12, 6023–6032. doi: 10.2147/OTT.S214579

Li, J., Xu, Q., Wang, W., and Sun, S. (2019). MIR100HG: a credible prognostic biomarker and an oncogenic lncRNA in gastric cancer. *Biosci. Rep.* 39:BSR20190171. doi: 10.1042/BSR20190171

Li, Y., Jiang, T., Zhou, W., Li, J., Li, X., Wang, Q., et al. (2020). Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers. *Nat. Commun.* 11:1000. doi: 10.1038/s41467-020-14802-2

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

Lv, X. B., Lian, G. Y., Wang, H. R., Song, E., Yao, H., and Wang, M. H. (2013). Long noncoding RNA HOTAIR is a prognostic marker for esophageal squamous cell carcinoma progression and survival. *PLoS One* 8:e63516. doi: 10.1371/journal.pone.0063516

Masugi, Y., Abe, T., Ueno, A., Fujii-Nishimura, Y., Ojima, H., Endo, Y., et al. (2019). Characterization of spatial distribution of tumor-infiltrating CD8(+) T cells refines their prognostic utility for pancreatic cancer survival. *Mod. Pathol.* 32, 1495–1507. doi: 10.1038/s41379-019-0291-z

Mimura, K., Yamada, L., Ujiie, D., Hayase, S., Tada, T., Hanayama, H., et al. (2018). Immunotherapy for esophageal squamous cell carcinoma: a review. *Fukushima J. Med. Sci.* 64, 46–53. doi: 10.5387/fms.2018-09

Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., et al. (2011). Tumour evolution inferred by single-cell sequencing. *Nature* 472, 90–94. doi: 10.1038/nature09807

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12, 453–457. doi: 10.1038/nmeth.3337

Pickup, M. W., Mouw, J. K., and Weaver, V. M. (2014). The extracellular matrix modulates the hallmarks of cancer. *EMBO Rep.* 15, 1243–1253. doi: 10.15252/embr.201439246

Ribatti, D. (2017). The discovery of plasma cells: an historical note. *Immunol. Lett.* 188, 64–67. doi: 10.1016/j.imlet.2017.06.006

Rinn, J. L., and Chang, H. Y. (2012). Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* 81, 145–166. doi: 10.1146/annurev-biochem-051410-092902

Shi, D., Qu, Q., Chang, Q., Wang, Y., Gui, Y., and Dong, D. (2017). A five-long non-coding RNA signature to improve prognosis prediction of clear cell renal cell carcinoma. *Oncotarget* 8, 58699–58708. doi: 10.18632/oncotarget.17506

Siegel, R. L., Miller, K. D., and Jemal, A. (2017). Cancer statistics, 2017. *CA Cancer J. Clin.* 67, 7–30. doi: 10.3322/caac.21387

So, J. S. (2018). Roles of endoplasmic reticulum stress in immune responses. *Mol. Cells* 41, 705–716. doi: 10.14348/molcells.2018.0241

Song, W., Wang, H., Tian, Y., Liu, S., Chen, X., Cui, J., et al. (2020). Refractory solitary cervical lymph node metastasis after esophageal squamous cell carcinoma surgery and its successful treatment with immune checkpoint inhibitor: a case report and literature review. *Medicine (Baltimore)* 99:e19440. doi: 10.1097/MD.0000000000019440

Sun, Y., Li, X., Jiang, C., Xiao, M., Zheng, Y., Zhang, J., et al. (2014). Data mining and bioinformatics of the expression data of esophageal squamous cell carcinoma. *Cell Biochem. Biophys.* 69, 481–485. doi: 10.1007/s12013-014-9821-y

Tamborero, D., Rubio-Perez, C., Muiños, F., Sabarinathan, R., Piulats, J. M., Muntasell, A., et al. (2018). A pan-cancer landscape of interactions between solid tumors and infiltrating immune cell populations. *Clin. Cancer Res.* 24, 3717–3728. doi: 10.1158/1078-0432.CCR-17-3509

Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Yang, T. H. O., et al. (2018). The immune landscape of cancer. *Immunity* 48, 812–830.e4. doi: 10.1016/j.immuni.2018.03.023

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global cancer statistics, 2012. *CA Cancer J. Clin.* 65, 87–108. doi: 10.3322/caac.21262

Vishnubalaji, R., Shaath, H., Elango, R., and Alajez, N. M. (2020). Noncoding RNAs as potential mediators of resistance to cancer immunotherapy. *Semin. Cancer Biol.* 65, 65–79. doi: 10.1016/j.semcancer.2019.11.006

Wang, J., Zhang, C., Wu, Y., He, W., and Gou, X. (2019). Identification and analysis of long non-coding RNA related miRNA sponge regulatory network in bladder urothelial carcinoma. *Cancer Cell Int.* 19:327. doi: 10.1186/s12935-019-1052-2

Wang, Y., Huang, L., Wang, Y., Luo, W., Li, F., Xiao, J., et al. (2020). Single-cell RNA-sequencing analysis identifies host long noncoding RNA MAMDC2-AS1 as a co-factor for HSV-1 nuclear transport. *Int. J. Biol. Sci.* 16, 1586–1603. doi: 10.7150/ijbs.42556

Xu, J., Shi, A., Long, Z., Xu, L., Liao, G., Deng, C., et al. (2018). Capturing functional long non-coding RNAs through integrating large-scale causal relations from gene perturbation experiments. *EBioMedicine* 35, 369–380. doi: 10.1016/j.ebiom.2018.08.050

Xue, J., Yu, X., Xue, L., Ge, X., Zhao, W., and Peng, W. (2019). Intrinsic beta-catenin signaling suppresses CD8(+) T-cell infiltration in colorectal cancer. *Biomed. Pharmacother.* 115:108921. doi: 10.1016/j.biopha.2019.108921

Zhan, C., Yan, L., Wang, L., Jiang, W., Zhang, Y., Xi, J., et al. (2016). Landscape of expression profiles in esophageal carcinoma by The Cancer Genome Atlas data. *Dis. Esophagus* 29, 920–928. doi: 10.1111/dote.12416

Zhang, X., Sun, S., Pu, J. K. S., Tsang, A. C. O., Lee, D., Man, V. O. Y., et al. (2012). Long non-coding RNA expression profiles predict clinical phenotypes in glioma. *Neurobiol. Dis.* 48, 1–8. doi: 10.1016/j.nbd.2012.06.004

Zhang, X., Xu, Y., He, C., Guo, X., Zhang, J., He, C., et al. (2015). Elevated expression of CCAT2 is associated with poor prognosis in esophageal squamous cell carcinoma. *J. Surg. Oncol.* 111, 834–839. doi: 10.1002/jso.23888

Zhang, X. Q., Sun, S., Lam, K. F., Kiang, K. M. Y., Pu, J. K. S., Ho, A. S. W., et al. (2013). A long non-coding RNA signature in glioblastoma multiforme predicts survival. *Neurobiol. Dis.* 58, 123–131. doi: 10.1016/j.nbd.2013.05.011

Zhu, X., Tian, X., Yu, C., Shen, C., Yan, T., Hong, J., et al. (2016). A long non-coding RNA signature to improve prognosis prediction of gastric cancer. *Mol. Cancer* 15:60. doi: 10.1186/s12943-016-0544-0

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.