

RESEARCH

Open Access



# A robust immune-related gene pairs signature for predicting the overall survival of esophageal cancer

Wei Zheng<sup>1</sup>, Gaofeng Fang<sup>2</sup>, Qiao Huang<sup>3</sup>, Dan Shi<sup>2\*</sup> and Biao Xie<sup>4\*</sup>

## Abstract

**Background** Identifying reliable biomarkers could effectively predict esophagus carcinoma (EC) patients with poor prognosis. In this work, we constructed an immune-related gene pairs (IRGP) signature to evaluate the prognosis of EC.

**Results** The IRGP signature was trained by the TCGA cohort and validated by three GEO datasets, respectively. Cox regression model together with LASSO was applied to construct the overall survival (OS) associated IRGP. 21 IRGPs consisting of 38 immune-related genes were included in our signature, according to which patients were stratified into high- and low-risk groups. The results of Kaplan-Meier survival analyses indicated that high-risk EC patients had worse OS than low-risk group in the training set, meta-validation set and all independent validation datasets. After adjustment in multivariate Cox analyses, our signature continued to be an independent prognostic factor of EC and the signature-based nomogram could effectively predict the prognosis of EC sufferers. Besides, Gene Ontology analysis revealed this signature is related to immunity. 'CIBERSORT' analysis revealed the infiltration levels of plasma cells and activated CD4 memory T cells in two risk groups were significantly different. Ultimately, we validated the expression levels of six selected genes from IRGP index in KYSE-150 and KYSE-450.

**Conclusions** This IRGP signature could be applied to select EC patients with high mortality risk, thereby improving prospects for the treatment of EC.

**Keywords** Esophageal cancer, Prognosis, Immune-related gene pairs, Bioinformatic analysis, Cells lines

\*Correspondence:

Dan Shi  
danshi@cqmu.edu.cn  
Biao Xie  
kybiao@cqmu.edu.cn

<sup>1</sup>Department of Cardiology, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China

<sup>2</sup>Department of Nutrition and Food Hygiene, Chongqing Medical University, Chongqing, China

<sup>3</sup>Anatomy Teaching and Research Section, Basic department, Chongqing Medical and Pharmaceutical College, Chongqing, China

<sup>4</sup>Department of Biostatistics, Chongqing Medical University, Chongqing, China

## Background

The new Esophageal cancer (EC) patients and deaths exceed 1,000,000 a year [1, 2]. The treatment for EC typically focuses on radiotherapy, chemotherapy, chemoradiotherapy endoscopic resection and surgery. Despite the advancements in diagnosis and treatment in recent years have improved the clinical outcomes of EC patients, the prognosis remained poor. The 5-year survival rate for EC is less than 20% [3]. Thus, identifying reliable prognostic biomarkers is necessary for the treatment of EC.

The immunity system is relevant to tumor formation and progression. The tumor microenvironment and



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

immune cells participated in the initiation and deterioration of cancers [4, 5]. Immunotherapy has become an effective antitumor strategy of various cancers [6–8]. Represented by programmed death ligand-1, immune checkpoint inhibitors (ICIs) can restrain tumor growth and improve overall survival (OS) through inhibiting the immune escape [9, 10]. Esophageal tumor cells possess the high potential of immune escape because EC is characterized by high heterogeneity and high tumor mutational burden [11]. A few preclinical studies and clinical trials researchers have found that the ICIs pembrolizumab, nivolumab, camrelizumab and toripalimab can be used as a form of immunotherapy for EC [12–15]. However, immunotherapy for EC leads to mixed results. Indeed, several patients were shown to deteriorate sharply after receiving immunotherapy [16].

Recently, researchers have attempted to study the association between immune-related genes (IRGs) and the prognosis of EC [17–21]. For example, Wang and Li et al. ascertained the prognostic value of IRGs and infiltrating immune cells for EC patients [20, 21]. Some researchers used advanced computational tools. For example, Zheng et al. identified predictive and prognostic factors for esophageal squamous cell carcinoma based on Similarity Network Fusion and Consensus Clustering Analysis [22]. A study from China identified CLDN6 as a molecular biomarker in pan-cancer through multiple omics data integrative analysis [23]. Zhang et al. explored the role of YTH domain containing 2 in epigenetic modification and immune infiltration of pan-cancer [24]. Unsatisfactorily, none of these signatures could be translated to clinical application due to issues such as technical bias and biological heterogeneity of different platforms [25, 26]. To eliminate the disadvantage of normalization and scale transformation, researchers have developed a method on the basis of relative rank of gene expression value. This new method has been shown to produce robust signatures in prognostic stratification of various tumors [27–30]. In our study, an immune-related gene pairs (IRGP) signature was conducted to estimate the OS of EC.

## Materials and methods

### Public datasets and study design

The mRNA expression and corresponding clinicopathologic data of EC patients in the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) were acquired. The detail information of these cohorts is shown in Table S1. Altogether, four studies were enrolled, including TCGA cohort (n=170), GSE13898 (n=60), GSE19417 (n=70) and GSE52625 (n=179). Patients from TCGA cohort were allocated to the training set and all the patients from three GEO datasets were assigned to the meta-validation cohort. Patients without complete survival information and clinicopathologic information

were excluded. In total, 479 patients were included in this study. 2483 IRGs were downloaded from ImmPort (<https://immport.niaid.nih.gov>) on 2022.05.01.

### Gene expression data processing

The function ‘normalizeBetweenArrays’ of R package ‘limma’ was first applied for the normalization of gene expression profiles, and then log-transformation was utilized for further normalization.

### Construction of predictive-related IRGP

Among the 2483 IRGs, only those existed on all platforms and with a relatively high median absolute deviation ( $MAD > 0.5$ ) were selected. The IRGP score was obtained by contrasting the mRNA level of two paired genes in each sample. An IRGP score is assigned to 1 if the value of IRG1 was higher than that of IRG2, conversely, the IRGP score was 0 [20]. IRGP with steady value in datasets were removed. The left IRGP served as initial candidates for further analyses.

Prognostic IRGPs were chosen by the following steps. First, the association between each initial candidate IRGP and patients’ OS in the training dataset was assessed by the univariate Cox regression model. After rough filtration, the LASSO was further carried out to minimize overfitting. To improve the robustness of selected IRGPs, we randomly divided the training set into new training sets and test sets through a ratio of 2:1, and then duplicated this procedure 30 times. The LASSO was then used in the 30 training sets to choose those IRGPs with a frequency  $> 15$ . Ultimately, we adopted the multivariate Cox regression analysis to construct the IRGP-based model, which consisted of relevant IRGP index (IRGPI) and respective coefficients. The appropriate cut-off of IRGPI for the high- and low- risk groups was determined by receiver operating characteristic (ROC) curve at 3 years.

### Validation of IRGPI

The accuracy of the IRGPI was evaluated in the training cohort, independent validation cohorts and meta-validation cohort by log-rank and ROC curve analyses. To explore whether the IRGPI was an independent variable, it was combined with other clinicopathologic factors (age, gender, tumor stage, smoking habits and alcohol intake) in multivariate Cox analysis. The prognostic value of IRGPI was further verified by log-rank analysis in different subgroups on the basis of clinical characteristics including age, sex and stage.

### Tumor-infiltrating immune cells

CIBERSORT analysis associated with reference mRNA expression values (LM22) were utilized to appraise the relative infiltrating level of 22 immune cells [31]. R package ‘CIBERSORT’ was utilized to analyze. Samples with

$p$ -value $<0.05$  were remained for the following analysis. The discrepant abundances of different immune cells between the two risk groups were contrasted through the Wilcoxon rank test.

#### Functional enrichment

Gene ontology (GO) enrichment analysis was implemented with the R package ‘cluster-Profiler’ to explore potential biological processes and enrichment pathways of the IRGPI. The thresholds were determined by a false discovery rate adjusted  $P$ -value (FDR) $<0.05$ .

#### Construction of an IRGPI-based predictive nomogram

Using IRGPI and clinicopathologic factors (age, gender, tumor stage, smoking habits and alcohol intake), a nomogram was employed to evaluate prognostic value at 1-, 2- and 3- year survival and the prognostic value was also verified in the meta-validation cohort. The risk score of each factor was computed, and the total score of all factors was taken as the sum of each factor’s risk score. Ultimately, we generated calibration plots to appraise its predictive performance.

#### Calculation of tumor immune dysfunction and exclusion (TIDE) score and small molecule drug analysis

The TIDE score (<http://tide.dfci.harvard.edu/>, accessed on 27 May 2023) was used to evaluate if there were different ICI treatment responses in patients between high-risk and low-risk groups. The CMAP (Connectivity Map; version 1.1.1.43; <https://clue.io>) database has comprehensive data on the transcriptome of drug interference therapies. We identified potential effective drugs in treating EC through this tool (accessed on 28 May 2023).

#### Cell culture

EC cell line KYSE-150 (TCHu236) was purchased from the Cell Bank of the Chinese Academy of Sciences, Shanghai, China. EC cell line KYSE-450 (GDC0633) was purchased from China Center for Type Culture Collection, Wuhan, China. Normal esophageal epithelial cell line Het-1A (CRL-2692) was acquired from American Type Culture Collection. EC cells lines (KYSE-150 and KYSE-450) were cultured with 1640 medium (RPML, 11875093, Gibco, USA) supplemented with 10% fetal bovine serum (10099141C, Gibco, USA). Normal esophageal epithelial cells (Het-1A) were treated with bronchial epithelial cell basal medium (CC-3170, Lonza, Switzerland). All cell lines were placed in an incubator which contains 5% CO<sub>2</sub> at 37°C.

#### Quantitative real-time PCR

As described by our previous research (PMID: 33440166), total RNA was extracted with TRIzol reagent (9109, Takara, Japan). RNA was reverse transcribed

to cDNA by the High-Capacity cDNA Reverse Transcription Kit (RR047A-5, Takara, Japan). The transcription levels of *IL13RA2*, *RORC*, *IL20*, *SAA2*, *FABP6*, *CHGB* were measured by Real-time PCR, SYBR Green PCR Master Mix (HY-KO511, Takara, Japan). Data were analyzed by normalization to the mRNA level of human  $\beta$ -actin. The gene primers were as follows: human  $\beta$ -actin: F primer, CCTGGCACCCAG-CACAAT, R primer, GGGCCGGACTCGTCATAC; *IL13RA2*: F primer, AAAGTTCAGGA-TATGGATT-GCGT, R primer, GAAGTACACCTATGCCAGGTTTC; *RORC*: F primer, AGATACCCTCACCTACACCTTG, R primer, CCGCTCAGGGCTGTATTCAA; *IL20*: F primer, ATGAAAGCCTCTAGTCTTGCCT, R primer, GCCCGTATCTCAGAAAATCC; *SAA2*: F primer, GCTTCTTTTCGTTCCCTGGCG, R primer, GCCGATGTAATT-GGCTTCTCTCA; *FABP6*: F primer, ACCGGCAAGTTCGAGATGG, R primer, CCTTTTCGATTACATCGCTGGA; *CHGB*: F primer, CAGCCAACGCTGCTTCTCA, R primer, GGTTCCT-GTTATCCACTGGCA. Three biological replicates were conducted during our qRT-PCR experiment.

#### Statistical analysis

All analyses were performed with R software (version 4.0.2). Log-rank test was utilized to contrast the OS between different risk groups. The Cox regression model was adopted to evaluate the prognostic role of IRGPI. For all analysis, two-sided  $p$ -value $<0.05$  was considered as significant.

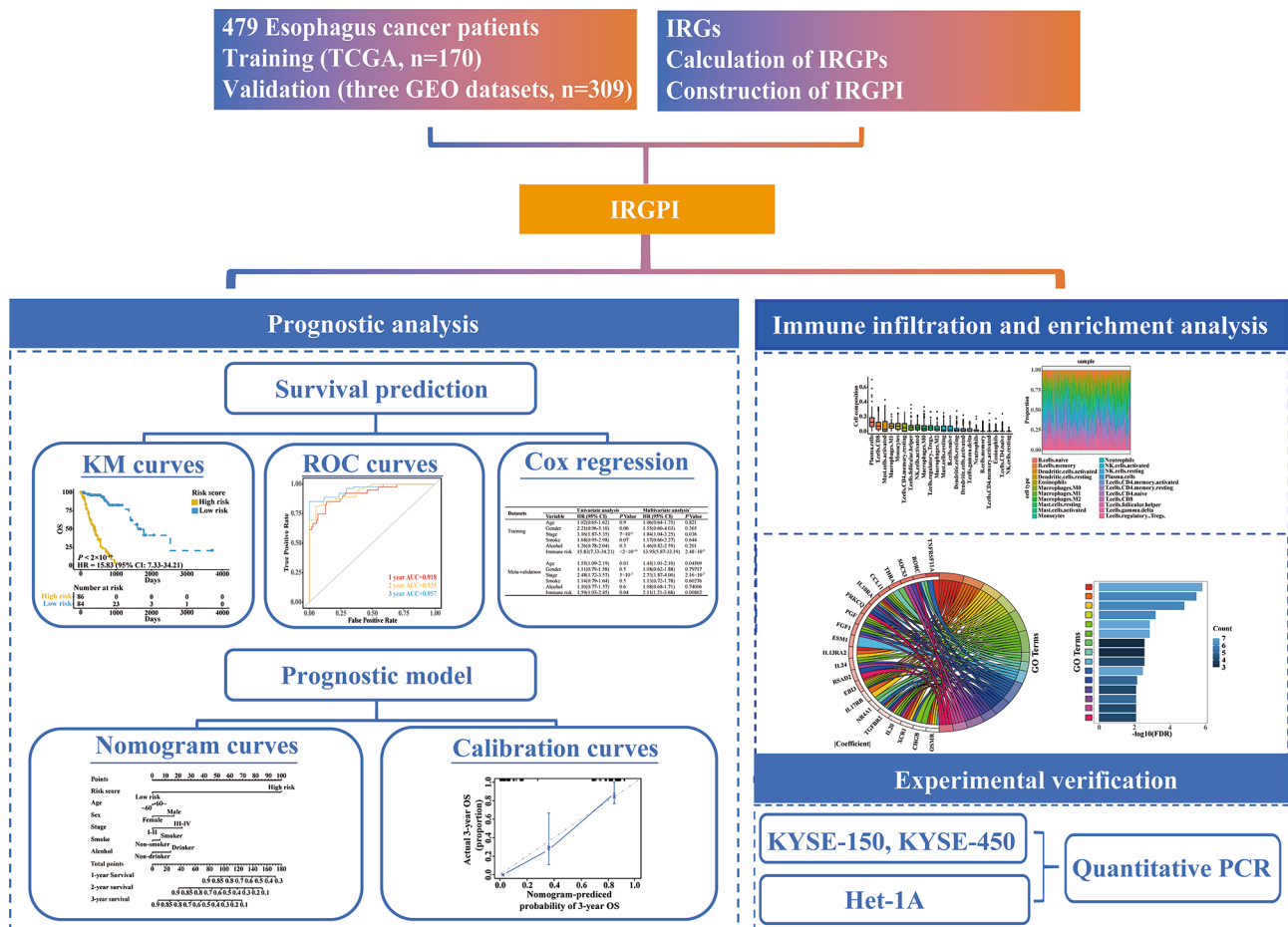
## Results

#### Construction of IRGPI

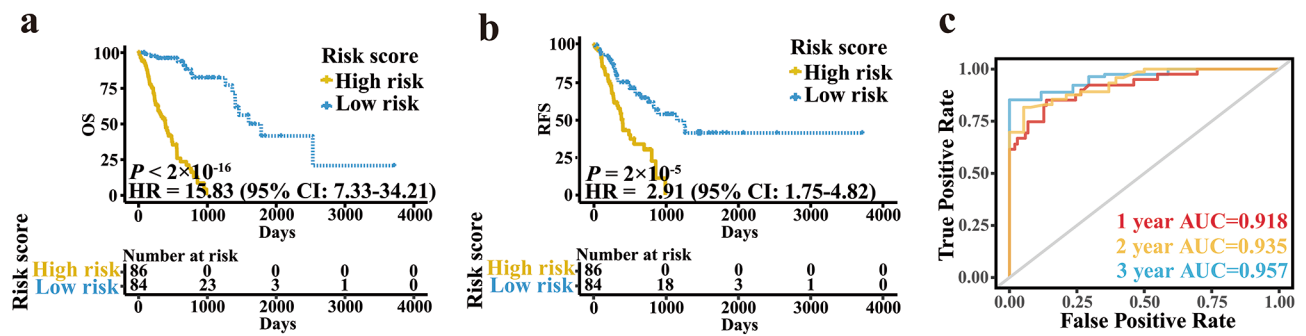
In total, 479 EC patients were involved in the study (Table 1). Patient data (170 EC patients) of TCGA dataset were allocated to the training set. Patient data (309 EC patients) from three independent GEO datasets were allocated to the meta-validation cohort (Fig. S1). The overview design in this study was shown in Fig. 1. In total, 694 of the 2483 IRGs measured by all platforms were considered to meet the criteria (MAD $>0.5$ ). Based on these 694 IRGs, 240,471 IRGP were established. After removing 223,503 IRGP with constant ordering and those were not shared in all tested datasets, 16,968 IRGPs were retained as initial candidates for subsequent analysis. The univariate Cox regression model was implemented to accomplish the rough filtration. 912 IRGPs selected by Cox regression model were further filtered out by LASSO analysis. Finally, 21 IRGPs that arose more than 15 times out of 30 LASSO analyses were selected. The 21 IRGPs consisted of 38 unique IRGs (Table S2). The results of multivariate Cox regression were implemented to confirm the IRGP-based prediction and generate IRGPI scores in the training cohort. IRGPI=1.07740610918684\*RSAD2\_

**Table 1** Patients' demographics and clinicopathologic characteristics in different cohorts

	Training cohort	Validation cohorts		
	TCGA (n=170)	GSE13898(n=60)	GSE19417(n=70)	GSE53625(n=179)
Age.mean	63.11 ± 11.62	63.17 ± 11.63	-	59.35 ± 9.03
Sex				
Male	144(84.71%)	55(91.67%)	46(65.71%)	146(81.56%)
Female	26(15.29%)	5(8.33%)	24(34.29%)	33(18.44%)
Pathologic stage				
I	18(10.59%)	23(38.33%)	-	10(5.59%)
II	72(42.35%)	16(26.67%)	-	77(43.02%)
III	50(29.41%)	4(6.67%)	-	92(51.40%)
IV	9(5.29%)	4(6.67%)	-	0(0%)
Unknown	21(12.35%)	13(21.67%)	-	0(0%)
Smoke				
Yes	152(89.41%)	43(71.67%)	-	114(63.69%)
No	17(10.00%)	17(28.33%)	-	65(36.31%)
Unknown	1(0.59%)	0(0%)	-	0(0%)
Alcohol				
Yes	117(68.82%)	46(76.67%)	-	106(59.22%)
No	50(29.41%)	14(23.33%)	-	73(40.78%)
Unknown	3(1.76%)	0(0%)	-	0(0%)



**Fig. 1** The overview design in this study



**Fig. 2** The Kaplan-Meier curves and time-dependent ROC curves in patients of training cohort. **(a)** OS in patients of training dataset classified by the signature. **(b)** RFS in patients of training dataset classified by the signature. **(c)** The time-dependent ROC curves of the IRGPI for 1-, 2- and 3-year OS of patients in training cohort

CRABP1 + 1.80900713722625 \* CCL11\_KLRC1 + 1.0760980156295 \* HCST\_EBI3 + 1.52141607156199 \* CLDN4\_PGF + 0.342479305941611 \* PPP3CCESM1 + 1.26035610621237 \* ROBO2ESM1 + 0.455689539588702 \* TGFBR2NR4A1 + 0.158188766234162 \* CTSG\_CHGB + 0.714247087056512 \* MC1R\_PROC + 0.327149276848437 \* XCR1\_IL20 - 1.47563060474962 \* ITGAL\_FGF1 - 0.0736174398078415 \* SOCS3\_OSMR - 0.791568426985749 \* FABP6\_OPRL1 - 1.3977742637501 \* NDRG1\_CSRP1 + 1.2367977698673 \* SAA2\_CRABP1 + 0.461612771669907 \* SLPI\_IL17RB - 2.51795133845012 \* SOCS3\_TNFRSF11A - 1.82699508603865 \* THRA\_RORC + 0.230471766652905 \* HMOX1\_ESM1 - 1.67908325639079 \* IL10RA\_PRKCQ - 1.09793933631461 \* IL24\_IL13RA2. The detailed IRGPI construction process was shown in Fig. S2.

#### Identification of the prognostic ability of IRGPI

The ROC curve analysis at 3 years was utilized to choose the optimal cut-off by which EC patients were split into two risk groups. To evaluate the clinical outcomes of the two risk groups, a log-rank test was implemented and the Kaplan-Meier curves indicated that the OS of EC patients in the high-risk group were worse than that in the low-risk group (Fig. 2a, HR 15.83, 95% CI 7.33–34.21;  $P < 2 \times 10^{-16}$ ). In addition, the relapse-free survival (RFS) of patients was also calculated and the result was similar to that of OS (Fig. 2b, HR 2.91, 95% CI 1.75–4.82;  $P = 2 \times 10^{-5}$ ). To identify the accuracy of the IRGPI in prediction, time-dependent ROC curves were analyzed. The area under the curve (AUC) of predicting 1-, 2- and 3-year OS in EC patients was 0.918, 0.935 and 0.957, respectively (Fig. 2c). We also evaluated the predictive ability of IRGPI in various subgroups of patients, which

**Table 2** Identifying the independent prognostic factors through univariate and multivariate Cox analyses

Cohorts	Variable	Univariate analysis		Multivariate analysis		
		HR (95% CI)	P Value	HR (95% CI)	P Value	
Training	Age	1.02(0.65–1.62)	0.9	1.06(0.64–1.75)	0.821	
	Gender	2.21(0.96–5.10)	0.06	1.55(0.60–4.03)	0.365	
	Stage	3.16(1.87–5.35)	$7 \times 10^{-6}$	1.84(1.04–3.25)	0.036	
	Smoke	1.68(0.95–2.98)	0.07	1.17(0.60–2.27)	0.644	
	Alcohol	1.26(0.78–2.04)	0.3	1.46(0.82–2.59)	0.201	
	Immune risk	15.83(7.33–34.21)	$< 2 \times 10^{-16}$	13.95(5.87–33.19)	$2.48 \times 10^{-9}$	
	Meta-validation	Age	1.55(1.09–2.19)	0.01	1.45(1.01–2.10)	0.04569
		Gender	1.11(0.79–1.58)	0.5	1.08(0.62–1.88)	0.79717
Stage		2.48(1.72–3.57)	$5 \times 10^{-7}$	2.73(1.87–4.00)	$2.16 \times 10^{-7}$	
Smoke		1.14(0.79–1.64)	0.5	1.13(0.72–1.78)	0.60276	
Alcohol		1.10(0.77–1.57)	0.6	1.08(0.68–1.71)	0.74006	
Immune risk		1.59(1.03–2.45)	0.04	2.11(1.21–3.68)	0.00862	

were stratified according to different tumor stage (early stage and late stage), gender (female and male) and age ( $\leq 60$  and  $> 60$ ). IRGPI also displayed effective predictive value in all these subgroups (Fig. S3). The results of multivariate Cox regression indicated that IRGPI was an independent prognostic factor when age, sex, stage, smoking habits and alcohol intake were adjusted (Table 2,  $P = 2.48 \times 10^{-9}$ ). Similarly, tumor stage was also an independent prognostic factor (Table 2). We further assessed the prognostic ability of the combination of IRGPI and tumor stage. As a result, patients with high IRGPI scores



and early stages had the highest OS among all patients (Fig. S4a).

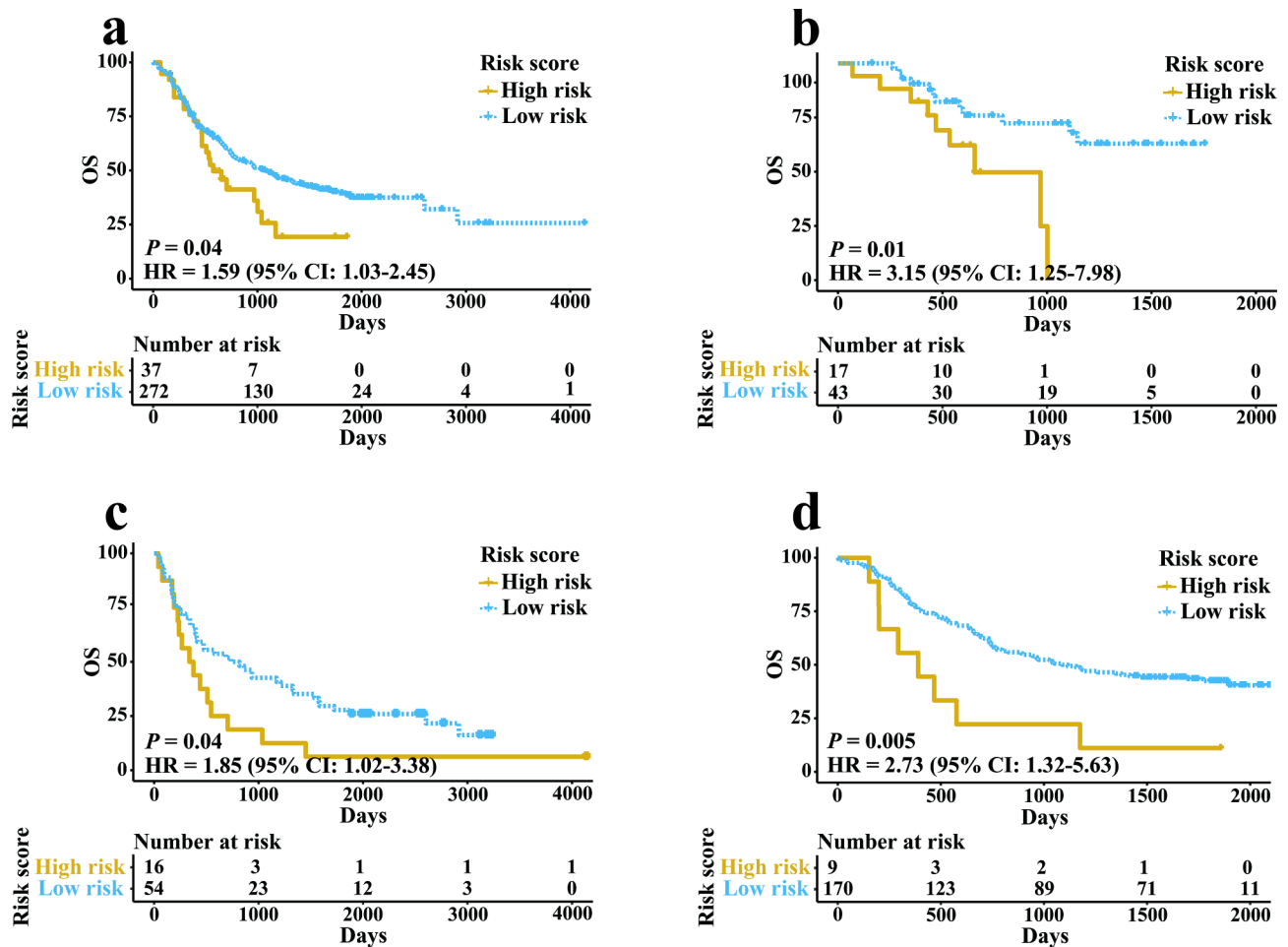
### Verification of the prognostic ability of IRGPI

We merged the three independent GEO datasets (GSE13898, GSE19417 and GSE53625) to a meta-validation cohort. Patients in the meta-validation set and each independent validation set were assigned to high- and low-risk groups on the basis of the result of ROC curves. The results of log-rank analysis were consistent with the findings from the training datasets. The OS in EC patients of the high-risk group were worse than that of the low-risk group in meta-validation cohort (Fig. 3a, HR 1.59, 95% CI 1.03–2.45;  $P=0.04$ ), GSE13898 cohort (Fig. 3b, HR 3.15, 95% CI 1.25–7.98;  $P=0.01$ ), GSE19417 cohort (Fig. 3c, HR 1.85, 95% CI 1.02–3.38;  $P=0.04$ ) and GSE53625 cohort (Fig. 3d, HR 2.73, 95% CI 1.32–5.63;  $P=0.005$ ). In GSE13898 cohort, IRGPI could predict not only OS but also RFS (Fig. S4b). In the meta-validation dataset, when stratifying patients by sex, age and tumor stage, the IRGPI prognostic value remained high for

patients  $\leq 60$  years old, patients with late-stage and male patients (Fig. S5). IRGPI could independently predict the OS of EC patients in the meta-validation cohort after adjusting several clinicopathologic factors (Table 2, HR 2.11, 95% CI 1.21–3.68;  $P=0.00862$ ). Patients in the low-risk group with early-stage cancer have the best OS in the meta-validation dataset (Fig. S4c).

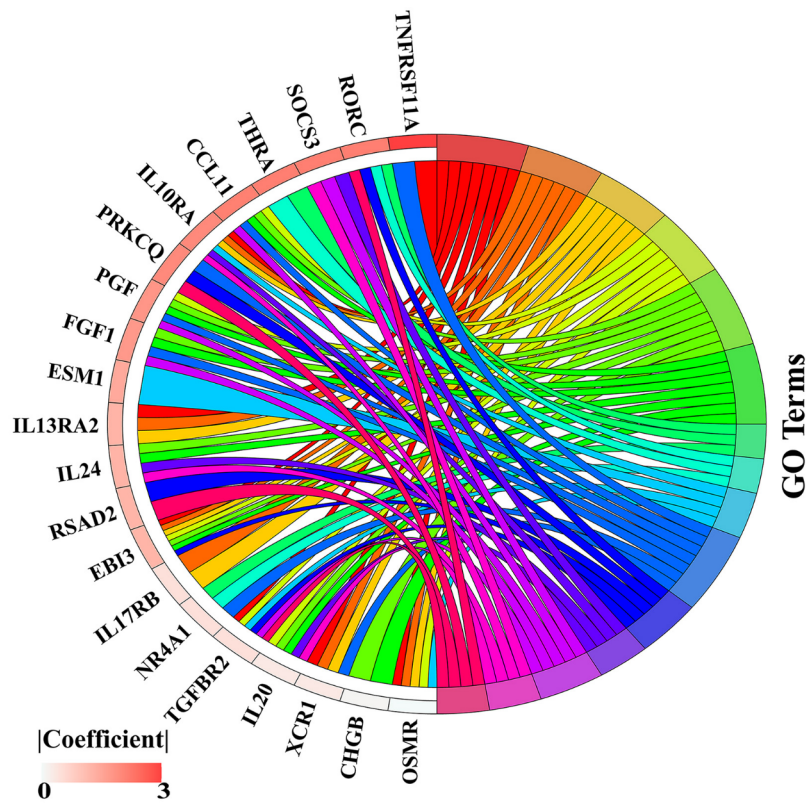
### Functional enrichment analysis and immune infiltration

To acquire the functional activities of these IRGs, we conducted GO analysis and immune infiltration. GO analysis revealed that IRGs in our signature were mainly contained in the chemotaxis, immune response, immune system process, defense response, inflammatory response and cytokine activity (Fig. 4, Table S3). 'CIBERSORT' was applied to calculate the abundances of 22 type immune cells in two risk groups. As a result, the number of plasma cells, activated CD4 memory T cells and resting mast cells were significantly higher in the low-risk group. Moreover, the number of activated mast cells and resting

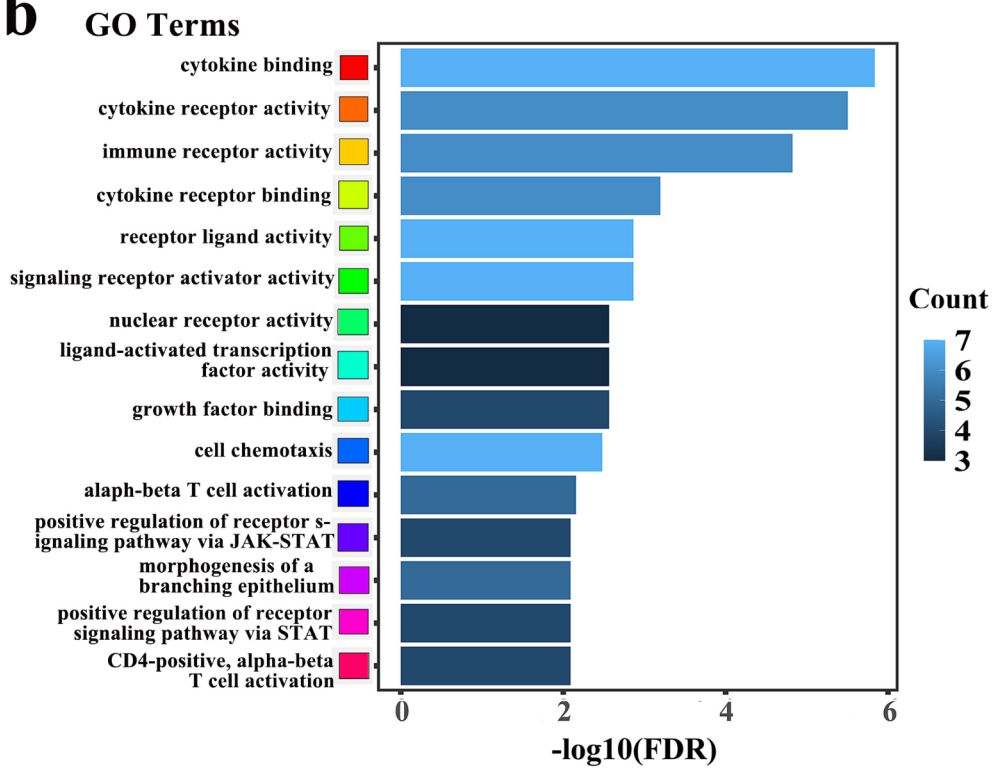


**Fig. 3** Kaplan-Meier curves in patients of validation sets classified by the signature. (a) OS of patients in meta-validation set. (b) OS of patients in GSE13898. (c) OS of patients in GSE19417. (d) OS of patients in GSE53625.

**a**



**b**



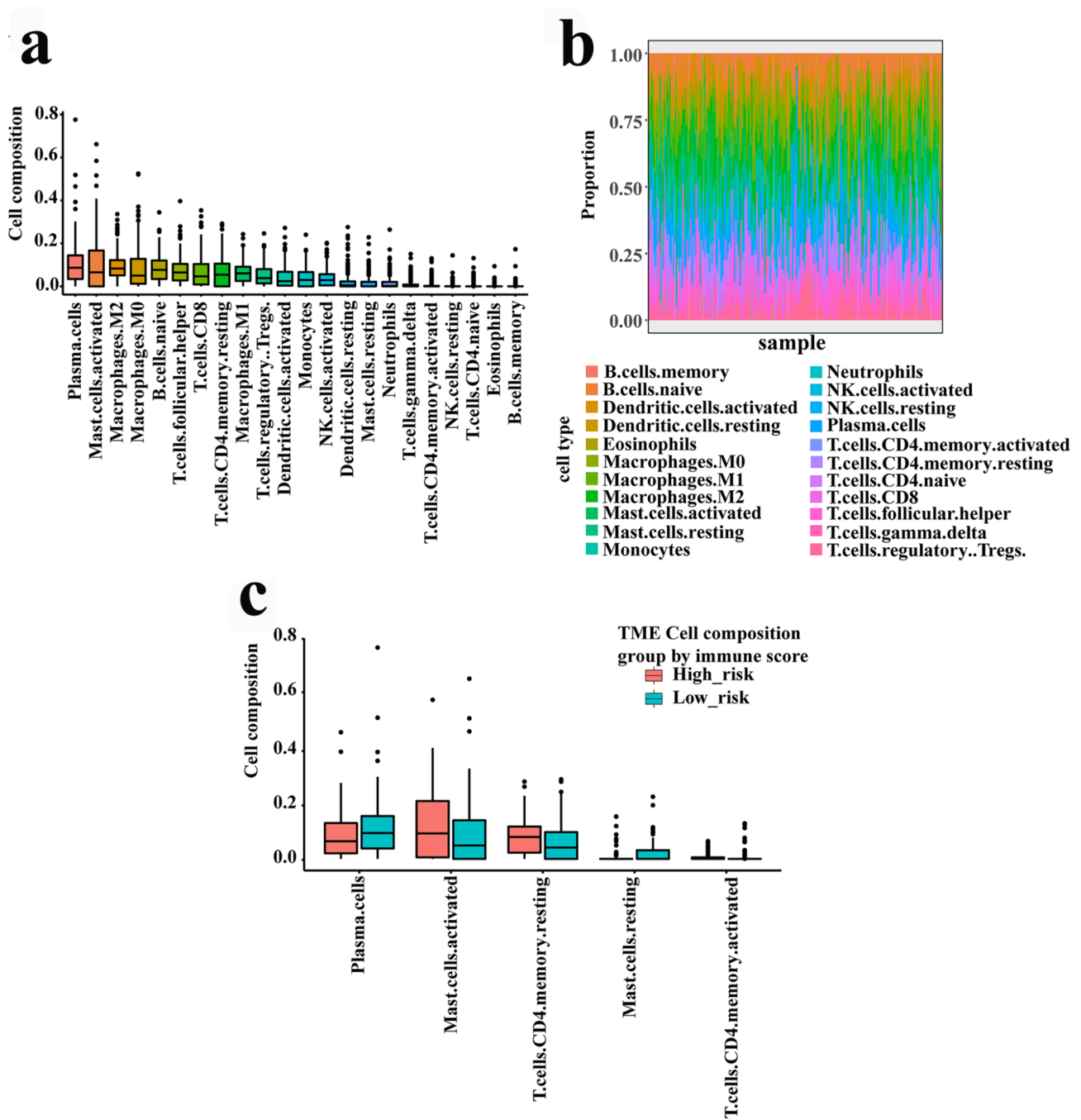
**Fig. 4** GO enrichment analysis on IRGs of the signature. The GO terms which were the top 15 ranked by FDR were listed. **(a)** The GO terms and their genes. **(b)** FDR and count of each GO term

CD4 memory T cells were higher in the high-risk group (Fig. 5).

**Integrated prognostic model based on IRGPI and clinicopathologic factors**

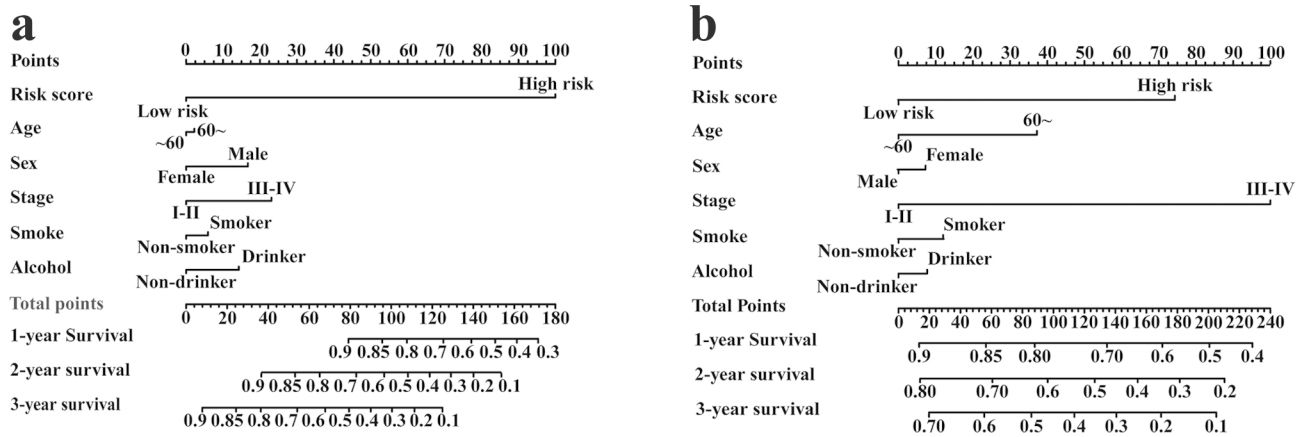
To obtain the IRGPI based prognostic model for 1-, 2- and 3-year OS rates, we built a nomogram model that incorporated IRGPI and clinicopathologic factors

[(age ≤ 60 or > 60), sex, tumor stage, smoking habits and alcohol intake] (Fig. 6a and 6b). The calibration plots were utilized to evaluate the predictive accuracy of the nomogram. The calibration plots revealed good prognostic prediction performance of nomogram in both training dataset and meta-validation dataset (Fig. S6).

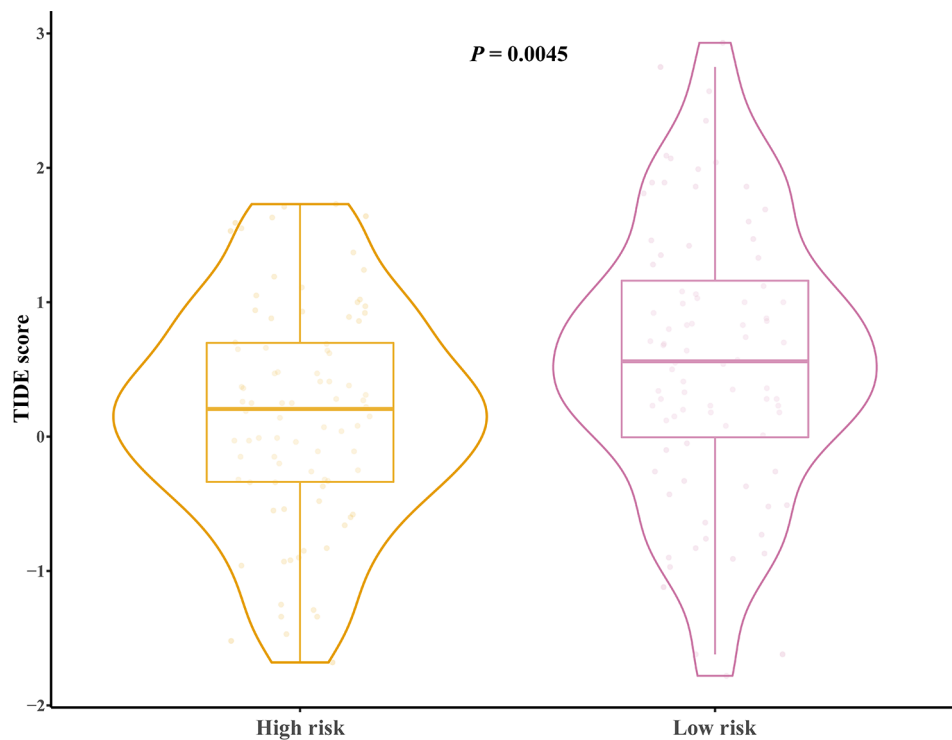


**Fig. 5** Infiltrating immune cell of the signature. (a) 22 immune cells abundance. (b) The percentage in 22 immune cells of different samples. (c) The immune cells abundance of different risk groups





**Fig. 6** The Nomogram of OS. (a) training cohort. (b) meta-validation cohort



**Fig. 7** The Correlation analysis between the signature and TIDE score

**The analysis of the TIDE score and small molecule drug inference**

The TIDE score of different risk groups were calculated and high-risk groups showed a significantly lower TIDE score ( $P=0.0045$ , Fig. 7), which suggested patients in high-risk groups have better immunotherapy responses. Moreover, we further identified five potential effective drugs in treating EC through the CMap and they were serotonin receptor agonist, vitamin D receptor agonist, CDC inhibitor, PDGFR receptor inhibitor and leucine rich repeat kinase inhibitor (Table 3).

**Transcription levels of genes involved in IRGPI**

To further ascertain the reliability of this IRGPs signature, we selected six genes (*IL13RA2*, *RORC*, *IL20*, *SAA2*, *FABP6* and *CHGB*) to conduct qRT-PCR experiments in Het-1A cells and EC cells lines (KYSE-150 and KYSE-450) based on the fold change, p-values between tumor and normal tissues gene expression in public databases, as well as high-frequency repeat among gene pairs (Table S4). The mRNA expressions of the majority of genes, including *IL13RA2*, *IL20* and *SAA2*, were persistently higher in tumor cells lines than Het-1A cells (Fig. 8a, 8c and 8d). *RORC* was persistently lower in EC cells lines than Het-1A cells (Fig. 8b). Transcription level of *FABP6*

**Table 3** The list of compounds which could be as potential drugs in treating EC.

Rank	Score	Name	Description	Target
8549	-87.34	quipazine	Serotonin receptor agonist	HTR2A, HTR3A, HTR3B, HTR1A, HTR1D, HTR2B, HTR2C, HTR6, SLC6A4
8548	-87.05	calcifediol	Vitamin D receptor agonist	VDR
8544	-85.69	KU-C103428N	CDC inhibitor	NFE2
8540	-83.55	linifanib	PDGFR receptor inhibitor	CSF1R, KDR, PDGFRB, FLT1, FLT3, FLT4, CSF1, KIT, PDGFRA, RET, TEK
8530	-76.3	XMD-1150	Leucine rich repeat kinase inhibitor	LRRK2

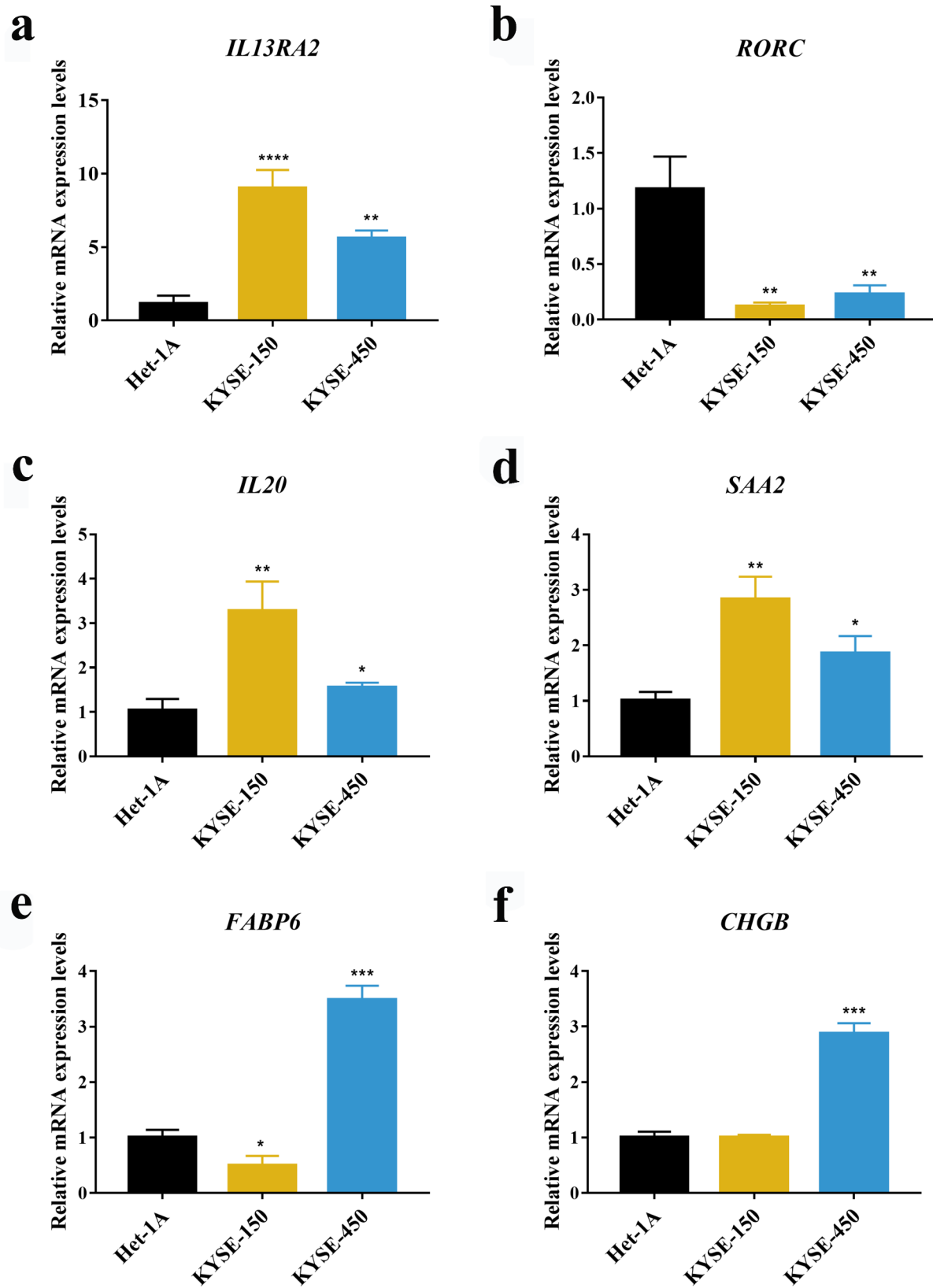
was increased in the KYSE-450 cells but not in the KYSE-150 cells in relative to the Het-1A cells (Fig. 8e). The transcription level of *CHGB* was only upregulated in the KYSE-450 cells (Fig. 8f).

## Discussion

EC, a lethal malignancy, become the sixth leading reason of tumor-related death [2]. Despite the improvement in therapies, the clinical outcomes of EC patients remain poor [4, 5]. Accumulating evidence has verified an association between prognosis and the tumor immune micro-environment. Immunotherapy has been administrated to combat various types of tumors, including EC; however, the cost of immunotherapy associated with EC is prohibitive and patient outcome has been mixed. As such, it's essential to recognize credible biomarkers to predict EC patients with high risk of mortality or who might be sensitive to immunotherapy. Many researchers have concentrated on the relevance between IRGs expression and tumorigenesis and progression. The prognostic abilities of IRGs were extensively studied [32–34]. Several researches have proposed IRGs-based signatures as prognostic predictors of EC [17–21]. However, the reliability of these signatures is limited. Specifically, one common drawback of these signatures is the lack of accurate normalization to reduce the biological heterogeneity when processing samples measured by different platforms. IRGPs signatures are on basis of the relative gene expression ranking of paired genes in each patient and eliminates the bias of normalization. IRGPs-based signatures have been proven to be effective prognostic biomarkers in many solid tumors, including gastric tumor, lung tumor, colorectal tumor, and renal cell cancer [27–30]. Based on the TCGA, GEO and ImmPort datasets, we constructed an IRGPs signature using 21 IRGPs. EC patients were segmented into two risk groups depending on the IRGPI

results above, with survival analyses revealing a significant difference between different groups. When patients were further stratified into different subgroups according to age, sex and tumor stage, differences remained regardless of these clinicopathologic features. The nomogram based on IRGPI and clinicopathologic features including age, sex and tumor stages could quantitatively predict the OS of patients with EC.

When it comes to the researches identifying cancer prognostic biomarkers, all strives aim to obtain robust results. To achieve this goal, we had done the following four things. Firstly, we validated the expression levels of six selected genes from IRGPI in KYSE-150 and KYSE-450. Our research is a combination of bioinformatics and experimentation. Based on these results and related literatures which showed that these genes were related to tumorigenesis [35–40], we infer that they are associated with esophageal tumorigenesis. After searching literatures in PubMed with “immune-related gene pair [Title/Abstract]” as keywords, we found 25 articles and none had their own experiments validation. Secondly, researches with multiple platforms could obtain relatively robust results [41]. This study included adequate external validation datasets with three GEO datasets obtained from multiple platforms, as shown in Fig. 3. Moreover, its abilities in distinguishing different OS in patients with different pathological stages, age and sex groups were verified in both the training cohort and validation cohorts, as shown in Fig. S3 and Fig. S5. Thirdly, the signature was developed based on a robust calculation process. At first, the association between each IRGP and patients' OS in the training dataset was assessed by the univariate Cox regression model. After rough filtration, the LASSO was further carried out to minimize overfitting. To improve the robustness of selected IRGPs, we randomly divided the training set into new training sets and test sets through a ratio of 2:1, and then duplicated this procedure 30 times. The LASSO was then used in the 30 training sets to choose those IRGPs with a frequency > 15. Then, we adopted the multivariate Cox regression analysis to construct the IRGP-based model, which consisted of relevant IRGPI and respective coefficients. Similarly, previous studies have utilized advanced machine learning methods to obtain more robust and accurate results [42–46]. Zhao et al. investigated cardiotoxicity related with hERG channel blockers using molecular fingerprints and graph attention mechanism [42]. Hu et al. conducted the association analysis between gene function and cell surface protein based on single-cell multi-omics data [43]. A study used a deep learning method to predict metabolite-disease associations via a graph neural network [44]. There were also some researches which focused on predicting lncRNA-miRNA interactions based on graph convolution network with



**Fig. 8** Validation of mRNA expression of genes generated from genes pairs. (a-f) Real-time PCR analysis of selected six genes in Het-1 A, KYSE-150 and KYSE-450 cells. Data were presented as the mean  $\pm$  SE.

conditional random field and network distance analysis [45, 46]. These interesting works would provide valuable directions for us. Fourthly, biological functions, immune infiltration and bioinformatics analysis were combined to confirm the reliability of our results. Pathway enrichment analysis showed that our signature-related genes were mainly involved in immune and inflammatory response, cytokine activity and chemotaxis. A significantly higher infiltration level of plasma cells and activated CD4 memory T cells was found in the low-risk group versus the high-risk group. Those results could provide clues for our future experimental researches.

GO analysis was conducted to obtain the functional activities of those IRGs consisting of IRGPs. The pathways associated with IRGPs mainly contained in cytokine, chemotaxis, T cell activation, receptor and ligand activation, morphogenesis of epithelium and angiogenesis (Fig. 4). Namely, most are involved in specific immune activities, and a few are receptor and ligand activation, and tumorigenesis. The results indicated that IRGs included in IRGPI participated in some specific immune and tumorigenesis activities. The results of immune cell abundance analysis revealed that the number of activated CD4 memory T cells, plasma cells and resting mast cells were significantly higher in the low-risk EC patients. Besides, resting CD4 memory T cells and activated mast cells have an inverse result. In recent years, evaluation of immune activity in multiple cancers has revealed that B cells and plasma cells may act as biomarkers to predict the effectiveness of immunotherapy. Plasma cells in tumor microenvironment are related to ameliorative outcomes [31]. Yao et al. also found an association between high tumor-infiltrating plasma cells and favorable OS prognosis in EC patients [31]. The results of our study were similar to these previous studies, whereby we found the abundance of activated CD4 memory T cells were lower in the high-risk group. Indeed, activated CD4 memory T cells was related to favorable outcome in several tumors [27, 33]. To a certain extent, the differences in immune cells composition in two groups may explain the prognostic value of the IRGPI in EC.

To identify the reliability of the IRGPs signature in EC, we conducted qRT-PCR to investigate the transcription levels of six selected genes. The *IL13RA2*, *SAA2* and *IL20* have a significantly higher expression level, whereas the transcriptional level of *RORC* was lower, in EC cells lines than that in Het-1A cells. *IL13RA2* participated in the regulation of *IL13* and *IL4* [32]. *IL13RA2* is overexpressed in various tumors, such as glioblastoma, pancreatic tumor, ovarian tumor and head and neck tumor [35, 36, 47, 48]. Thus, we surmise that *IL13RA2* possesses an essential role in the tumorigenesis of EC, and further work is needed to investigate the mechanism of *IL13RA2* in EC. Human *SAA1* and *SAA2* encode identical proteins.

The *SAA2* protein belongs to the *SAA* protein [49] and participates in the progression of several tumors. It is reported that *SAA2* is significantly increased when renal cell carcinoma (RCC) patients have the highest Fuhrman grade [50]. *SAA2* could also be a potential predictor of OS in RCC [50]. Indeed, *SAA2* was significantly increased in lung cancer and had effective diagnostic value in lung, endometrial and colorectal tumors [37, 38, 51]. Wang et al. showed that *SAA* levels were related with poor survival [41, 52]. According to our results and previous studies, it is a reasonable inference that *SAA2* is related to tumor tumorigenesis and poor survival of EC patients. However, this assumption needs further study. *RORC*, a member of *ROR* subfamily, has been studied in various cancer cells and their corresponding tumor microenvironment in recent years, with results showing that it might possess effective prognostic value in both lung and breast cancers. *RORC* could also be as a promising molecular target for the therapy of prostate tumor [39]. The results of our research indicated that *RORC* was downregulated in EC cells when compared with Het-1A cells. *IL20* is a part of *IL10* family and plays important roles in various immunopathological diseases [53]. Previous studies have indicated that *IL20* takes part in cancer progression in several tumors [54–57]. However, researches of *IL20* in EC were limited. Increased expression of *IL20* was found in EC of previous microarray analysis [40], but the potential mechanisms of *IL20* in EC is unknown yet. Our in vitro experiment showed that *IL20* has a significantly higher expression in EC cells than in Het-1A cells. This result is in accordance with the increased expression of *IL20* in several cancers including EC. Further study is warranted to explore the role of *IL20* on EC tumorigenesis.

It should be noted that there are some limitations in our research. First, despite four unique datasets being included, only 479 samples were analyzed and more samples are needed for further validation. Second, considering the results of this study were obtained from retrospective analyses, prospective study is required to confirm the efficiency of this IRGPs signature model.

## Conclusion

In conclusion, The IRGPI, constructed in this study, can both stratified EC into different risk groups and predict the prognosis of EC, effectively. The results of our study provide promising perspectives for selecting patients who are sensitive to immunotherapy and pave the way for EC treatment.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-023-09496-x>.

**Fig. S1.** Overview of the construction and validation of immune-related gene pairs signature (IRGPI). Four datasets were included in this study. TCGA dataset was used for training cohort, and GSE13898, GSE19417, GSE53625 were merged to the meta-validation cohort. The training cohort was used to build an IRGPI. The IRGPI was verified on the meta-validation cohort and independent validation cohorts.

**Fig. S2.** Flow chart of IRGPI construction.

**Fig. S3.** Kaplan-Meier curves in different subgroups' cases of training set. (a) OS of cases with early stage. (b) OS of male cases. (c) OS of cases  $\leq$  60 years old. (d) OS of cases with late stage. (e) OS of female cases. (f) OS of cases  $>$  60 years old

**Fig. S4.** The Kaplan-Meier curves. (a) OS among patients in training cohort stratified by IRGPI and tumor stage. (b) Relapse-free survival among patients in validation cohort of GSE13898. (c) OS among patients in meta-validation cohort stratified by IRGPI and tumor stage.

**Table S1.** Details about datasets used in this study.

**Table S3.** The significant biological processes enriched by genes consisted in the IRGPI.

**Table S2.** Model information about the IRGPI

**Table. S4.** The expression levels of six genes in training and validation dataset.

**Fig. S5.** Kaplan-Meier curves in different subgroups' cases of meta-validation dataset. (a) OS of cases in early stage. (b) OS of male cases. (c) OS of cases  $\leq$  60 years old. (d) OS of cases in late stage. (e) OS of female cases. (f) OS of cases  $>$  60 years old.

**Fig. S6.** Nomogram evaluation for predicting 1-, 2- and 3-year OS. (a-c) Calibration plots of the nomogram for predicting the probability of OS at 1 (a), 2 (b) and 3 years (c) in training cohort. (d-f) Calibration plots of the nomogram for predicting the probability of OS at 1 (d), 2 (e), 3 years (f) in meta-validation cohort.

## Acknowledgements

We express heartfelt thanks to the researchers of TCGA project, GSE13898 project, GSE19417 project and GSE53625 project.

## Authors' contributions

Biao Xie and Wei Zheng constructed the research and wrote the paper. Biao Xie and Dan Shi did data analyses and figure plotted. Gaofeng Fang, Qiao Huang and Dan Shi did experiments. All authors reviewed the manuscript.

## Funding

Postdoctoral Science Foundation of Chongqing (cstc2021jcyj-bshX0220) and National Natural Science Foundation of China (82204159, 8210120502).

## Data availability

The data of bioinformatics analysis is available at TCGA and GEO database, data of experimental research will be provided upon request. Cancer Genome Atlas (TCGA): <https://xenabrowser.net/datapages/>; GSE13898: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE13898>; GSE19417: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE19417>; GSE52625: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE52625>. Further information and requests for resources and reagents should be directed to and will be fulfilled by Biao Xie, kybiao@cqmu.edu.cn.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

Received: 23 February 2023 / Accepted: 30 June 2023

Published online: 10 July 2023

## References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68(6):394–424.
2. Smyth EC, Lagergren J, Fitzgerald RC, Lordick F, Shah MA, Lagergren P, Cunningham D. Oesophageal cancer. *Nat Rev Dis Primers.* 2017;3:17048.
3. Lagergren J, Smyth E, Cunningham D, Lagergren P. Oesophageal cancer. *Lancet.* 2017;390(10110):2383–96.
4. Bruni D, Angell HK, Galon J. The immune contexture and immunoscore in cancer prognosis and therapeutic efficacy. *Nat Rev Cancer.* 2020;20(11):662–80.
5. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, Nair VS, Xu Y, Khuong A, Hoang CD, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med.* 2015;21(8):938–45.
6. Ettinger DS, Wood DE, Aisner DL, Akerley W, Bauman JR, Bharat A, Bruno DS, Chang JY, Chirieac LR, D'Amico TA, et al. NCCN guidelines insights: non-small cell lung cancer, version 2.2021. *J Natl Compr Canc Netw.* 2021;19(3):254–66.
7. Larkin J, Chiarion-Sileni V, Gonzalez R, Grob JJ, Rutkowski P, Lao CD, Cowey CL, Schadendorf D, Wagstaff J, Dummer R, et al. Five-year survival with combined nivolumab and ipilimumab in advanced melanoma. *N Engl J Med.* 2019;381(16):1535–46.
8. Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, Tykodi SS, Sosman JA, Procopio G, Plimack ER, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med.* 2015;373(19):1803–13.
9. He X, Xu C. Immune checkpoint signaling and cancer immunotherapy. *Cell Res.* 2020;30(8):660–9.
10. Xia Y, Medeiros LJ, Young KH. Immune checkpoint blockade: releasing the brake towards hematological malignancies. *Blood Rev.* 2016;30(3):189–200.
11. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL, et al. Signatures of mutational processes in human cancer. *Nature.* 2013;500(7463):415–21.
12. Huang J, Xu B, Mo H, Zhang W, Chen X, Wu D, Qu D, Wang X, Lan B, Yang B, et al. Safety, activity, and biomarkers of SHR-1210, an Anti-PD-1 antibody, for patients with advanced esophageal carcinoma. *Clin Cancer Res.* 2018;24(6):1296–304.
13. Janjigian YY, Bendell J, Calvo E, Kim JW, Ascierto PA, Sharma P, Ott PA, Peltola K, Jaeger D, Evans J, et al. CheckMate-032 study: efficacy and safety of nivolumab and nivolumab plus ipilimumab in patients with metastatic esophagogastric cancer. *J Clin Oncol.* 2018;36(28):2836–44.
14. Shah MA, Kojima T, Hochhauser D, Enzinger P, Raimbourg J, Hollebecque A, Lordick F, Kim SB, Tajika M, Kim HT, et al. Efficacy and safety of pembrolizumab for heavily pretreated patients with advanced, metastatic adenocarcinoma or squamous cell carcinoma of the esophagus: the phase 2 KEYNOTE-180 study. *JAMA Oncol.* 2019;5(4):546–50.
15. Xing W, Zhao L, Fu X, Liang G, Zhang Y, Yuan D, Li Z, Gao Q, Zheng Y. A phase II, single-centre trial of neoadjuvant toripalimab plus chemotherapy in locally advanced esophageal squamous cell carcinoma. *J Thorac Dis.* 2020;12(11):6861–7.
16. Zhang W, Wang P, Pang Q. Immune checkpoint inhibitors for esophageal squamous cell carcinoma: a narrative review. *Ann Transl Med.* 2020;8(18):1193.
17. Fei Z, Xie R, Chen Z, Xie J, Gu Y, Zhou Y, Xu T. Establishment of a novel risk score system of immune genes associated with prognosis in esophageal carcinoma. *Front Oncol.* 2021;11:625271.
18. Gao J, Tang T, Zhang B, Li G. A prognostic signature based on immunogenomic profiling offers guidance for esophageal squamous cell cancer treatment. *Front Oncol.* 2021;11:603634.
19. Guo X, Cheng Y, Zhang H, Qin C, Cheng A, Liu J, Dai X, Wang Z. Identification of the prognostic value of immune-related genes in esophageal cancer. *Front Genet.* 2020;11:989.
20. Li Y, Lu Z, Che Y, Wang J, Sun S, Huang J, Mao S, Lei Y, Chen Z, He J. Immune signature profiling identified predictive and prognostic factors for esophageal squamous cell carcinoma. *Oncoimmunology.* 2017;6(11):e1356147.
21. Wang L, Wei Q, Zhang M, Chen L, Li Z, Zhou C, He M, Wei M, Zhao L. Identification of the prognostic value of immune gene signature and infiltrating



- immune cells for esophageal cancer patients. *Int Immunopharmacol*. 2020;87:106795.
22. Zheng Y, Gao Q, Su X, Xiao C, Yu B, Huang S, Sun Y, Wu S, Wo Y, Xu Q et al. Genome-wide DNA methylation and gene expression profiling characterizes molecular subtypes of esophagus squamous cell carcinoma for predicting patient survival and immunotherapy efficacy. *Cancers (Basel)*. 2022;14(20).
  23. Zhang C, Guo C, Li Y, Liu K, Zhao Q, Ouyang L. Identification of claudin-6 as a molecular biomarker in pan-cancer through multiple omics integrative analysis. *Front Cell Dev Biol*. 2021;9:726656.
  24. Zhang C, Guo C, Li Y, Ouyang L, Zhao Q, Liu K. The role of YTH domain containing 2 in epigenetic modification and immune infiltration of pan-cancer. *J Cell Mol Med*. 2021;25(18):8615–27.
  25. Franks JM, Cai G, Whitfield ML. Feature specific quantile normalization enables cross-platform classification of molecular subtypes using gene expression data. *Bioinformatics*. 2018;34(11):1868–74.
  26. Shen Y, Rahman M, Piccolo SR, Gusenleitner D, El-Chaar NN, Cheng L, Monti S, Bild AH, Johnson WE. ASSIGN: context-specific genomic profiling of multiple heterogeneous biological pathways. *Bioinformatics*. 2015;31(11):1745–53.
  27. Li B, Cui Y, Diehn M, Li R. Development and validation of an individualized immune prognostic signature in early-stage nonsquamous non-small cell lung cancer. *JAMA Oncol*. 2017;3(11):1529–37.
  28. Tao Z, Zhang E, Li L, Zheng J, Zhao Y, Chen X. A united risk model of 11 immunerelated gene pairs and clinical stage for prediction of overall survival in clear cell renal cell carcinoma patients. *Bioengineered*. 2021;12(1):4259–77.
  29. Wu J, Zhao Y, Zhang J, Wu Q, Wang W. Development and validation of an immune-related gene pairs signature in colorectal cancer. *Oncoimmunology*. 2019;8(7):1596715.
  30. Zhao E, Zhou C, Chen S. A signature of 14 immune-related gene pairs predicts overall survival in gastric cancer. *Clin Transl Oncol*. 2021;23(2):265–74.
  31. Weiner AB, Vidotto T, Liu Y, Mendes AA, Salles DC, Faisal FA, Murali S, McFarlane M, Imada EL, Zhao X, et al. Plasma cells are enriched in localized prostate cancer in black men and are associated with improved outcomes. *Nat Commun*. 2021;12(1):935.
  32. David M, Ford D, Bertoglio J, Maizel AL, Pierre J. Induction of the IL-13 receptor alpha2-chain by IL-4 and IL-13 in human keratinocytes: involvement of STAT6, ERK and p38 MAPK pathways. *Oncogene*. 2001;20(46):6660–8.
  33. Xue YN, Xue YN, Wang ZC, Mo YZ, Wang PY, Tan WQ. A novel signature of 23 immunity-related gene pairs is prognostic of cutaneous melanoma. *Front Immunol*. 2020;11:576914.
  34. Yao C, Yu H, Zhou G, Xu J, Gu D, Yin L, He X, Xia H. Tumor-infiltrating plasma cells are the promising prognosis marker for esophageal squamous cell carcinoma. *Esophagus*. 2021;18(3):574–84.
  35. Kawakami M, Kawakami K, Kasperbauer JL, Hinkley LL, Tsukuda M, Strome SE, Puri RK. Interleukin-13 receptor alpha2 chain in human head and neck cancer serves as a unique diagnostic marker. *Clin Cancer Res*. 2003;9(17):6381–8.
  36. Kioi M, Kawakami M, Shimamura T, Husain SR, Puri RK. Interleukin-13 receptor alpha2 chain: a potential biomarker and molecular target for ovarian cancer therapy. *Cancer-Am Cancer Soc*. 2006;107(6):1407–18.
  37. Wong YF, Cheung TH, Lo KW, Yim SF, Siu NS, Chan SC, Ho TW, Wong KW, Yu MY, Wang VW, et al. Identification of molecular markers and signaling pathway in endometrial cancer in Hong Kong Chinese women by genome-wide gene expression profiling. *Oncogene*. 2007;26(13):1971–82.
  38. Zhang FF, Han B, Xu RH, Zhu QQ, Wu QQ, Wei HM, Cui ZL, Zhang SL, Meng MJ. Identification of plasma SAA2 as a candidate biomarker for the detection and surveillance of non-small cell lung cancer. *Neoplasma*. 2021;68(6):1301–9.
  39. Fan J, Lv Z, Yang G, Liao TT, Xu J, Wu F, Huang Q, Guo M, Hu G, Zhou M, et al. Retinoic acid receptor-related orphan receptors: critical roles in tumorigenesis. *Front Immunol*. 2018;9:1187.
  40. Hsing CH, Ho CL, Chang LY, Lee YL, Chuang SS, Chang MS. Tissue microarray analysis of interleukin-20 expression. *Cytokine*. 2006;35(1–2):44–52.
  41. Gratten J, Wray NR, Keller MC, Visscher PM. Large-scale genomics unveils the genetic architecture of psychiatric disorders. *Nat Neurosci*. 2014;17(6):782–90.
  42. Wang T, Sun J, Zhao Q. Investigating cardiotoxicity related with hERG channel blockers using molecular fingerprints and graph attention mechanism. *Comput Biol Med*. 2023;153:106464.
  43. Hu H, Feng Z, Lin H, Cheng J, Lyu J, Zhang Y, Zhao J, Xu F, Lin T, Zhao Q, et al. Gene function and cell surface protein association analysis based on single-cell multiomics data. *Comput Biol Med*. 2023;157:106733.
  44. Sun F, Sun J, Zhao Q. A deep learning method for predicting metabolite-disease associations via graph neural network. *Brief Bioinform*. 2022;23(4):bbac266.
  45. Wang W, Zhang L, Sun J, Zhao Q, Shuai J. Predicting the potential human lncRNA-miRNA interactions based on graph convolution network with conditional random field. *Brief Bioinform*. 2022;23(6):bbac463.
  46. Zhang L, Yang P, Feng H, Zhao Q, Liu H. Using Network Distance Analysis to predict lncRNA-miRNA interactions. *Interdiscip Sci*. 2021;13(3):535–45.
  47. Fujisawa T, Joshi B, Nakajima A, Puri RK. A novel role of interleukin-13 receptor alpha2 in pancreatic cancer invasion and metastasis. *Cancer Res*. 2009;69(22):8678–85.
  48. Lian X, Kats D, Rasmussen S, Martin LR, Karki A, Keller C, Berlow NE. Design considerations of an IL13Ralpha2 antibody-drug conjugate for diffuse intrinsic pontine glioma. *Acta Neuropathol Commun*. 2021;9(1):88.
  49. Malle E, Sodin-Semrl S, Kovacevic A. Serum amyloid A: an acute-phase protein involved in tumour pathogenesis. *Cell Mol Life Sci*. 2009;66(1):9–26.
  50. Cooley LS, Rudewicz J, Souleyreau W, Emanuelli A, Alvarez-Arenas A, Clarke K, Falciani F, Dufes M, Lambrechts D, Modave E, et al. Experimental and computational modeling for signature and biomarker discovery of renal cell carcinoma progression. *Mol Cancer*. 2021;20(1):136.
  51. Brock R, Xiong B, Li L, Vanbogelen RA, Christman L. A multiplex serum protein assay for determining the probability of colorectal cancer. *Am J Cancer Res*. 2012;2(5):598–605.
  52. Wang JY, Zheng YZ, Yang J, Lin YH, Dai SQ, Zhang G, Liu WL. Elevated levels of serum amyloid a indicate poor prognosis in patients with esophageal squamous cell carcinoma. *BMC Cancer*. 2012;12:365.
  53. Rutz S, Wang X, Ouyang W. The IL-20 subfamily of cytokines-from host defense to tissue homeostasis. *Nat Rev Immunol*. 2014;14(12):783–95.
  54. Ding WZ, Han GY, Jin HH, Zhan CF, Ji Y, Huang XL. Anti-IL-20 monoclonal antibody suppresses hepatocellular carcinoma progression. *Oncol Lett*. 2018;16(5):6156–62.
  55. Hsu YH, Hsing CH, Li CF, Chan CH, Chang MC, Yan JJ, Chang MS. Anti-IL-20 monoclonal antibody suppresses breast cancer progression and bone osteolysis in murine models. *J Immunol*. 2012;188(4):1981–91.
  56. Lee SJ, Cho SC, Lee EJ, Kim S, Lee SB, Lim JH, Choi YH, Kim WJ, Moon SK. Interleukin-20 promotes migration of bladder cancer cells through extracellular signal-regulated kinase (ERK)-mediated MMP-9 protein expression leading to nuclear factor (NF-kappaB) activation by inducing the up-regulation of p21(WAF1) protein expression. *J Biol Chem*. 2013;288(8):5539–52.
  57. Lu SW, Pan HC, Hsu YH, Chang KC, Wu LW, Chen WY, Chang MS. IL-20 antagonist suppresses PD-L1 expression and prolongs survival in pancreatic cancer models. *Nat Commun*. 2020;11(1):4611.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.