

RESEARCH

Open Access



An immune-related pseudogene signature to improve prognosis prediction of endometrial carcinoma patients

Shanshan Tang and Yiyi Zhuge*

*Correspondence:

zgyy1991@163.com
Department of Gynecology,
Hangzhou Women's
Hospital, No. 369 Kunpeng
Road, Shangcheng District,
Hangzhou 310008, Zhejiang,
China

Abstract

Background: Pseudogenes show multiple functions in various cancer types, and immunotherapy is a promising cancer treatment. Therefore, this study aims to identify immune-related pseudogene signature in endometrial cancer (EC).

Methods: Gene transcriptome data of EC tissues and corresponding clinical information were downloaded from The Cancer Genome Atlas (TCGA) through UCSC Xena browser. Spearman correlation analysis was performed to identify immune-related pseudogenes (IRPs) between the immune genes and pseudogenes. Univariate Cox regression, LASSO, and multivariate were performed to develop a risk score signature to investigate the different overall survival (OS) between high- and low-risk groups. The prognostic significance of the signature was assessed by the Kaplan–Meier curve, time-dependent receiver operating characteristic (ROC) curve. The abundance of 22 immune cell subtypes of EC patients was evaluated using CIBERSORT.

Results: Nine IRPs were used to build a prognostic signature. Survival analysis revealed that patients in the low-risk group presented longer OS than those in the high-risk group as well as in multiple subgroups. The signature risk score was independent of other clinical covariates and was associated with several clinicopathological variables. The prognostic signature reflected infiltration by multiple types of immune cells and revealed the immunotherapy response of patients with anti-programmed death-1 (PD-1) and anti-programmed cell death 1 ligand 1 (PD-L1) therapy. Function enrichment analysis revealed that the nine IRPs were mainly involved in multiple cancer-related pathways.

Conclusion: We identified an immune-related pseudogene signature that was strongly correlated with the prognosis and immune response to EC. The signature might have important implications for improving the clinical survival of EC patients and provide new strategies for cancer treatment.

Keywords: Uterine corpus endometrial cancer, Immune, Pseudogene, Survival, Signature



Background

Endometrial cancer (EC) was the mainly gynecological malignancy and ranked fourth in women malignancies worldwide behind breast, lung, and colorectal cancers, and there were approximately 61,880 new patients and 12,160 deaths of uterine corpus endometrial cancer in 2019, according to global cancer statistics [1, 2]. Early-stage patients presented a favorable clinical outcome with a 5-year survival rate of 95%. However, patients in advanced stage had a reduced 5-year survival rate of lower than 20% for stage IV [3]. Surgery, adjuvant chemotherapy, radiation, immunotherapy, and hormonal therapy are common models of treatment for patients with EC and have reduced the mortality of patients to a certain extent in recent years. However, despite rapid advances in the treatments of UCEC, the incidence and mortality rates are still increasing since it is a clinically heterogeneous disease characterized by different genetic background and pathogenesis [4]. Thus, to improve the survival rates of UCEC patients, it is imperative to identify mechanistic differences at the molecular level and develop novel predictive biomarkers to predict patient outcomes accurately.

Pseudogenes were initially considered as nonfunctional gene fossils or junk genes, which usually originating from their parent genes but have lost the capacity to encode functional proteins due to the accumulation of gene mutations [5, 6]. In recent years, accumulating evidence has strongly demonstrated that aberrant expression of pseudogenes plays vital roles in many human diseases such as cancer [7, 8]. Previous studies reported that pseudogenes mainly regulate gene expression at the post-transcriptional level and function as positive or negative regulators in tumor initiation or progression via two possible pathways. Numerous pseudogenes can act as competitive endogenous RNAs to competitively bind miRNAs with the protein-coding gene, therefore positively regulating gene expression [9, 10]. In addition, pseudogenes can also play a negative role through competing with their parent genes to destabilize RNA binding proteins, which results in a reduction in parent gene expression [11]. A previous study has revealed that LDHAP5 pseudogene was associated with the poor prognosis of ovarian serous cystadenocarcinoma via its targeting of EGFR [12]. Abnormally activated OCT4 pseudogene 5 (OCT4-pg5) contributed to enhanced cell proliferation by competing with miR-145 in EC via upregulating OCT4 expression [9]. Although a number of studies have identified multiple differentially expressed pseudogenes in various cancers, the generally predictive roles of immune-related pseudogenes in EC remain unclear. Since no studies have systematically evaluated immune-related pseudogenes in EC, discovering a number of promising prognostic biomarkers and exploring the underlying molecular mechanisms are eagerly needed. Therefore, for the first time, we identified an immune-related pseudogene signature in EC.

In this study, we screened the immune-related pseudogenes found in EC and explored the relationship between the screened immune-related pseudogenes and the prognosis of EC. A signature of immune-related pseudogenes in EC was developed using multivariate Cox regression analysis. Also, we investigated the prognostic value

of this signature in various clinical groups and the potential role of the immune-related pseudogenes signature in immune checkpoint inhibitors immunotherapy, with the aim of using it as a promising immune therapeutic target.

Discussion

With the rapid development of next-generation sequencing, growing transcriptomic data from the public database such as TCGA could be easily acquired. Through mining the big data from public databases, increasing studies have demonstrated pseudogenes, lncRNAs, miRNAs, and circRNAs play important roles in various cellular functions, including proliferation, cell differentiation, DNA stability, and tumorigenesis [9, 13, 14]. As a special group of lncRNAs, pseudogenes are remnants of their parental genes that lost the ability to encode proteins [15]. Accumulating evidence demonstrates crucial roles for pseudogenes in multiple cellular processes and various cancers [16].

In this study, we systematically gathered data from the TCGA portal and extracted immune-related pseudogenes via Spearman correlation analysis. Subsequently, we identified nine prognostic immune-related pseudogenes through univariate, LASSO, and multivariate Cox regression analyses and used them to develop a signature risk score, which was capable of classifying patients into the high-risk and low-risk groups with significantly different OS. Patients in the high-risk group presented shorter OS than patients in the low-risk group. As a heterogeneous disease with multiple clinicopathological features and risk factors, stratification analyses should be performed to confirm whether the signature was robust. The results illustrated that signature could well distinguish patients with all subgroups. In clinical practice, if we obtained the expression of the nine immune-related pseudogenes, the risk scores can be computed based on the coefficients. As a result, whether the patients are classified as low or high risk could be determined; thus the prognosis of patients could be predicted. Compared with four previous signatures, the AUC for 5-year survival and principal components analysis (PCA) illustrated that the present signature had powerful predictive ability. The risk signature was strongly correlated with age, grade, stage, and neoplasm status. The risk score remained an independent prognostic factor by combination with other clinicopathological characteristics via univariate and multivariate Cox regression analyses. Besides, KEGG enrichment analysis illustrated that the immune-related pseudogene signature might be associated with multiple well-known cancer-related pathways, including the choline metabolism in cancer, endometrial cancer, cell cycle, prostate cancer pathways. Previous studies demonstrated obvious differences in choline and lipid metabolism and protein expression patterns between breast and prostate cancer cells in culture and in cancers derived from these cells [17, 18]. A recent study formed a five-gene signature related to the cell cycle that can predict OS for gastric cancer, which was useful for elaborating cell cycle mechanisms and for identifying patients with gastric cancer with poor OS [19]. As for EC, Liu et al. identified five cell cycle-related genes that were notably dysregulated between EC and normal tissues, and the signature based on the 5-cell cycle genes could predict EC prognosis exactly and independently [20]. These further indicated the important role of immune-related pseudogenes in EC.

Tumor microenvironment immune cells infiltration has been regarded as vital information for predicting the outcome and immunotherapy response in various malignancies according to the clinical trials with immune checkpoint inhibitors [21, 22]. The complex interaction between tumor cells and tumor microenvironment not only plays a critical role during the development of tumor, but also has significant effects on the efficacy of immunotherapy and the overall survival rate of patients [23]. Therefore, by applying the newly developed algorithm “CIBERSORT”, the immune cell infiltration levels of patients between high-risk and low-risk groups were assessed. We uncovered a significant difference between two risk groups in terms of CD8⁺ T cell, follicular helper T cells, regulatory T cell, gamma delta T cells, resting dendritic cells, and activated dendritic cells. We further analyzed the relationship between the signature risk score and immune cell infiltration abundance. The results demonstrated a significant positive/negative correlation between the risk score and multiple immune cells, such as activated myeloid dendritic cell, B cell naïve, activated mast cell, gamma delta T cell, regulatory T cell, and CD8⁺ T cell. This is consistent with the viewpoint that T cell and B cell responses play critical roles in diagnosis, prognosis, and survival of cancer patients [24, 25]. Immune checkpoint inhibitors have opened a new era of tumor immunotherapy. Immune regulation against immune checkpoints can result in cancer cell death by providing immune response signals to T cells [26]. The expression of immune checkpoint genes (PD-1, PD-L1, and CTLA-4) has been commonly used as predictive biomarkers for immune checkpoint inhibitors response [27, 28]. PD-L1 and PD-1 are expressed on human B cells and validated to act immunosuppressive roles in cancer progression [29]. In our study, CTLA4 and PD1 were higher expressed in a low-risk group than those in the high-risk group. Furthermore, the risk score was significantly negatively correlated with CTLA4 and PD1 expression. These findings, taken together with survival analysis, suggested that CTLA4 and PD1 have closely participated in immunosuppression, and their high expression is correlated with poor prognosis. Furthermore, we also found that patients with higher IPSs in the low-risk group tended to be candidates for immune checkpoint inhibitor. This demonstrated that the immune-related pseudogene signature is associated with an immunosuppressive phenotype and could be a potential predictive marker for immune checkpoint inhibitor response.

The nine immune-related pseudogenes were novel biomarkers of EC and had vital prognostic significance. The immune-related pseudogenes signature may serve as a new potential biomarker for immunotherapy and contribute to new therapeutic strategies. Furthermore, this study has opened many avenues for future research to pursue this topic. However, there are several limitations that should be noted. First, the number of the clinicopathological parameters released in publicly available datasets is limited and not comprehensive. The signature should be adequately validated in more independent cohorts with larger amounts of patients. Besides, the main research analysis approach we used is based on bioinformatics technology and further cellular experiments and animal studies (in vitro and in vivo) should be conducted to explore the predictive accuracy of the signature and to identify potential immune-related mechanisms in EC.

These conclusions need to be further validated for clinical application with additional experimental data, such as flow cytometry or immunohistochemistry. Despite these limitations, to our knowledge, this is the first study to focus on the immune-related pseudogenes signature of EC.

Conclusion

In conclusion, we proposed an immune-related pseudogenes signature, which can be used as an independent prognostic biomarker in stratifying risk subgroups in terms of OS for patients with EC. The nine-pseudogene signature had a superior performance for risk stratification compared to two existing signatures. The immune-related pseudogenes signature can assess survival and immune checkpoint inhibitor response of patients with anti-PD-1 and anti-CTLA4 therapy, potentially enabling more personalized and precise tumor immunotherapy in the future.

Methods

Acquisition of EC expression and clinical data

The three-level RNA-Seq expression profiles data and the corresponding clinical data were retrieved from the public TCGA portal (<https://cancergenome.nih.gov/>). The probe IDs were converted to the corresponding gene symbols based on their annotation files. When several probes matched to an identical gene symbol, we averaged them for further analysis. Briefly, genes were identified as protein-coding genes or pseudogenes based on their Ensembl IDs in the Ensembl database (<https://www.ensembl.org/>). The data of genes or pseudogenes expression matrix were obtained. Patients without complete overall survival (OS) time were excluded from this study. Finally, we enrolled expression and clinical data of 541 EC patients.

Immune-associated pseudogenes acquisition

We acquired 1793 unique immune-related genes from the ImmPort database (<https://immport.niaid.nih.gov/>). Subsequent, the immune-related genes matrix was extracted from the mRNA expression profile. A total of 13,602 pseudogenes were downloaded from the HUGO Gene Nomenclature Committee (<https://www.genenames.org/>). The relationship was calculated based on the expression value between pseudogenes and immune-related genes. Next, Spearman correlation analysis was conducted between immune-related genes matrix and pseudogenes expression levels in samples to identify immune-related pseudogenes according to the correlation coefficients and P values ($|\text{correlation coefficient}| > 0.4$, $P < 0.001$). Then, the expression matrix of immune-related pseudogenes in the TCGA database was extracted.

Construction of a risk score based on immune-related pseudogenes

To confirm the survival of immune-related pseudogenes, a risk score was designed to construct a unitive signature for EC. First, the univariate Cox regression analysis was performed for all possible immune-related pseudogenes to identify prognostic pseudogenes with significant prognostic value. Pseudogenes with $P < 0.01$ were screened

for subsequent analysis. We further narrowed the gene range after the univariate analysis by performing least absolute shrinkage and selection operator (LASSO)-penalized Cox regression analysis with 10-times cross-validations using the *glmnet* package in R. An optimal risk signature was constructed by performing the stepwise regression multivariate Cox analysis with Akaike information criteria (AIC) algorithm among the prognostic immune-related pseudogenes using the “*glmnet*” and “*survival*” packages. Final immune-related pseudogenes and their corresponding coefficients with the smallest AIC value were identified to form the prognostic signature in EC. The signature risk score of each sample was calculated using the following algorithm: Risk score = $\beta_{\text{pseudogene}_1} * \text{exprpseudogene}_1 + \beta_{\text{pseudogene}_2} * \text{exprpseudogene}_2 + \dots + \beta_{\text{pseudogene}_n} * \text{exprpseudogene}_n$. The risk score was computed by a linear combination of the expression level of lncRNAs weighted by the regression coefficient (β). *Exprgene* refers to the expression of immune-related pseudogenes in the sample, and β indicates the regression coefficient derived from multivariate Cox analysis. The β was calculated by log-transformed hazard ratio (HR) derived from multivariate Cox regression analysis. Based on the median risk score of the signature, all patients were separated into high- or low-risk groups. Principal component analysis (PCA) was carried out to profile expression patterns of grouped samples. Kaplan–Meier curve analysis, time-dependent ROC analysis, and patients’ survival distribution were performed.

Independence of the signature risk score and clinical relationship with other clinicopathological features

We compared the relationship between a single pseudogene’s expression level in the signature and clinicopathological variables to explore the impact of pseudogene on EC deeply. Next, to explore whether the signature risk score and clinicopathological variables were independent prognostic factors, we conducted univariate and multivariate Cox regression analyses for each factor.

Immune cell subtypes and its correlation with signature risk score

To further explore the differences between two risk groups in the abundance of infiltrating immune cells from gene expression profiles in EC, we used the CIBERSORT algorithm (<https://cibersortx.stanford.edu/>) coupled with a set of gene expression matrix features of 22 leukocyte subtypes (LM22) that distinguished 22 immune cell subpopulations from CIBERSORT to calculate immune cell infiltrations [9, 30]. We used the mRNA expression matrix data as the input files to evaluate the immune infiltrations of each sample through the CIBERSORT algorithm. The number of permutations was set to 1,000, and samples with a CIBERSORT output $P < 0.05$ were used for further analysis. For each sample, the sum of all estimated subpopulations is equal to 1.

Analysis of immune-checkpoint inhibitors response through immunophenoscore analysis

Immunomodulators or checkpoints, major histocompatibility complex molecules, effector cells, and immunosuppressive cells are four leading categories that determine the immunogenicity of cancer, which could be evaluated as immunophenoscore (IPS). IPS is calculated without bias using machine learning algorithms with z-scores ranged from 0 to 10 based on the gene expression in representative cell types, where higher z-scores are positively correlated to enhanced immunogenicity. IPS is computed using a scale with a range of 0–10 based on representative cell type gene expression z-scores, where higher scores are associated with increased immunogenicity. The IPS of patients with EC was retrieved from The Cancer Immunome Atlas (TCIA) (<https://tcia.at/home>) [31].

Functional enrichment analysis

We examined the co-expressed protein-coding genes between identified pseudogenes and genes expression matrix by calculating the Spearman correlation coefficient through expression profiles in 541 patients with EC ($|\text{correlation coefficient}| > 0.4$, $P < 0.001$). KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analysis of immune-related pseudogenes was performed to explore potential biological pathways that immune-related pseudogenes may be participated in.

Statistical analysis

Wilcoxon rank-sum test was performed to compare the differential abundances of immune infiltrates between low- and high-risk groups, which were presented with P -value by “vioplot” package. The Student’s t-test was used to test the expression changes of immune-checkpoint genes between two risk groups. We used the Chi-squared test to investigate the survival differences between the high-risk and low-risk groups. Kaplan–Meier (KM) survival curves and log-rank test were used to assess differences in survival between two risk groups using the “survminer” package. A time-dependent ROC analysis was used to compare the predictive accuracy. Univariate and multivariate Cox regression analyses were performed based on the “survival” package to identify EC’s independent prognostic factors. A P value < 0.05 was thought to be significant.

Results

Identification of immune-related pseudogenes and establishment of a signature

By conducting the correlation analysis based on $|\text{correlation coefficient}| > 0.4$ and $P < 0.001$ between immune-related gene and pseudogenes expression matrixes, a total of 312 immune-related pseudogenes were screened. Univariate Cox regression was used to analyze the prognostic pseudogenes of EC. A total of 20 pseudogenes were significantly relevant to the OS of EC (Fig. 1A). After minimizing overfitting by the LASSO regression algorithm, 16 pseudogenes were entered into the candidate pool for further analysis (Fig. 1B). To determine the optimal prognostic pseudogenes,

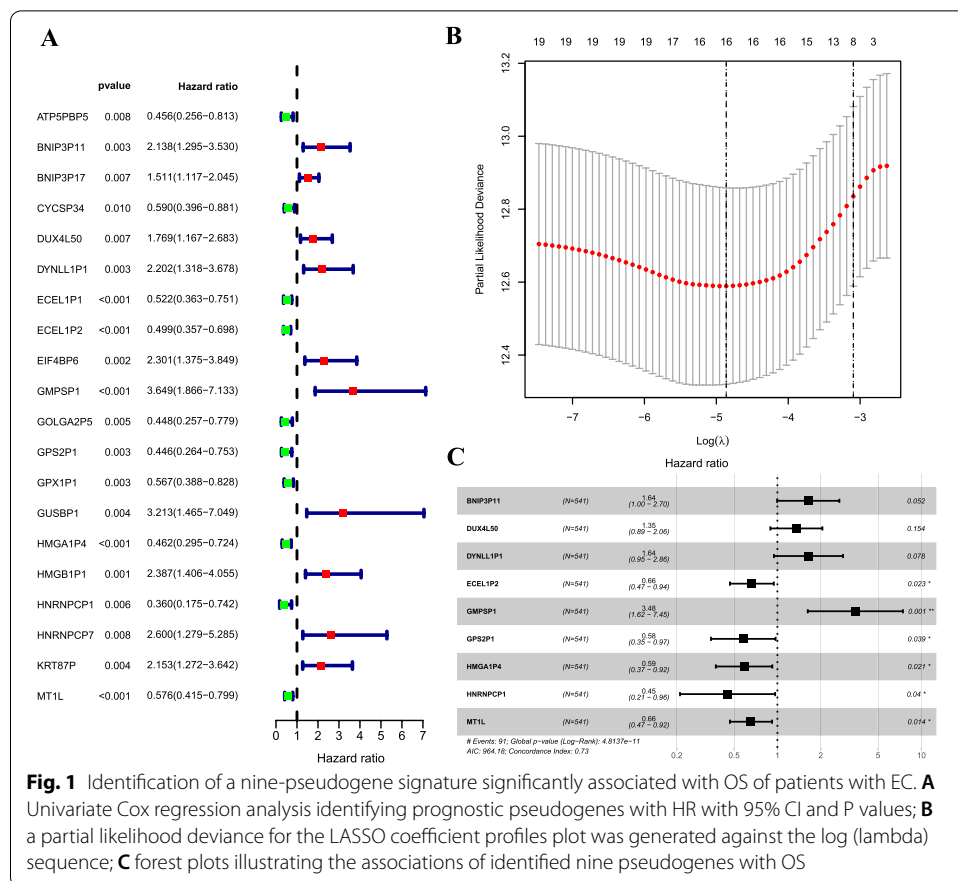


Fig. 1 Identification of a nine-pseudogene signature significantly associated with OS of patients with EC. **A** Univariate Cox regression analysis identifying prognostic pseudogenes with HR with 95% CI and P values; **B** a partial likelihood deviance for the LASSO coefficient profiles plot was generated against the log (lambda) sequence; **C** forest plots illustrating the associations of identified nine pseudogenes with OS

we adopted stepwise multivariate Cox proportional hazards regression to establish the risk score. Nine pseudogenes entered the final model (Fig. 1C, Table 1). The risk score of the signature for each sample was calculated as the followed equation: risk score = 0.493652865* expression of BNIP3P11 + 0.3035749* expression of DUX4L50 + 0.497475173* expression of DYNLL1P1 - 0.408662104* expression of ECEL1P2 + 1.246821769* expression of GMPSP1 - 0.545480258* expression of GPS2P1 - 0.533330167* expression of HMGA1P4 - 0.797024483* expression of HNRNPCP1 - 0.422476955* expression of MT1L. The risk score of each sample was computed according to the equation. Based on the median risk score of risk score, all patients were divided into low-risk and high-risk groups. The clinicopathological variables in the two risk groups are displayed in Table 2.

The performance of pseudogenes signature

The mRNA expression heatmap, signature risk score and survival status distribution of each patient are illustrated in Fig. 2A, B. With the risk score increased, more patients died. As revealed in Fig. 2C, patients in the high-risk group have a greater risk of mortality than those in the low-risk group ($P < 0.0001$). Additionally, PCA was used to explore the different distribution patterns between low- and high-risk

Table 1 Demographics and clinicopathological characteristics in high-risk and low-risk groups

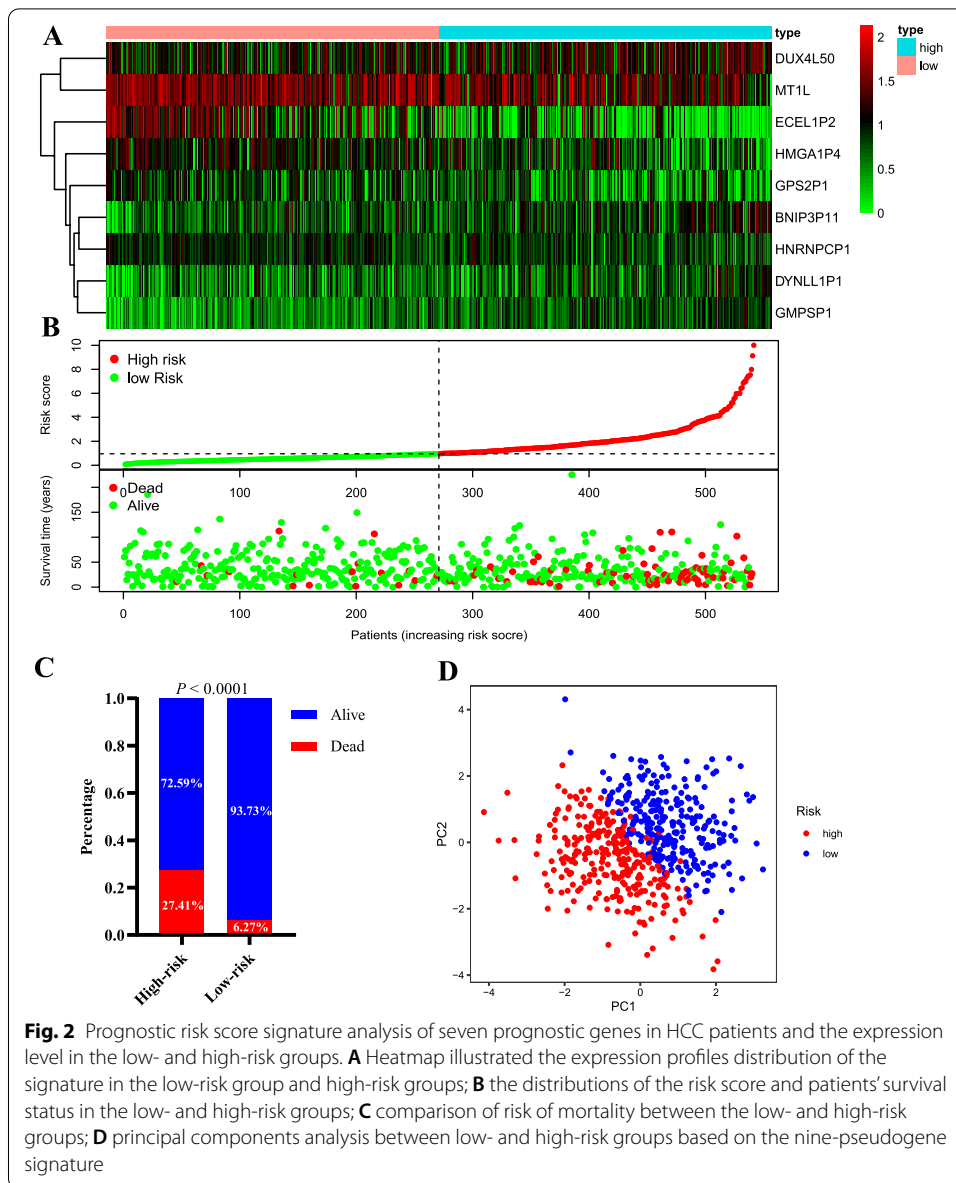
Variables	Subgroup	High-risk group (n = 270)	Low-risk group (n = 271)
Age	< 65 years	116	171
	> = 65 years	153	99
	Unknown	1	1
Stage	Stage I	141	197
	Stage II	28	23
	Stage III	81	42
	Stage IV	20	9
Grade	Grade I	18	80
	Grade II	39	81
	Grade III	203	109
	Unknown	10	1
Neoplasm	With neoplasm	60	18
	Without neoplasm	188	238
	Unknown	22	15
Survival	Alive	196	254
	Dead	74	17

Table 2 The nine pseudogenes identified from multivariate Cox regression analyses

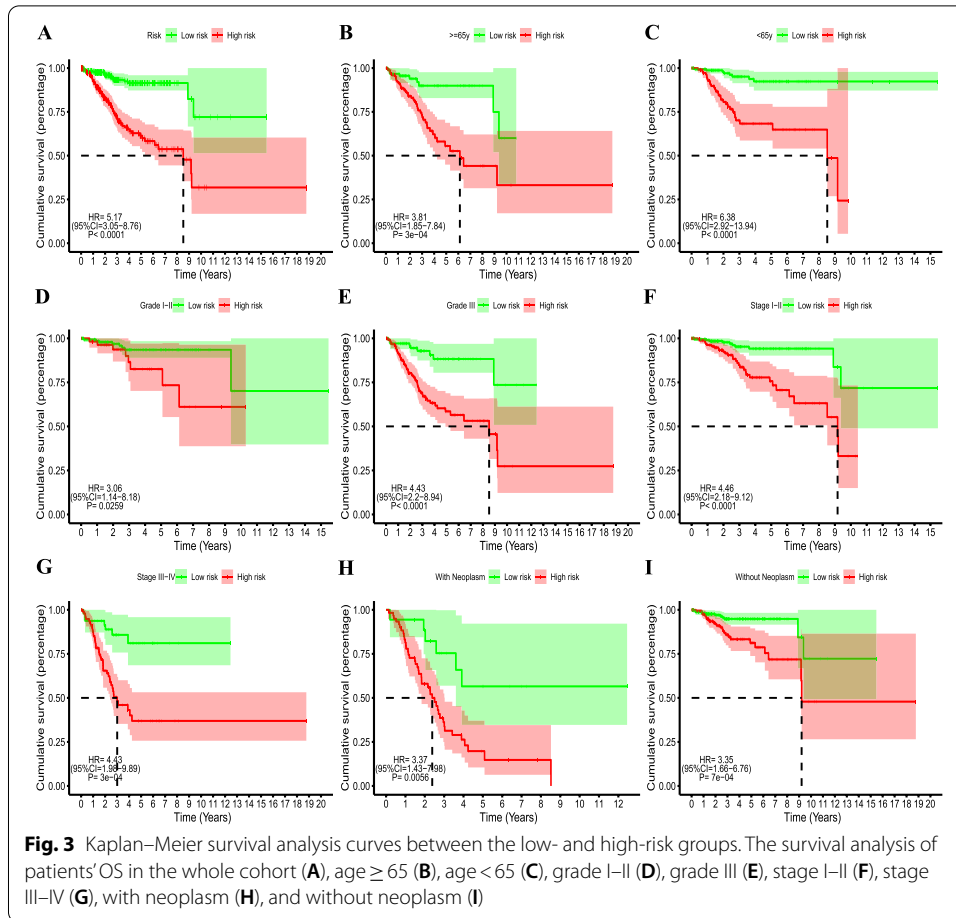
Gene symbol	Coefficient	HR	95% CI	P-value
BNIP3P11	0.493652865	1.6383	0.9956–2.6959	0.0521
DUX4L50	0.3035749	1.3547	0.8929–2.0554	0.1535
DYNLL1P1	0.497475173	1.6446	0.9466–2.8573	0.0776
ECEL1P2	−0.408662104	0.6645	0.4675–0.9446	0.0227
GMPSP1	1.246821769	3.4793	1.6245–7.4516	0.0013
GPS2P1	−0.545480258	0.5796	0.3455–0.9721	0.0387
HMGA1P4	−0.533330167	0.5866	0.3731–0.9224	0.0209
HNRNPCP1	−0.797024483	0.4507	0.2105–0.9648	0.0402
MT1L	−0.422476955	0.6554	0.4676–0.9186	0.0142

HR: hazard ratio; CI: confidence interval

groups on the basis of the immune-related pseudogenes. PCA analysis of the samples indicated that the clustering of the samples demonstrating a significant distinction between high- and low-risk groups (Fig. 2D). Therefore, immune-related pseudogenes were used to divide the EC patients into two categories, revealing that the immune status of the EC patients in the high-risk group was distinguishable from that in the low-risk group. Patients in the high-risk group illustrated



shorter overall survival (OS) than the low-risk group (HR = 5.17, 95% CI 3.05–8.76; $P < 0.0001$; Fig. 3A). These findings were further validated in multiple subgroups. As shown in Fig. 3B-I, the Kaplan–Meier curves demonstrated that significantly worse OS was observed in high-risk patients stratified by age (< 65 years or ≥ 65 years), grade (male or female), stage (stage I–II or stage III–IV), and neoplasm status (with or without), indicating that the signature was a stable prognostic biomarker for patients with EC.

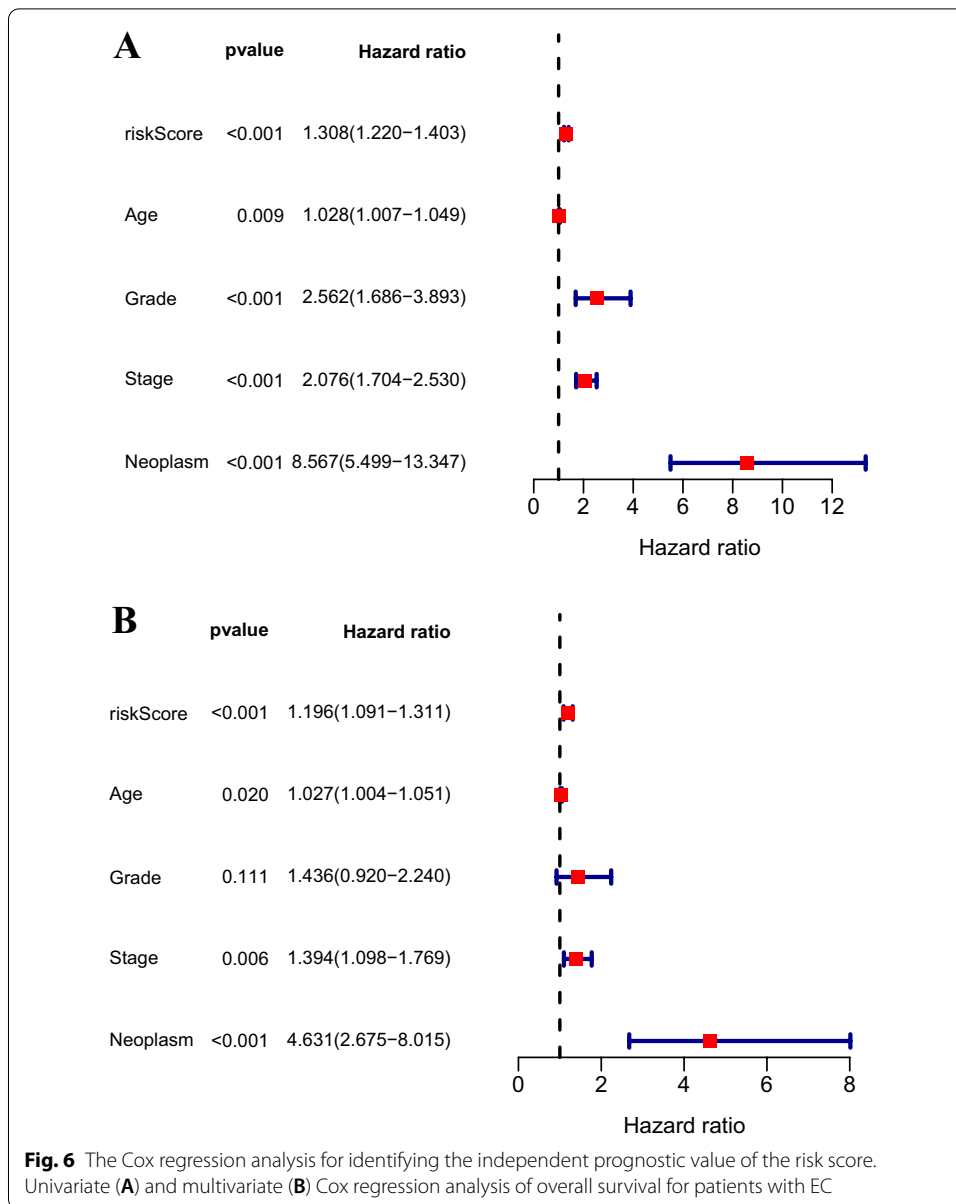


Correlation between the risk score and clinicopathological characteristics

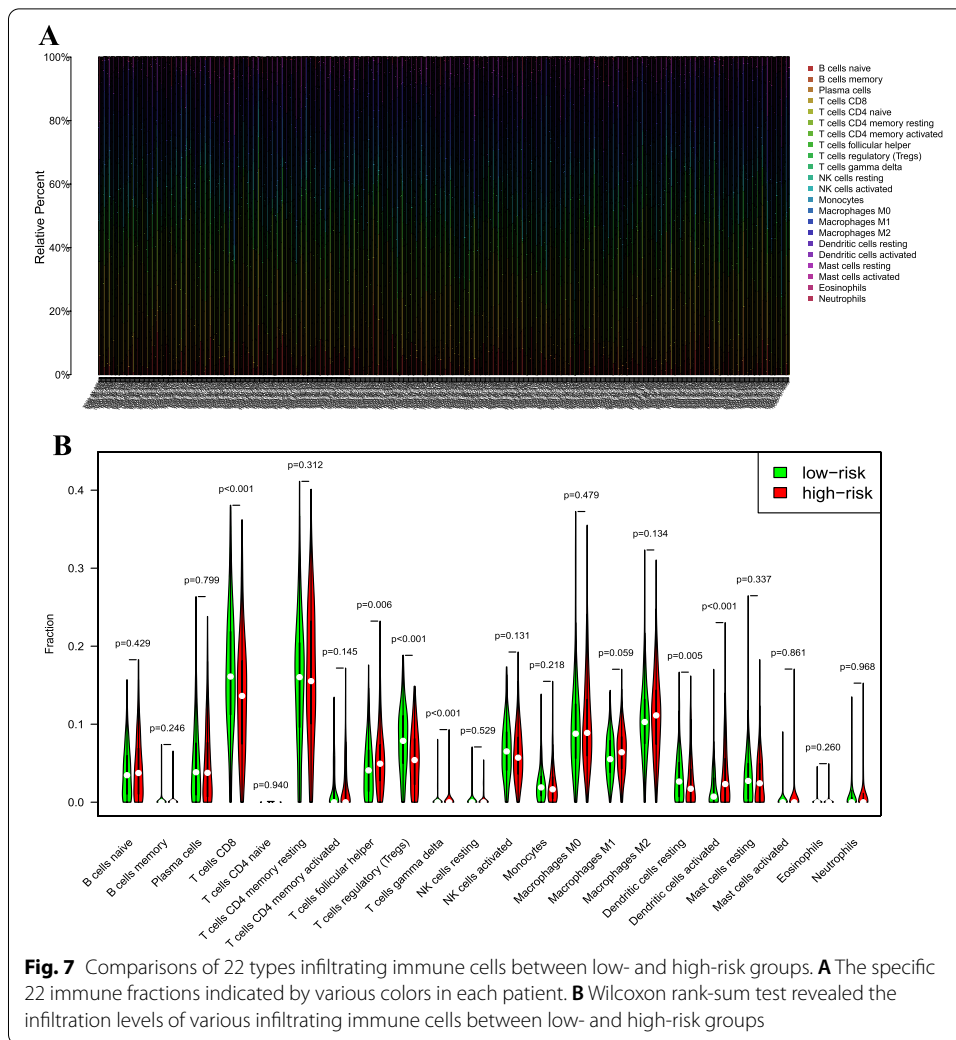
We analyzed the association between the immune-related pseudogene signature and clinicopathological factors. Significant differences were observed in age (< 65 years or ≥ 65 years), grade (male or female), stage (stage I–II or stage III–IV), and neoplasm status (with or without) (all $P < 0.0001$; Fig. 4A). Next, we compared the correlation between the expression level of a single pseudogene in the signature and clinicopathological characteristics to investigate the impact of pseudogenes on EC. There was a significant difference in the distribution of expression levels of multiple pseudogenes among age (< 65 years or ≥ 65 years), grade (male or female), stage (stage I–II or stage III–IV), and neoplasm status (with or without) (Fig. 4B), especially for the pseudogene ECEL1P2.

Performance comparison of the pseudogene signature with four existing signatures in survival prediction

We compared the prediction performance of the novel immune-related pseudogene signature with four recently published signatures: nine-lncRNA signature



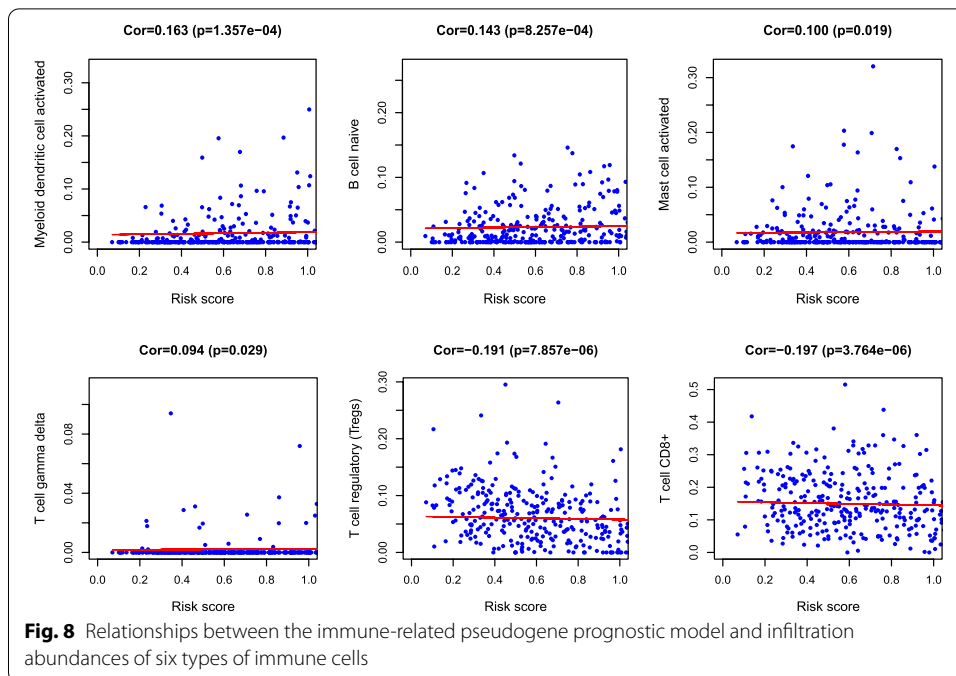
(LINC02387, FUT8-AS1, UBXN10-AS1, LINC00473, AL353194.1, FAM222A-AS1, AP002761.3, AL731566.2, and AP001021.2) derived from Xu’s research [9], nine-mRNA signature (CYP4F3, LYPLA2, CEL, PHGDH, GPAT3, HNMT, UCK2, CKM, and ACACB) derived from Jiang’s study [32], nine-mRNA signature (TP53, RAE1, RFC2, TAF10, DDB2, UMPS, TAF12, ERCC2, SEC61A1) derived from Liu’s research [33], and five-lncRNA signature (AL121906.2, BOLA3-AS1, LINC01833, AC016405.3, and RAB11B-AS1) from Jiang’s study [34] using the same TCGA EC patient cohort. As illustrated in Fig. 5, the AUC at 5-year of OS for our pseudogene signature is 0.780, which is significantly higher than that of Jiang’s mRNA



signature (AUC = 0.7), Xu’s lncRNA signature (AUC = 0.595), Liu’s mRNA signature (AUC = 0.748), and Jiang’s lncRNA signature (AUC = 0.712). Additionally, the number of pseudogenes included in our signature is smaller than that included in the two existing signatures. These results revealed the better prognostic performance of the immune-related pseudogene signature in predicting OS than previously published signatures.

Independence of the immune-related pseudogenes signature from clinicopathological characteristics

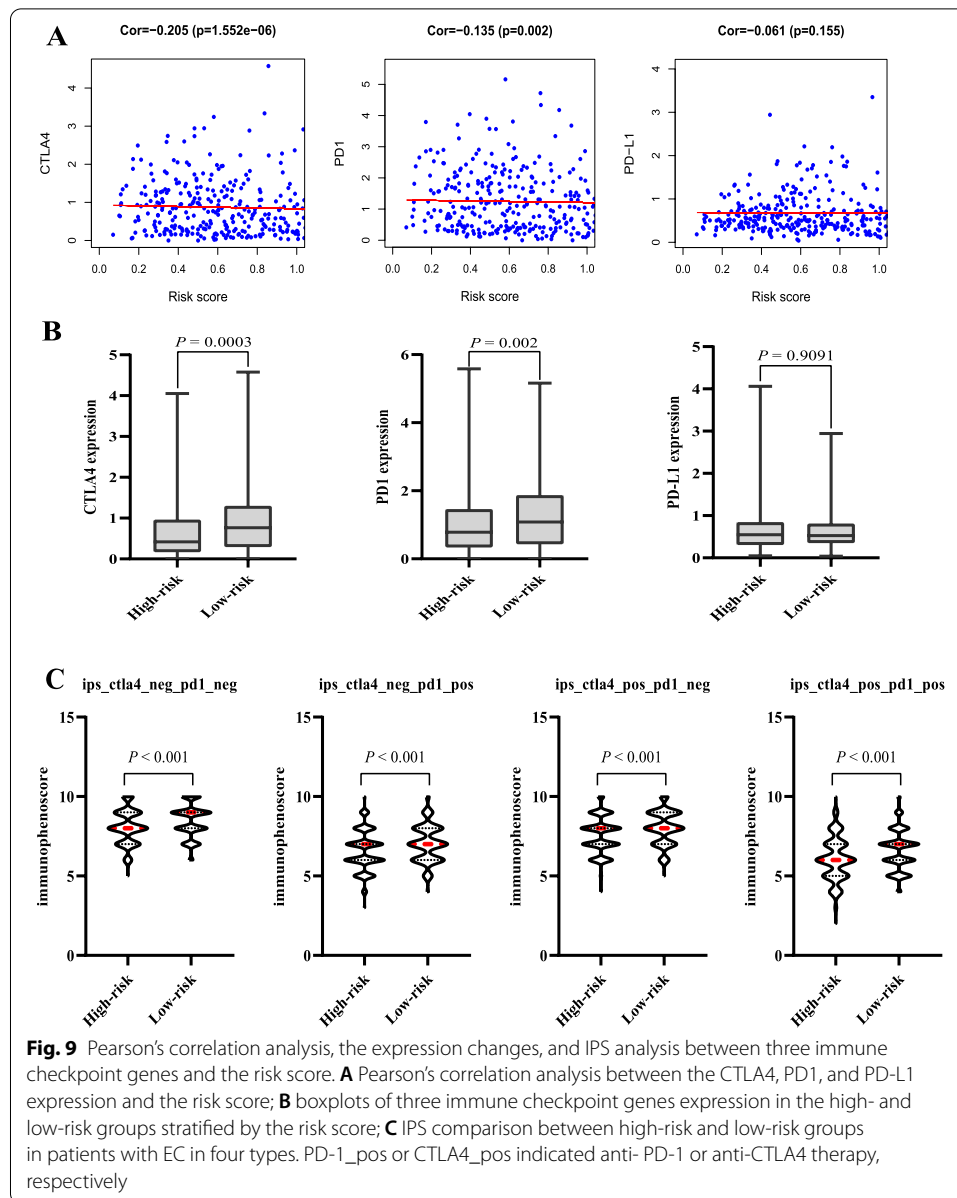
To further explore whether the pseudogene signature could predict OS independently of other clinicopathological factors, univariate and multivariate Cox regression analyses were conducted. We first conducted a univariate analysis of pseudogene signature and other potentially clinicopathological characteristics, and the results illustrated that all the factors were significantly associated with OS (all $P < 0.05$; Fig. 6A). Then, we



performed multivariate Cox regression analysis of them, and the results revealed that the pseudogene signature risk score (HR = 1.196, 95% CI 1.091–1.311, $P < 0.001$), stage (HR = 1.394, 95% CI 1.098–1.769, $P = 0.006$), age (HR = 1.027, 95% CI 1.004–1.051, $P = 0.02$), and neoplasm (HR = 4.631, 95% CI 2.675–8.015, $P < 0.001$) could serve as independent prognostic factors for EC patients (Fig. 6B).

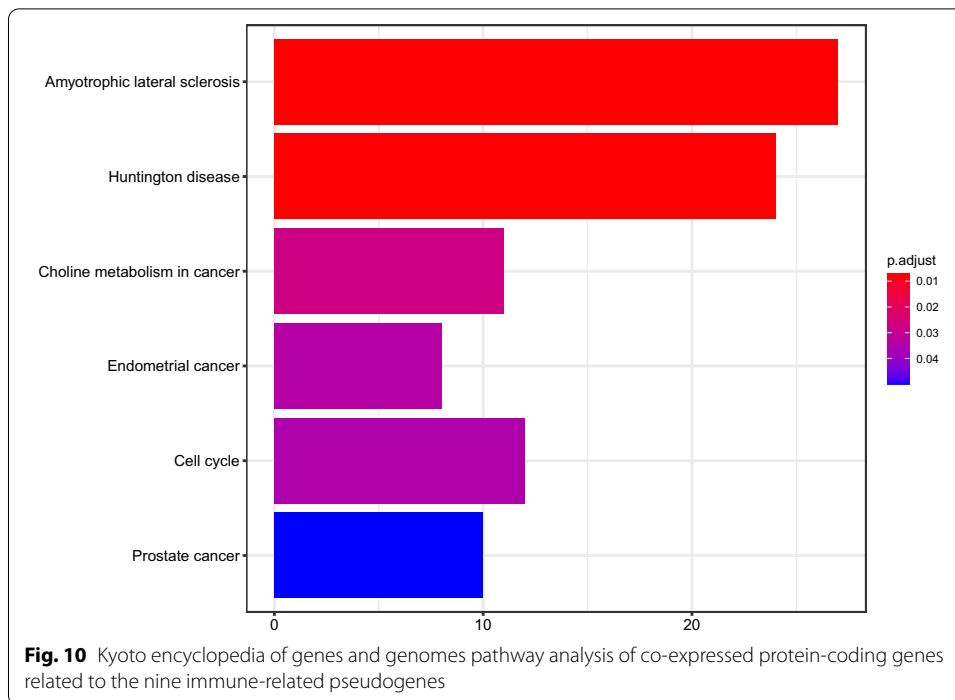
Estimation of immune cell type fractions

Previous studies have demonstrated that infiltrating immune cells are strongly related to the prognosis and treatment of malignant cancers [35–37]. Using CIBERSORT, we compared the differences in the immune infiltration of 22 immune cell types between low- and high-risk groups. First, a bar plot was used to illustrate the fractions of 22 immune cells in each sample (Fig. 7A). Next, we analyzed the tumor-infiltrating cells between high- and low-risk subgroups. The results exhibited that CD8⁺ T cell, follicular helper T cells, regulatory T cell, gamma delta T cells, resting dendritic cells, and activated dendritic cells were significantly different between the two risk groups (all $P < 0.05$; Fig. 7B), which might help to predict immune response and survival. To explore whether immune-related pseudogene signature effectively reflected the tumor immune microenvironment status, the associations between the risk score and infiltration abundances of 22 types of immune cells were analyzed (Fig. 8). The risk score was significantly correlated with activated myeloid dendritic cell (correlation = 0.163, $P = 1.357e-04$), B cell naïve (correlation = 0.143, $P = 8.257e-04$), activated mast cell (correlation = 0.10, $P = 0.019$), gamma delta T cell (correlation = 0.094, $P = 0.029$), regulatory T cell (correlation = -0.191 , $P = 7.857e-06$), CD8⁺ T cell (correlation = -0.197 , $P = 3.764e-06$).



The immune-related pseudogene signature predicts responses of immunotherapy

The association between the risk score and expression levels of three immune checkpoint genes was explored. The risk score was significantly negatively correlated with CTLA4 (Spearman correlation coefficient = -0.205 , $P = 1.552e-06$) and PD1 (Spearman correlation coefficient = -0.135 , $P = 0.002$; Fig. 9A). The expression changes of three immune checkpoint genes were compared between the high- and low-risk groups, and the results illustrated that patients with low-risk showed significantly higher expression levels compared with those in the high-risk group ($P = 0.0003$ for CTLA4, and $P = 0.002$ for PD1; Fig. 9B). These findings were consistent with previous views that the immune checkpoint genes closely participated in immunosuppression, and their



high expression is correlated with poor prognosis [38]. The relationship between IPS and immune-related pseudogene signature was further explored. The IPS, IPS-CTLA4, IPS-PD1, and IPS-PD1-CTLA4 scores were computed to investigate the signature and predict potential effects on immune checkpoint inhibitor for patients with EC. In the low-risk group, the scores were notably higher than high-risk group (IPS, $P < 0.001$; IPS-CTLA4, $P < 0.001$; IPS-PD1, $P < 0.001$; and IPS-PD1-CTLA4, $P < 0.001$; Fig. 9C). These results revealed that patients presenting higher IPSs in the low-risk group promised to be candidates for immune checkpoint inhibitor.

Functional annotation of immune-related pseudogenes

We computed the expression correlation between identified pseudogenes and potential protein-coding genes by calculating the Spearman correlation coefficient through expression profiles in 541 patients with EC. Co-expression analysis results identified 857 genes that correlated with nine pseudogenes. KEGG pathway enrichment illustrated that immune-related pseudogenes were mainly involved in choline metabolism in cancer, endometrial cancer, cell cycle, prostate cancer pathways (Fig. 10).

Acknowledgements

The authors sincerely acknowledge the publicly available ImmPort database (<https://immport.niaid.nih.gov>) and TCGA (<https://portal.gdc.cancer.gov/>) database.

Authors' contributions

ST and YZ is the principal investigator. ST conducted statistical analysis and data management. ST and YZ edited and revised the manuscript. Both authors read and approved the final manuscript.

Funding

None.

Availability of data and materials

The raw data of this study are derived from ImmPort database (<https://immport.niaid.nih.gov>) and the TCGA database (<https://portal.gdc.cancer.gov/>), which are publicly available databases.

Declarations**Ethics approval and consent to participate**

Not necessary.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 15 December 2020 Accepted: 20 June 2021

Published online: 30 June 2021

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *Cancer J Clin*. 2019;69(1):7–34.
2. Brooks RA, Fleming GF, Lastra RR, Lee NK, Moroney JW, Son CH, Tatebe K, Veneris JL. Current recommendations and recent progress in endometrial cancer. *Cancer J Clin*. 2019;69(4):258–79.
3. Colombo N, Creutzberg C, Amant F, Bosse T, González-Martín A, Ledermann J, Marth C, Nout R, Querleu D, Mirza MR, Sessa C. ESMO-ESGO-ESTRO consensus conference on endometrial cancer: diagnosis, treatment and follow-up. *Ann Oncol*. 2016;27(1):16–41.
4. Lortet-Tieulent J, Ferlay J, Bray F, Jemal A. International patterns and trends in endometrial cancer incidence, 1978–2013. *J Natl Cancer Inst*. 2018;110(4):354–61.
5. Ding W, Lin L, Chen B, Dai J. L1 elements, processed pseudogenes and retrogenes in mammalian genomes. *IUBMB Life*. 2006;58(12):677–85.
6. Milligan MJ, Harvey E, Yu A, Morgan AL, Smith DL, Zhang E, Berengut J, Sivananthan J, Subramaniam R, Skoric A, et al. Global intersection of long non-coding RNAs with processed and unprocessed pseudogenes in the human genome. *Front Genet*. 2016;7:26.
7. Pink RC, Wicks K, Caley DP, Punch EK, Jacobs L, Carter DR. Pseudogenes: pseudo-functional or key regulators in health and disease? *RNA*. 2011;17(5):792–8.
8. Han L, Yuan Y, Zheng S, Yang Y, Li J, Edgerton ME, Diao L, Xu Y, Verhaak RGW, Liang H. The Pan-Cancer analysis of pseudogene expression reveals biologically and clinically relevant tumour subtypes. *Nat Commun*. 2014;5:3963.
9. Bai M, Yuan M, Liao H, Chen J, Xie B, Yan D, Xi X, Xu X, Zhang Z, Feng Y. OCT4 pseudogene 5 upregulates OCT4 expression to promote proliferation by competing with miR-145 in endometrial carcinoma. *Oncol Rep*. 2015;33(4):1745–52.
10. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell*. 2011;146(3):353–8.
11. Hu X, Yang L, Mo YY. Role of pseudogenes in tumorigenesis. *Cancers*. 2018;10(8):256.
12. Lin S, Meng Y, Cao C, Wu P, Gao P, Zhi W, Peng T, Wu P. Comprehensive analysis of LDHAP5 pseudogene expression and potential pathogenesis in ovarian serous cystadenocarcinoma. *Cancer Cell Int*. 2020;20:229.
13. Chen X, Wan L, Wang W, Xi WJ, Yang AG, Wang T. Re-recognition of pseudogenes: from molecular to clinical applications. *Theranostics*. 2020;10(4):1479–99.
14. Cheetham SW, Faulkner GJ, Dinger ME. Overcoming challenges and dogmas to understand the functions of pseudogenes. *Nat Rev Genet*. 2020;21(3):191–201.
15. An Y, Furber KL, Ji S. Pseudogenes regulate parental gene expression via ceRNA network. *J Cell Mol Med*. 2017;21(1):185–92.
16. Thomson DW, Dinger ME. Endogenous microRNA sponges: evidence and controversy. *Nat Rev Genet*. 2016;17(5):272–83.
17. Mori N, Wildes F, Takagi T, Glunde K, Bhujwalla ZM. The tumor microenvironment modulates choline and lipid metabolism. *Front Oncol*. 2016;6:262.
18. Glunde K, Jacobs MA, Bhujwalla ZM. Choline metabolism in cancer: implications for diagnosis and therapy. *Expert Rev Mol Diagn*. 2006;6(6):821–9.
19. Zhao L, Jiang L, He L, Wei Q, Bi J, Wang Y, Yu L, He M, Zhao L, Wei M. Identification of a novel cell cycle-related gene signature predicting survival in patients with gastric cancer. *J Cell Physiol*. 2019;234(5):6350–60.
20. Liu J, Mei J, Li S, Wu Z, Zhang Y. Establishment of a novel cell cycle-related prognostic signature predicting prognosis in patients with endometrial cancer. *Cancer Cell Int*. 2020;20:329.
21. Lee JM, Lee MH, Garon E, Goldman JW, Salehi-Rad R, Baratelli FE, Schae D, Wang G, Rosen F, Yanagawa J, et al. Phase I Trial of intratumoral injection of CCL21 gene-modified dendritic cells in lung cancer elicits tumor-specific immune responses and CD8(+) T-cell infiltration. *Clin Cancer Res*. 2017;23(16):4556–68.
22. Hegde PS, Karanikas V, Evers S. The where, the when, and the how of immune monitoring for cancer immunotherapies in the era of checkpoint inhibition. *Clin Cancer Res*. 2016;22(8):1865–74.
23. Fu Y, Liu S, Zeng S, Shen H. From bench to bed: the tumor immune microenvironment and current immunotherapeutic strategies for hepatocellular carcinoma. *J Exp Clin Cancer Res*. 2019;38(1):396.

24. Schmidt M, Hellwig B, Hammad S, Othman A, Lohr M, Chen Z, Boehm D, Gebhard S, Petry I, Lebrecht A, et al. A comprehensive analysis of human gene expression profiles identifies stromal immunoglobulin κ C as a compatible prognostic marker in human solid tumors. *Clin Cancer Res*. 2012;18(9):2695–703.
25. Zhou C, Li C, Yan F, Zheng Y. Identification of an immune gene signature for predicting the prognosis of patients with uterine corpus endometrial carcinoma. *Cancer Cell Int*. 2020;20(1):541.
26. Romani M, Pistillo MP, Carosio R, Morabito A, Banelli B. Immune checkpoints and innovative therapies in glioblastoma. *Front Oncol*. 2018;8:464.
27. Duffy MJ, Crown J. Biomarkers for predicting response to immunotherapy with immune checkpoint inhibitors in cancer patients. *Clin Chem*. 2019;65(10):1228–38.
28. Havel JJ, Chowell D, Chan TA. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. *Nat Rev Cancer*. 2019;19(3):133–50.
29. Sun X, Zhang T, Li M, Yin L, Xue J. Immunosuppressive B cells expressing PD-1/PD-L1 in solid tumors: a mini review. *QJM*. 2019. <https://doi.org/10.1093/qjmed/hcz162>.
30. Zhao E, Xie H, Zhang Y. Predicting diagnostic gene biomarkers associated with immune infiltration in patients with acute myocardial infarction. *Front Cardiovasc Med*. 2020;7:586871.
31. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H, Trajanoski Z. Pan-cancer immunogenomic analyses reveal genotype–immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep*. 2017;18(1):248–62.
32. Jiang P, Sun W, Shen N, Huang X, Fu S. Identification of a metabolism-related gene expression prognostic model in endometrial carcinoma patients. *BMC Cancer*. 2020;20(1):864.
33. Liu J, Jiang P, Chen X, Shen Y, Cui G, Ma Z, Zhao S, Zhang Y. Construction of a nine DNA repair-related gene prognostic classifier to predict prognosis in patients with endometrial carcinoma. *BMC Cancer*. 2021;21(1):29.
34. Jiang Y, Chen J, Ling J, Zhu X, Jiang P, Tang X, Zhou H, Li R. Construction of a glycolysis-related long noncoding RNA signature for predicting survival in endometrial cancer. *J Cancer*. 2021;12(5):1431–44.
35. Jochems C, Schlom J. Tumor-infiltrating immune cells and prognosis: the potential link between conventional cancer therapy and immunity. *Exp Biol Med*. 2011;236(5):567–79.
36. Zhou H, Chen L, Lei Y, Li T, Li H, Cheng X. Integrated analysis of tumor mutation burden and immune infiltrates in endometrial cancer. *Curr Probl Cancer*. 2020;45:100660.
37. Zhao E, Bai X. Identification of key genes of prognostic value in clear cell renal cell carcinoma microenvironment and a risk score prognostic model. *Dis Markers*. 2020;2020:8852388.
38. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer*. 2012;12(4):252–64.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

