

METHODOLOGY

Open Access



Construction of a new immune-related lncRNA model and prediction of treatment and survival prognosis of human colon cancer

Sicheng Liu, Xingyu Peng, Xun Wu, Fanqin Bu, Zhonglin Yu, Jinfeng Zhu, Chen Luo, Wenjun Zhang, Jiang Liu and Jun Huang*

Abstract

Background: An increasing number of studies have shown that immune-related long noncoding RNAs (lncRNAs) do not require a unique expression level. This finding may help predict the survival and drug sensitivity of patients with colon cancer.

Methods: We retrieved original transcriptome and clinical data from The Cancer Genome Atlas (TCGA), sorted the data, differentiated mRNAs and lncRNAs, and then downloaded immune-related genes. Coexpression analysis predicted immune-related lncRNAs (irlncRNAs) and univariate analysis identified differentially expressed irlncRNAs (DEirlncRNAs). We have also amended the lasso penalizing region. Next, we compared the areas under the curve (AUCs), counted the Akaike information standard (AIC) value of the 3-year receiver operating characteristic (ROC) curve, and determined the cutoff point to establish the best model to differentiate the high or low disease risk group of colon cancer patients.

Results: We reevaluated the patients regarding the survival rate, clinicopathological features, tumor-infiltrating immune cells, immunosuppressive biomarkers, and chemosensitivity. A total of 155 irlncRNA pairs were confirmed, 31 of which were involved in the Cox regression model. After the colon cancer patients were regrouped according to the cutoff point, we could better distinguish the patients based on adverse survival outcomes, invasive clinicopathological features, the specific tumor immune cell infiltration status, high expression of immunosuppressive biomarkers, and low chemosensitivity.

Conclusions: In this study, we established a characteristic model by pairing irlncRNAs to better predict the survival rate, chemotherapy efficacy, and prognostic value of patients with colon cancer.

Keywords: lncRNA, Colon cancer, Tumor-infiltrating immune cell, Immunotherapy

Background

Colon cancer is a universal malignant tumor of the digestive tract. The incidence and death rates in China have increased in the last ten years [1]. Although surgery, chemotherapy, and targeted therapy have

significantly improved the overall survival rate of patients, approximately 50% of colon cancer patients have distant metastasis, which is the most common cause of treatment failure [2]. With tumor metastasis, the 5-year survival rate is reduced to 8.1% [3]. Immunotherapy in the treatment of metastatic colorectal cancer has recently improved. Currently, the main diagnostic methods of colon cancer include biopsy, laboratory examination, and colonoscopy [4]. Immunotherapy

*Correspondence: huangjun0727@163.com

Department of Gastrointestinal Surgery, Second Affiliated Hospital of Nanchang University, Nanchang 330006, Jiangxi, China



(including antagonists of immune checkpoints, such as PD-1 and CTLA4) works best when combined with tumor vaccines or therapeutic agents that induce immune cell death [5]. For example, in some rodent models of colon cancer, OXP causes immunogenic cell death. To date, the immunogenicity of cell death induced by OXP is governed by the same rules as those induced by anthracyclines, involving CRT exposure, HMGB1 release, and the presence of functional TLR4 in the immune system [6].

LncRNAs constitute a class of noncoding RNAs (ncRNAs) greater than 200 nucleotides in length that play a vital role in cancer development [7]. LncRNAs are abundant, accounting for approximately 80% of the human transcriptome. They adjust 70% of human gene expression and do not act universally because they could affect DNA, RNA, and protein and show enhanced or inhibitory effects. Long noncoding RNAs (lncRNAs) play a key role in the carcinogenesis and progression of various human malignant tumors, including colon cancer [8]. The latest research findings show that lncRNA BC200 can be used as a new oncogene and therapeutic target for colon cancer [9]. A previous study found that lncRNA HOTAIR acts by cutting miRNA34a to promote the development of colon cancer [10]. LncRNA RAMS11 directly affects mCRC biology, including promoting an aggressive phenotype and being related to treatment response and resistance [11]. These studies indicate that the lncRNA expression level affects the development and survival of colon cancer.

Therefore, a credible risk evaluation model must be established to evaluate the survival prognosis of patients with colon cancer and optimize clinical decision-making. LncRNAs can help these signal models [12]. Identifying immune-related 9-lncRNA signaling can not only improve the ability of colon cancer patients to predict prognosis but also facilitate better clinical strategies and elucidate its potential mechanisms [13]. We found 18 immune-related genes to predict the prognosis of patients with pancreatic cancer characteristics, proving the validity of the IRGP signature, and validating the model and relationship between the immune cells [14]. Bioinformatics analysis integrates three immune-related long-term noncoding RNA (mRNA, microRNA, lncRNA) signals, which can be used as a predictive model for clear cell renal cell carcinoma [15].

The combination of two biomarkers is better than a single gene [16] regarding the accuracy of cancer diagnosis and prognosis models. For this study, we used a new modeling algorithm, pairing, and iterative operation, to build an irlncRNA signal without any fixed expression

level. Finally, we evaluated its prognostic value in patients with colon cancer and its diagnostic efficacy, drug sensitivity, and tumor immune invasion.

Materials and methods

Obtaining colon cancer data from TCGA

The transcriptome data of the fragments per kilobase million (FPKM) format and clinicopathological information of colon cancer were downloaded from TCGA (<https://portal.gdc.cancer.gov/>). The data were derived from 41 normal patients and 373 tumor patients. Ensembl (<http://asia.ensembl.org>) was used to download the gene transfer format (GTF) file to differentiate mRNAs from lncRNAs for subsequent analysis. Immune genetic databases (<http://www.immport.org>) were used to download immune-related genes and were coexpressed to screen irlncRNAs. The correlation between immune genes and all lncRNAs was analyzed. The absolute value of the immune gene correlation coefficient was greater than 0.4, and the *p* value was less than 0.001. To determine the DEirlncRNAs, we used the limma R package related to immune lncRNA differential expression analysis. The threshold was set to a log fold change (FC) > 2 and false discovery rate (FDR) < 0.05.

Pairing DEirlncRNAs

If the expression quantity of lncRNA A is lower than that of lncRNA B, C is defined as 0; otherwise, C is defined as 1. Next, the constructed 0 or 1 matrix continued to be filtered. When the number of lncRNAs with 0 or 1 expression was more than 20% and less than 80% of the total, it was considered an effective match. If the expression level of an lncRNA pair is 0 or 1, the relationship between the lncRNA pair and survival prognosis is not considered because no specific level of lncRNA pairs can correctly predict the survival outcome of patients.

Build risk model to assess risk score

First, univariate factor analysis was performed, and then lasso regression was executed to screen immune gene pairs. In lasso regression, 1000 cycles were run. For each loop, random stimulation was set to 1000 times, and the frequency of each pair in 1000 repetitive lasso regression models was recorded. Next, cross-validation was performed. Finally, the immune pairs were analyzed by Cox proportional hazards regression, and the model was constructed. The AUC values of each model were calculated and plotted. The larger was the AUC, the higher was the accuracy of predicting survival. Given that the curve reached the zenith,

indicating that the AUC value was the largest, the calculation process was terminated, and the model was deemed the optimal candidate. The 1-, 2-, and 3-year ROC curves of the model were drawn. The established risk model was used to calculate the risk score of all clinical cases. Risk score = $\sum_{i=1}^n Coef_i \times x_i$, n, x_i and $Coef_i$ represents the number of genes, gene expression level, and coefficient value respectively. The AIC value of each point in the 3-year ROC curve was calculated to determine the maximum inflection point, which had the best specificity and sensitivity.

Risk assessment model validation

To demonstrate this cutoff point, we performed Kaplan–Meier analysis and visualized the survival curve to show the distinction in survival between low-risk and high-risk groups. The risk scores of all the samples in the risk model were visualized using the R tool. The R packages used in these steps include survival, glmnet, survminer, survivalROC, pbapply, and P Heatmap. Additionally, to prove the clinical application value of the model, we used chi squared test to analyze the connection between the risk model and

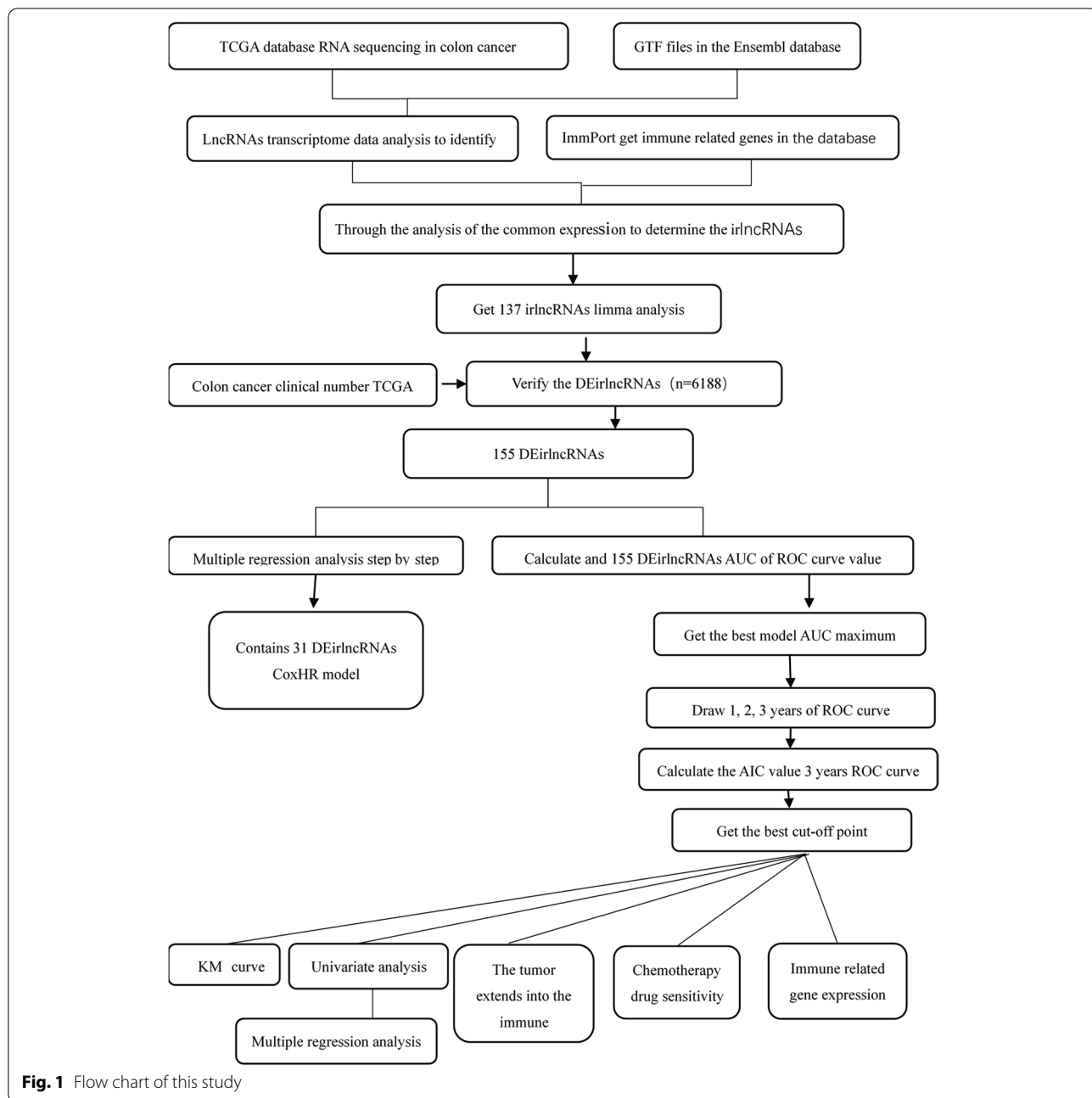


Fig. 1 Flow chart of this study

clinicopathological characteristics. The results of the analysis are displayed in the strip diagram, marked as described below: * < 0.05, ** < 0.01, *** < 0.001. The Wilcoxon signed-rank test was used to calculate the difference in the risk score between groups with different clinicopathological characteristics. Block diagram visualization was used to show the analysis results. To determine whether the risk model can be used as an independent predictor of clinical prognosis, univariate and multivariate Cox regression analyses were compared between the risk score and clinicopathological features. Forest map visualization was used to demonstrate the results. The R tools used in these operations are ggpubr, p-heatmap, and survival.

Correlation analysis of tumor-infiltrating immune cells

To identify the relationship between immune cell characteristics and risk scores, we used the following six credible methods to analyze the immune cell infiltration status of patients with colon cancer from the TCGA database: TIMER, CIBERSORT, EPIC, MCP-COUNTER, XCELL, and CIBERSORT. Next, the difference in the content of immune infiltrating cells between the high-risk and low-risk groups was analyzed using the Wilcoxon signed-rank test. Spearman correlation analysis was used to analyze the relationship between the risk score and immune infiltrating cells. The correlation coefficient of the results showed that in the lollipop diagram, the implication threshold was set as $p < 0.05$. The program was executed using the rggplot2 package.

To explore the significance of the model in clinical treatment

To evaluate the application of this model in the clinical treatment of colon cancer, we calculated the IC50 of common administrating chemotherapeutic drugs for colon cancer patients in TCGA. The Wilcoxon signed-rank test was used to compare the IC50 differences between the high-risk and low-risk groups. The results are shown in the block diagram, which was obtained using the prphetic and ggplot2 packages in R.

Model- and immune checkpoint inhibitor-related genes

To evaluate the risk model using the immune checkpoint inhibitor-related gene expression level, we used the ggstatsplot package for data visualization.

Result

Extraction of DElncRNA

The process is shown in Fig. 1. We retrieved and downloaded the transcriptome and clinical datasets of colon

cancer from the colon cancer project of the TCGA database including 41 normal samples and 473 tumor samples. Next, the data were annotated according to the GTF file of Ensembl to distinguish mRNA from lncRNA. The immune genes were downloaded from the immune report database, and coexpression analysis was executed between the known immune genes and lncRNAs. We identified 1093 irlncRNAs (Suppl. Fig. 1); 137 were DElncRNAs (Fig. 2A), 126 were upregulated, and 11 were downregulated (Fig. 2B).

Construction of the immune prognosis model and risk prediction model

Making iterative cycle of 0 or 1 matrix extract 137 DElncRNA, 6188 effective DElncRNA pairs were identified. After univariate testing and improved lasso region analysis, 155 pairs of irlncRNAs were obtained, among which 31 pairs were involved in the Cox risk assessment model in a step-by-step manner (Fig. 2 C). We counted the AUC of each subject operating characteristic curve of 155 pairs of subjects, drew the curve, and calculated a value of 0.911. We also used the AIC value to identify the maximum inflection point as the cutoff point of the 3-year ROC curve, collected the data of 426 patients with acceptable colon cancer from TCGA for the risk score, and established the cutoff point to redifferentiate the low-risk and high-risk groups in the cohort for testing (Fig. 3A). The 3-year ROC curve was compared with other clinical data (Fig. 3B). To verify the optimality, we plotted ROC curves for 1, 2, and 3 years, indicating that all the AUC values exceeded 0.91 (Fig. 3C).

Clinical correlation analysis based on the risk prediction model

Based on the cutoff point of AUC determined before the risk assessment model for clinical relevance evaluation, 110 subjects were classified as the high-risk group, and 316 patients were classified as the low-risk group. The risk score and survival are shown in Fig. 4A and Fig. 4B, respectively. These data suggest that with increasing value at risk, the death toll rises. Kaplan–Meier analysis showed that the survival time of patients in the high-risk group was shorter than that of patients in the low-risk group ($p < 0.001$) (Fig. 4C). Next, we used chi-squared test to evaluate the risk of colon cancer and the relationship between the clinicopathological features. The Wilcoxon signed-rank test strip chart (Fig. 5A) and scatter diagram show that T (Fig. 5B), M (Fig. 5C) and N (Fig. 5D) stages and clinical staging (Fig. 5E) were associated with a significant risk. Next, we performed univariate Cox regression

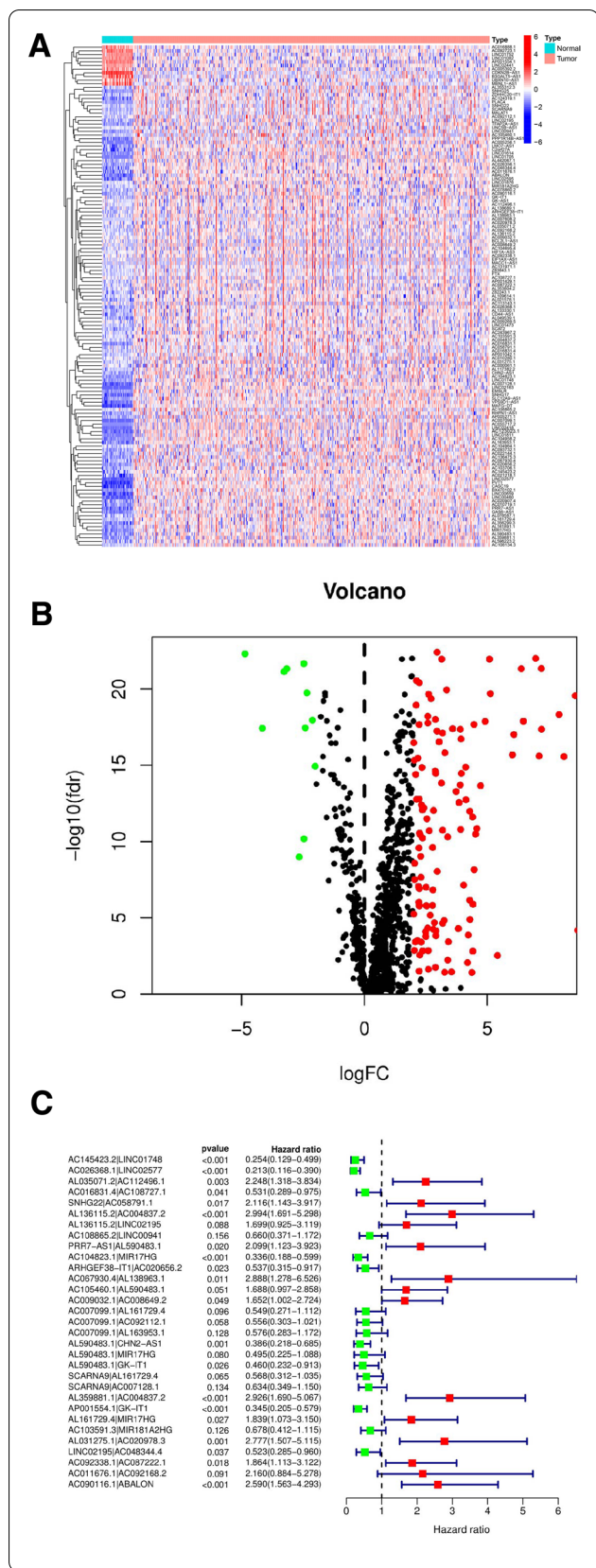


Fig. 2 Extract has the difference of immune-related genes. Using the TCGA data set and immune-related Ensembl annotation to identify differentially expressed genes. Shows the heat map figure (A) and the volcano (B). (C) The forest figure shows the COX proportional hazards regression method of stepwise regression method and the DEIRlncRNAs

analysis (Fig. 5F) and proved that the clinical stage ($p < 0.001$, HR = 2.208, 95% CI [1.722 2.831]) and risk score ($p < 0.001$, HR = 1.004, 95% CI [1.003 1.005]) showed statistical significance, and multivariate Cox regression analysis (Fig. 5G) proved that age ($p < 0.001$, HR = 1.035, 95% CI [1.015 1.056]) and clinical stage ($p < 0.001$, HR = 2.260, 95% CI [1.750 2.918]), HR = 1.003, 95% CI [1.002 1.005]) (the detailed numerical values are in Table 1).

A risk assessment model can predict immunosuppressive molecules and tumor-infiltrating immune cells

Because immune-related genes and lncRNAs initially correspond, we evaluated whether the model is relevant to the tumor-immune microenvironment. The high-risk group was negatively correlated with tumor-infiltrating immune cells such as hematopoietic stem cells and neutrophils and positively correlated with CD4 T cells and monocytes. Spearman correlation analysis was performed, and the results are shown in Fig. 6A, B, and Table 2. Immune checkpoint inhibitors were used to treat colon cancer in clinical medication. We studied whether the risk model is relevant to immune checkpoint inhibitor-relevant biomarkers and found that a high stake score was positively correlated with high expression of PLD2 ($p < 0.05$; Fig. 6C) and negatively correlated with low expression of MLH1 ($p < 0.05$; Fig. 6D).

Correlation analysis between the risk model and chemotherapy drugs

In addition to checkpoint blockade treatment, we determined the association between the risk and efficacy of common chemotherapy drugs for colon cancer in TCGA colon cancer patient data. We show that high-risk scores are associated with a low IC50 of chemotherapeutic drugs such as oxaliplatin ($p = 0.00089$); however, it is related to the high half inhibitory concentration (IC50) of the protease inhibitor (MG132) and anaplastic lymphoma kinase inhibitor (NVP-TAE684). This model can be used as a potential predictor of chemosensitivity (Fig. 6E).

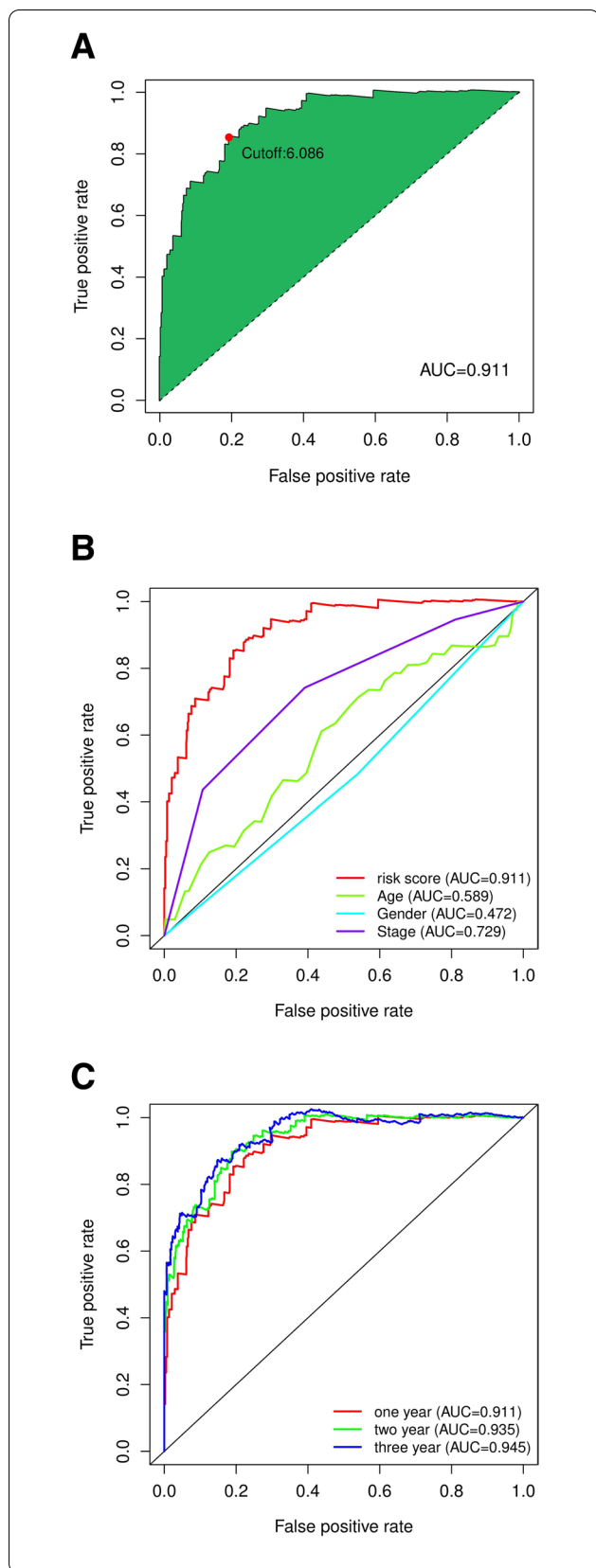


Fig. 3 Immune-related genes were used to establish a risk assessment model. **(A)** Draw 155 DEIRlncRNA each generated on the model of ROC AUC value curve, and the highest point of the AUC was determined. ROC ideal DEIRlncRNA model associated with the AUC value of maximum. Of 426 patients with colon cancer risk score, the biggest inflection point is made by AIC cutoff point. **(B)** 3 years compared with other common clinical features of ROC curve shows the superiority of risk score. **(C)** The optimal model of the ROC 1 year, 2 years, and 3 years shows that all the AUC values are greater than 0.91

Discussion

Surgery is a common treatment for colon cancer, but many shortcomings exist. Recent studies have found that the immune environment of primary tumors has a profound impact on immunotherapy [17]. Pembrolizumab monotherapy has shown long-lasting antitumor activity in non-small-cell lung cancer (NSCLC) expressing advanced programmed death ligand 1 (PD-L1, [18]). Some tumor cells can use SNHG12 to escape immune-mediated attack and enhance the immune response [19]. The expression of mir155HG was significantly correlated with the infiltration level of immune cells and immune molecules [20]. These studies indicate that immune intervention therapy has broad application prospects in tumor therapy. Therefore, new methods of treatment and prognosis of colon cancer must be identified. Additionally, lncRNAs were identified as essential regulators of the cancer immune response. In the current study, enlightened by the tactics of immune-relevant gene pairing, we attempted to build a viable model of double genome pairing.

First, we collected transcriptome data and clinical data and downloaded immune-related gene datasets from TCGA. Differential coexpression analysis was executed to classify the data, and we verified the nonrepetitive sequence pairs using the improved circular single-pair method and 1 or 0 matrix. Next, we performed univariate analysis, combined with the improved lasso penalized regression, to increase the accuracy and effectiveness of risk prediction [21]. We obtained prognosis-related immune-related lncRNAs to construct a prognostic model of immune-related lncRNAs. We not only calculated every AUC value on the ROC curve to obtain the best model but also obtained the AIC value of all points on the AUC curve to identify the best cutoff point, which was used to distinguish the high-risk and low-risk groups in colon cancer patients. Finally, we evaluated the new model in a clinical setting and included the survival rate, clinicopathological features, tumor-infiltrating immune cells, immunosuppressive biomarkers, and chemosensitivity.

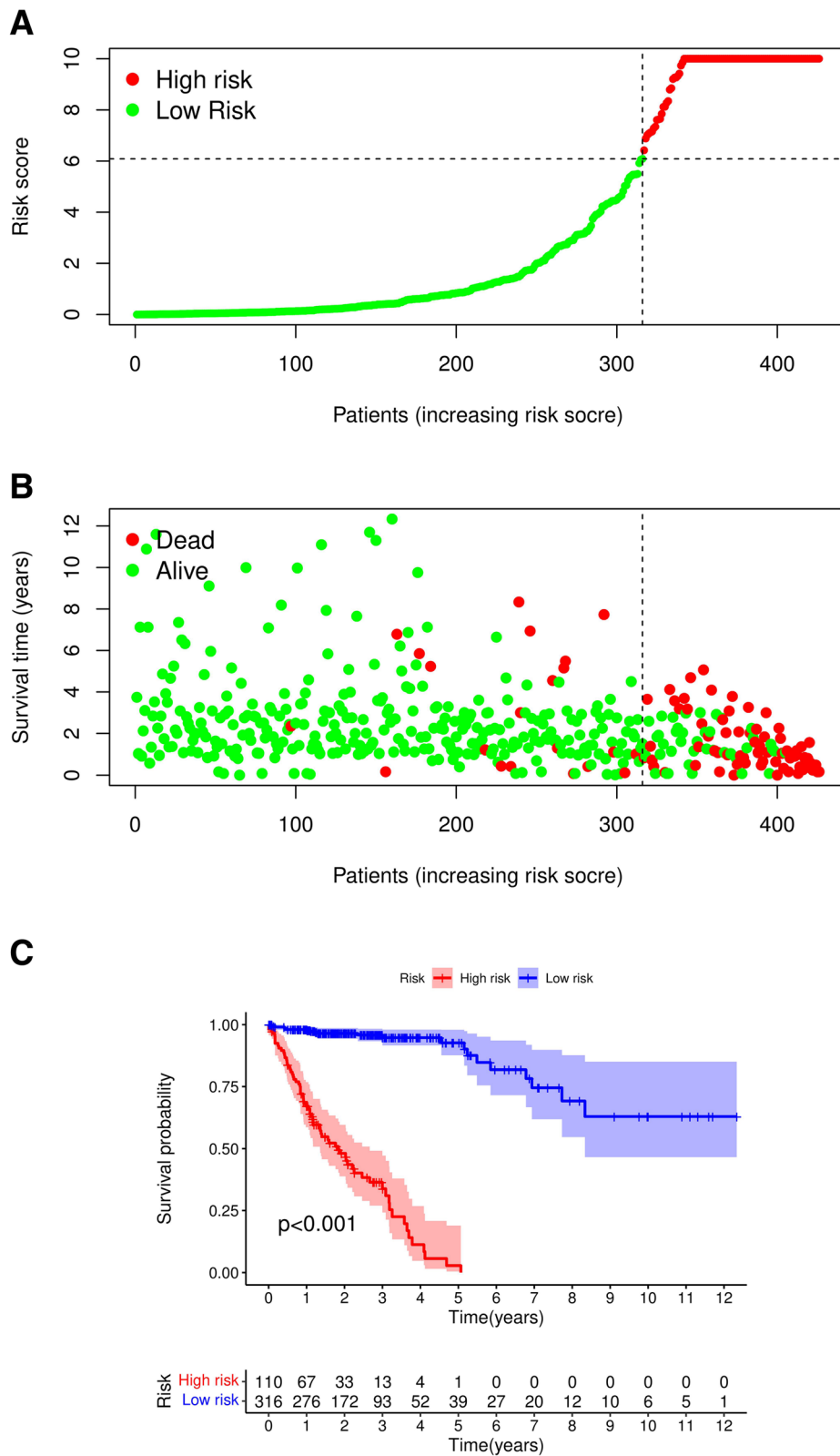
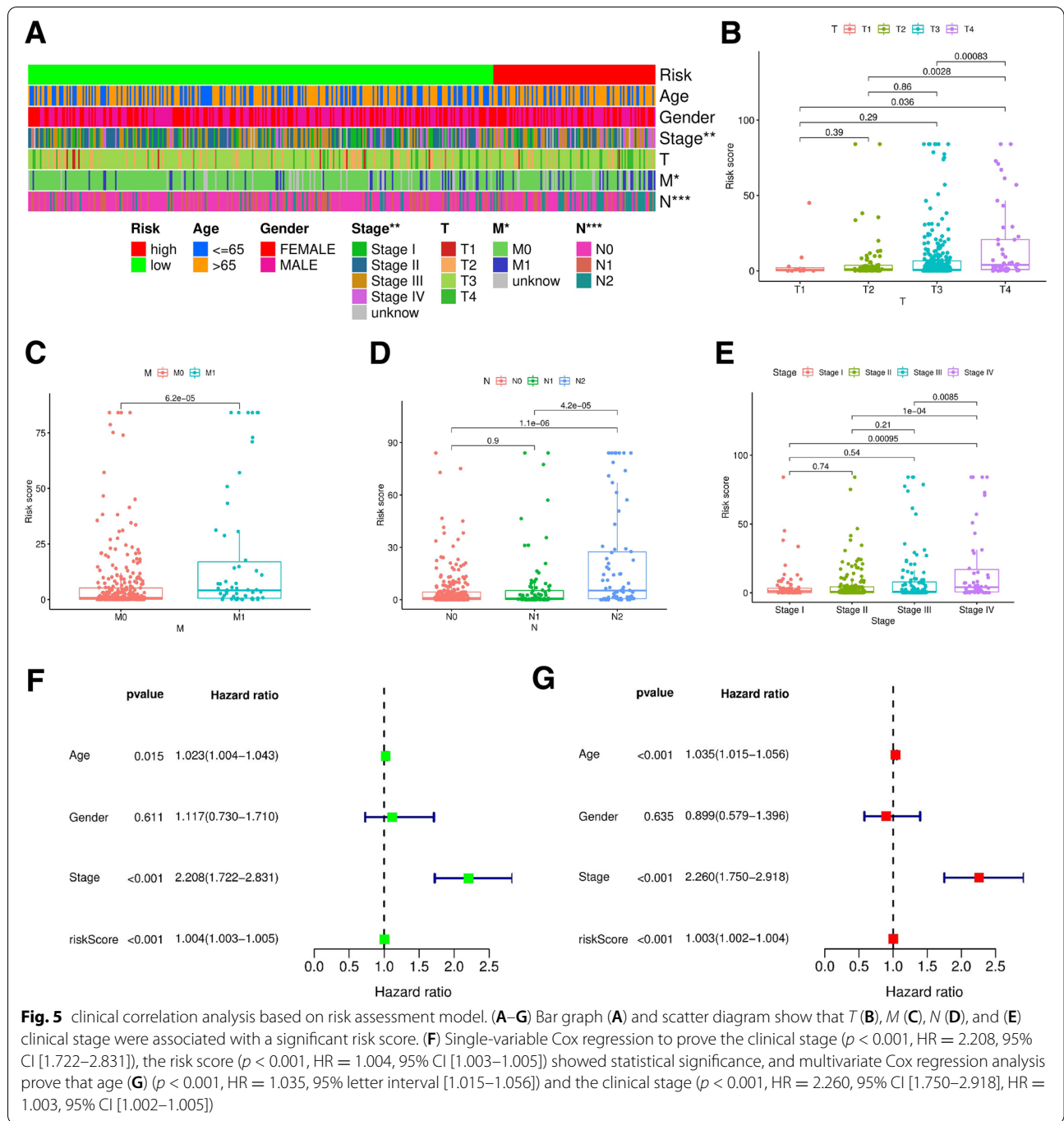


Fig. 4 Risk assessment model of survival prognosis. (A) and (B) shows risk score and survival results of each case. (C) By the Kaplan-Meier test, high-risk group of patients' survival time is short



The present study explored a signal based on seven irlncRNAs to predict the overall survival rate of colon cancer [22]. Generally, the high expression of lncRNAs shows important biological functions. Our data analysis can identify the DEirlncRNAs and identify the most valuable irlncRNAs. Therefore, detecting the value of each specific irlncRNA was not required, only the high or low expression of irlncRNA pairs. The prediction model can

differentiate the high risk and low risk of patients with clinical disease. In this study, some of the DEirlncRNAs in the process of modeling that have been already identified play an important role in malignant phenotypes of various cancer types, such as SNHG22 [23, 24], PRR7-AS1 [25], and LINC00941 [26, 27], especially for colon cancer, while others were revealed for the first time. For instance, SNHG22 promoted CRC tumorigenesis and metastasis

Table 1 univariate and multivariate Cox regression analysis of the specific value

ID	Univariate analysis			Multivariate analysis		
	HR	95% CI	P-Value	HR	95% CI	P-Value
Age	1.02	1.004 ~1.043	0.015	1.03	1.015 ~1.056	0.001
Gender	1.12	0.730 ~1.710	0.611	1.03	0.579 ~1.396	0.896
Stage	2.21	1.722 ~2.831	0.000	2.49	1.750 ~2.918	0.000
Risk-Score	1.00	1.003 ~1.005	0.000	1.00	1.002 ~1.004	0.000

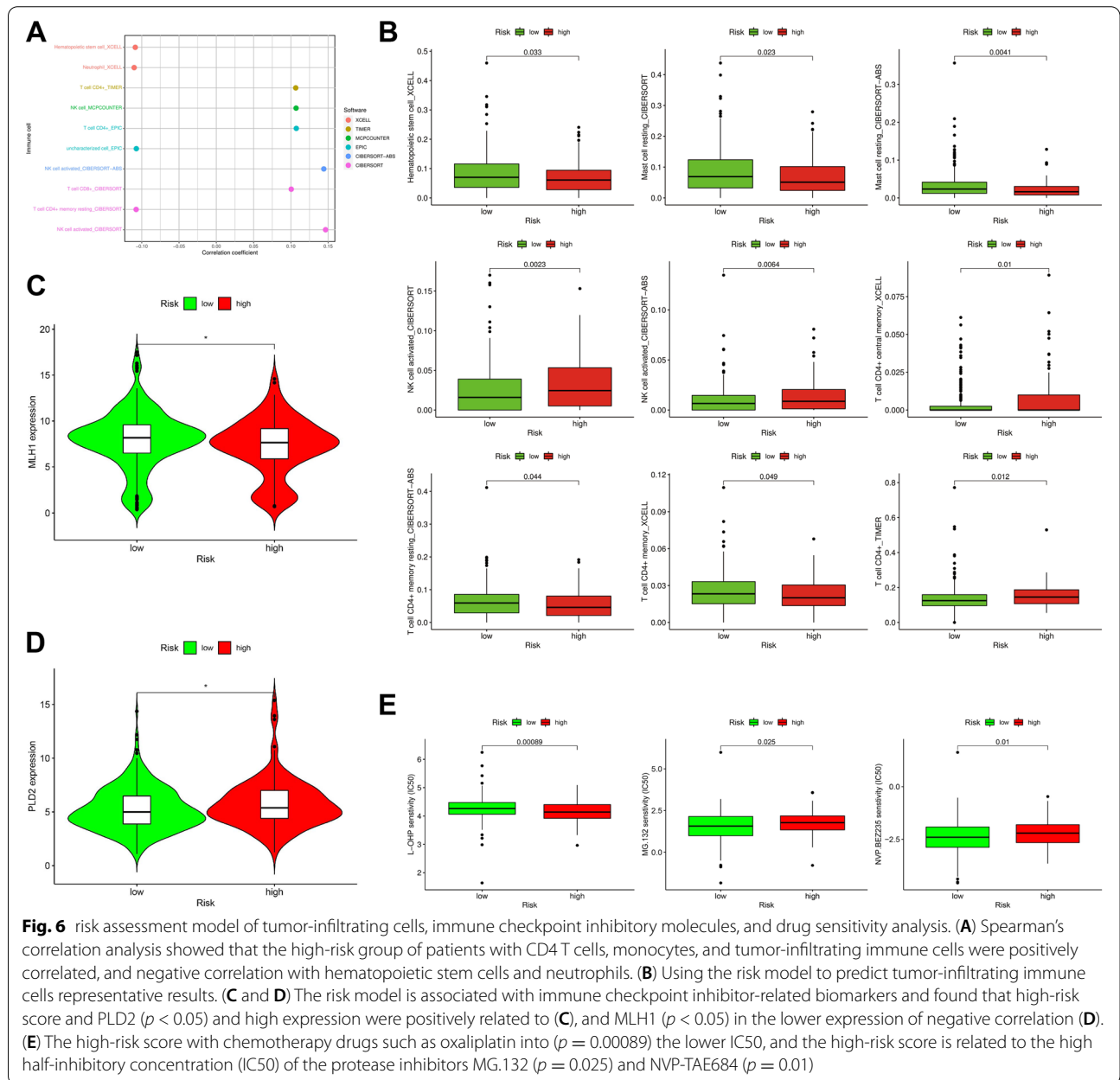


Table 2 Tumor infiltration of immune cells and risk scores of the comparison results

symbol	type	correlation	p-Value
T cell CD4+_TIMER	TIMER	0.1065	0.0279
T cell CD8+_CIBERSORT	CIBERSORT	0.1005	0.0381
T cell CD4+ memory resting-CIBERSORT	CIBERSORT	-0.1075	0.0265
NK cell activated-CIBERSORT	CIBERSORT	0.1466	0.0024
NK cell activated-CIBERSORT-ABS	CIBERSORT-ABS	0.1442	0.0029
NK cell-MCPCOUNTER	MCPCOUNTER	0.1070	0.0272
Hematopoietic stem cell-XCELL	XCELL	-0.1087	0.0248
Neutrophil-XCELL	XCELL	-0.1102	0.0229
T cell CD4+_EPIC	EPIC	0.1073	0.0268
uncharacterized cell-EPIC	EPIC	-0.1073	0.0268

by sponging miR-128-3p [28]. Reveal the essential role of LINC00941 [29] in metastatic CRC via activation of the TGF- β /SMAD2/3 axis. Demonstrate that MIR17HG [30] plays an oncogenic role in colorectal cancer. Long noncoding RNA (lncRNA) imbalance has been found in many human cancers, including colon cancer. Therefore, identifying potential lncRNA biomarkers with prognostic value is crucial [31]. Our prediction model can distinguish new markers for later study.

To continue to improve the modeling process, we counted each AUC value to determine the peak value of the best model and then compared it with related clinical data. We used the AIC value to determine the best cutoff point for model fitting. The ROC curve shows that the risk score can provide a better AUC than other clinical features. Second, the connection between the risk assessment score and clinical pathology was analyzed. After using this prognostic model to distinguish the low-risk group from the high-risk group, we reappraised the survival results, performed multivariate and univariate analyses of the clinicopathological characteristics, and explored tumor cell immune infiltration, chemotherapeutic drug sensitivity to colon cancer, and biomarkers related to immunosuppressive agents, indicating that this model is useful for the treatment and prognosis analysis of colon cancer.

According to the current research findings, immune cell infiltration can be used as a biomarker for the diagnosis and prognosis of stage I–III colon cancer [32]. To assess the connection between the risk score and tumor-infiltrating immune cells, we used six common methods to evaluate immune infiltrating cells—TIMER [33], CIBERSORT [34], XCELL [35, 36], MCPCOUNTER [37], EPIC [38], and CIBERSORT-ABS [39]. The results showed that lncRNAs with different expression levels were positively correlated with tumor-infiltrating immune cells such as CD4+ T

cells, CD8+ T cells, and NK cells but negatively correlated with hematopoietic stem cells and neutrophils. Studies have explained that immune risk scores based on immunohistochemical analysis can demonstrate the therapeutic benefits of immunotherapy and chemotherapy [40]. Stromal cell PD-L1 inhibits the CD8+ T cell antitumor immune response and promotes colon cancer. The MiR-448 targets *ido1* and regulates the CD8+ T cell response in human colon cancer [41, 42]. NKILA lncRNA promotes tumor immune escape by sensitizing T cells to activation-induced cell death [43]. Our drug sensitivity analysis indicates that high risk is related to sensitivity to oxaliplatin, MG132, and NVP-TAE684 chemotherapy. Oxaliplatin added to the fluorouracil and folic acid regimens improves adjuvant therapy for colon cancer [44]. MG132 is a potential therapeutic and preventive agent for cancer cachexia [45]. NVP-TAE684 is an anaplastic lymphoma kinase (ALK) inhibitor that inhibits the proliferation of human pancreatic adenocarcinoma cells [46]. We hypothesized that immunotherapy is more beneficial than chemotherapy by removing cancer cells, producing more new antigens, and inhibiting tumor progression. According to the model, we calculated immunosuppressant-related biomarkers such as PLD2 and MLH1. MMR gene (dMMR) deficiency, usually hMSH2 or hMLH1, promotes the development of colon cancer because of mutation or silencing [47]. MLH1 deletion induces the activation of Her-2/PI3K/Akt signaling and leads to cetuximab resistance in colon cancer [48]. The phospholipase D (PLD) family is widely expressed in cells. PLD2 is mainly found on the plasma membrane under nonstimulated conditions. PLD2 can also be used as a downstream effector of various cell surface receptors to trigger and regulate the transmission of intracellular signals during tumorigenesis and metastasis [49]. PLD2 is abnormally expressed in various cancers. Inhibition or elimination of PLD activity has been proven to

reduce tumor growth and metastasis [50]. Immunotherapy targeting the ICIs PD-1 and PD-L1 has been widely defined as ICIS, showing an obvious improvement in lifetime, and the combination of immunomodulatory therapy can be further optimized to better block the immunosuppressive pathway of the TME and stimulate antitumor immunity [51].

However, this study has shortcomings. For example, our data were only derived from the TCGA database. However, the model must collect clinical cases and increase the sample size for experimental demonstration because the expression level of different samples may make the model unreliable. We hope that further experiments can be performed in the future, such as Western blotting or drug sensitivity tests.

Conclusion

lincRNAs have independent prognostic meaning in colon cancer. Our findings offer a way to predict the prognosis and survival of patients with colon cancer and help to identify which colon cancer patients are more suitable for antitumor immunotherapy. At the same time, our study provides a reference to evaluate other tumor prognosis models.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12957-022-02508-2>.

Additional file 1: Table S1. A total of 1093 immune-related lincRNAs were identified by co-expression analysis.

Additional file 2: Figure S2. (A) and (C) train queue lasso region, (B) and (D) test queue lasso region, (E), train queue forest map shows 14 DElincRNA pairs determined by Cox proportional risk regression in the stepwise method, and (F) test queue forest map shows 13 DElincRNA pairs determined by Cox proportional risk regression in the stepwise method. **Figure S3.** (A) Train queue Risk Score for 213 patients with COAD, the maximum inflection point is the cut-off point obtained by the AIC. (B), Test queue Risk Score for 213 patients with colon cancer, the maximum inflection point is the cut-off point obtained by the AIC. (C) The comparison of the 3-year ROC curve of the train cohort with other common clinical features shows the superiority of risk score. (D) The comparison of the 3-year ROC curve of the test cohort with other common clinical features shows the superiority of risk score. (E) Analysis of time-dependent receiver operating characteristic (ROC) curve in train queue. (F) Time dependent ROC curve analysis of test queue. **Figure S4:** The prognostic model was validated in the train and test cohorts. (A) Overall survival (OS) of the train queue. (B) Total lifetime of test queue. (C) Risk score distribution in the train queue. (D) Risk score distribution in test queue. (E) Scatter diagram of survival status of train queue. (F) Scatter diagram of survival state of test queue.

Acknowledgements

We deeply appreciate Fanqin Bu and Zhonglin Yu for their kind suggestions and generous help in data extraction and analysis.

Authors' contributions

Project design and design: Sicheng Liu and Jun Huang. Research and data collection: Sicheng Liu, Xingyu Peng, Xun Wu, Zhonglin Yu, Jinfeng Zhu, and Fanqin Bu. Statistics and data analysis: Sicheng Liu, Chen Luo, Wenjun Zhang and Jiang Liu. Writing: Sicheng Liu. All authors have read and approved the final version of the manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (Grant No. 82060445), the key research project of Jiangxi Province (Grant No. 2018 bbg70019), and the translational and precision medicine project of the Second Affiliated Hospital of Nanchang University (Grant No. 2016 ynzj12004).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 10 November 2021 Accepted: 20 January 2022

Published online: 06 March 2022

References

- 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-a Cancer J Clin*. 2018;68(6):394–424.
- Stein U, Walther W, Arlt F, Schwabe H, Smith J, Fichtner I, et al. MACC1, a newly identified key regulator of HGF-MET signaling, predicts colon cancer metastasis. *Nat Med*. 2009;15(1):59–67.
- O'Connell JB, Maggard MA, Ko CY. Colon cancer survival rates with the new American joint committee on cancer sixth edition staging. *J National Cancer Inst*. 2004;96(19):1420–5.
- Cross AJ, Wooldrage K, Robbins EC, Kralj-Hans I, MacRae E, Piggott C, et al. Faecal immunochemical tests (FIT) versus colonoscopy for surveillance after screening and polypectomy: a diagnostic accuracy and cost-effectiveness study. *Gut*. 2019;68(9):1642–52.
- Sasikumar PG, Ramachandra RK, Adurthi S, Dhudashiya AA, Vadlamani S, Vemula K, et al. A rationally designed peptide antagonist of the PD-1 signaling pathway as an immunomodulatory agent for cancer therapy. *Mole Cancer Ther*. 2019;18(6):1081–91.
- Tesniere A, Schlemmer F, Boige O, Martins I, Ghiringhelli F, et al. Immunogenic death of colon cancer cells treated with oxaliplatin. *Oncogene*. 2010;29(4):482–91.
- Geisler S, Collier J. RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat Rev Mole Cell Biol*. 2013;14(11):699–712.
- Chen S, Shen X. Long noncoding RNAs: functions and mechanisms in colon cancer. *Mole Cancer*. 2020;19(1):167.
- Wu K, Xu K, Liu K, Huang J, Chen J, Zhang J, et al. Long noncoding RNA BC200 regulates cell growth and invasion in colon cancer. *Int J Biochem Cell Biol*. 2018;99:219–25.
- Peng CL, Zhao XJ, Wei CC, Wu JW. LincRNA HOTAIR promotes colon cancer development by down-regulating miRNA-34a. *Eur Rev Med Pharmacol Sci*. 2019;23(13):5752–61.
- Silva-Fisher JM, Dang HX, White NM, Strand MS, Krasnick BA, Rozycki EB, et al. Long non-coding RNA RAMS11 promotes metastatic colorectal cancer progression. *Nat Commun*. 2020;11(1):2156.
- Wei C, Liang Q, Li X, Li H, Liu Y, Huang X, et al. Bioinformatics profiling utilized a nine immune-related long noncoding RNA signature as a prognostic target for pancreatic cancer. *J Cell Biochem*. 2019;120(9):14916–27.
- Lin Y, Pan X, Chen Z, Lin S, Chen S. Identification of an immune-related nine-lincRNA signature predictive of overall survival in colon cancer. *Front Genet*. 2020;11:318.
- Bu F, Nie H, Zhu X, Wu T, Lin K, Zhao J, et al. A signature of 18 immune-related gene pairs to predict the prognosis of pancreatic cancer patients. *Immunity Inflammation Dis*. 2020;8(4):713–26.
- Jiang Y, Gou X, Wei Z, Tan J, Yu H, Zhou X, et al. Bioinformatics profiling integrating a three immune-related long non-coding RNA signature as

- a prognostic model for clear cell renal cell carcinoma. *Cancer Cell Int.* 2020;20(1):166.
16. Lv Y, Lin S-Y, Hu F-F, Ye Z, Zhang Q, Wang Y, et al. Landscape of cancer diagnostic biomarkers from specifically expressed genes. *Briefings Bioinform.* 2020;21(6):2175–84.
 17. Li W, Wang H, Ma Z, Zhang J, Ou-yang W, Qi Y, et al. Multi-omics analysis of microenvironment characteristics and immune escape mechanisms of hepatocellular carcinoma. *Front Oncol.* 2019;9:1019.
 18. Garon EB, Hellmann MD, Rizvi NA, Carcereny E, Leighl NB, Ahn M-J, et al. Five-year overall survival for patients with advanced non-small-cell lung cancer treated with pembrolizumab: results from the phase I KEY-NOTE-001 study. *J Clin Oncol.* 2019;37(28):2518–27.
 19. Tamang S, Acharya V, Roy D, Sharma R, Aryaa A, Sharma U, et al. SNHG12: An lncRNA as a Potential Therapeutic Target and Biomarker for Human Cancer. *Frontiers. Oncology.* 2019;9:901.
 20. Peng L, Chen Z, Chen Y, Wang X, Tang N. MIR155HG is a prognostic biomarker and associated with immune infiltration and immune checkpoint molecules expression in multiple cancers. *Cancer Med.* 2019;8(17):7161–71.
 21. Sveen A, Agesen TH, Nesbakken A, Meling GI, Rognum TO, Liestol K, et al. ColoGuidePro: a prognostic 7-gene expression signature for stage III colorectal cancer patients. *Clin Cancer Res.* 2012;18(21):6001–10.
 22. Li Z, Wang D, Yin H. A seven immune-related lncRNA signature predicts the survival of patients with colon adenocarcinoma. *Am J Transl Res.* 2020;12(11):7060–78.
 23. Zhang PF, Wu J, Luo JH, Li KS, Wang F, Huang W, et al. SNHG22 overexpression indicates poor prognosis and induces chemotherapy resistance via the miR-2467/Gal-1 signaling pathway in epithelial ovarian carcinoma. *Aging (Albany NY).* 2019;11(19):8204–16.
 24. Mao X, Ji T, Liu A, Weng Y. ELK4-mediated lncRNA SNHG22 promotes gastric cancer progression through interacting with EZH2 and regulating miR-200c-3p/Notch1 axis. *Cell Death Dis.* 2021;12(11):957.
 25. Lu Y, Wang W, Liu Z, Ma J, Zhou X, Fu W. Long non-coding RNA profile study identifies a metabolism-related signature for colorectal cancer. *Mol Med.* 2021;27(1):83.
 26. Wang J, He Z, Xu J, Chen P, Jiang J. Long noncoding RNA LINC00941 promotes pancreatic cancer progression by competitively binding miR-335-5p to regulate ROCK1-mediated LIMK1/Cofilin-1 signaling. *Cell Death Dis.* 2021;12(1):36.
 27. Chang L, Zhou D, Luo S. Novel lncRNA LINC00941 promotes proliferation and invasion of colon cancer through activation of MYC. *Onco Targets Ther.* 2021;14:1173–86.
 28. Yao J, Wang C, Dong X, Zhang Y, Li Y, Zhou H, et al. lncRNA SNHG22 sponges miR1283p to promote the progression of colorectal cancer by upregulating E2F3. *Int J Oncol.* 2021;59(3):71.
 29. Wu N, Jiang M, Liu H, Chu Y, Wang D, Cao J, et al. LINC00941 promotes CRC metastasis through preventing SMAD4 protein degradation and activating the TGF-beta/SMAD2/3 signaling pathway. *Cell Death Differ.* 2021;28(1):219–32.
 30. Xu J, Meng Q, Li X, Yang H, Xu J, Gao N, et al. Long noncoding RNA MIR17HG promotes colorectal cancer progression via miR-17-5p. *Cancer Res.* 2019;79(19):4482–895.
 31. Lv J, Guo Y, Yan L, Lu Y, Liu D, Niu J. Development and validation of a five-lncRNA signature with prognostic value in colon cancer. *J Cell Biochem.* 2019;8:3780–93.
 32. Zhou R, Zhang J, Zeng D, Sun H, Rong X, Shi M, et al. Immune cell infiltration as a biomarker for the diagnosis and prognosis of stage I-III colon cancer. *Cancer Immunol Immunother.* 2019;68(3):433–42.
 33. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* 2017;77(21):E108–10.
 34. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT, in *Cancer Systems Biology: Methods and Protocols*, L. VonStechow; 2018. p. 243–259.
 35. Aran D. Cell-type enrichment analysis of bulk transcriptomes using xCell. *Methods Mole Biol (Clifton, N.J.).* 2020;2120:263–76.
 36. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 2017;18(1):220.
 37. Dienstmann R, Villacampa G, Sveen A, Mason MJ, Niedzwiecki D, Nesbakken A, et al. Relative contribution of clinicopathological variables, genomic markers, transcriptomic subtyping and microenvironment features for outcome prediction in stage II/III colorectal cancer. *Ann Oncol.* 2019;30(10):1622–9.
 38. Racle J, de Jonge K, Baumgaertner P, Speiser DE, Gfeller D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *Elife.* 2017;6:e26476.
 39. Tamminga M, Hiltermann TJN, Schuurin E, Timens W, Fehrmann RSN, Groen HJM. Immune microenvironment composition in non-small cell lung cancer and its association with survival. *Clin Transl Immunol.* 2020;9(6):e1142.
 40. Dai G-P, Wang L-P, Wen Y-Q, Ren X-Q, Zuo S-G. Identification of key genes for predicting colorectal cancer prognosis by integrated bioinformatics analysis. *Oncol Letters.* 2020;19(1):388–98.
 41. O'Malley G, Treacy O, Lynch K, Naicker SD, Leonard NA, Lohan P, et al. Stromal Cell PD-L1 Inhibits CD8(+) T-cell antitumor immune responses and promotes colon cancer. *Cancer Immunol Res.* 2018;6(11):1426–41.
 42. Lou Q, Liu R, Yang X, Li W, Huang L, Wei L, et al. miR-448 targets IDO1 and regulates CD8(+) T cell response in human colon cancer. *J Immunother Cancer.* 2019;7(1):210.
 43. Huang D, Chen J, Yang L, Ouyang Q, Li J, Lao L, et al. NKILA lncRNA promotes tumor immune evasion by sensitizing T cells to activation-induced cell death. *Nat Immunol.* 2018;19(10):1112–25.
 44. Andre T, Boni C, Mounedji-Boudiaf L, Navarro M, Tabernero J, Hickish T, et al. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. *N Engl J Med.* 2004;350(23):2343–51.
 45. Zhang L, Tang H, Kou Y, Li R, Zheng Y, Wang Q, et al. MG132-mediated inhibition of the ubiquitin-proteasome pathway ameliorates cancer cachexia. *J Cancer Res Clin Oncol.* 2013;139(7):1105–15.
 46. Duong HQ, Than VT, Nguyen HT, Nguyen PT, Thi HTH, Bui TNH, et al. Anaplastic lymphoma kinase inhibitor NVPTAE684 suppresses the proliferation of human pancreatic adenocarcinoma cells. *Oncol Rep.* 2021;45(4):28.
 47. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. *Science.* 2013;339(6127):1546–58.
 48. Han Y, Peng Y, Fu Y, Cai C, Guo C, Liu S, et al. MLH1 deficiency induces cetuximab resistance in colon cancer via Her-2/PI3K/AKT signaling. *Adv Sci (Weinh).* 2020;7(13):2000112.
 49. Hiroshima M, Exton JH. Localization and regulation of phospholipase D2 by ARF6. *J Cell Biochem.* 2005;95(1):149–64.
 50. Yao Y, Wang X, Li H, Fan J, Qian X, Li H, et al. Phospholipase D as a key modulator of cancer progression. *Biol Rev Camb Philos Soc.* 2020;95(4):911–35.
 51. Buoncervello M, Gabriele L, Toschi E. The Janus face of tumor microenvironment targeted by immunotherapy. *Int J Mol Sci.* 2019;20(17):4320.

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

