Development of an Immune-Related Gene Pairs Signature for Predicting Clinical Outcome in Lung Adenocarcinoma

Chunlei Wu
Taizhou Hospital of Zhejiang Province
Quanteng Hu
Taizhou Hospital of Zhejiang Province
Dehua Ma (madh_tzyy@sina.com)
Taizhou Hospital of Zhejiang Province

Research article

Keywords: immune, gene pairs, lung adenocarcinoma

Posted Date: July 20th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-43696/v1

Version of Record: A version of this preprint was published at Scientific Reports on February 11th, 2021.
See the published version at https://doi.org/10.1038/s41598-021-83120-4.

Abstract

Background: Lung adenocarcinoma (LUAD) is the main pathological subtype of Non-small cell lung cancer. The aim of this study was to establish an immune-related gene pairs (IRGPs) signature for predicting the prognosis of LUAD patients.

Methods: We downloaded the gene expression profile and immune-related gene set from TCGA and ImmPort database, respectively, to establish IRGPs. Then, IRGPs were subjected to univariate Cox regression analysis, LASSO regression analysis and multivariable Cox regression analysis to screen and develop a IRGPs signature. The receiver operating characteristic curve (ROC) was applied for evaluating the predicting accuracy of this signature by calculating the area under ROC (AUC) and data from GEO was used to validate this signature.

Results: A IRGPs signature with 8 IRGPs was constructed. The AUC for 1- and 3-year overall survival in TCGA set was 0.867 and 0.870, respectively. Similar result was observed in the AUC of GEO set and Total set (GEO set [1-year: 0.819; 3-years: 0.803]; Total set [1-year: 0.845; 3-years: 0.801]). Survival analysis of three sets demonstrated high-risk LUAD patients exhibited poorer prognosis. The multivariable Cox regression indicated that risk score was independent prognostic factors.

Conclusions: We developed a novel IRGPs signature for predicting prognosis of LUAD.

Background

Lung cancer (LC) is the most common cancer globally [1]. There was estimated that approximate 234,000 new cases were diagnosed as LC per year, which account for 14% and 13% new malignant tumor cases in men and women, respectively [1, 2]. Additionally, LC is the main cause of cancer-related deaths and result in over 170,000 deaths annually. Non-small cell lung cancer (NSCLC) is the most common LC (85%) and lung adenocarcinoma (LUAD) is the main pathological subtype of NSCLC (50%) [2, 3]. What is worse, incidence and mortality rate of LUAD continue to rise. TNM staging (AJCC) is the most commonly used parameters for clinical decision and assessment of outcome in LUAD [4, 5]. However, emerging studies have shown that although patients with same TNM stage and treatment strategy, the prognosis regimen different, indicating that TNM staging alone may not provide adequate information for prognosis assessment in LC.

Recently, researchers have come to realize that immune system plays a vital role in the development and progression of malignant tumors [6, 7]. Immune cells recognize malignant cells and eradicate them through immune surveillance [8]. However, tumors could manipulate the immune system to avoid recognition of tumor-associated antigens and to facilitate their own development [9]. Based on this theory, immunotherapy which acts via harness the immune system against tumors has been approved for the treatment of manifold tumors and revolutionized cancer treatment.

Aberrations of gene expression are universal events in malignancies and may facilitate tumor progression [10]. Omics technology provide a novel opportunity to understand gene changes and potential mechanisms in numerous cancers. In addition, bioinformatics analysis could secondary analyze the result of high throughput sequencing to identify new tumor biomarkers and provide more accurate prognosis prediction and clinical decision.

Immune-related gene pairs (IRGPs) signature has been established in several cancers including colorectal cancer [11], liver cancer [12] and ovarian cancer [13], and shown well accurate prognosis prediction. So far, there is no research using IRGPs to establish a prognosis signature in LUAD. Thence, in this study, we downloaded gene expression profile from The Cancer Genome Altas (TCGA, https://cancergenome.nih.gov) and immune-related gene set from ImmPort (https://www.immport.org/home), respectively, to performed systematic and comprehensive analysis on the characteristics of IRGPs and develop a IRGPs signature in LUAD. Then, we validated this IRGPs signature with data from Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov) and evaluated the predictive accuracy by calculating the area under curve (AUC) of receiver operating characteristic curve (ROC). Finally, we compared this signature with clinical characteristics to prove the predictive accuracy and effectiveness of this IRGPs signature.

Methods

Patient Data Sets

The FPKM level gene expression matrix was taken from TCGA. The raw data of mRNA expression matrix of GSE68465 was downloaded from GEO dataset. The platform for GSE68465 was GPL96 (Affymetrix Human Genome U133A Array). In addition, relevant clinical characteristics of patients were also collected. Moreover, Patients lacking of survival time and survival state or with follow-up time < 30 days would be removed. A total of 896 LUAD patients were collected from two independent data-sets (TCGA set: 465 cases; GEO set: 431 cases). Then, the raw data of mRNA expression matrix in GSE68465 were transformed with $\log_2$ and normalized with “affy” package in R 3.6.3 (https://www.r-project.org).

Construction of prognostic IRGPs signature

The IRGPs signature was constructed as described by previous study [12]. In a specific IRGP, if the first IRG (immune-related gene) expression level was lower than the second IRG expression level, the score of this IRGP was 0; otherwise, the score was 1. Moreover, if the score of an IRGP in more than 80% cases was 0 or 1, the IRGP would be discarded. Then, we performed univariate Cox regression analysis to preliminary filtrate IRGPs with $p < 0.001$ as a cut-off criterion in TCGA set. Next, LASSO regression analysis was used to further screen out IRGPs with iteration = 1000. Finally, multivariate Cox regression analysis was carried out to identify top overall survival (OS)-related IRGPs and to establish a prognostic IRGPs signature. An immune risk score formula was also formed, based on which immune risk score of each patient was calculated. According to the median value of immune risk score, patients were divided into low- and high- risk score group.

Validation and evaluation of prognostic IRGPs signature

With above risk score formula, immune risk score of patients in GEO set was also calculated and classified patients into low- and high- risk score group. The AUC both in TCGA set, GEO set and Total set were calculated to assess the predictive accuracy and effectiveness of this prognostic IRGPs signature.

Correlation between IRGPs signature and tumor-infiltrating immune cells (TIICs)

The CIBERSORT algorithm is novel accurate way that can determine 22 TIICs simultaneously in the tumor microenvironment (TME). With this algorithm, we quantified the proportions of 22 TIICs in all samples. CIBERSORT $P < 0.05$ was considered as cut-off value. In addition, the different landscape of 22 TIICs between low- and high- group was also compared. Meanwhile, several key immunomodulators were also quantified.

Gene set enrichment analysis (GSEA)

To determine the biological processes and signaling pathways altered by this IRGPs signature, GSEA was performed. $FDR < 0.05$ was set as the cut-off value.

Statistical analysis

All statistical analyses were performed with R 3.6.3 software. All clinical information of LUAD patients were present as number (No.) and percentage (%) in Table 1. For categorical data, chi-square test was performed to compare the differences among different groups, whereas, for measurement data, t test or one-way ANOVA was used. Survival curves were performed by the Kaplan-Meier method, and survival rates were compared with log-rank test. Moreover, univariate Cox regression analysis and multivariate Cox regression analysis were also performed to identify independent prognostic factors.

Table 1
The baseline characteristics of lung adenocarcinoma patients in this study.

Parameter	TCGA set	GEO set	Total set
Gender			
Female	254(54.62%)	216(50.12%)	470(52.46%)
Male	211(45.38%)	215(49.88%)	426(47.54%)
Age			
≤ 65	232(49.89%)	226(52.44%)	458(51.11%)
> 65	233(50.11%)	205(47.56%)	438(48.89%)
EGFR mutation			
No	174(37.42%)	NA	174(19.42%)
Yes	69(15.05%)	NA	69(7.70%)
NA	221(47.53%)	431(100%)	653(72.88%)
KRS mutation			
No	34(7.32%)	NA	34(3.79%)
Yes	17(3.66%)	NA	17(1.90%)
NA	414(89.02%)	431(100%)	845(94.31%)
Smoking			
Never	62(13.33%)	48(11.14%)	110(12.28%)
Ever	391(84.09%)	295(68.45%)	686(76.56%)
NA	12(2.58%)	88(20.41%)	100(11.16%)
Radiotherapy			
No	336(9.38%)	353(81.90%)	689(76.90%)
Yes	53(85.62%)	64(14.85%)	117(13.06%)
NA	76(5.00%)	14(3.25%)	90(10.04%)
Chemotherapy			
No	461(99.14%)	329(76.33%)	790(88.17%)
Yes	3(0.64%)	89(20.65%)	92(10.27%)

Abbreviations: TCGA, The Cancer Genome Altas; GEO, Gene Expression Omnibus; NA, represents information not available.

Parameter	TCGA set	GEO set	Total set
NA	1(0.22%)	13(3.02%)	14(1.56%)
Histologic grade			
Poor	NA	161(37.35%)	161(17.97%)
Moderate	NA	203(47.10%)	203(22.66%)
Well	NA	60(13.92%)	60(6.70%)
NA	465(100%)	7(1.63%)	472(52.68%)
TNM stage			
I	261(56.12%)	270(62.65%)	531(59.26%)
II	106(22.80%)	99(22.97%)	205(22.88%)
III	73(5.70%)	60(13.92%)	133(14.84%)
IV	84(18.06%)	0	84(9.38%)
NA	1(0.22%)	2(0.46%)	3(0.34%)
Tumor size			
T1	159(34.19%)	145(33.64%)	304(33.93%)
T2	248(53.33%)	244(56.61%)	492(54.91%)
T3	40(8.60%%)	27(6.26%)	67(7.48%)
T4	18(3.87%)	11(2.55%)	29(3.14%)
NA	0	4(0.94%)	4(0.44%)
Lymph node			
N0	309(66.45%)	292(67.75%)	601(67.08%)
N1-3	151(32.47%)	137(31.79%)	288(32.14%)
NA	5(1.08%)	2(0.46%)	7(0.78%)
Metastasis			
M0	441(94.84%)	429(99.54%)	870(97.10%)
M1	24(5.16%)	0	24(2.68%)
NA	0	2(0.46%)	2(0.22%)

Abbreviations: TCGA, The Cancer Genome Altas; GEO, Gene Expression Omnibus; NA, represents information not available.

Parameter	TCGA set	GEO set	Total set
Survival status			
Alive	310(66.67%)	202(46.87%)	512(57.14%)
Dead	155(33.33%)	229(53.13%%)	384(42.86%%)
Risk score			
Low	233(50.11%)	248(57.54%)	481(53.68%)
High	232(49.89%)	183(42.26%)	415(46.32%)
Total	465(100%)	431(100%)	896(100%)

Abbreviations: TCGA, The Cancer Genome Altas; GEO, Gene Expression Omnibus; NA, represents information not available.

Results

Construction of prognostic IRGPs signature

After removing IRGPs with score of 0 or 1 in more than 80% of LUAD cases, a total of 12334 IRGPs were pared. After screened by univariate Cox regression analysis, 54 IRGPs were subjected to LASSO regression analysis and 21 IRGPs were filter out (Fig. 1a). Then, 8 IRGPs were identified and used to develop a prognostic IRGPs signature with multivariate Cox regression analysis (Fig. 1b, Table 2). Risk score = (1.139* Score $_{BIRC5|BPHL}$) + (-0.658* Score $_{CCL2|OAS1}$) + (-0.461* Score $_{CD19|PI3}$) + (-0.557* Score $_{CD3G|IL7}$) + (0.723* Score $_{DKK1|IKBKB}$) + (0.448* Score $_{F2RL1|LTB}$) + (-0.428* Score $_{PIK3CD|S100A2}$) + (-0.606* Score $_{SERPIND1|VEGFC}$).

Table 2
Information on the 8 immune-related gene pairs (IRGPs).

IRG 1	Immune processes	IRG 2	Immune processes	Coefficient
BIRC5	Antimicrobials	BPHL	Antimicrobials	1.139
CCL2	Antimicrobials	OAS1	Antimicrobials	-0.658
CD19	BCR Signaling Pathway	PI3	BCR Signaling Pathway	-0.461
CD3G	TCR signaling Pathway	IL7	Cytokines	-0.557
DKK1	Cytokines	IKBKB	TCR signaling Pathway	0.723
F2RL1	Antimicrobials	LTB	Cytokines	0.448
PIK3CD	BCR Signaling Pathway	S100A2	Antimicrobials	-0.428
SERPIND1	Antimicrobials	VEGFC	Cytokines	-0.606
Abbreviations: IRGPs, immune-related gene pairs; IRG, immune-related gene.				

Validation and evaluation of prognostic IRGPs signature

According to cut-off risk score (0.99), patients both in TCGA set and GEO set were divided into low- and high- risk group. Moreover, we calculated the AUC for predicting 1- and 3- year OS to evaluate the predictive capacity of this IRGPs signature. The AUC for 1- and 3- year OS in TCGA set was 0.867 and 0.870, respectively (Fig. 2a-b). And, in GEO set was 0.819 and 0.803, respectively (Fig. 2c-d). In addition, the AUC in Total set was 0.845 and 0.801, respectively (Fig. 2e-f).

The IRGPs signature is an independent prognostic factor of survival

We performed survival analysis to compare the survival difference between low- and high- risk group. All of Kaplan-Meier plot in three sets demonstrated that high-risk LUAD patients exhibited poorer prognosis than low-risk LUAD patients (TCGA set: $p < 0.001$, Fig. 3a; GEO set: $p < 0.001$, Fig. 3b; Total set: $p < 0.001$, Fig. 3c). Furthermore, stratification analyses showed the clinical outcome of high-risk LUAD patients in each stratum of age, gender, TNM stage, tumor size, lymph node metastasis and distance metastasis was poorer than that of low-risk patients (Additional file 1: Fig. S1).

Then, we took advantage of univariate and multivariate Cox regression model to compare the immune risk score with clinical parameters (age, gender, smoking, histologic grade, TNM grade, tumor size, lymph node metastasis and distance metastasis). Univariable Cox regression analysis indicated that immune risk score was an important factor of patients' prognosis (TCGA set: HR = 4.227, 95%CI [2.952, 6.053], $p < 0.001$, Fig. 3d; GEO set: HR = 2.484, 95%CI [1.911, 3.229], $p < 0.001$, Fig. 3f; Total set: HR = 2.975, 95%CI [2.416, 3.662], $p < 0.001$, Fig. 3h). Moreover, multivariable Cox regression demonstrated that immune risk score was independent prognostic factors (TCGA set: HR = 3.856, 95%CI [2.621, 5.673], $p < 0.001$, Fig. 5d;

GEO set: HR = 2.473, 95%CI [1.789, 3.436], p < 0.001, Fig. 5g; Total set: HR = 2.876, 95%CI [2.264, 3.654], p < 0.001, Fig. 5i).

Correlation between IRGPs signature and clinical characteristics

As shown in Fig. 4a, between low- and high- risk group, the distribution of gender (p = 0.001), histologic grade (p < 0.001), TNM grade (p < 0.001), tumor size (p < 0.001), lymph node metastasis (p < 0.001) and KRS mutation (0.012) was significantly different. Meanwhile, compared with female, the immune risk score in male was significantly increased (p = 0.001, Fig. 4b). Similar result was observed in patients with lymph node metastasis (p < 0.001, Fig. 4f) and distance metastasis (p = 0.024, Fig. 4g). In addition, with the increase of TNM grade (p < 0.001, Fig. 4c) and tumor size (p < 0.001, Fig. 4e), the immune risk score was also increased. The immune risk score was significantly different among three histologic grades (poor differentiation, Moderate differentiation and well differentiation) (p < 0.001, Fig. 4d).

Relationship between IRGPs and TIICs

A total of 823 LUAD cases met the cut-off value: CIBERSORT P < 0.05. Macrophages (M0, M1, M2) (32.98%) was the most abundant immune cell, followed by Plasma cells (17.23%) and resting memory CD4 T cells (9.70%) (Additional file 1: Fig. S2a). Compared with normal lung tissues, LUAD tissues had more proportions of naive B cells (p = 0.001), memory B cells (p < 0.001), Plasma cells (p < 0.001), activated memory CD4 T cells (p < 0.001), helper follicular T cells (p < 0.001), regulatory T cells (Tregs) (p < 0.001), gamma delta T cells (p = 0.011), Macrophages M1 (p < 0.001) and less proportions of T resting memory cells CD4 (p = 0.001), resting NK cells (p < 0.001), Monocytes (p < 0.001), resting Mast cells (p = 0.004), Eosinophils (p < 0.001) (Additional file 1: Fig. S2b). In addition, the proportions of Macrophages M0 (p < 0.001), Macrophages M1 (p = 0.001), activated memory CD4 T cells (p < 0.001) were significantly increased in high- risk group, whereas, the proportions of memory B cells (p < 0.001), Plasma cells (p = 0.008), naive T cells CD4 (p = 0.010), gamma delta T cells (p = 0.012), resting NK cells (p < 0.001), Monocytes (p < 0.001), resting Dendritic cells (p < 0.001), resting Mast cells (p < 0.001) and Neutrophils (p = 0.002) were significantly decreased (Fig. 5a). Furthermore, survival analysis demonstrated that low level of Monocytes (p = 0.020, Fig. 5b), Plasma cells (p = 0.020, Fig. 5d) and memory B cells (p = 0.010, Fig. 5e) as well as high level of Macrophages M1 (p = 0.048, Fig. 5c) were highly related with poor prognosis of LUAD patients.

Expression profile of immunomodulators.

In the present study, we quantified 11 immunomodulators (CTLA4, ICOS, ICOSLG, IFN- γ, LAG3, NKG2A, PD − 1, PD − L1, TIGIT, TIM3 and VISTA). The expression of CTLA4 (p < 0.001), IFN- γ (p = 0.002), LAG3 (p < 0.001), PD − 1 (p < 0.001) and TIGIT (p < 0.001) was significantly up-regulated in LUAD tissues compared with that in normal lung tissues, whereas, ICOS (p < 0.001), PD − L1 (p < 0.001), TIM3 (p < 0.001) and VISTA (p < 0.001) was significantly down-regulated (Additional file 1: Fig. S3a). Moreover, in

high- risk group, 8 immunomodulators (CTLA4 ($p < 0.001$), ICOS ($p < 0.001$), IFN- γ ($p = 0.032$), LAG3 ($p = 0.003$), PD − 1 ($p < 0.001$), TIGIT ($p < 0.001$), TIM3 ($p = 0.036$) and VISTA($p < 0.001$)) were significantly decreased (Additional file 1: Fig. S3b).

GSEA analysis

To explore the basic biological mechanisms of the IRGPs signature, we carried out GSEA analysis. A total of 27 KEGG pathways were enriched between high- risk and low- risk group, among which, various KEGG pathways were highly related with immune system, such as “Chemokine signaling pathway”, “Intestinal immune network for IgA production”, “T cell receptor signaling pathway”, “B cell receptor signaling pathway”, “Primary immunodeficiency”, “Leukocyte transendothelial migration”, “MAPK signaling pathway” and “PPAR signaling pathway” (Fig. 6).

Discussion

Immune system plays a vital role in the development and progression of tumors [14]. In addition, immunotherapy has revolutionized cancer treatment and has been approved for the treatment of manifold tumors recently [15], which has shed new light on the therapy of LUAD. Thence, in the current study, we collected immune-related gene matrix from TCGA to construct immune-related gene pairs (IRGPs). A total of 12334 IRGPs were paired and a prognostic IRGPs signature based on 8 IRGPs was established with multivariate Cox regression analysis. According to immune risk score, LUAD patients were divided into high- and low- risk group. Survival analysis demonstrated that high- risk patients predicted poorer clinical outcome compared with low- risk patients. Moreover, the result of multivariate Cox regression analysis showed that the immune risk score was an independent prognostic factor for LUAD patients. Then, we further evaluated the predictive effective and accuracy of this prognostic IRGPs signature for 1- and 3-year OS and validated this finding. The AUC of this signature in TCGA set for predicting 1- and 3- year OS was 0.867 and 0.870, respectively, which was significantly higher than the AUC of clinical parameters (Age [1-year: 0.528; 3-years: 0.687]; Gender [1-year: 0.621; 3-years: 0.408]; TNM stage [1-year: 0.724; 3-years: 0.664]; EGFR mutation [1-year: 0.501; 3-years: 0.465]; KRS mutation [1-year: 0.505; 3-years: 0.268]; Radiotherapy [1-year: 0.569; 3-years: 0.568]; Chemotherapy [1-year: 0.528; 3-years: 0.504]; Smoking [1-year: 0.511; 3-years: 0.410]). Similar phenomenon was observed in GEO set and Total set (GEO set: IRGPs signature [1-year: 0.819; 3-years: 0.803]; Age [1-year: 0.589; 3-years: 0.565]; Gender [1-year: 0.532; 3-years: 0.545]; TNM stage [1-year: 0.759; 3-years: 0.718]; Radiotherapy [1-year: 0.509; 3-years: 0.555]; Chemotherapy [1-year: 0.487; 3-years: 0.549]; Smoking [1-year: 0.529; 3-years: 0.521]; Histologic grade [1-year: 0.420; 3-years: 0.427]; Total set: IRGPs signature [1-year: 0.845; 3-years: 0.801]; Age [1-year: 0.556; 3-years: 0.551]; Gender [1-year: 0.578; 3-years: 0.530]; TNM stage [1-year: 0.742; 3-years: 0.716]; EGFR mutation [1-year: 0.501; 3-years: 0.465]; KRS mutation [1-year: 0.505; 3-years: 0.268]; Radiotherapy [1-year: 0.536; 3-years: 0.558]; Chemotherapy [1-year: 0.500; 3-years: 0.519]; Smoking [1-year: 0.518; 3-years: 0.512]; Histologic grade [1-year: 0.420; 3-years: 0.427]). All data suggested that this prognostic IRGPs signature was suitable for estimating 1- and 3- year survival probability of LUAD patients. In recent years, several prognostic signatures have been established for exploring prognosis-

related biomarkers and predicting the 1- and/or 3- year OS of LUAD. For example, Guo et al. have built an immune signature for 1- and 3- year survival rate of LUAD [16]. The AUC for 1- and 3- year of the immune signature in training cohort was 0.70 and 0.68, respectively, and in validation cohort was 0.72 and 0.73, respectively, all of which were inferior to that in this study. Similarly, a study reported an immune-related signature [17], which AUC of 1- (0.78) and 3- (0.76) year was also lower than that of this study. In addition, Zhang et al. constructed a glycolysis-related gene prognostic signature with the AUC = 0.72 [18]. Meanwhile, a research developed a autophagy-related gene prognostic signature with the AUC = 0.615 [19]. Both the AUC of these two studies were inferior to that of this IRGPs signature.

Nowadays, emerging studies show TME is critical for the initiation, progression and metastasis of cancers and therapy targeting the TME seem to be an encouraging method to overthrow therapeutic escape issues. In this study, we also calculated the proportions of 22 TIICs in TME of LUAD and found that macrophages were the most abundant immune cell, which was in line with previous finding. In addition, in high- risk patients, the proportions of macrophages M0 and M1 were significantly increased. Furthermore, survival analysis showed high level of macrophages M1 patients exhibited poorer prognosis than low level patients. Macrophages M0 may polarized into different types: M1 and M2 [20, 21]. Several studies have revealed that macrophages M1 could promote LUAD cells proliferation, invasion and metastasis [22–24] and an increased density of macrophages M1 in TME was associated with poor prognosis in LUAD [25, 26], which was in accord with the finding in this study. Previous researches demonstrated accumulating memory B cells were strongly correlated with favorable clinical outcomes in various tumors [27]. In TME, B cells could produce antibodies and present antigen to regulate innate immunity and promote antigen-specific immune responses to repress tumor development [28–30]. Here, we found that low- risk patients had higher proportions of memory B cells and patients with high level of memory B cells had better prognosis.

Although this prognostic IRGPs signature showed a well predictive accuracy and effectiveness for LUAD patients in this study, there are still some limitations needed to be addressed. Firstly, owing to all samples were collected from public database, the potential selection bias could not be excluded. Secondly, due to the signature was constructed with microarray expression and RNA-seq data which is costly and time-consuming, it is difficult to popularize in clinical applications. Finally, there was no experimental research conducted to testify the finding in this study. Hence, further investigation is demanded to examine the discovery of this research both in vitro and in vivo.

Conclusions

Take together, in the current study, we developed a prognostic IRGPs signature with 8 immune-related genes pairs for predicting 1- and 3- year overall survival in LUAD. This signature will be an available predictive tool to identify patients who might benefit from immunotherapy and provide a convenient tool for risk assessment and prognosis assessment.

Abbreviations

LC: Lung cancer; NSCLC: Non-small cell lung cancer; LUAD: lung adenocarcinoma; IRGPs: Immune-related gene pairs; TCGA: The Cancer Genome Altas; GEO: Gene Expression Omnibus; ROC: the receiver operating characteristic curve; AUC: the area under curve of ROC; OS: overall survival; TME: the tumor microenvironment; TIICs: tumor-infiltrating immune cells; GSEA: Gene set enrichment analysis.

Declarations

Acknowledgements

Not applicable.

Funding

Not applicable.

Availability of data and materials

All data were gathered from TCGA dataset (The Cancer Genome Altas, https://cancergenome.nih.gov) and GEO dataset (Gene Expression Omnibus, https://www.ncbi.nlm.nih.gov/geo/).

Authors' contributions

Research design: DM. Conducted experiments: CW, QT. Data collection and analysis: CW. Writing and editing manuscript: CW, QT. Revised the manuscript: CW, DM. All authors read and approved the manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

All authors confirm that there is no conflict of interest regarding this study.

References

1. Barta JA, Powell CA, Wisnivesky JP. Global Epidemiology of Lung Cancer. Ann Glob Health. 2019; 85(1): pii: 8.
2. 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.
3. Li L, Liu YD, Zhan YT, Zhu YH, Li Y, Xie D, et al. High levels of CCL2 or CCL4 in the tumor microenvironment predict unfavorable survival in lung adenocarcinoma. Thorac Cancer. 2018; 9(7):775-784.
4. Zeng D, Zhou R, Yu Y, Luo Y, Zhang J, Sun H, et al. Gene expression profiles for a prognostic immunoscore in gastric cancer. Br J Surg. 2018;105(10):1338-1348.
5. Li L, Ouyang Y, Wang W, Hou D, Zhu Y. The landscape and prognostic value of tumor-infiltrating immune cells in gastric cancer. 2019; 7:e7993.
6. Chen Y, Chen H, Mao B, Zhou Y, Shi X, Tang L, et al. Transcriptional Characterization of the Tumor Immune Microenvironment and Its Prognostic Value For Locally Advanced Lung Adenocarcinoma In A Chinese Population. Cancer Manag Res. 2019; 11:9165-9173.
7. Kurbatov V, Balayev A, Saffarzadeh A, Heller DR, Boffa DJ, Blasberg JD, et al. Digital Inference of Immune Microenvironment Reveals “Low Risk” Subtype of Early Lung Adenocarcinoma. Ann Thorac Surg. 2020; 109(2):343-349.
8. Corthay, A. Does the immune system naturally protect against cancer? Front Immunol.. 2014; 5:197.
9. Lakshmi Narendra B, Eshvendar Reddy K, Shantikumar S, Ramakrishna S. Immune system: a double-edged sword in cancer. Inflamm. Res. Inflamm Res. 2013; 62(9):823-34.
10. Zhao J, Wang, Hu G, Wei B. A 6-Gene Risk Signature Predicts Survival of Glioblastoma Multiforme. Biomed Res Int. 2019; 2019:1649423.
11. 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.
12. Sun XY, Yu SZ, Zhang HP, Li J, Guo WZ, Zhang SJ. A signature of 33 immune-related gene pairs predicts clinical outcome in hepatocellular carcinoma. Cancer Med. 2020; 9(8):2868-2878.
13. Zhang L, Zhu P, Tong Y, Wang Y, Ma H, Xia X, et al. An immune-related gene pairs signature predicts overall survival in serous ovarian carcinoma. Onco Targets Ther. 2019; 12:7005-7014.
14. McGranahan T, Therkelsen KE, Ahmad S, Nagpal S. Current state of immunotherapy for glioblastoma. Curr Treat Options Oncol. 2019; 20(3):24.
15. Hsu PC, Wang CW, Kuo SC, Lin SM, Lo YL, Huang AC, et al. The Co-Expression of Programmed Death-Ligand 1 (PD-L1) in Untreated EGFR-Mutated Metastatic Lung Adenocarcinoma. 2020; 8(2). pii: E36.
16. Guo D, Wang M, Shen Z, Zhu J. A new immune signature for survival prediction and immune checkpoint molecules in lung adenocarcinoma. J Transl Med. 2020; 18(1):123.

17. Zhang M, Zhu K, Pu H, Wang Z, Zhao H, Zhang J, et al. An immune-related signature predicts survival in patients with lung adenocarcinoma. Front Oncol. 2019; 9:1314.
18. Zhang L, Zhang Z, Yu Z. Identification of a novel glycolysis-related gene signature for predicting metastasis and survival in patients with lung adenocarcinoma. J Transl Med. 2019;17(1):423.
19. Zhu J, Wang M, Hu D. Development of an autophagy-related gene prognostic signature in lung adenocarcinoma and lung squamous cell carcinoma. 2020; 8:e8288
20. Martinez FO, Gordon S, Locati M, Mantovani A. Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: new molecules and patterns of gene expression. J Immunol. 2006;177(10):7303–7311.
21. Martinez FO, Sica A, Mantovani A, Locati M. Macrophage activation and polarization. Front Biosci. 2008; 13:453–461.
22. Wang A, Lu C, Ning Z, Gao W, Xie Y, Zhang N, et al. Tumor-associated macrophages promote Ezrin phosphorylation- mediated epithelial-mesenchymal transition in lung adenocarcinoma through FUT4/LeY up-regulation. 2017; 8(17):28247-28259.
23. Hegab AE, Ozaki M, Kagawa S, Hamamoto J, Yasuda H, Naoki K, et al. Tumor associated macrophages support the growth of FGF9-induced lung adenocarcinoma by multiple mechanisms. Lung Cancer. 2018; 119:25-35.
24. Guilbaud E, Gautier EL, Yvan-Charvet L. Macrophage Origin, Metabolic Reprogramming and IL-1 Signaling: Promises and Pitfalls in Lung Cancer. Cancers (Basel). 2019; 11(3). pii: E298.
25. Mei J, Xiao Z, Guo C, Pu Q, Ma L, Liu C, et al. Prognostic impact of tumor-associated macrophage in filtration in non-small cell lung cancer: A systemic review and meta-analysis. 2016; 7(23):34217-28.
26. Ohtaki Y, Ishii G, Nagai K, Ashimine S, Kuwata T, Hishida T, et al. Stromal macrophage expressing CD204 is associated with tumor aggressiveness in lung adenocarcinoma. J Thorac Oncol. 2010; 5(10):1507–15.
27. Ho KH, Chang CJ, Huang TW, Shih CM, Liu AJ, Chen PH, et al. Gene landscape and correlation between B- cell infiltration and programmed death ligand 1 expression in lung adenocarcinoma patients from The Cancer Genome Atlas data set. PLoS One. 2018; 13(12):e0208459.
28. Nelson BH. CD20+ B cells: the other tumor-infiltrating lymphocytes. J Immunol. 2010; 185(9):4977–82.
29. Linnebacher M, Maletzki C. Tumor-infiltrating B cells: The ignored players in tumor immunology. Oncoimmunology. 2012; 1(7):1186–8.
30. Tsou P, Katayama H, Ostrin EJ, Hanash SM. The Emerging Role of B Cells in Tumor Immunity. Cancer Res. 2016; 76(19):5597–601.

Figures

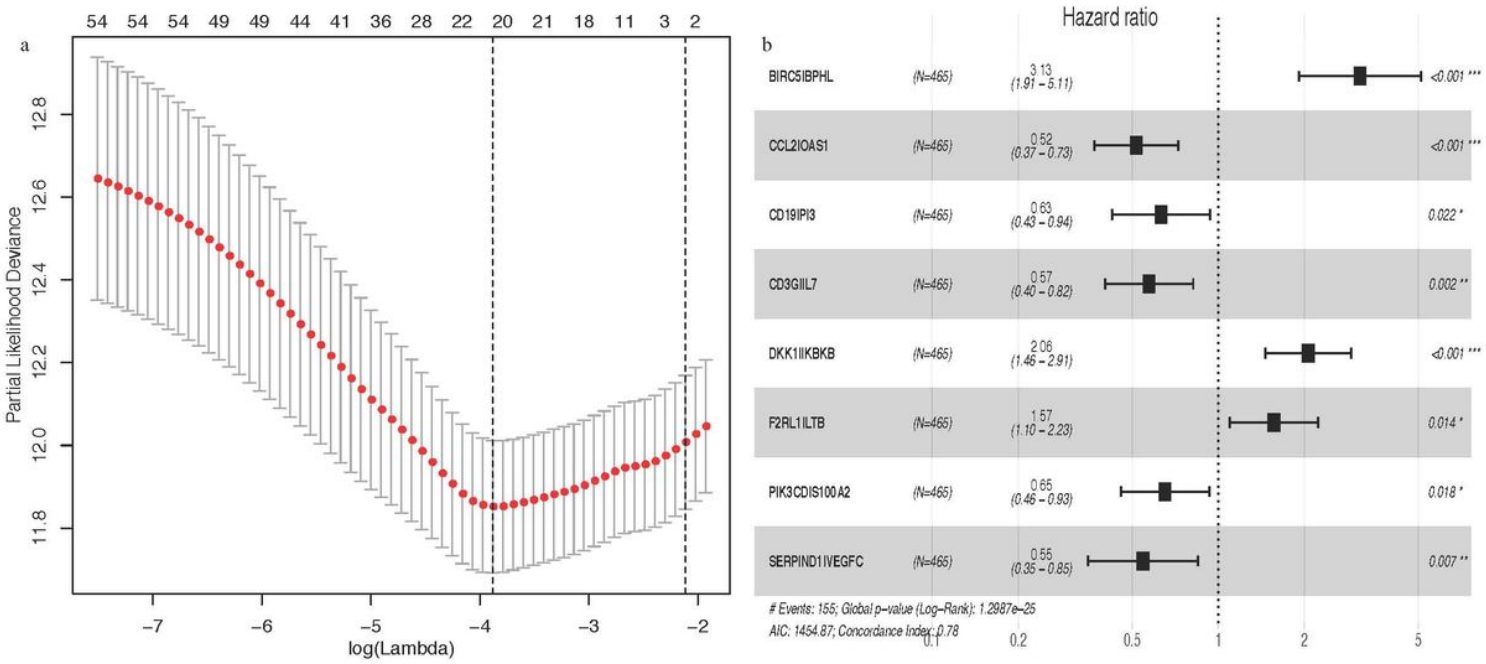

Figure 1

Construction of IRGPs signature. (a) “Leave- one-out-cross-validation” for parameter selection in LASSO regression. (b) The forest map of multivariate Cox regression analysis.

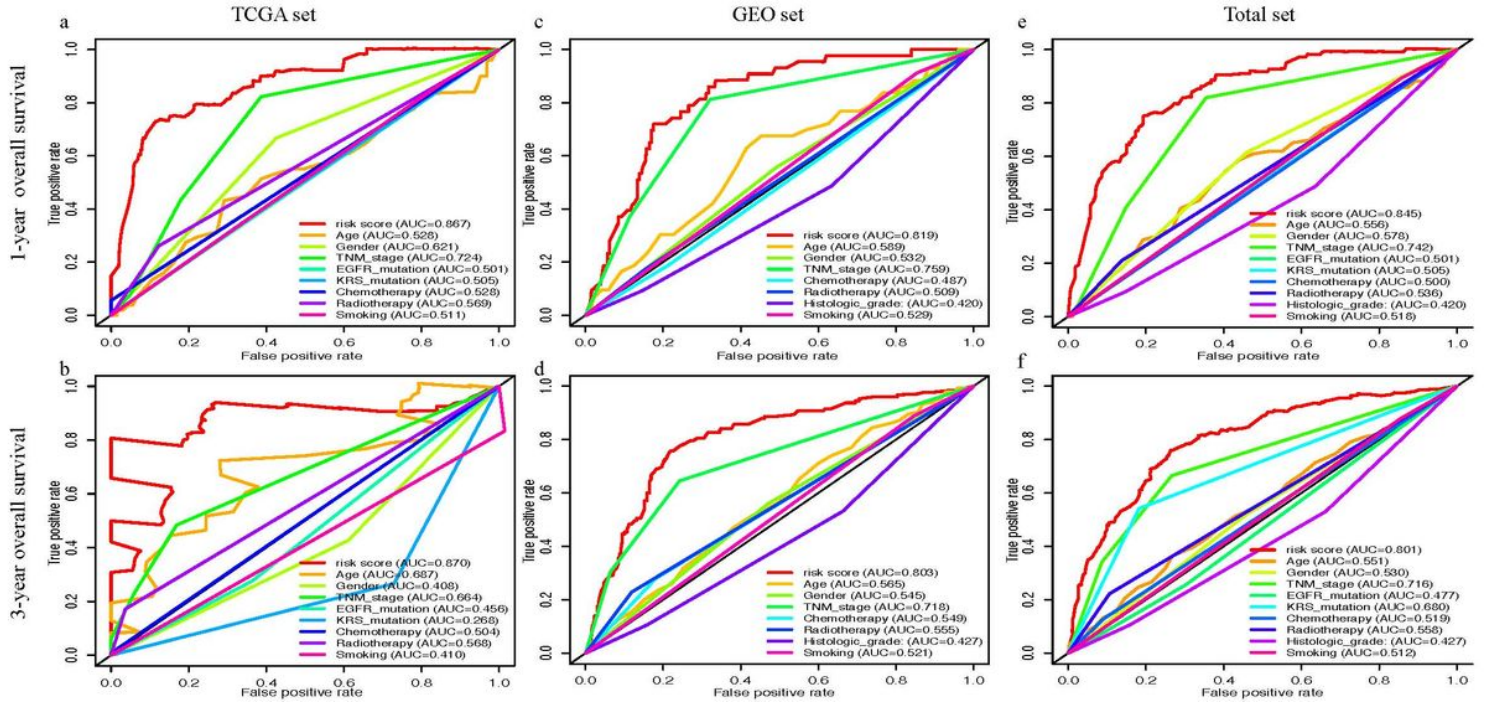

Figure 2

Evaluation of IRGPs signature. (a) The area under the receiver operating characteristic (ROC) curve (AUC) for 1-year overall survival of LUAD patients in TCGA set. (b) 3-year overall survival in TCGA set. (c) 1-year overall survival in GEO set. (d) 3-year overall survival of in GEO set. (e) 1-year overall survival in Total set. (f) 3-year overall survival in Total set.

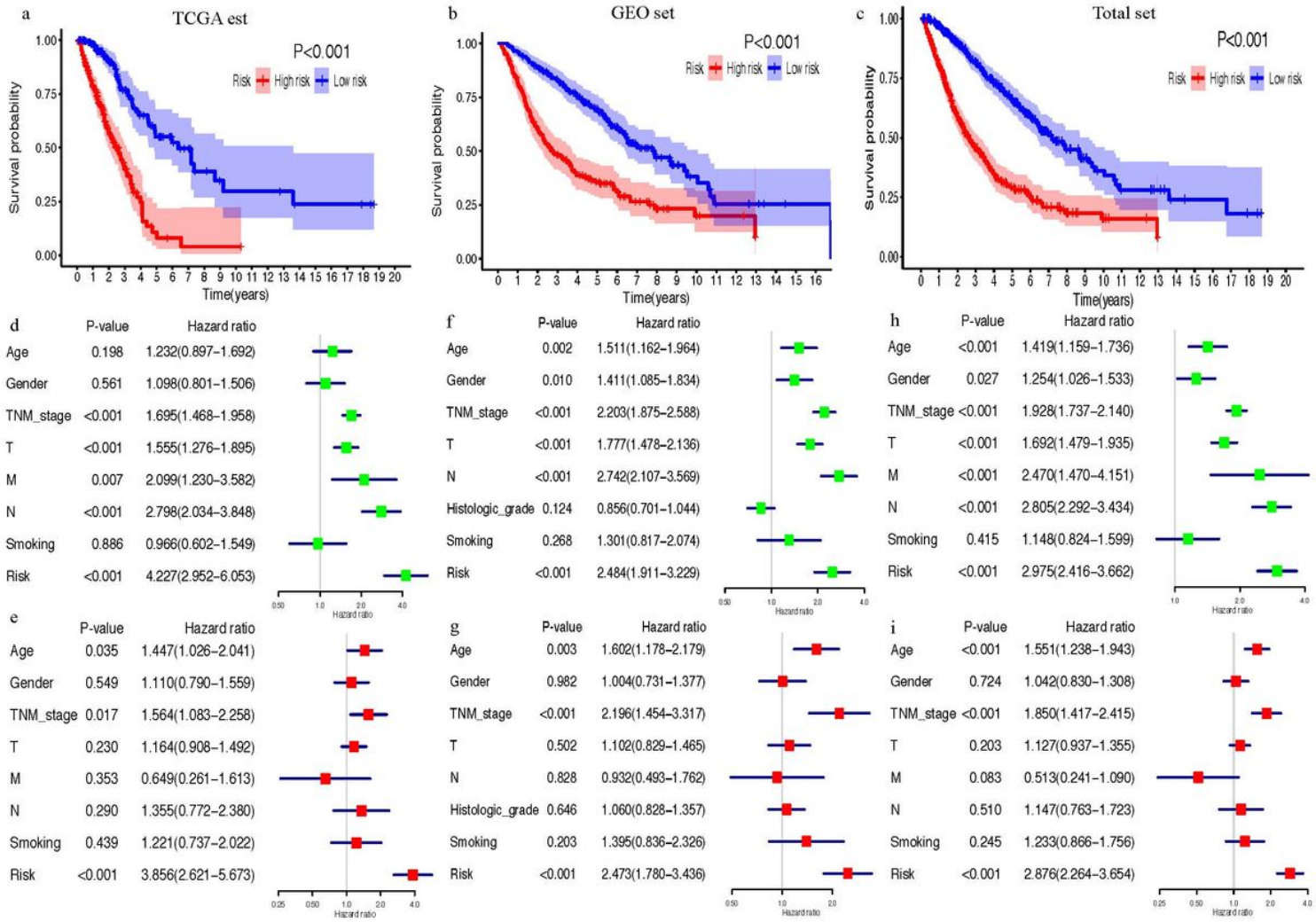

Figure 3

The IRGPs signature is an independent prognostic factor of survival. Survival difference between high- and low- risk group (a) In TCGA set. (b) In GEO set. (c) In Total set. (d) The result of univariable Cox regression analysis in TCGA set. (e) The result of multivariable Cox regression analysis in TCGA set. (f) The result of univariable Cox regression analysis in GEO set. (g) The result of multivariable Cox regression analysis in GEO set. (h) The result of univariable Cox regression analysis in Total set. (i) The result of multivariable Cox regression analysis in Total set.

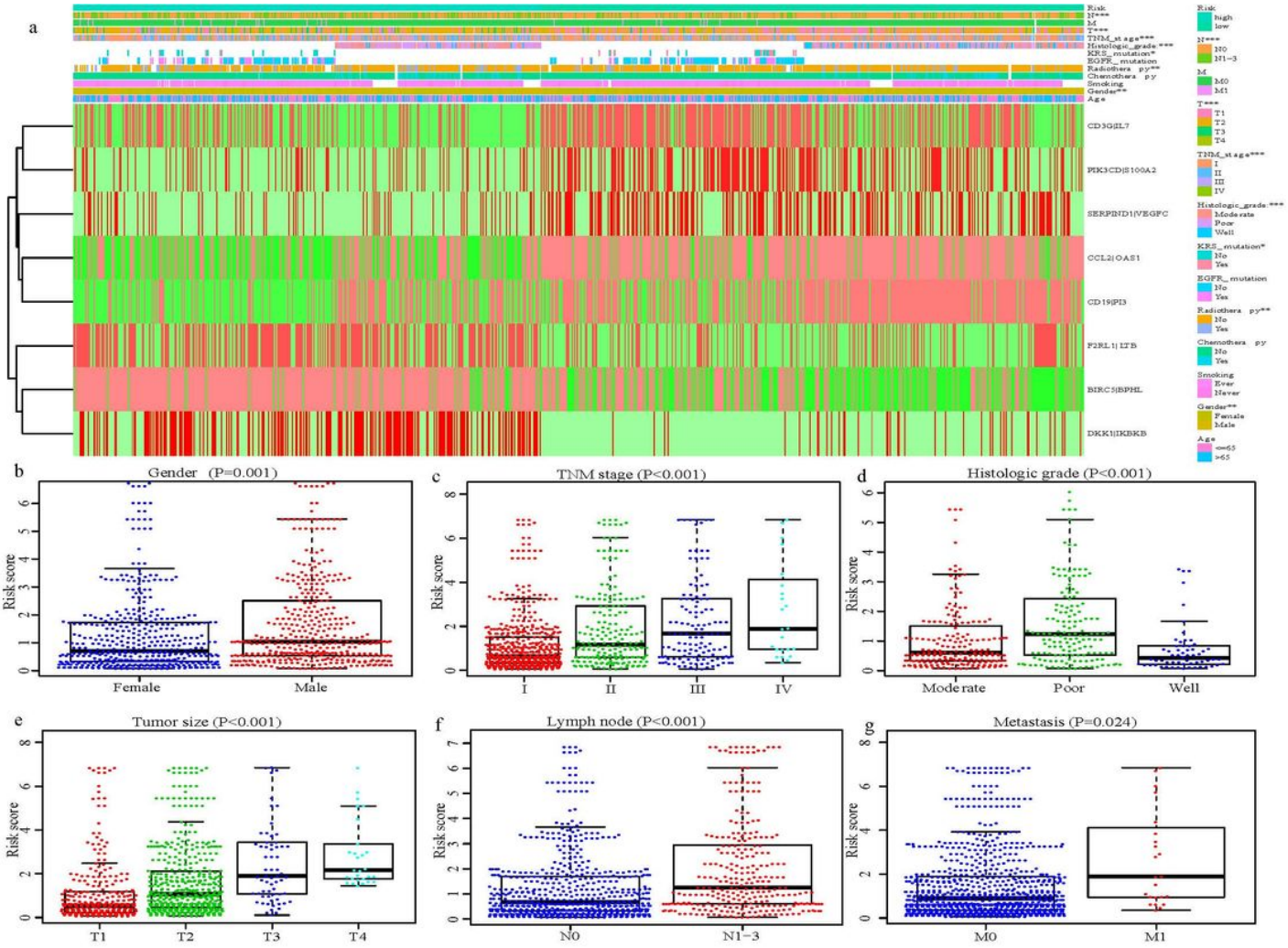

Figure 4

Correlation between risk score and clinical characteristics. (a) Heat map for the distribution of clinicopathological features between high- and low- risk group. (b) The difference of risk score among different gender. (c) among different TNM grades. (d) among different Histologic grades. (e) among different tumor size. (f) between with and without lymph node metastasis. (g) between with and without distance metastasis. *P<0.05, ** P<0.01, ***P<0.001.

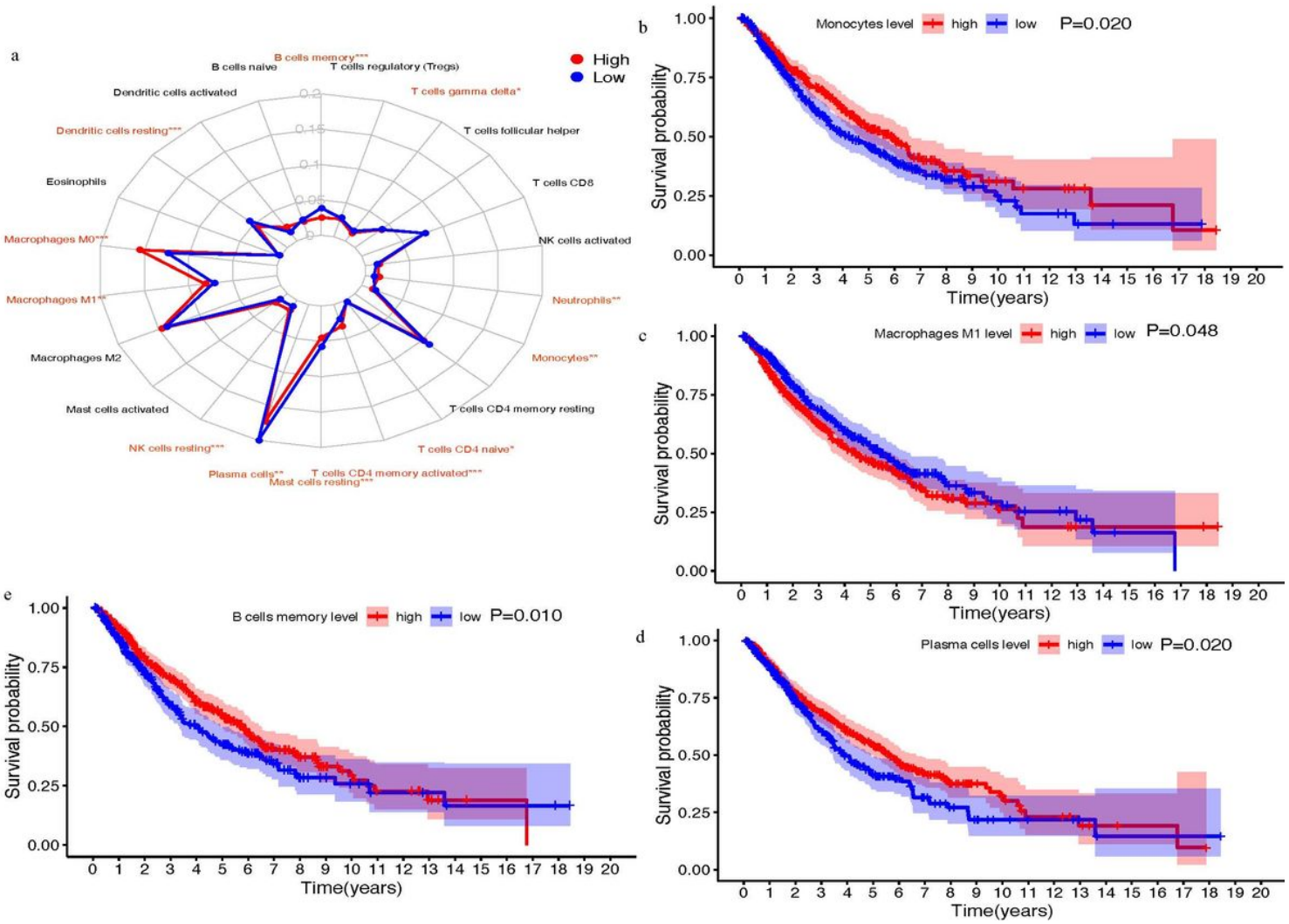

Figure 5

Relationship between IRGPs and tumor-infiltrating immune cells. (a) The difference of 22 TIICs between high- and low- risk group. (b) Survival difference between high- and low- proportions of Monocytes. (c) Macrophages M1, (d) Plasma cells. (e) memory B cells. *P<0.05, ** P<0.01, ***P<0.001.

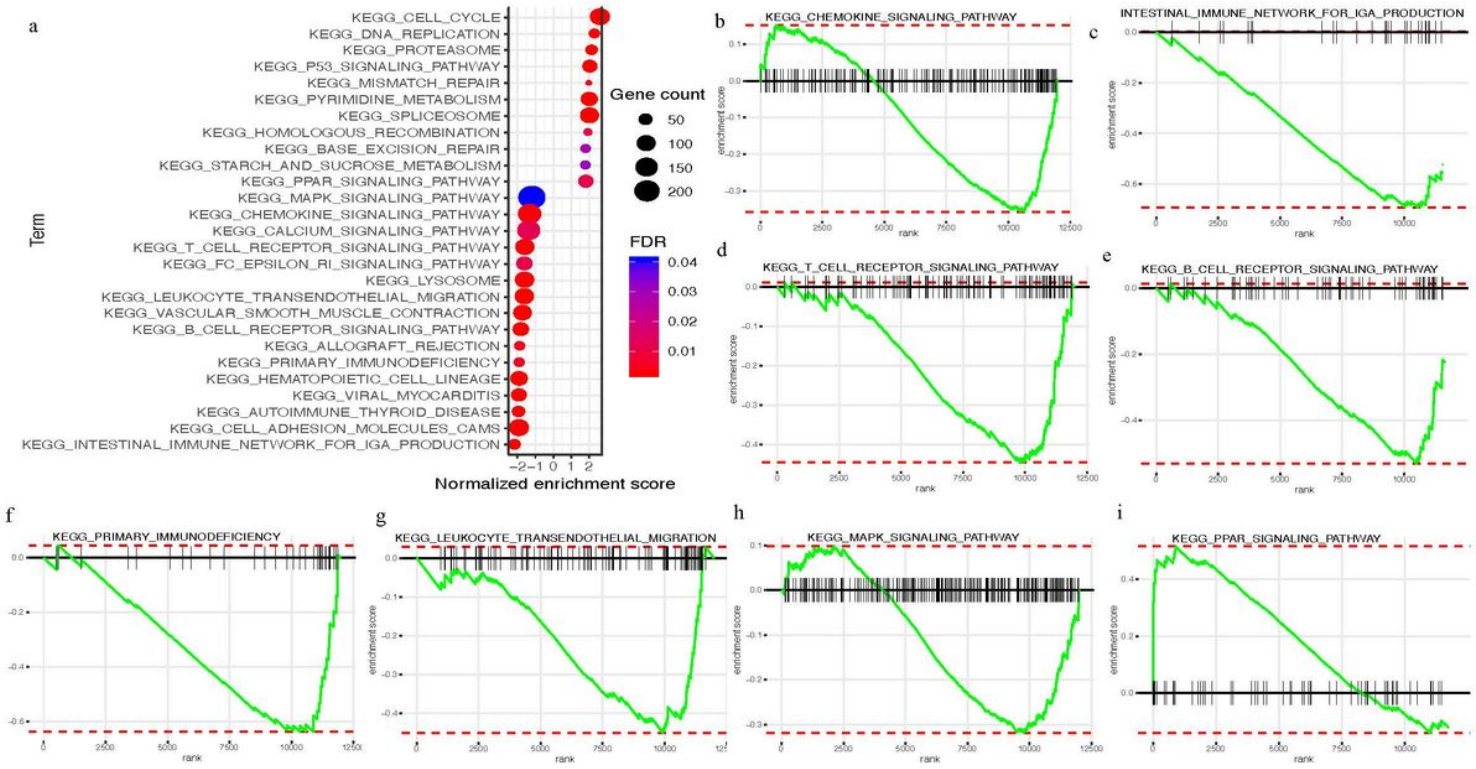

Figure 6

Gene set enrichment analysis (GSEA) between high and low immune risk groups. (a) 27 KEGG pathway-related gene sets. (b) “Chemokine signaling pathway”. (c) “Intestinal immune network for IgA production”. (d) “T cell receptor signaling pathway”, (e) “B cell receptor signaling pathway”. (f) “Primary immunodeficiency”. (g) “Leukocyte transendothelial migration”. (h) “MAPK signaling pathway”. (i) “PPAR signaling pathway”.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx