



FULL LENGTH ARTICLE

# The combination of novel immune checkpoints HHLA2 and ICOSLG: A new system to predict survival and immune features in esophageal squamous cell carcinoma

Chaoqi Zhang <sup>a,1</sup>, Feng Wang <sup>a,1</sup>, Nan Sun <sup>a,1</sup>, Zhen Zhang <sup>b</sup>, Guochao Zhang <sup>a</sup>, Zhihui Zhang <sup>a</sup>, Yuejun Luo <sup>a</sup>, Yun Che <sup>a</sup>, Hong Cheng <sup>a</sup>, Jiagen Li <sup>a</sup>, Jie He <sup>a,\*</sup>

<sup>a</sup> Department of Thoracic Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, PR China

<sup>b</sup> Biotherapy Center, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, PR China

Received 23 March 2020; received in revised form 23 June 2020; accepted 14 August 2020

Available online 21 August 2020

## KEYWORDS

Esophageal cancer;  
HHLA2;  
ICOSLG;  
Immune checkpoint;  
Immunotherapy

**Abstract** Studies on immune checkpoint inhibitors targeting B7-CD28 family pathways in esophageal squamous cell carcinoma (ESCC) have shown promising results. However, a comprehensive understanding of B7-CD28 family members in ESCC is still limited. This study aimed to construct a novel B7-CD28 family-based prognosis system to predict survival in patients with ESCC. We collected 179 cases from our previously published microarray data and 86 cases with qPCR data. Specifically, 119 microarray data (GSE53624) were used as a training set, whereas the remaining 60 microarray data (GSE53622), all 179 microarray data (GSE53625) and an independent cohort with 86 qPCR data were used for validation. The underlying mechanism and immune landscape of the system were also explored using bioinformatics and immunofluorescence. We examined 13 well-defined B7-CD28 family members and identified 2 genes (ICOSLG and HHLA2) with the greatest prognostic value. A system based on the combination HHLA2 and ICOSLG (B7-CD28 signature) was constructed to distinguish patients as high- or low-risk of an unfavorable outcome, which was further confirmed as an independent

\* Corresponding author. Department of Thoracic Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 17 Panjiayuananli, Beijing 100021, PR China.

E-mail address: [prof.jiehe@gmail.com](mailto:prof.jiehe@gmail.com) (J. He).

Peer review under responsibility of Chongqing Medical University.

<sup>1</sup> These authors contributed equally to this work.

prognostic factor. As expected, the signature was well validated in the entire cohort and in the independent cohort, as well as in different clinical subgroups. The signature was found to be closely related to immune-specific biological processes and pathways. Additionally, high-risk group samples demonstrated high infiltration of Tregs and fibroblasts and distinctive immune checkpoint panels. Collectively, we built the first, practical B7-CD28 signature for ESCC that could independently identify high-risk patients. Such information may help inform immunotherapy-based treatment decisions for patients with ESCC.

Copyright © 2020, Chongqing Medical University. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

#### List of abbreviations

EC	Esophageal carcinoma
ESCC	Esophageal squamous cell carcinoma
TNM	Tumor-node-metastasis
lncRNA	Long non-coding RNA
APC	Antigen-presenting cells
PD-1	Programmed cell death protein-1
CTLA-4	Cytotoxic T-lymphocyte-associated protein 4
PD-L1	PD-1 ligand 1
PD-L2	PD-1 ligand 2
HHLA2	HERV-H LTR-associating 2
ICOSLG	Inducible T cell costimulatory ligand
OS	Overall survival
RFS	Relapse-free survival
qPCR	Quantitative real-time polymerase chain reaction
GO	Gene ontology
KEGG	Kyoto encyclopedia of genes and genomes
LN	Lymph node
ROC	Receiver operating characteristic
AUC	Area under ROC curve
MHC	Major histocompatibility complex
CAMs	Cell adhesion molecules
GSA	Gene sets variation analysis
BTLA	B and T lymphocyte attenuator
HCK	Hematopoietic cellular kinase
LCK	Lymphocyte-specific protein-tyrosine kinase
STAT1	Signal transducer and activator of transcription 1
TNFRSF4	Tumor necrosis factor receptor superfamily member 4
LGALS9	Interest in Galectin 9
LUAD	Lung adenocarcinoma

## Introduction

Esophageal carcinoma (EC) is one of the most common malignancies worldwide, and esophageal squamous cell carcinoma (ESCC) is major subtype of EC, especially in China.<sup>1</sup> There were 455,800 new cases and 400,200 deaths attributable to ESCC worldwide in 2012.<sup>2</sup> With the development of traditional therapies, including surgical therapy, chemotherapy, and radiotherapy, the prognosis of ESCC has changed over time; however, survival remains poor with

five-year survival rates ranging from 22.7 to 29.7%.<sup>2–4</sup> Although the tumor-node-metastasis (TNM) system is well known and widely used in clinical practice for prognostication, the survival of patients in the same clinical stage still varies.<sup>5</sup> Therefore, we continue to require methods for early prediction and determination of high-risk patients who may need adjuvant therapies.

We have been working on the discovery of prognosis biomarkers for ESCC,<sup>6</sup> built the first long non-coding RNA (lncRNA)-based prognosis signature for ESCC,<sup>7</sup> and uploaded the original microarray data (GSE53625). Although we, and other groups, have revealed multiple biomarkers linked to clinical outcomes, most biomarkers were screened out from the whole genome or transcriptome. This process ignored the intrinsic biological connection of cancer.

Immunotherapy has now become a novel therapeutic strategy for patients with ESCC.<sup>8</sup> Immunotherapeutic treatments of ESCC have garnered increased attention over the past few decades. The phase III Keynote 181 study demonstrated that pembrolizumab improves overall survival (OS) in patients with higher PD-L1 expression compared to chemotherapy.<sup>9</sup> The target of pembrolizumab is PD-1, one of the most important members of the B7-CD28 family. The B7-CD28 family—including 13 well-defined members—belongs to the B7 family (CD80, CD86, PD-L1, PD-L2, ICOSLG, B7-H3, B7x, and HHLA2). The CD28 family (CD28, CTLA4, ICOS, PD-1, and TMIGD2)<sup>10</sup> is a co-signaling superfamily (both co-inhibition and costimulation) which is involved in the initial stages of interaction between APC and T cells.<sup>11</sup> The most famous members of the CD28 family are programmed cell death protein-1 (PD-1) and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4), which play important roles in the immunotherapy responsiveness of malignant tumors. PD-1 was reported to be an inhibitory receptor in TCR signaling,<sup>12</sup> cytokine production,<sup>13,14</sup> and proliferation of T cells.<sup>15</sup> PD-1 ligand 1 (PD-L1) and PD-1 ligand 2 (PD-L2)—which belong to the B7 family and are expressed in tumor cells—can suppress the antitumor immune response. What's more, PD-L1 is associated with prognosis in patients with ESCC.<sup>16</sup> Various studies have indicated a connection between B7-CD28 members and ESCC progression. This observation inspired us to build a novel B7-CD28 family-based prognosis system.

HHLA2, short for HERV-HLTR-Associating 2, is recognized as a new member of B7 family in the past decade.<sup>17,18</sup> It's a type I transmembrane molecular with three extracellular Ig domains.<sup>17</sup> High expression of HHLA2 in several malignant

tumors indicated that it may be a potential biomarker and therapy target for tumor.<sup>19,20</sup> Inducible T cell Costimulatory Ligand (ICOSLG), another B7 family member, is a critical member for individual B cells to competitively participate in germinal center reaction.<sup>21</sup> As the ligand of ICOS, ICOSLG also involved in type 2 innate lymphoid cells function.<sup>22</sup> The combination of two critical member molecular of B7-CD28 family may reveal the interaction between the tumor and the immune system.

In this study, we analyzed all the B7-CD28 family members of ESCC and screened out the genes with the greatest prognostic value. Then, we built a system based on the combination HHLA2 and ICOSLG (B7-CD28 signature), that was well validated in different cohorts. The signature could independently classify patients with ESCC who were at high risk for a poor prognosis. In these individuals, the local immune response was enhanced with high infiltration of immune cells and active inflammatory activities. The different immune microenvironments among the ESCC subgroups, defined by our novel combination of HHLA2 and ICOSLG, may provide deep insights into the immunotherapeutic strategy.

## Methods

### Patients and study design

We enrolled 265 cases in this study, including 179 cases from our previous public data and 86 frozen tumor tissues from an independent cohort. All 179 cases, and their associated microarray data and corresponding clinical characteristics, are publicly available (GSE53625). This population includes two cohorts: 119 cases in GSE53624 and 60 cases in GSE53622. The original microarray data used in this study were processed as described previously.<sup>7</sup> In short, mRNA expression data of the samples were extracted by quantile normalization and then log 2-scaled transformed. For genes with more than one probe, the mean expression was calculated and used. Additionally, we

updated the recurrence-free survival (RFS) data for these patients. The 86 frozen ESCC tumor tissue samples were obtained from the First Affiliated Hospital of Zhengzhou University from 2011 to 2014. This research was approved by the Ethics Committee Board of the First Affiliated Hospital of Zhengzhou University. We used 119 cases from GSE53624 as the discovery cohort, and 60 cases from GSE53622 and the entire cohort of 179 for validation. Then, the clinical application value of this signature was validated in the independent cohort with qPCR data. All the clinical features of these cases are displayed in Table 1.

### Quantitative real-time polymerase chain reaction (qPCR) analysis

The total RNA of 86 samples was obtained from the frozen tissues mentioned above. cDNA was synthesized according to the manufacturers' instructions and diluted as templates for quantitative PCR. We used a 10  $\mu$ l volume system, containing 5  $\mu$ l SYBR Green Master Mix (Invitrogen), 3  $\mu$ l nuclease-free water, 1  $\mu$ l template, and 1  $\mu$ l of each PCR primer in the Agilent Mx3005P Real-Time PCR system. The expression values of ICOSLG and HHLA2 were normalized to the expression of GAPDH and then log<sub>2</sub> transformed for further validation. The primer sequences of the target genes and GAPDH are shown in Table S1.

### Biological pathway analysis and xCell analysis

We performed a Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of the most-related B7-CD28 family-based signature genes in the training cohort using DAVID 6.8 (<http://david.abcc.ncifcrf.gov>). To estimate the cellular heterogeneity in the tumor microenvironment, xCell,<sup>23</sup> which contains 64 different immune and stromal cell types, was used to evaluate the populations of different cell types in the ESCC tissues.

**Table 1** Clinical characteristics of the patients in training cohort, GSE53265 and independent cohort.

Characteristics	Training cohort (N = 119)	Validation cohort (N = 60)	GSE53625 (N = 179)	Independent cohort (N = 86)
Gender, No. (%)				
Male	98 (82.4%)	48 (80%)	146 (81.6%)	63 (73.3%)
Female	21 (17.6%)	12 (20%)	33 (18.4%)	23 (16.7%)
Age, No. (%)				
≤60	61 (51.3%)	30 (50.0%)	99 (55.3%)	24 (27.9%)
>60	58 (48.7%)	30 (50.0%)	80 (44.7%)	62 (72.1%)
Subtype, No. (%)				
LN+	65 (54.6%)	31 (51.7%)	96 (53.6%)	37(43.0%)
LN-	54 (45.45)	29 (48.3%)	83 (46.4%)	49 (57%)
TNM stage, No. (%)				
1-2	53 (44.5%)	34 (56.7%)	87 (48.6%)	17 (19.8%)
3-4	66 (55.5%)	26 (43.3%)	92 (51.4%)	69 (80.1%)
OS state, No. (%)				
Alive	46 (38.7%)	27(45.0%)	73 (40.8%)	60 (69.8%)
Dead	73 (61.3%)	33(55.0%)	106 (59.2%)	26 (30.2%)

## Immunofluorescence technique

Samples were fixed in 10% neutral-buffered formalin and embedded in paraffin. The tumor tissue sections (3  $\mu$ m) were then deparaffinized and blocked for preparation. Then, the primary and matched secondary antibodies were diluted in PBS containing 2% BSA and used to stain SMA AND Foxp3. Next, we washed the cells with PBS three times and stained the cell nuclei with 4, 6-diamidino-2-phenylindole (DAPI). Three independent experiments were carried out.

## Signature generation and statistical analysis

A univariate Cox regression analysis was first used to evaluate the significance of the 13 well-defined B7-CD28 members in GSE53624. Then, two genes were found to be correlated with OS ( $P < 0.05$ ). Next, the signature was generated based on a linear combination of the expression value of two genes, weighted with regression coefficients from a multivariable Cox proportional hazards regression model. Patients in the training and all validation cohorts were divided into high- and low-risk groups based on the signature score with an optimal cutoff point. The Kaplan–Meier method was used to evaluate the OS and RFS between the high- and low-risk groups, and a log-rank test was used to assess the difference in prognosis between the two groups. The Mann–Whitney  $U$ -test was used to calculate the distribution of xCells between high- and low-risk patients. Independent prognostic factors in this study were calculated by Cox proportional hazards regression model. Other statistical computations and the figures, including heatmap, boxplots, Receiver operating characteristic (ROC) curves, and survival curves, were realized by using several packages (ggplot2, ggrepel, ggthemes, pheatmap, cowplot, pROC, and survival) in the statistical software environment R version 3.5.1 (<https://www.r-project.org>). For all statistical methods, a  $P$  value less than 0.05 was regarded as significant.

## Results

### Construction and internal validation of a B7-CD28 family-based signature in ESCC

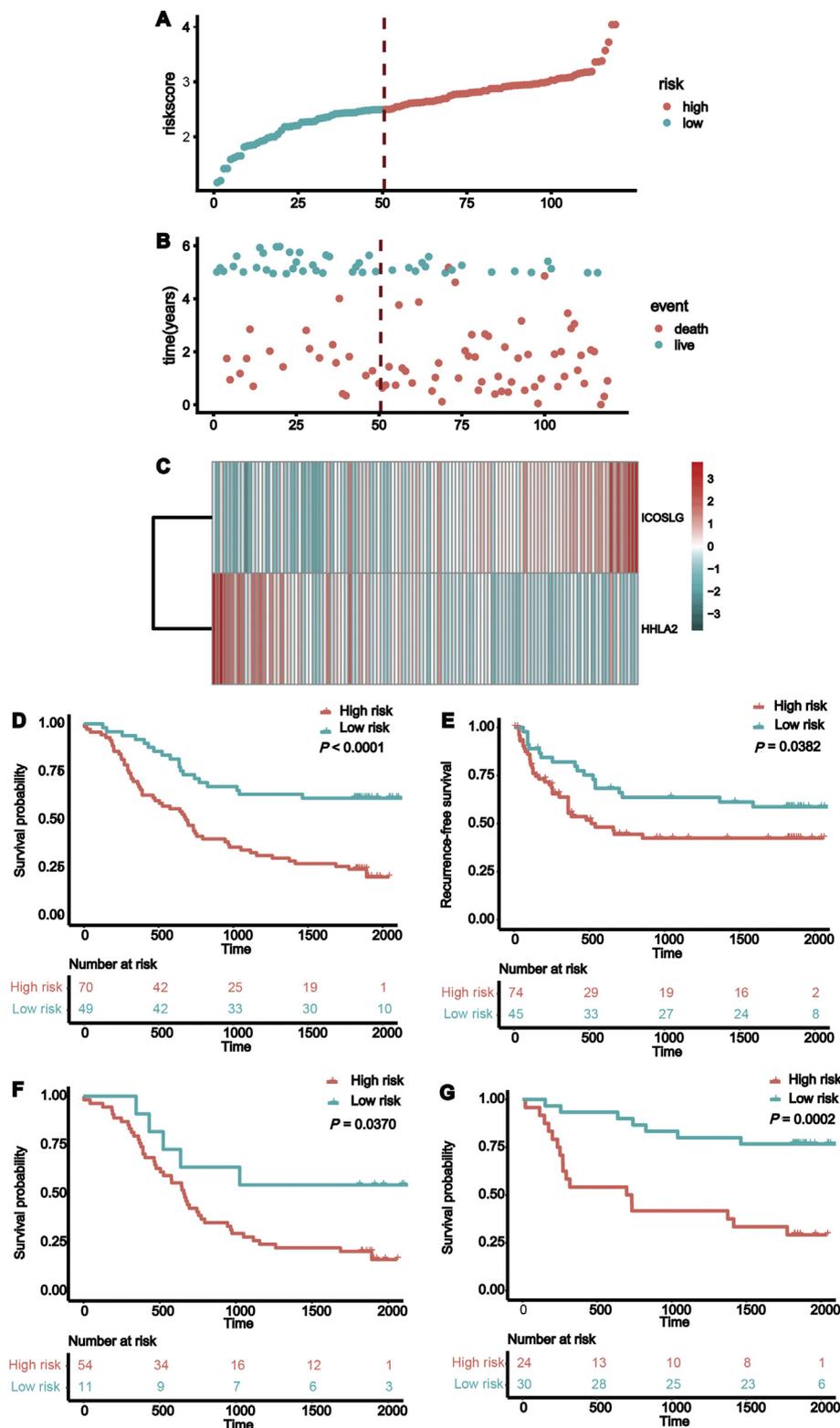
To explore the relationship between the B7-CD28 superfamily and prognosis in ESCC, we constructed a model to describe the inner relationship. Data from 119 patients with ESCC in the GSE53623 dataset was used as a training cohort, and another 60 patients with ESCC in the GSE53622 dataset was used as a validation cohort. Detailed clinical

characteristics of the patients are shown in Table 1. The expression of 13 well-defined B7-CD28 genes in the training cohort was analyzed by univariate Cox regression, and the details of the genes are shown in Table S2. Then, the significant genes were incorporated into a multivariable Cox proportional hazards regression model (Table 2). Ultimately, a two-gene prognostic model was established based on the combination of HHLA2 and ICOSLG. The model was called the B7-CD28 based signature. Risk scores were calculated as follows: risk score =  $0.5419 \times \text{ICOSLG} - 0.4976 \times \text{HHLA2}$ . Risk scores were calculated for each patient, and the optimal cutoff point was used as a cutoff to divide all patients into high- and low-risk groups (Fig. 1A–C). The OS ( $P < 0.0001$ ) and RFS ( $P = 0.0382$ ) were both shorter in the high-risk group when compared to the low-risk group (Fig. 1D, E). From the result of ROC analysis of B7-CD28 signature for prediction of mortality risk at 1, 3, and 5 years in the training cohort, we found that area under ROC curve (AUC) of 1, 3 and 5 years were all above 0.5, which further confirmed the effectiveness of the model (Fig. S1). The risk score remained an independent prognostic factor even after it was incorporated into a multivariate Cox proportional hazards regression model together with important clinical variables (Table 3).

As lymph node metastasis plays a crucial role in the prognostic outcomes of ESCC,<sup>24</sup> we further explored the connection of OS and risk score for both lymph node metastasis positive (LN+) and lymph node metastasis negative (LN-) patients within the 119 person cohort (Fig. 1F, G). The patients in the high-risk group showed shorter OS compared to the low-risk group in both subgroups, with  $P$  values of 0.0370 and 0.0002, respectively. We also found that the signature in the high-risk group indicated shorter survival in subtypes of the training cohort including TNM1-2 (Fig. S2A), TNM3-4 (Fig. S2B), older (age  $\geq 60$ ) (Fig. S2C), younger (age  $< 60$ ) (Fig. S2D), and male (Fig. S2E) patients. We did not find a significant difference in risk scores and survival for female patients (Fig. S2F). In the validation cohort (Fig. 2A), the signature only showed a borderline difference between the high- and low-risk patients, with  $P$  values of 0.2729 and 0.0908 for OS and RFS, respectively (Fig. 2B, C). This may have been caused by the small size of this cohort. Therefore, the entire cohort (GSE53625) was further explored (Fig. 2D). As expected, we found that both the OS and RFS were shorter in the high-risk group, with  $P$ -values of 0.0002 and 0.0167, respectively. Similarly, the signature was found still have prognostic value in different clinical subtypes of the entire cohort, regardless of LN status, TNM stage, age, or sex (Fig. S3). Moreover, the results of the multivariable Cox regression analysis further confirmed that the risk score was an independent factor for ESCC in the entire cohort (Table 3).

**Table 2** Individual B7-CD28 gene multivariate Cox proportional hazards models in training cohort.

Gene	Aliases	HR	95% CI	$P$ -value	Family
ICOSLG	B7-H2, CD275	1.719296	1.211862–2.439203	0.002391	B7 family
HHLA2	B7-H5, B7-H7	0.607967	0.41155–0.898127	0.012431	B7 family



**Figure 1** Risk score distribution and survival of patients in the training cohort. (A) The risk scores for 119 patients in the training cohort (GSE53624) are plotted in ascending order and marked as high risk (red) and low risk (green). (B) Survival of each patient in the cohort. Death is indicated by the color red, and alive patients are indicated by the color green. (C) Two genes in the B7-CD28 family expression distribution in the training cohort. Red: higher expression. Green: lower expression. (D, E) Kaplan–Meier curves of OS and RFS in 119 patients in the training cohort. (F, G) Kaplan–Meier curves of OS in LN+ and LN– patients.  $P < 0.05$  is regarded as statistically significant.

**Table 3** Univariable and multivariable Cox regression analysis of the signature and survival in the training cohort ( $N = 119$ ) and GSE53265 ( $N = 179$ ).

		Univariable analysis		Multivariable analysis	
		HR (95% CI)	P-value	HR (95% CI)	P-value
Training cohort					
Risk Score	High/Low	2.718 (1.655–4.466)	0.0001	2.723 (1.502–4.937)	0.0010
Age	>60/≤60	1.421 (0.896–2.252)	0.1352	1.127 (0.681–1.865)	0.6429
Sex	Female/Male	0.827 (0.468–1.461)	0.5126	1.003 (0.434–2.317)	0.9940
Tobacco Use	Y/N	0.860 (0.532–1.388)	0.5361	0.557 (0.287–1.083)	0.0845
Alcohol Use	Y/N	1.052 (0.656–1.688)	0.8333	1.901 (0.929–3.889)	0.0786
Tumor Location	Upper, middle/lower	1.104 (0.755–1.614)	0.6090	1.344 (0.870–2.077)	0.1825
Tumor Grade	Moderately differentiated, poorly/well differentiated	1.220 (0.855–1.739)	0.2733	1.117 (0.759–1.646)	0.5745
T stage	T3, T4/T1.T2	1.127 (0.839–1.515)	0.4277	1.074 (0.694–1.662)	0.7483
N stage	N1, N2, N3/N0	2.159 (1.319–3.535)	0.0022	1.197 (0.517–2.775)	0.6743
TNM stage	III/I, II	1.901 (1.226–2.948)	0.0041	1.460 (0.620–3.436)	0.3868
GSE53625					
Risk Score	High/Low	1.774 (1.273–2.472)	0.0007	1.693 (1.159–2.473)	0.0064
Age	>60/≤60	1.575 (1.073–2.312)	0.0203	1.365 (0.902–2.066)	0.1410
Sex	Female/Male	0.783 (0.489–1.252)	0.3066	1.068 (0.570–2.001)	0.8378
Tobacco Use	Y/N	0.749 (0.508–1.105)	0.1452	0.590 (0.349–0.995)	0.0478
Alcohol Use	Y/N	0.863 (0.588–1.269)	0.4553	1.302 (0.749–2.263)	0.3490
Tumor Location	Upper, middle/lower	1.253 (0.925–1.697)	0.1452	1.417 (1.008–1.990)	0.0447
Tumor Grade	Moderately differentiated, poorly/well differentiated	1.352 (1.002–1.823)	0.0482	1.211 (0.892–1.645)	0.2201
T Stage	T3, T4/T1.T2	1.187 (0.910–1.549)	0.2054	1.081 (0.732–1.596)	0.6948
N Stage	N1, N2, N3/N0	2.130 (1.420–3.193)	0.0002	1.402 (0.663–2.965)	0.3766
TNM Stage	III/I, II	1.994 (1.398–2.846)	0.0001	1.547 (0.728–3.290)	0.2565

### Biological process and pathway analysis of the signature

To investigate the biological features of ESCC that are related to the risk score, we investigated genes that strongly related to the risk score (Pearson  $|R| > 0.35$ ) in the training cohort. The details of the risk scores of the 622 most related genes for RFS status, OS status, age, pathology stage, and alcohol history are shown in Figure 3A. The significantly related genes were entered into the DAVID website for Gene Ontology analysis. The related genes were more involved in antigen processing and presentation of peptides or polysaccharide antigens via major histocompatibility complex (MHC) class II, T cell costimulation, and immune response (Fig. 3B). A KEGG pathway analysis was also used to find the common pathway of the involved genes. As a result, we found that genes participated more with Cell Adhesion Molecules (CAMs), the intestinal immune network for IgA production, *Staphylococcus aureus* infection, and asthma (Fig. 3C). These results indicated that immune-related biological pathways may play an important role in the underline mechanisms of the signature.

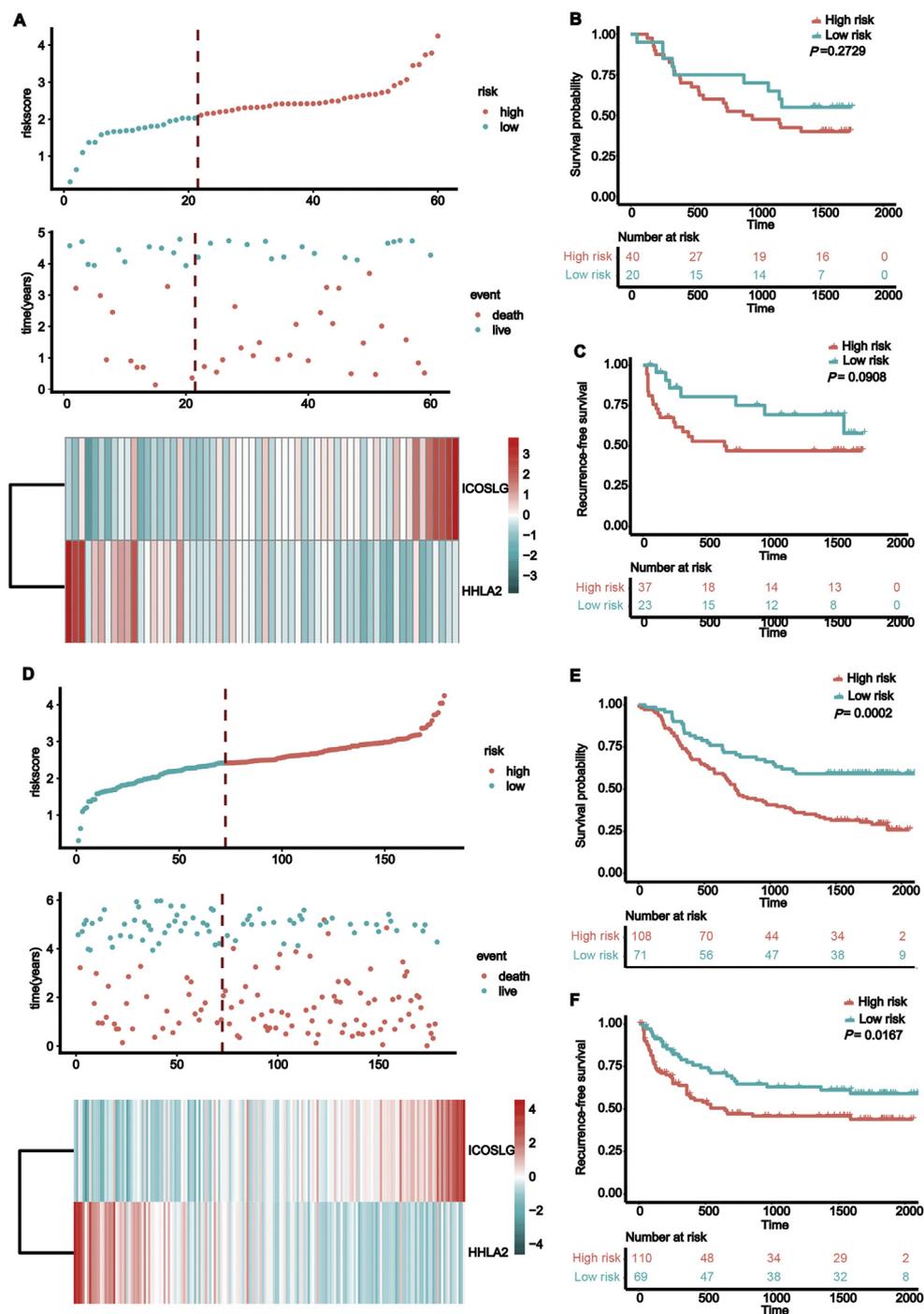
### Signature-specific immune cells infiltration

To further investigate the relationship between the immune landscape and the signature, xCell (<http://xCell.ucsf.edu/>) was used in the training cohort. This was

implemented using a single sample gene set enrichment analysis (ssGSEA) algorithm. This algorithm allows enrichment scores to be calculated for 64 immune and stromal cell types from gene expression profiles.<sup>23</sup> We used the xCell method to estimate infiltration of different immune cells in the high- and low-risk groups. Details pertaining to risk scores of immune cell infiltration with OS status, pathology status, sex, age, and alcohol history are shown in Figure 4A. Differential expression of the cells is shown in Figure 4B and C. The results indicated that several immune cells, such as Treg cells, CD4<sup>+</sup> T-cells, and CD8<sup>+</sup> T-cells, exhibited statistically significant infiltration patterns between the high- and low-risk groups. In particular, patients in the high-risk group showed a significantly higher proportion of Treg cells and fibroblasts. Foxp3 and  $\alpha$ -SMA are the specific biomarkers of Treg cells and fibroblasts, respectively.<sup>25,26</sup> Thus, to primary verify the estimated data, we stained the two markers by immunofluorescence assay in two corresponding tumor samples slices from the high- and low-risk groups. The pictured results include Case 1 (a high-risk patient) and Case 2 (a low-risk patient). The foxp3 is marked in red, and the  $\alpha$ -SMA is marked in green (Fig. 4D). This image confirms that high-risk patients exhibit high infiltration of Tregs and fibroblasts.

### Risk score related to inflammatory metagenes and immune checkpoints

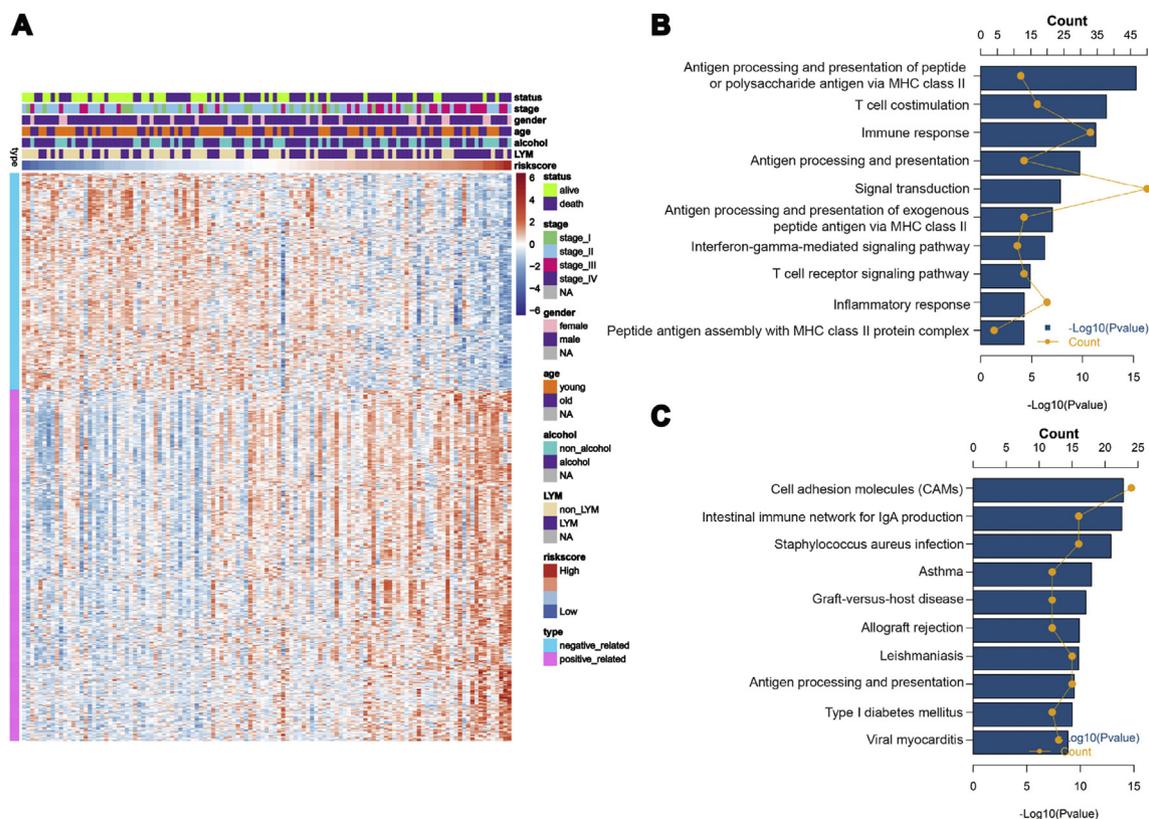
To explore the relationship between risk score and inflammatory activities, we defined seven clusters of 104 genes as



**Figure 2** Risk score distribution and survival of patients in the validation cohort. (A) The risk scores for 60 patients of GSE56322, the survival of each patient, and gene expression distribution in the validation cohort. (B, C) Kaplan–Meier curve of OS and RFS in validation cohort: (D) Risk scores for the entire cohort (GSE53265), the survival of each patient, and gene expression distribution (E, F) Kaplan–Meier curve of OS and RFS in the entire cohort.  $P < 0.05$  is regarded as statistically significant.

immune metagenes, which indicate different types of inflammatory and immune responses.<sup>27</sup> The details of these genes and the risk scores are shown in Fig. 5A. We then generated seven metagenes using the results of the Gene Sets Variation Analysis (GSVA) of corresponding clusters of genes and obtained a risk score expression of corrgram in the training cohort (Fig. 5B). The high-risk score was

positively associated with hematopoietic cell kinase (HCK), IgG, interferon, lymphocyte-specific protein-tyrosine kinase (LCK), MHC-I, MHC-II, and signal transducer and activator of transcription 1 (STAT1). Associations were closest between high-risk score and HCK, LCK, and MHC-II. This indicates that a high-risk score was associated with activation of macrophages and T cells signaling transduction in



**Figure 3** Relationship between risk score and most related genes and biological pathways. **(A)** Details of risk score and the most-related genes. **(B)** Gene enrichment with Go terms of the selected genes. **(C)** Gene enrichment with KEGG terms of the selected genes.

patients with ESCC. Other immune checkpoint members also played a vital role in the communication systems that regulate the antitumor immune response.<sup>28</sup> We detected a correlation between risk score and several immune checkpoints. B and T lymphocyte attenuator (BTLA) is a lymphocyte inhibitory receptor that is similar to CTLA-4 and PD-1. BTLA-deficient T cells showed increased proliferation, and BTLA-deficient mice had increased specific antibody responses.<sup>29</sup> Our study showed that a high score was tightly associated with BTLA (Fig. 5C). A previous study has demonstrated that CD27, a TNF receptor, activated costimulatory pathways of the T cell response. CD27 signaling also promoted tumor growth.<sup>30</sup> Our results showed that a high-risk score was also positively associated with CD27 (Fig. 5D). What's more, the expressions of tumor necrosis factor receptor superfamily member 4 (TNFRSF4) and Galectin 9 (LGALS9) were linked to the high-risk group (Fig. 5E, F).

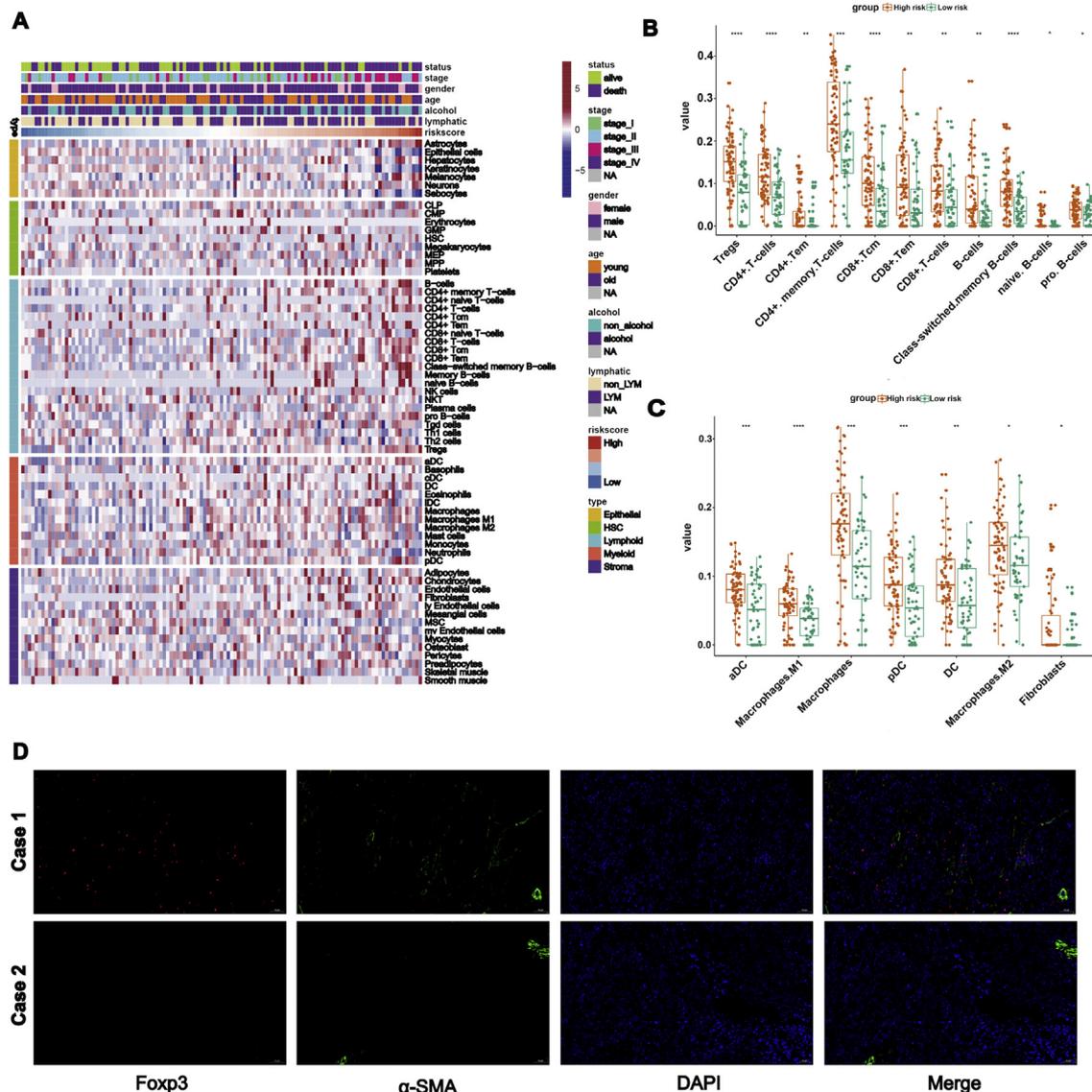
### The prognosis value of signature was validated in an independent cohort

Eighty-six patients formed an independent cohort to validate the prognostic value of the signature. The HHLA2 and ICOSLG expressions in the mRNA level were detected in each patient, and a risk score was then calculated (Fig. 6A). OS ( $P < 0.001$ ) and RFS ( $P = 0.0217$ ) were both shorter in the high-risk group when compared to the low-risk group

(Fig. 6B, C). Unsurprisingly, the high-risk group also showed shorter OS compared to the low-risk group for the LN+ and LN- subgroups, with  $P$ -values of 0.0012 and 0.0004, respectively. Importantly, the risk score was also an independent following use of a multivariate Cox proportional hazards regression model together with other clinical variables such as sex, age, and pathology stage (Table 4).

### Discussion

With the rapid development of genomic research of malignant tumors, various prognosis markers and therapeutic targets have been found and validated. ESCC remains a type of cancer with a relatively poor prognosis. The identification of ESCC biomarkers gives hope to patients, but the transfer of these biomarkers into clinical use has been difficult. Preclinical and clinical immunotherapy for ESCC has shown promising results. However, the specific therapy target and optimal patient selection methods remain unclear. As mentioned above, our team is committed to the creation of a signature to predict prognosis in patients with ESCC and thus determine candidacy for novel target therapy. The signature we built in this study was specific to ESCC for a Chinese population. Several signatures have been built to predict overall survival in patients with ESCC.<sup>31,32</sup> A B7-CD28-based signature is able to significantly predict prognosis in patients with lung adenocarcinoma (LUAD).<sup>33</sup> However, the B7-CD28-based gene signature was

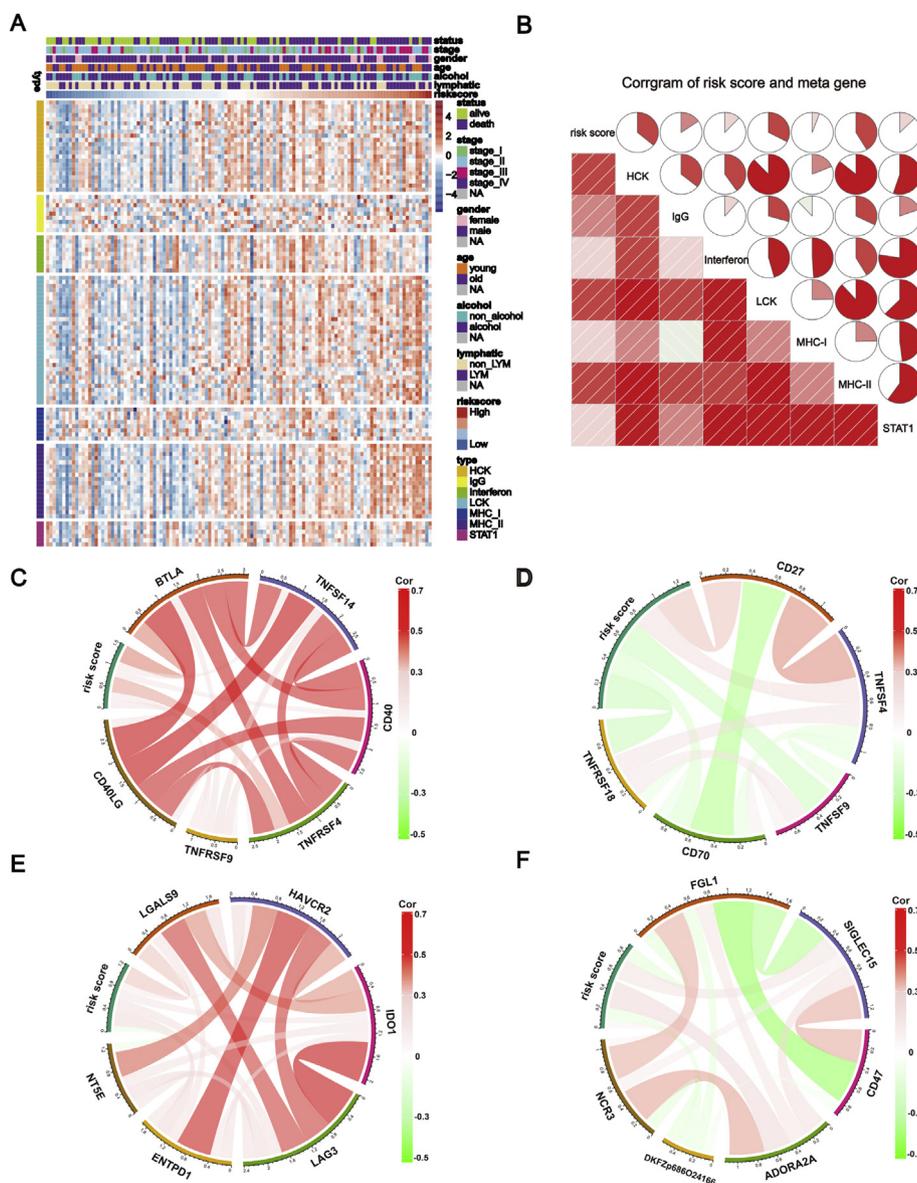


**Figure 4** Relationship between risk score and different cells estimated by xCell. **(A)** The cellular landscape of tumor immune microenvironment in high and low risk scores group of ESCC. **(B, C)** Different distribution of estimated cells that are in two groups. **(D)** Immunofluorescence image of Treg and fibroblasts in tissues in a patient from the high-risk group (Case 1) and low-risk group (Case 2). The foxp3 is marked as red, and  $\alpha$ -SMA is marked as green. (200 $\times$ ). \*, \*\*, \*\*\* and \*\*\*\* represent  $P < 0.05$ ,  $P < 0.01$ ,  $P < 0.001$  and  $P < 0.0001$ , respectively.

not able to predict prognosis in patients with ESCC. In this research, we improved upon current methods of analysis and incorporated additional patients with ESCC. Thus, we developed a novel B7-CD28 family gene prognostic signature. By analyzing the association between gene expression profiles and clinical outcomes in patients with ESCC, we identified a B7-CD28 family gene signature that exhibited a significant relationship to OS in patients with ESCC. The signature was closely associated with RFS and OS in different subsets of patients with ESCC, regardless if the patients were from the validation cohort or the independent cohort. We also found that the signature had a significant relationship to pathways of immune and lymphocyte infiltration in patients with ESCC. Patients in the high-risk score group also exhibited high expression of

critical immune checkpoints. Above all, this B7-CD28-based signature enhances our understanding of immunotherapy in patients with ESCC.

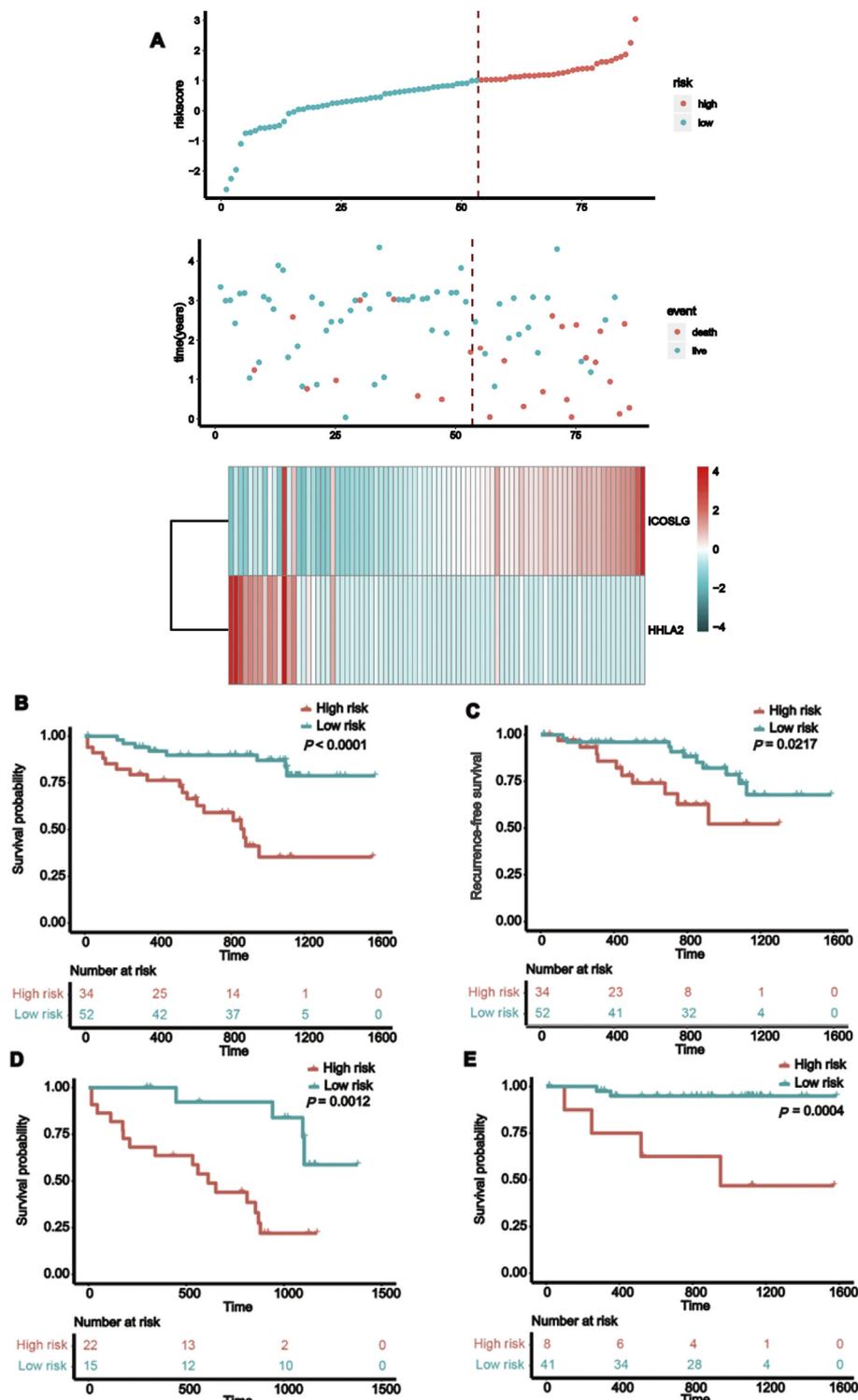
The B7-CD28 superfamily genes help activate critical immune responses. The ability of various tumors to withstand the immune response is one reason for poor prognosis and a low responsiveness to therapy. This observation inspired us to explore the prognostic signature using the B7-CD28 family genes. T cell activation is regulated by both positive and negative costimulatory molecules. Members of the B7 family are critical for T cell responses. The CD80/CD86-CD28/CTLA-4 pathway is involved in T cell activation and tolerance.<sup>34</sup> CD80 (B7-1) and CD86 (B7-2) provide costimulatory signals that augment and sustain the T cell response via their interaction with CD28.<sup>35,36</sup> CD28 transmits a signal that



**Figure 5** Relationship between risk scores and immune metagenes and immune checkpoints. (A) Relationship between risk scores and seven clusters of metagenes in the training cohort. (B) Corgrams were derived based on Pearson  $r$  value between risk values and seven clusters of metagenes in the training cohort. (C, D) Correlations between immune checkpoints from the TNF family and risk scores. (E, F) Correlation between immune checkpoints from other family members and risk scores.

synergizes with the TCR signal to promote T-cell activation. CD28 signaling also reduces the number of TCR, which are essential in T-cell activation.<sup>37</sup> CTLA-4, as an inhibitory receptor, can combine with the B7 family molecule due to its higher affinity than CD28.<sup>38</sup> CTLA-4 delivers a negative signal to inhibit the TCR and CD28 signals. The immune checkpoint inhibitor targeted to CTLA-4 is effective in several tumors.<sup>39</sup> ICOS is induced on T cells after the engagement of TCR, and its expression is related to CD28.<sup>40</sup> ICOSL, the ligand of ICOS is a potential target for cancer immunotherapy.<sup>41</sup> ICOSL overexpression is associated with tumor progress and poor overall survival.<sup>42,43</sup> Our findings also indicated the similar conclusion. ICOSL is a co-stimulatory signal for T-cell proliferation and cytokine secretion and plays an important role in mediating local response to inflammatory condition.<sup>44</sup> The

activation of ICOSL was found in gastric cancer to sustain immunosuppressive CD4<sup>+</sup> T cell subsets, especially Tregs.<sup>45</sup> Tregs are considered responsible for an immunosuppressive environment and tumor immune escape.<sup>46</sup> Dual use of vaccine and ICOSL blockade could deplete infiltrated Tregs with the possibility to enhance the vaccine-induced immunity.<sup>47</sup> The finding is accordance with our result that higher risk score cases were featured with higher infiltration of Tregs. HHLA2, also known as B7-H5, is a novel immune checkpoint member of the B7 family.<sup>10</sup> HHLA2 was first proposed in 2013 and serves as an inhibitory molecule in human CD4 and CD8 T-cell function.<sup>17</sup> HHLA2 has a higher expression in lung cancer,<sup>48</sup> colorectal carcinoma<sup>49</sup> and malignant glioma.<sup>19</sup> Yan H et al found that HHLA2 was detected in 77.17% of the pancreatic dual adenocarcinoma and associated with better



**Figure 6** Performance of B7-CD28 signature in the independent cohort with qPCR data. (A) The risk scores, the survival of each patient, and gene expression distribution in the cohort. (B, C) Kaplan–Meier curve of OS and RFS in the cohort. (D, E) Kaplan–Meier curve of OS in subgroups of LN+ and LN– patients.  $P < 0.05$  is regarded as statistically significant.

prognosis. Their findings suggest that HHLA2 may behave as a costimulatory ligand in pancreatic cancer.<sup>50</sup> What’s more, HHLA2 was found highly expressed in clear cell renal cell carcinoma predicts a favorable survival outcome.<sup>51</sup> These results indicate that HHLA2 might be a valuable biomarker in

various types of cancer. In our system, we found that HHLA2 was a protective factor with HR less than 1 and predicted a longer OS and RFS in ESCC. The role of HHLA2 in the immune environment of ESCC and its value of predicting prognosis in patients with ESCC is a potential trend for future exploration.

**Table 4** Univariable and multivariable Cox regression analysis of the signature and survival in the independent cohort ( $N = 86$ ).

		Univariable analysis		Multivariable analysis	
		HR (95% CI)	P-value	HR (95% CI)	P-value
Risk score	High/Low	3.239 (1.778–5.899)	0.0001	2.518 (1.391–4.558)	0.002284
Gender	Female/Male	1.078 (0.432–2.690)	0.8722	1.199 (0.454–3.162)	0.714195
Age	>60/≤60	2.964 (1.018–8.630)	0.0463	2.328 (0.784–6.910)	0.127882
TNM stage	III/I, II	2.089 (1.288–3.388)	0.0028	1.351 (0.641–2.841)	0.428379
N stage	N1, N2, N3/N0	5.361 (2.149–13.370)	0.0003	3.486 (1.170–10.385)	0.024932

The signature we built in this study indicated that a high-risk score was associated with poor prognosis in different subtypes of ESCC. The subsequent exploration of immune-related pathways and immune checkpoints provided potential therapeutic targets for high-risk patients. Rodziewicz-Motowidlo et al revealed that a higher ratio of BTLA/CD8 was an independent predictor of unfavorable outcomes in patients with GBC, and the upregulation of BTLA in cancer tissues was involved in the inhibition of antitumor immunity.<sup>52</sup> The higher expression of BTLA was also correlated with a higher level of PD-L1 and shorter RFS in patients with NSCLC.<sup>53</sup> The blockade of BTLA on mice, with downregulation of interleukin 6 and 10, enhanced the immune therapy of B lymphocytes in patients with epithelial ovarian carcinoma (EOC).<sup>54</sup> The clinical trial of an inhibitor of BTLA (TAB004) is ongoing (Tri No.: NCT04137900). Similarly, the CD27 expression level is associated with prognosis in patients with lung cancer<sup>55</sup> and colorectal carcinoma.<sup>56</sup> The anti-CD27 monoclonal antibody enhanced the antitumor efficacy of a dendritic cell-based vaccine in prostate tumor-bearing mice.<sup>57</sup>

Several limitations to this study warrant consideration. First, the risk score of the validation cohort was not significant, unlike the training cohort. However, the entire cohort exhibited statistical significance. This may have been due to the small sample size of the validation cohort; however, the whole-cohort results illustrate the specific performance of the signature. Second, our study focused on B7-CD28 super genes, not genes from the whole genome. Consequently, the predictive value of the signature may be limited. Finally, most of the relationship between risk score and the immune landscape was fulfilled by bioinformatics, which may be affected by some noise.

In conclusion, combined with our previous published microarray data and external validation data, we identified a novel, practical prognostic system based on the combination HHLA2 and ICOSLG (B7-CD28 signature) for ESCC. This signature divided patients into high- and low-risk groups with distinct immunity features, immune cell proportions, and immune checkpoint molecules. With prospective validation, these findings may enable a more accurate prediction of survival while increasing our understanding of the candidacy for immunotherapy in patients with ESCC.

### Author contributions

Jie He and Chaoqi Zhang conceived and designed the study. Chaoqi Zhang contributed to the conduction of related

statistical analysis. Feng Wang and Nan Sun drafted the manuscript and contributed to the Immunofluorescence of clinical samples. Zhen Zhang performed the validation in the independent cohort. Guochao Zhang, Zhihui Zhang and Yuejun Luo developed the inclusion criteria and normalized the expression profile data. Yun Che, Hong Cheng, Jiagen Li contributed to preparing the figures and tables. All authors reviewed the manuscript.

### Conflict of interests

The authors declare that they have no conflicts of interest.

### Funding

This work was supported by the CAMS Innovation Fund for Medical Sciences (No. 2017-I2M-1-005, 2016-I2M-1-001), the National Key R&D Program of China (No. 2016YFC1303201), the National Natural Science Foundation of China (No. 81802299, 81502514), the Fundamental Research Funds for the Central Universities (No. 3332018070), and the National Key Basic Research Development Plan (No. 2018YFC1312105).

### Acknowledgements

All authors would like to thank the specimen donors for this collection.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.gendis.2020.08.003>.

### References

1. Abnet CC, Arnold M, Wei WQ. Epidemiology of esophageal squamous cell carcinoma. *Gastroenterology*. 2018;154(2):360–373.
2. Arnold M, Soerjomataram I, Ferlay J, Forman D. Global incidence of oesophageal cancer by histological subtype in 2012. *Gut*. 2015;64(3):381–387.
3. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. *CA: Cancer Journal Clin*. 2016;66(2):115–132.
4. Allemani C, Matsuda T, Di Carlo V, et al. Patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet*. 2018;391(10125):1023–1075.

5. Chen Y, Hao D, Wu X, et al. Neoadjuvant versus adjuvant chemoradiation for stage II-III esophageal squamous cell carcinoma: a single institution experience. *Dis Esophagus*. 2017;30(7):1–7.
6. Chen Z, Li J, Tian L, et al. MiRNA expression profile reveals a prognostic signature for esophageal squamous cell carcinoma. *Cancer Lett*. 2014;350(1–2):34–42.
7. Li J, Chen Z, Tian L, et al. LncRNA profile study reveals a three-lncRNA signature associated with the survival of patients with oesophageal squamous cell carcinoma. *Gut*. 2014;63(11):1700–1710.
8. Thrumurthy SG, Chaudry MA, Thrumurthy SSD, Mughal M. Oesophageal cancer: risks, prevention, and diagnosis. *BMJ*. 2019;366:l4373.
9. Metges J, Francois E, Shah M, et al. The phase 3 KEYNOTE-181 study: pembrolizumab versus chemotherapy as second-line therapy for advanced esophageal cancer. *Ann Oncol*. 2019;30(suppl 4):iv130.
10. Janakiram M, Chinai JM, Zhao A, Sparano JA, Zang X. HHLA2 and TMIGD2: new immunotherapeutic targets of the B7 and CD28 families. *Oncoimmunology*. 2015;4(8):e1026534.
11. Schildberg FA, Klein SR, Freeman GJ, Sharpe AH. Coinhibitory pathways in the B7-CD28 ligand-receptor family. *Immunity*. 2016;44(5):955–972.
12. Mizuno R, Sugiura D, Shimizu K, et al. PD-1 primarily targets TCR signal in the inhibition of functional T cell activation. *Front Immunol*. 2019;10:630.
13. Batyrova B, Luwaert F. PD-1 expression affects cytokine production by ILC2 and is influenced by peroxisome proliferator-activated receptor- $\gamma$ . *Immunity*. 2020;8(1):8–23.
14. Hoeres T, Holzmann E, Smetak M, Birkmann J, Wilhelm M. PD-1 signaling modulates interferon-gamma production by Gamma Delta (gammadelta) T-Cells in response to leukemia. *Immun Inflamm Dis*. 2019;8(3):1550618.
15. Leung CS, Yang KY, Li X, et al. Single-cell transcriptomics reveal that PD-1 mediates immune tolerance by regulating proliferation of regulatory T cells. *Genome Med*. 2018;10(1):71.
16. Guo W, Zhang F, Shao F, et al. PD-L1 expression on tumor cells associated with favorable prognosis in surgically resected esophageal squamous cell carcinoma. *Hum Pathol*. 2019;84:291–298.
17. Zhao R, Chinai JM, Buhl S, et al. HHLA2 is a member of the B7 family and inhibits human CD4 and CD8 T-cell function. *Proc Natl Acad Sci USA*. 2013;110(24):9879–9884.
18. Flajnik MF, Tlapakova T, Criscitiello MF, Krylov V, Ohta Y. Evolution of the B7 family: co-evolution of B7H6 and Nkp30, identification of a new B7 family member, B7H7, and of B7's historical relationship with the MHC. *Immunogenetics*. 2012;64(8):571–590.
19. Qi Y, Deng G, Xu P, et al. HHLA2 is a novel prognostic predictor and potential therapeutic target in malignant glioma. *Oncol Rep*. 2019;42(6):2309–2322.
20. Wang B, Ran Z, Liu M, Ou Y. Prognostic significance of potential immune checkpoint member HHLA2 in human tumors: a comprehensive analysis. *Front Immunol*. 2019;10:1573.
21. Liu D, Xu H, Shih C, et al. T-B-cell entanglement and ICOSL-driven feed-forward regulation of germinal centre reaction. *Nature*. 8 2015;517(7533):214–218.
22. Maazi H, Patel N, Sankaranarayanan I, et al. ICOS: ICOS-ligand interaction is required for type 2 innate lymphoid cell function, homeostasis, and induction of airway hyperreactivity. *Immunity*. 2015;42(3):538–551.
23. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol*. 2017;18(1):220.
24. Kakegawa T. Forty years' experience in surgical treatment for esophageal cancer. *Int J Clin Oncol*. 2003;8(5):277–288.
25. Korn T, Muschaweckh A. Stability and maintenance of Foxp3(+) Treg cells in non-lymphoid microenvironments. *Front Immunol*. 2019;10:2634.
26. Zhan S, Liu Z, Zhang M, et al. Overexpression of B7-H3 in alpha-SMA-positive fibroblasts is associated with cancer progression and survival in gastric adenocarcinomas. *Front Oncol*. 2019;9:1466.
27. Rody A, Holtrich U, Pusztai L, et al. T-cell metagene predicts a favorable prognosis in estrogen receptor-negative and HER2-positive breast cancers. *Breast Cancer Res*. 2009;11(2):R15.
28. Nurieva R, Thomas S, Nguyen T, et al. T-cell tolerance or function is determined by combinatorial costimulatory signals. *EMBO J*. 2006;25(11):2623–2633.
29. Watanabe N, Gavrieli M, Sedy JR, et al. BTLA is a lymphocyte inhibitory receptor with similarities to CTLA-4 and PD-1. *Nat Immunol*. 2003;4(7):670–679.
30. Claus C, Riether C, Schurch C, Matter MS, Hilmenyuk T, Ochsenbein AF. CD27 signaling increases the frequency of regulatory T cells and promotes tumor growth. *Cancer Res*. 2012;72(14):3664–3676.
31. Kim HS, Lee SE, Bae YS, et al. PIK3CA amplification is associated with poor prognosis among patients with curatively resected esophageal squamous cell carcinoma. *Oncotarget*. 2016;7(21):30691–30701.
32. Sun LL, Wu JY, Wu ZY, et al. A three-gene signature and clinical outcome in esophageal squamous cell carcinoma. *Int J Cancer*. 2015;136(6):E569–E577.
33. Zheng S, Luo X, Dong C, et al. A B7-CD28 family based signature demonstrates significantly different prognoses and tumor immune landscapes in lung adenocarcinoma. *Int J Cancer*. 2018;143(10):2592–2601.
34. Chambers CA, Kuhns MS, Egen JG, Allison JP. CTLA-4-mediated inhibition in regulation of T cell responses: mechanisms and manipulation in tumor immunotherapy. *Annu Rev Immunol*. 2001;19:565–594.
35. Lenschow DJ, Walunas TL, Bluestone JA. CD28/B7 system of T cell costimulation. *Annu Rev Immunol*. 1996;14:233–258.
36. Rudulier CD, McKinstry KK, Al-Yassin GA, Kroeger DR, Bretscher PA. The number of responding CD4 T cells and the dose of antigen conjointly determine the TH1/TH2 phenotype by modulating B7/CD28 interactions. *J Immunol*. 2014;192(11):5140–5150.
37. Lanzavecchia A, Iezzi G, Viola A. From TCR engagement to T cell activation: a kinetic view of T cell behavior. *Cell*. 1999;96(1):1–4.
38. Pentcheva-Hoang T, Egen JG, Wojnoonski K, Allison JP. B7-1 and B7-2 selectively recruit CTLA-4 and CD28 to the immunological synapse. *Immunity*. 2004;21(3):401–413.
39. Rowshanravan B, Halliday N, Sansom DM. CTLA-4: a moving target in immunotherapy. *Blood*. 2018;131(1):58–67.
40. Wikenheiser DJ, Stumhofer JS. ICOS Co-stimulation: friend or foe? *Front Immunol*. 2016;7:304.
41. Marinelli O, Nabissi M, Morelli MB, Torquati L, Amantini C, Santoni G. ICOS-L as a potential therapeutic target for cancer immunotherapy. *Curr Protein Pept Sci*. 2018;19(11):1107–1113.
42. Tamura H, Dan K, Tamada K, et al. Expression of functional B7-H2 and B7.2 costimulatory molecules and their prognostic implications in de novo acute myeloid leukemia. *Clin Cancer Res*. 2005;11(16):5708–5717.
43. Faget J, Bendriss-Vermare N, Gobert M, et al. ICOS-ligand expression on plasmacytoid dendritic cells supports breast cancer progression by promoting the accumulation of immunosuppressive CD4+ T cells. *Cancer Res*. 2012;72(23):6130–6141.
44. Wang F, Zhu W, Liu T, et al. The expression analysis of ICOS-L on activated T cells and immature dendritic cells as well as malignant B cells and Grave's-disease-derived thyroid tissues

- by two novel mAbs against human ICOS-L. *Tissue Antigens*. 2007;69(1):62–72.
45. Nagase H, Takeoka T, Urakawa S, et al. ICOS(+) Foxp3(+) TILs in gastric cancer are prognostic markers and effector regulatory T cells associated with *Helicobacter pylori*. *Int J Cancer*. 2017;140(3):686–695.
  46. Yang L, Li A, Lei Q, Zhang Y. Tumor-intrinsic signaling pathways: key roles in the regulation of the immunosuppressive tumor microenvironment. *J Hematol Oncol*. 2019;12(1):125.
  47. Mo L, Chen Q, Zhang X, et al. Depletion of regulatory T cells by anti-ICOS antibody enhances anti-tumor immunity of tumor cell vaccine in prostate cancer. *Vaccine*. 2017;35(43):5932–5938.
  48. Janakiram M, Pareek V, Cheng H, Narasimhulu DM, Zang X. Immune checkpoint blockade in human cancer therapy: lung cancer and hematologic malignancies. *Immunotherapy*. 2016;8(7):809–819.
  49. Zhu Z, Dong W. Overexpression of HHLA2, a member of the B7 family, is associated with worse survival in human colorectal carcinoma. *OncoTargets Ther*. 2018;11:1563–1570.
  50. Yan H, Qiu W, Koehne de Gonzalez AK, et al. HHLA2 is a novel immune checkpoint protein in pancreatic ductal adenocarcinoma and predicts post-surgical survival. *Cancer Lett*. 2019;442:333–340.
  51. Zhang Z, Liu J, Zhang C, et al. Over-expression and prognostic significance of HHLA2, a new immune checkpoint molecule, in human clear cell renal cell carcinoma. *Front Cell Dev Biol*. 2020;8:280.
  52. Rodziewicz-Motowidlo S, Oguro S, Ino Y, et al. Clinical significance of tumor-infiltrating immune cells focusing on BTLA and Cbl-b in patients with gallbladder cancer. *PLoS One*. 2015;10(12):1750–1760.
  53. Li X, Xu Z, Cui G, Yu L, Zhang X. BTLA expression in stage I-III non-small-cell lung cancer and its correlation with PD-1/PD-L1 and clinical outcomes. *OncoTargets Ther*. 2020;13:215–224.
  54. Chen YL, Lin HW, Chien CL, et al. BTLA blockade enhances Cancer therapy by inhibiting IL-6/IL-10-induced CD19(high) B lymphocytes. *J Immunother Cancer*. 2019;7(1):313.
  55. Kashima J, Okuma Y, Hosomi Y, Hishima T. High serum soluble CD27 level correlates with poor performance status and reduced survival in patients with advanced lung cancer. *Oncology*. 2019;97(6):365–372.
  56. Li L, Han C, Chen FX, Lu XT, Liu JQ, Fei SJ. Expression of CD27, CD28 and IL-17A in peripheral blood from patients with colorectal carcinoma. *Eur Rev Med Pharmacol Sci*. 2016;20(4):642–651.
  57. Wei SM, Fei JX, Tao F, et al. Anti-CD27 antibody potentiates antitumor effect of dendritic cell-based vaccine in prostate cancer-bearing mice. *Int Surg*. 2015;100(1):155–163.