



Using Immune-Related Long Non-coding Ribonucleic Acids to Develop a Novel Prognosis Signature and Predict the Immune Landscape of Colon Cancer

Xu Wang[†], Ke Chen[†], Zhenglin Wang[†], Yuanmin Xu, Longfei Dai, Tao Bai, Bo Chen, Wenqi Yang and Wei Chen*

Department of General Surgery, The First Affiliated Hospital of Anhui Medical University, Hefei, China

OPEN ACCESS

Edited by:

Feng Wang,
Affiliated Hospital of Nantong
University, China

Reviewed by:

Fei Mao,
Jiangsu University, China
Shukui Wang,
Nanjing Medical University, China

*Correspondence:

Wei Chen
chenwei366@ahmu.edu.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Molecular and Cellular Oncology,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 31 July 2021

Accepted: 08 September 2021

Published: 30 September 2021

Citation:

Wang X, Chen K, Wang Z, Xu Y,
Dai L, Bai T, Chen B, Yang W and
Chen W (2021) Using
Immune-Related Long Non-coding
Ribonucleic Acids to Develop a Novel
Prognosis Signature and Predict
the Immune Landscape of Colon
Cancer.
Front. Cell Dev. Biol. 9:750709.
doi: 10.3389/fcell.2021.750709

Purpose: This study aimed to construct a novel signature to predict the survival of patients with colon cancer and the associated immune landscape, based on immune-related long noncoding ribonucleic acids (lncRNAs).

Methods: Expression profiles of lncRNAs in 457 patients with colon cancer were retrieved from the TCGA database (<https://portal.gdc.cancer.gov>). Differentially expressed (DE) lncRNAs were identified and lncRNA pairs were recognized using Lasso regression and Cox regression analyses. Akaike information criterion (AIC) values of receiver operating characteristic (ROC) curve were calculated to identify the ideal cut-off point for dividing patients into two groups and constructing the prognosis signature. Quantitative real-time polymerase chain reaction (qRT-PCR) was performed to validate the expression of LINC02195 and SCARNA9 in colon cancer.

Results: We identified 22 lncRNA pairs and patients were divided into high-risk and low-risk groups based on the calculated risk score using these 22 lncRNA pairs. The lncRNA pairs were significantly related to patient survival. Low-risk patients had a significantly longer survival time than high-risk patients ($p < 0.001$). The area under the curve of the signature to predict 5-year survival was 0.951. The risk score correlated with tumor stage, infiltration depth, lymph node metastasis, and distant metastasis. The risk score remained significant after univariate and multivariate Cox regression analyses. A nomogram model to predict patient survival was developed based on the results of Cox regression analysis. Immune cell infiltration status, expression of some immune checkpoint genes, and sensitivity to chemotherapeutics were also related to the risk score. The results of qRT-PCR revealed that LINC02195 and SCARNA9 were significantly upregulated in colon cancer tissues.

Conclusion: The constructed prognosis signature showed remarkable efficiency in predicting patient survival, immune cell infiltration status, expression of immune checkpoint genes, and sensitivity to chemotherapeutics.

Keywords: immune related long noncoding RNAs, prognosis signature, colon cancer, tumor-infiltrating immune cell, immune checkpoint genes, chemotherapeutics

INTRODUCTION

Despite the rapid development of medical treatments, the trends in cancer incidence and death rates have been increasing worldwide. Specifically, the incidence and mortality rate of colon cancer are relatively high (Bray et al., 2018). Patients with early colon cancer can be treated by surgery; however, most patients with advanced colon cancer experience cancer recurrence and metastasis, and their 5-year survival rate is lower than 10% (Bhandari et al., 2017; Doonan et al., 2017; Russo et al., 2019). With the development of chemotherapy and targeted medicine, the overall survival rate of patients with colon cancer is significantly higher now than before. In recent years, advances in tumor immune therapy and application of immune checkpoint inhibitors have led to improvements in cancer treatment.

Programmed cell death protein 1 (PD-1), first discovered in 1992, is a 288 amino acid protein expressed on the surface of T cells and related to apoptosis (Ishida et al., 1992). When PD-1 is bound to its ligand programmed cell death ligand 1 (PD-L1), the anti-tumor effect of T cells is inhibited (Kouo et al., 2015). PD-1-blocking antibodies, pembrolizumab and nivolumab, were approved by the United States Food and Drug Administration (FDA) for the treatment of refractory melanoma and advanced non-small cell lung cancer in 2014 and 2015, respectively. Atezolizumab, the first anti-PD-L1 antibody, was approved for the treatment of urothelial cancers in 2016. With the rapid development of tumor immunotherapy, several immune checkpoint inhibitors have been used in the treatment of various types of malignant tumors. According to a recent meta-analysis (He et al., 2020), anti-PD-1 inhibitors have high efficacy and have led to a better prognosis in patients with deficient mismatch repair (dMMR)/microsatellite instability high (MSI-H) metastatic colorectal cancer (mCRC) (dMMR/MSI-H mCRC).

Long non-coding ribonucleic acids (lncRNAs), defined as RNAs longer than 200 nucleotides, are not translated into functional proteins (Iyer et al., 2015). The completely spliced lncRNA is transported into the cytoplasm or other organelles through a mechanism similar to that of mRNA. Once in the cytoplasm, lncRNAs transregulate gene expression at the post-transcriptional level, such as regulating mRNA translation and degradation, or participating in the regulation of intracellular signaling pathways (Statello et al., 2021). Recent studies have shown that lncRNAs not only change the genome or transcriptome, but also modify the immune microenvironment to contribute to the malignant phenotype of cancer. lncRNAs direct the expression of genes related to immune cell activation, leading to tumor immune cell infiltration (Atianand et al., 2017; Chen et al., 2017). Several lncRNAs are differentially expressed (DE) in various types of tumor tissues (Schmitt and Chang, 2016).

In recent years, many researchers have focused on constructing signatures using lncRNAs for predicting prognosis of patients with cancer based on the cancer Genome Atlas (TCGA) database. Qi et al. (2021) identified eight immune-related (ir) lncRNAs and developed a signature for the prognosis of patients with pancreatic adenocarcinoma. Zhou et al. (2021) identified an irlncRNA signature to predict the prognosis,

immune cell infiltration, and immunotherapy response in patients with hepatocellular carcinoma and validated the expression of the six lncRNAs using the quantitative real-time polymerase chain reaction (qRT-PCR) method. Zhang et al. (2021) built a signature based on ten hypoxia-related lncRNAs that showed promising predictive effect for patient prognosis. Ma et al. (2021) identified a metabolism-related lncRNA signature for prediction of risk of recurrence in patients with breast cancer. Compared with using single biomarkers, combinations of two biomarkers lead to much more accurate diagnostic models for cancers (Lv et al., 2020). Moreover, these combinations do not require any specific expression levels or testing methods. However, not many studies have focused on the development of signatures based on lncRNA pairs for diagnosis or survival prediction in patients with colon cancer. We constructed a prognosis signature using irlncRNA pairs, which showed remarkable efficiency in predicting patient survival, immune cell infiltration status, expression of immune checkpoint genes, and sensitivity to chemotherapeutics.

MATERIALS AND METHODS

Immune-Related Long Noncoding Ribonucleic Acids Identification and Expression

Expression patterns of irlncRNAs of patients with colon cancer were downloaded from the Genomic Data Commons (GDC) Data Portal.¹ lncRNAs and mRNAs were distinguished using GTF files, which were downloaded from Ensembl.² A list of human immune-related genes was prepared using the ImmPort database,³ a publicly available repository containing up-to-date information on human genes and proteins that are involved in immunity. Co-expression analysis was performed between immune-related genes and all lncRNAs to identify irlncRNAs. The screening criteria used was correlation coefficient >0.4 and *p*-value < 0.001. *limma* package of R software (version 4.0.3) was used to screen out differentially expressed immune-related lncRNAs (DEirlncRNAs) with criteria of log | fold change (FC)| >2.5 and *p*-value < 0.01.

Gene Oncology and Kyoto Encyclopedia of Genes and Genomes Analyses

Pearson correlation coefficient analysis was performed to analyze the relationship between the expression of DEirlncRNAs and mRNAs and the top 10 mRNAs were considered to be associated with lncRNAs. To better understand the biological functions and pathways involved in DEirlncRNAs, Gene Oncology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed using *ggplot2*, *Bioconductor*, and *org.Hs.eg.db* R packages. *P*-values and *q*-values < 0.05 were considered statistically significant.

¹<https://portal.gdc.cancer.gov>

²<http://asia.ensembl.org>

³<http://www.immport.org>

Construction of Differentially Expressed Immune-Related Long Non-coding Ribonucleic Acid Pairs

A DElncRNA pair was constructed using two DElncRNAs, for example, lncRNA A and lncRNA B. A 0-or-1 matrix was constructed. If the ratio of lncRNA A to lncRNA B was higher than 1, the expression of the lncRNA pair was defined as 1, otherwise, it was defined as 0. If the proportion of lncRNA pairs with expression defined as 0 or 1 was less than 20% or more than 80%, the pair was considered invalid.

Clinical Data of Colon Cancer Patients

Clinical information of patients with colon cancer was downloaded from colon adenocarcinoma (COAD) project of TCGA database. Patients with no follow-up or incomplete clinical information were excluded. After selection, 393 cases of colon cancer were included in this study.

Development and Validation of the Prognosis Signature Using Differentially Expressed Immune-Related Long Non-coding Ribonucleic Acid Pairs

Least Absolute Shrinkage and Selection Operator (LASSO) regression analysis and Cox regression analysis were performed to screen out prognosis-related lncRNA pairs using *survival*, *survminer*, and *glmnet* R packages. lncRNA pairs with $p < 0.05$ were considered significant. After selection, 22 lncRNA pairs were included to construct the prognosis signature. Risk scores were calculated based on the following formula: Risk score = $\sum_{i=1}^n \beta_i^* \lambda_i$, where n represents the numbers of lncRNA pairs included to construct the signature and β_i and λ_i represent the regression coefficient and expression value of lncRNA pairs, respectively. According to the maximum akaike information criterion (AIC) value of 5-year receiver operating characteristic (ROC) curve, patients were divided into high-risk and low-risk groups. ROC and AUC were used to test the prediction efficiency and compare the constructed signature and other clinical variables. The Kaplan-Meier (KM) method and log rank test were used to evaluate survival differences between high-risk and low-risk groups. Univariate and multivariate analyses were used to determine whether risk score was an independent predictor of prognosis in patients with colon cancer. Chi-square test was performed to analyze the relationship between the signature and clinical variables and the Wilcoxon signed-rank test was used to show the risk score differences between different groups for these clinical characteristics. A nomogram model was developed based on three independent prognosis factors that were significant in both the univariate and multivariate analyses ($p < 0.05$). Calibration graphs were constructed to show the differences between nomogram-predicted and actual survival rates of patients with colon cancer.

Immune Cell Infiltration Analysis

To better understand the relationship between the calculated risk score and tumor immune cell infiltration status, datasets

including XCELL, TIMER, QUANTISEQ, EPIC, CIBERSORT-ABS, and CIBERSORT were used to analyze the immune cell infiltration status. The lollipop diagram was drawn to show the correlation between risk score and immune infiltrated cells *via* Spearman correlation method. The differences of immune cell content in high-risk and low-risk groups were shown as boxplots using Wilcoxon signed-rank test.

Expression of Immune Checkpoint Genes in High-Risk and Low-Risk Groups

To understand the differences in the expression levels of immune checkpoint genes in the high-risk and low-risk groups, six immune checkpoint genes were selected, including CTLA4, HAVCR2, IDO1, lymphocyte-activation gene 3 (LAG3), PD-1, and PD-L1. Violin plots were drawn to show the results using *ggpubr* R package.

Evaluating the Differences in Chemosensitivity Between COAD Patients in High-Risk and Low-Risk Groups

IC50, half of the maximum inhibitory concentration, represents the concentration of drug required for 50% inhibition of cancer cells. IC50 was calculated to evaluate the significance of lncRNA-based signature in six types of chemotherapeutic drugs, including camptothecin, doxorubicin, erlotinib, gemcitabine, paclitaxel, and rapamycin, which have been used in the treatment of patients with colon cancer. Wilcoxon signed-rank test was performed to analyze the differences in IC50 in high-risk and low-risk groups. The results are shown as boxplots using *ggpubr*, *PRRphetic*, and *ggplot2* R packages.

The Verification of LINC02195 and SCARNA9 by Quantitative Real-Time Polymerase Chain Reaction

Ten pairs of colon cancer tissues and adjacent non-cancer tissues were collected from The First Affiliated Hospital of Anhui Medical University, and approved by the Ethics Committee. All participants signed an informed consent form. These samples were collected after surgical resection from colon cancer patients who had never received preoperative chemotherapy or radiotherapy. The HiPure Universal RNA Kit (Shanghai, Magen) was used to extract total RNA from the colon cancer and adjacent non-cancer tissues stored in liquid nitrogen following the manufacturer's instructions. The concentration and purity of RNA samples was measured using NanoDrop 2000 (Thermo Fisher Scientific, United States). Extracted RNA was reverse transcribed into cDNA using the PrimeScript RT kit (Vazyme, Nanjing, China) according to the protocol. Finally, the concentration of cDNA was measured using TB Green Premix Ex Taq II (GenStar, China) under the LightCycler480 System (Applied Biosystems, Waltham, MA, United States) according to the manufacturer's instruction. The primer sequences for PCR amplification were as follows: LINC02195, forward: 5'-GTCA

CACAGCAAGCCTAAAGAAACG-3', reverse: 5'-TCAGCCA TAGAGGAGACAGCAAGG-3'; SCARNA9, forward: 5'-AAGG GCATATGTCTGGTGTGTGTG-3', reverse: 5'-CCCCACCCTC AATCTCATTTCATCC-3'; GAPDH, forward: 5'-GGGAAGG TGAAGGTCGGAGT-3', reverse: 5'-GGGGTCATTGATGGCA ACA-3'. GAPDH was used as an internal control, and each sample was repeated three times. The relative expression levels of LINC02195 and SCARNA9 were calculated using the $2^{-\Delta\Delta Ct}$ method. The differences in LINC02195 and SCARNA9 expression between colon cancer tissues and adjacent non-cancer tissues were tested by *t*-tests. The graphs were drawn using GraphPad Prism software (version 9.2.0).

RESULTS

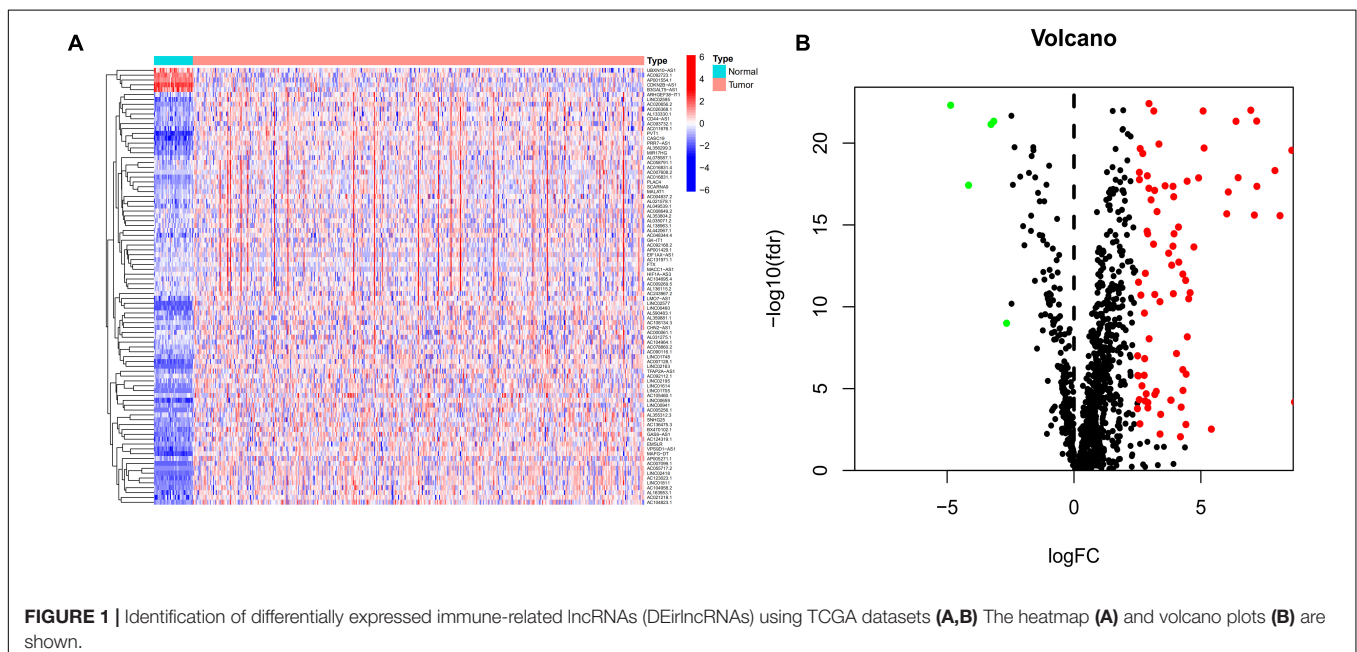
Differentially Expressed Immune-Related Long Noncoding Ribonucleic Acids

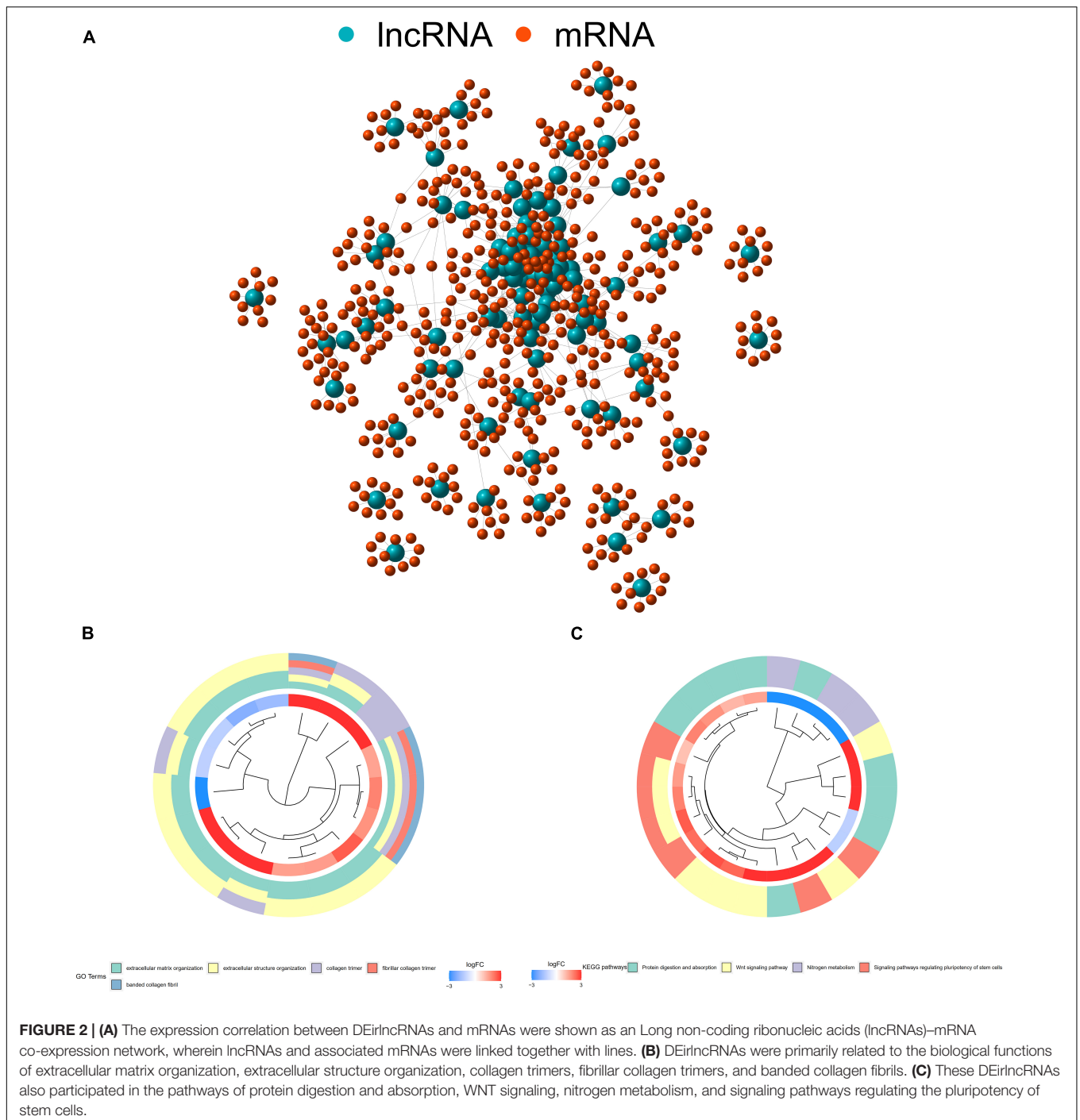
Expression patterns and clinical information from 457 COAD patients were downloaded from TCGA database. A total of 393 patients with complete clinical information and follow-up time >0 days were included in the study. Detailed clinical characteristics of the 393 cases are shown in **Supplementary Table 1**. Human immune-related genes were identified using the ImmPort database. Co-expression analysis was performed to identify irlncRNAs based on immune-related genes. As shown in **Figure 1A**, there were 90 irlncRNAs with $\log |FC| > 2.5$ and *p* value < 0.01, among which 85 lncRNAs were upregulated while five were downregulated (**Figure 1B**). To understand the biological functions and pathways involved in the 90 DEirlncRNAs, the expression correlation between the DEirlncRNAs and mRNAs were shown as a lncRNA-mRNA co-expression network, and lncRNAs and associated

mRNAs were linked together with lines (**Figure 2A**). As seen in **Figure 2B**, the 90 lncRNAs were primarily related to the biological functions of extracellular matrix organization, extracellular structure organization, collagen trimer, fibrillar collagen trimer, and banded collagen fibril (**Figure 2B**). These DEirlncRNAs also participated in the pathways of protein digestion and absorption, WNT signaling, nitrogen metabolism, and regulating pluripotency of stem cells (**Figure 2C**).

Construction of Differentially Expressed Immune-Related Long Non-coding Ribonucleic Acid Pairs and Prognosis Signature

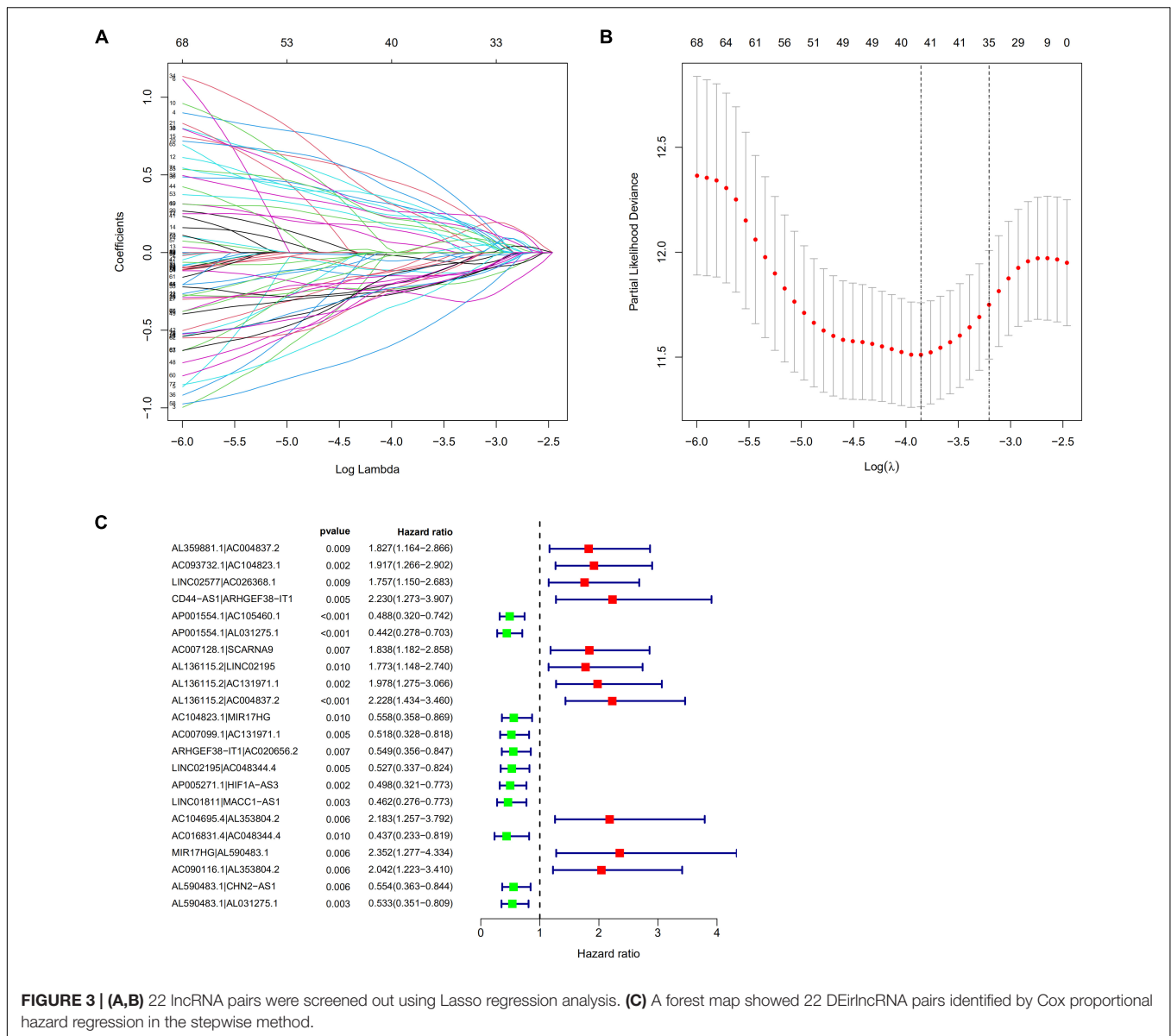
A total of 2,720 valid lncRNA pairs were constructed based on the 90 DElncRNAs and 22 were screened out using LASSO regression analysis (**Figures 3A,B**). As shown in **Figure 3C**, all the 22 lncRNA pairs were significant in univariate Cox regression analysis (*p* < 0.05). The risk score was defined as $\sum_{i=1}^n \beta_i^* \lambda_i$. Then, 5-year AUC for ROC curves of the 22 lncRNA pairs were calculated. The maximum AUC value of the signature to predict 5-year survival was 0.951; the AIC value was also calculated to identify the ideal cut-off point to divide patients into high-risk and low-risk groups (**Figure 4A**). As shown in **Figure 4B**, the AUC value of the signature obtained in this study was much higher than that of three other lncRNA-based signatures of patients with colon cancer from three other studies, including LiLncSig (Li Z. et al., 2020) (AUC = 0.721), LinLncSig (Lin et al., 2020) (AUC = 0.796), and XingLncSig (Xing et al., 2018) (AUC = 0.665). The AUC for 1-, 3-, and 5-year survival were 0.851, 0.893, and 0.951, respectively (**Figure 4C**). AUC of other clinical-pathological features were also presented. In the constructed signature, risk score had a higher efficiency for predicting 1-, 3-, and 5-year survival than other variables





in COAD patients (Figures 4D–F). Based on the risk score, patients were divided into high-risk and low-risk groups using the calculated cut-off point (Figure 5A), and patients with higher risk scores had a higher risk of mortality (Figure 5B). As seen in Figure 5C, survival curve was plotted to show the survival differences of COAD patients in the two groups. Patients in the high-risk group had a significantly lower probability of survival than those in the low-risk group ($p < 0.001$). Chi-square tests were performed to investigate the relationship between risk score

and other clinical-pathological features. A heatmap was plotted, showing that age, clinical stage, T stage, N stage, and M stage were significantly related to the risk score (Figure 6A). Univariate and multivariate Cox regression analyses were performed to identify prognosis-related factors in COAD patients (Figures 6B,C). Factors with p value < 0.05 in the univariate analysis were included in the multivariate analysis. Forest maps showed that age ($p < 0.001$, HR = 1.054, 95% CI [1.031–1.078]), stage ($p = 0.035$, HR = 2.344, 95% CI [1.060–5.182]), and the risk

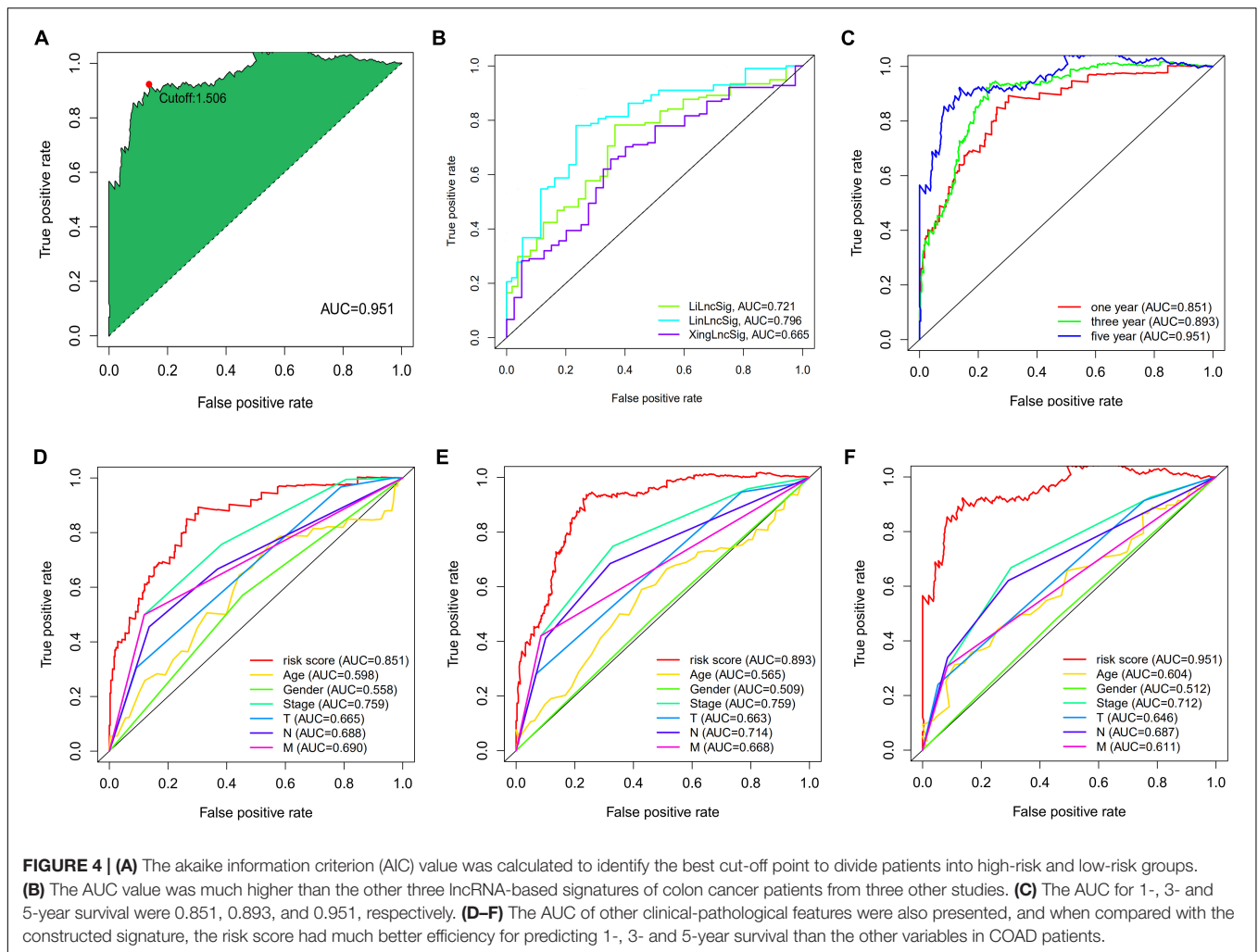


score ($p < 0.001$, HR = 1.042, 95% CI [1.029–1.056]) were still significant after multivariate analysis. Therefore, the risk score was independently associated with the prognosis of COAD patients. Wilcoxon signed-rank test showed that clinical stage (Figure 6D), T stage (Figure 6E), N stage (Figure 6F), and M stage (Figure 6G) were significantly related to the calculated risk score. To better predict 1-, 3-, and 5-year survival rates of COAD cases, a nomogram model was constructed based on the results of univariate and multivariate Cox regression analyses (Figure 7A). Age, clinical stage, and risk score were included in the nomogram model. Accordingly, a 60-year-old patient with stage IV colon cancer and risk score of 45 has an estimated 5-year survival rate of less than 10%. Moreover, calibration plots depicting the differences between nomogram-predicted and actual survival probabilities of COAD patients showed that the predicted 1-, 3-, and 5-year survival probabilities were close to the actual survival

probabilities (Figures 7B–D), indicating that this nomogram model accurately predicted survival of COAD patients.

Evaluation of the Relationship Between Risk Score and Immune Cell Infiltration Status

To better understand the correlation between risk score and tumor immune microenvironment, the Spearman correlation and Wilcoxon signed-rank tests were performed. As Figure 8A shows, the risk score correlated with many types of immune cells, including granulocyte-monocyte progenitors, neutrophils, CD4+ T cells, myeloid dendritic cells, cancer associated fibroblasts, and activated natural killer (NK) cells. High risk score was positively associated with tumor infiltrating immune cells including CD4+ T cells (Figures 8C,F), neutrophils (Figure 8D),



and activated NK cells (Figures 8H,I), and negatively associated with hematopoietic stem cells (Figure 8B), myeloid dendritic cells (Figure 8E), and uncharacterized cells (Figure 8G).

Expression of Immune Checkpoint Genes in High-Risk and Low-Risk Groups

To use the risk score to predict potential checkpoint blockade therapy, violin plots were drawn to show the differences of immune checkpoint gene expression in high-risk and low-risk groups. LAG3 (Figure 9D, $p < 0.05$) and PD-1 (Figure 9E, $p < 0.01$) expressions were significantly different in the two groups, whereas CTLA4 (Figure 9A, $p > 0.05$), HAVCR2 (Figure 9B, $p > 0.05$), IDO1 (Figure 9C, $p > 0.05$), and PD-L1 (Figure 9F, $p > 0.05$) expressions showed no significant difference between the groups.

Using the Risk Score to Predict Chemosensitivity of COAD Patients

The differences in chemosensitivity evaluated using IC50 values in high-risk and low-risk groups were analyzed using

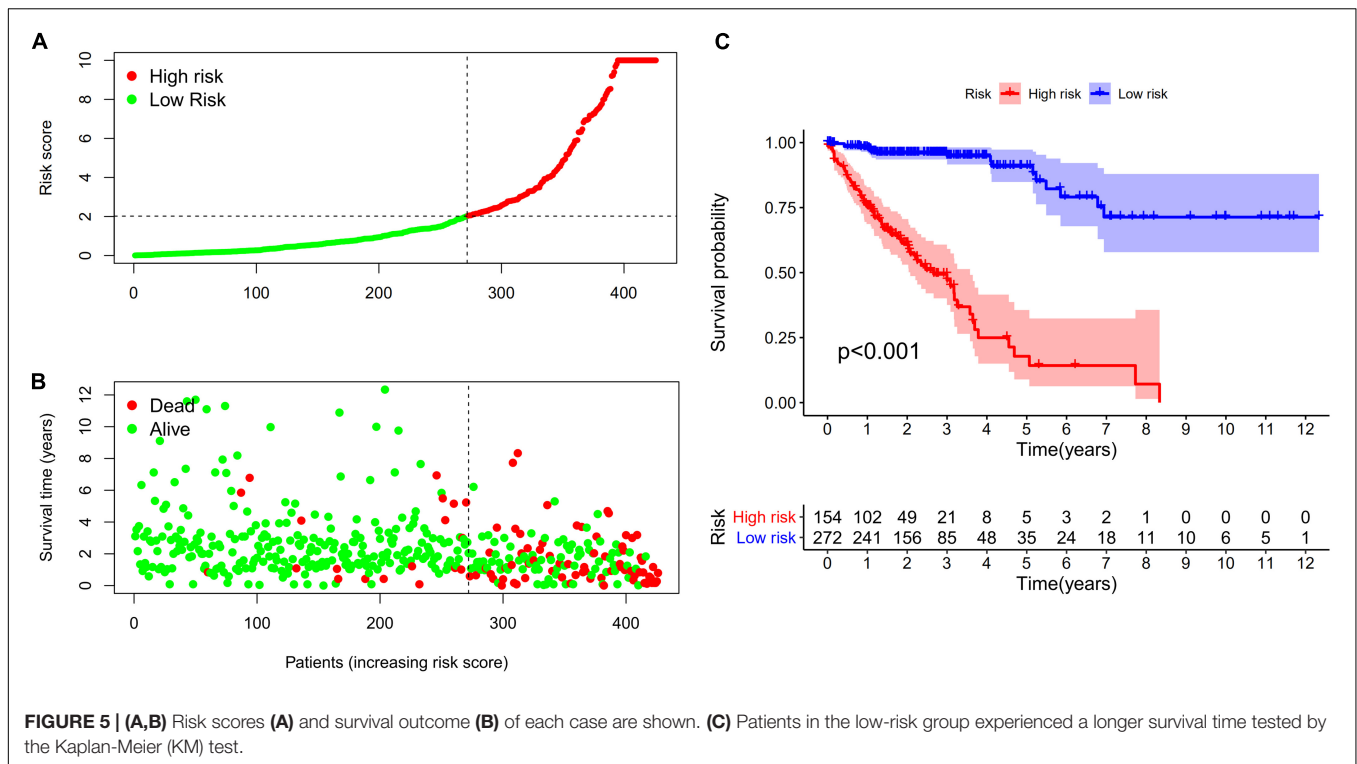
Wilcoxon signed-rank test. The results indicated that lower IC50 values of camptothecin, doxorubicin, erlotinib, gemcitabine, paclitaxel, and rapamycin were associated with higher risk scores (Figure 10). These results might provide reference for clinical treatment of COAD.

Validating Expression Levels of LINC02195 and SCARNA9 via Quantitative Real-Time Polymerase Chain Reaction

To explore the expression levels of LINC02195 and SCARNA9, these lncRNAs were tested in colon cancer and adjacent non-cancer tissues using qRT-PCR method. As Figure 11 shows, the expression levels of LINC02195 and SCARNA9 in colon cancer were significantly higher than that in adjacent non-cancer tissues.

DISCUSSION

Recent studies have presented signatures for cancer diagnosis and survival prediction based on the specific expression



levels of coding genes or noncoding RNAs (Liu et al., 2021; Wang R. et al., 2021; Wang X. et al., 2021; Wu et al., 2021), which required certain testing methods. In this study, lrlncRNA pairs were constructed and used for the development of prognosis signature. The combination of two lncRNAs did not require the exact expression quantities to be measured by certain methods, making it much more feasible and convenient in clinical use.

First, the expression and clinical data were retrieved from TCGA database. Differential co-expression analysis was performed to identify DElncRNAs and GO and KEGG analyses were performed to explore relevant pathways and molecular biological functions. The results suggested that DElncRNAs were related to the biological functions of the organization of cellular and noncellular components. Malignant tumors are surrounded by extracellular matrix and stromal cells and these cellular and non-cellular components build up the tumor microenvironment (Wang et al., 2017). The interactions between tumor microenvironment and cancer cells have great importance in cancer progression and metastasis (Joyce and Pollard, 2009; Quail and Joyce, 2013). Results showed that DElncRNAs participated in the pathways of protein digestion and absorption, WNT signaling, nitrogen metabolism, and regulating pluripotency of stem cells. The WNT signaling pathway is related to several cancer types, especially colorectal cancer (CRC; Cancer Genome Atlas Network, 2012). Wnt- β -catenin signal activation leads to the accumulation of β -catenin in the nucleus, which has been detected in over 80% of CRC tumor tissues (Wanitsuwan et al., 2008). In addition, high levels of nuclear β -catenin are associated with poor prognosis in patients with CRC (Baldus et al., 2004). A recent study indicated

that nitrogen metabolism in was changed in various types of cancer, which was detectable in body fluids and might cause new mutations in cancer tissues (Lee et al., 2018). Crespo et al. (2018) suggested that colonic organoids from human-induced pluripotent stem cells can be used for modeling CRC.

A total of 2,720 valid lncRNA pairs were constructed using 90 DElncRNAs. To explore the impact of DElncRNA pairs on prognosis in colon cancer patients, 22 prognosis-related DElncRNA pairs were identified using LASSO regression analysis and Cox regression analysis. Some of the DElncRNAs used for modeling have already been demonstrated to play an important role in CRC and other types of malignant tumors. Anirban et al. (Li H. et al., 2020) demonstrated that LINC02195 is a regulator of histocompatibility complex class I molecules and a prognosis biomarker for head and neck squamous cell carcinoma. Wang et al. (2020) showed that SCARNA9 is associated with the prognosis of patients with endometrial cancer. Xu et al. (2019) suggested that MIR17HG is an immune-related lncRNA, as it was upregulated in CRC tissues compared with normal tissues. Moreover, MIR17HG also contributed to tumorigenesis and metastasis in CRC cells both *in vitro* and *in vivo*. Cao et al. (2019) also showed that MIR17HG is upregulated in glioma tissues and cell lines and that downregulation of MIR17HG is related to inhibition of glioma cell progression. Therefore, the constructed signature can identify new biomarkers for further studies.

Subsequently, every AUC value was calculated to construct a signature with the maximum AUC value. The 1-, 3- and 5-year AUCs of the prognosis signature were also compared with other clinical variables. AIC value was calculated to find the

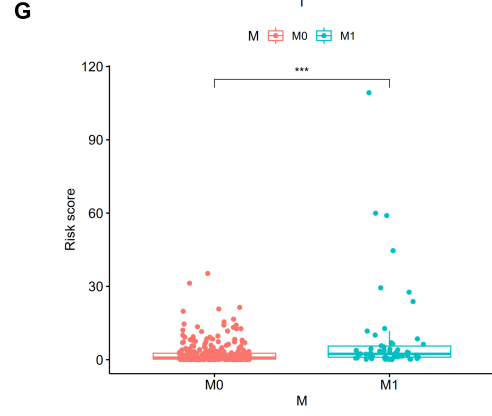
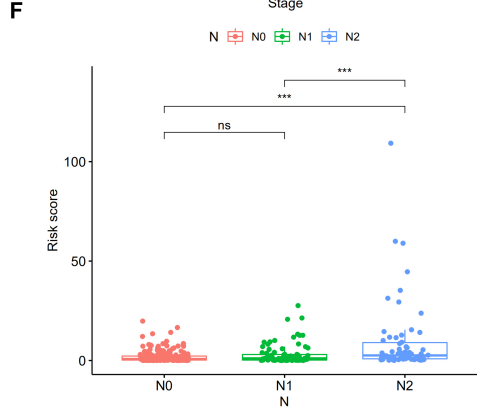
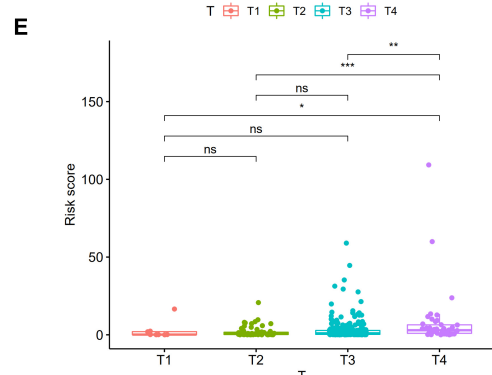
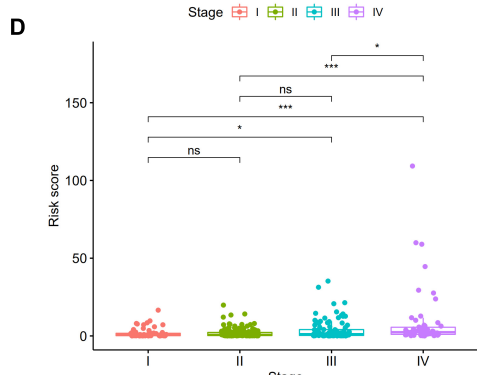
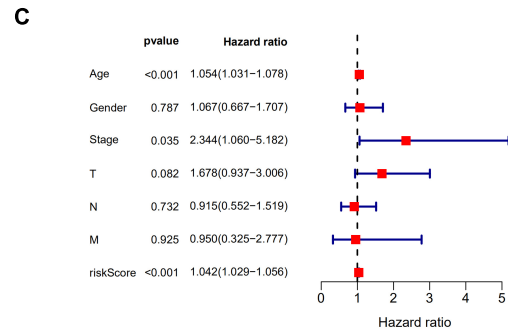
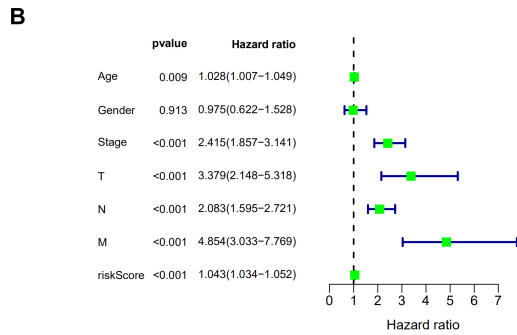
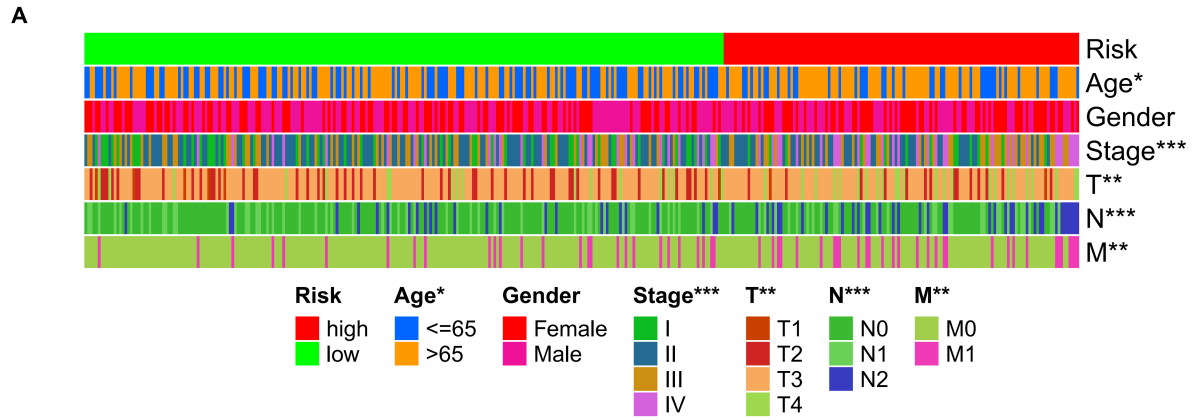
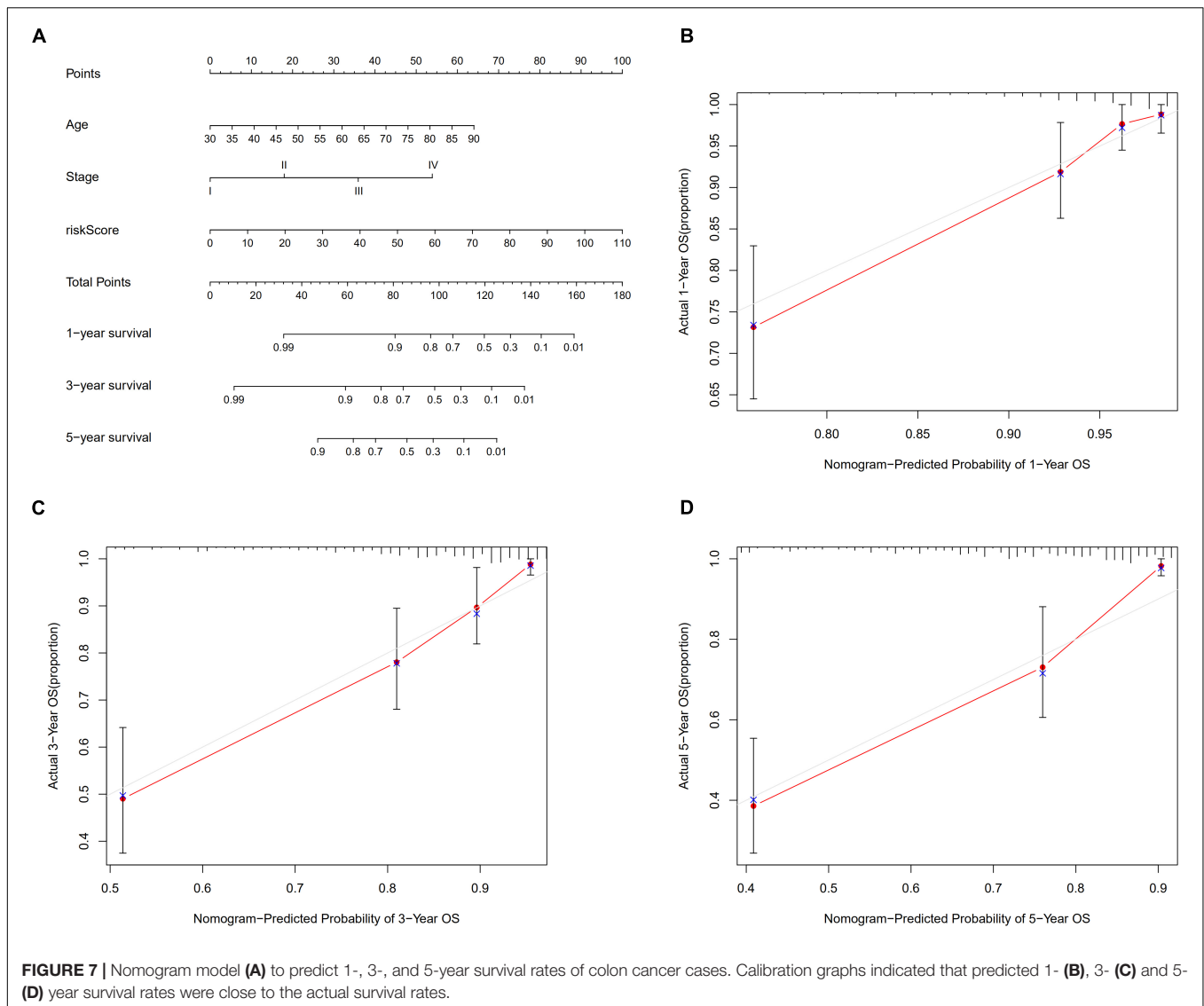


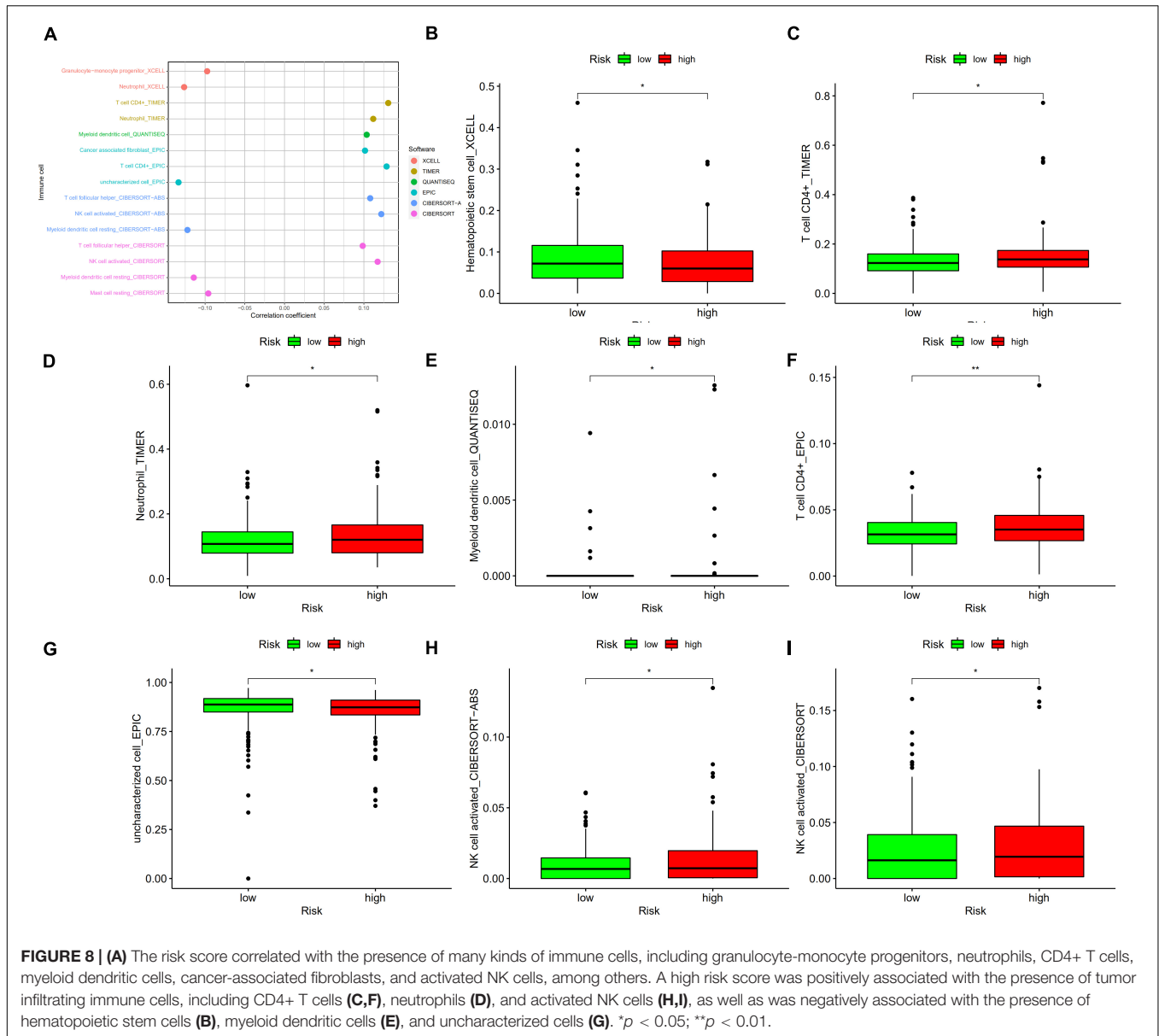
FIGURE 6 | Strip chart (A) showed that age, clinical stage, tumor infiltration depth, lymph node metastasis, and distant metastasis were significantly associated with the risk score. Forest plots of univariate (B) and multivariate (C) Cox regression analyses in colon cancer. Scatters diagram also showed that (D) clinical stage, (E) tumor infiltration depth, (F) lymph node metastasis, and (G) distant metastasis status significantly correlated with the risk score. * $p < 0.05$; ** $p < 0.01$; and *** $p < 0.001$.



ideal cut-off point to divide patients into high- and low-risk groups. KM curve showed that survival rates of low-risk patients were much higher than those of high-risk patients. To better understand the utility of clinical variables and risk score on predicting patient outcomes, univariate and multivariate Cox analyses were performed. These results showed that the risk score remained significant after these analyses, indicating that the calculated risk score was an independent predictor of patient prognosis. The relationship between the risk score and other clinical features was also determined. The results suggested that the risk score correlated with tumor stage, tumor infiltration depth, lymph node metastasis, and distant metastasis, indicating that the risk score might be related to the development and migration of colon cancer. Nomogram is a prediction tool in oncology, especially for cancer prognosis (Iasonos et al., 2008; Balachandran et al., 2015). A nomogram model was established to visualize the effects of clinical features and risk score on 1-, 3- and 5-year survival probabilities of patients. Calibration graphs

showed that the nomogram-predicted survival rates were close to actual survival rates, indicating that the nomogram model had high prediction efficiency.

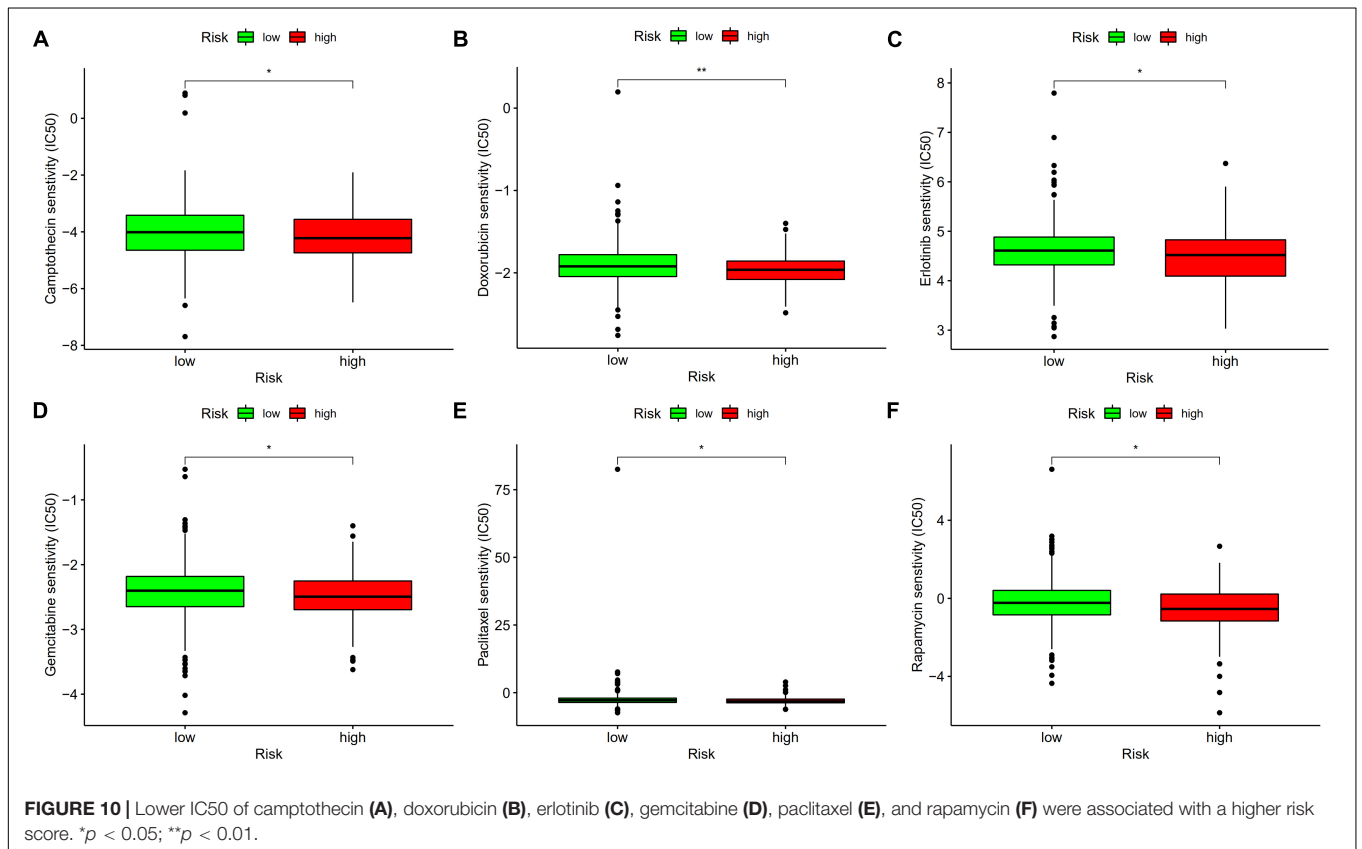
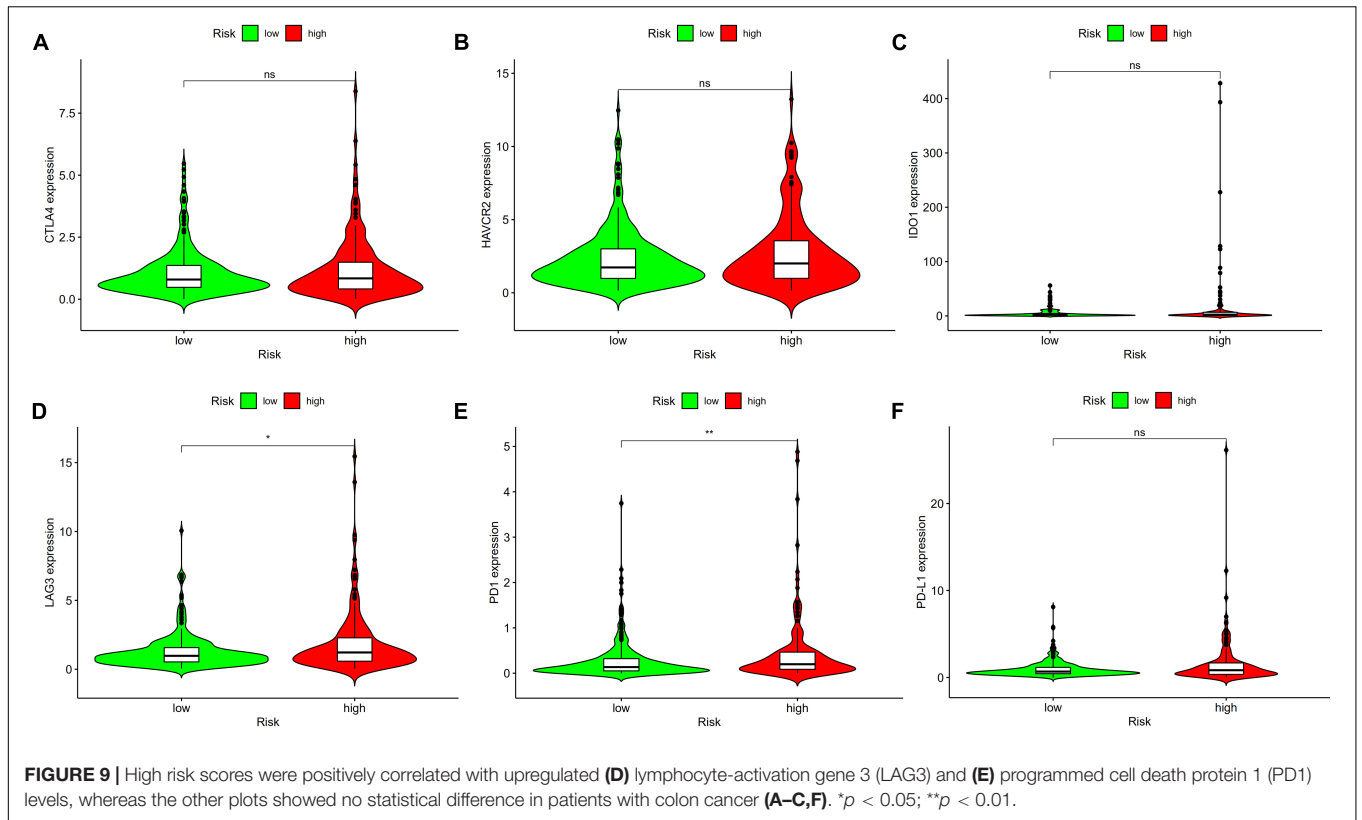
Tumor-infiltrating immune cells affect the response to tumor anti-checkpoint blockades. Tumor-infiltrating CD4⁺ T cells upregulated PD-1, cytotoxic T-lymphocyte-associated protein-4 (CTLA-4), T cell immunoglobulin and mucin domain-3 (TIM-3) and LAG-3 (Toor et al., 2019). Six common methods were used to evaluate the relationship between tumor infiltration immune cells and risk score, including XCELL (Aran et al., 2017), TIMER (Li et al., 2017), QUANTISEQ (Finotello et al., 2019), EPIC (Van Veldhoven et al., 2011), CIBERSORT-A (Tamminga et al., 2020), and CIBERSORT (Newman et al., 2015). The results revealed that the risk score was positively related to CD4⁺ T cells, neutrophils, and activated NK cells, and negatively related to hematopoietic stem cells and myeloid dendritic cells. A previous study showed that immune scores based on immune genomic analysis can indicate the therapeutic benefits

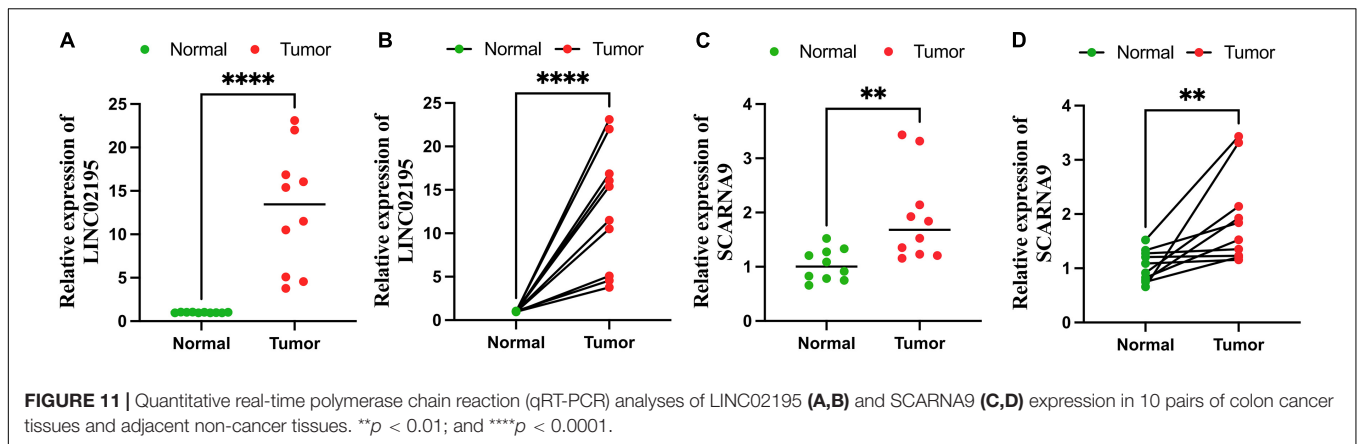


of immunotherapy and chemotherapy (Dai et al., 2020). To predict potential checkpoint blockade therapy, the relationship between risk score and immune checkpoint gene expression was explored and the results showed that PD-1 and LAG-3 expressions were significantly different in high- and low-risk groups. Anti-PD-1 inhibitor is effective in treating patients with dMMR/MSI-H mCRC (He et al., 2020). Xiao et al. (Xiao and Freeman, 2015) suggested that LAG-3 creates an immunosuppressive microenvironment in MSI-H CRC, possibly helping MSI-H tumors escape immune destruction by infiltrating immune cells. Our study also revealed that the risk score was associated with sensitivity to chemotherapeutics such as camptothecin, doxorubicin, erlotinib, gemcitabine, paclitaxel, and rapamycin. The findings our study can be applied for guiding clinical immunotherapy and chemotherapy in

patients with colon cancer. Among lrlncRNAs that were used to construct the signature, LINC02195 and SCARNA9 were found to be associated with neck squamous cell carcinoma and endometrial cancer in previous researches (Li H. et al., 2020; Wang et al., 2020), however, the effects of these two lrlncRNAs on colon cancer remains unknown. We further validated the expression levels of LINC02195 and SCARNA9 using qRT-PCR method, the results showed that LINC02195 and SCARNA9 were significantly upregulated in colon cancer compared with adjacent non-cancer tissues, indicating that the two lrlncRNAs may be potential biomarkers for diagnosis and therapy in colon cancer.

However, the study has some shortcomings and limitations. All data in our study were downloaded from TCGA database because we could not find other datasets that simultaneously





included lincRNA expression levels, clinicopathological characteristics, and survival outcomes for patients with colon cancer. The individual data source might result in unreliable results. Subsequent molecular biological experiments are needed to further examine the function of DElincRNAs in colon cancer development and to better understand carcinogenic mechanisms. In addition, clinical cases are required in further studies to increase stability and the predictive ability of our established signature.

CONCLUSION

In conclusion, our analysis of lincRNA expression profiles and clinical features identified DElincRNAs in colon cancer. lincRNA pairs were constructed and used for the development of prognosis signature, which did not require the exact expression quantities tested by certain methods. The constructed signature could effectively evaluate the prognosis of patients with colon cancer and guide clinical therapy. Additional studies are needed to validate the findings of this study and provide a basis for individualized treatment of patients with colon cancer.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Committee on Medical Ethics of The First Affiliated Hospital of Anhui Medical University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

XW and KC are responsible for writing and submitting the manuscript. ZW, BC, and YX are responsible for data collection and analysis. LD and TB are responsible for the production of pictures. WY and WC are responsible for the ideas and guidance. All authors have read and approved the final manuscript.

FUNDING

The Key Research and Development Plan Projects of Anhui Province (No. 201904a07020055).

ACKNOWLEDGMENTS

We acknowledged TCGA database for providing their platform and contributors for uploading their meaningful datasets. We would like to thank editage (www.editage.cn) for English language editing of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.750709/full#supplementary-material>

Supplementary Table 1 | Clinicopathologic characteristics of TCGA colon cancer patients.

Supplementary Table 2 | List of 90 DElincRNAs.

Supplementary Table 3 | Univariate Cox regression analysis of the 22 identified lincRNA pairs.

Supplementary Table 4 | Univariate Cox regression analysis of risk score and other clinical features.

Supplementary Table 5 | Multivariate Cox regression analysis of risk score and other clinical features.

Supplementary Table 6 | The correlation between risk score and immune infiltration of cells via the Spearman correlation method.

Supplementary Table 7 | Raw data of qRT-PCR.

REFERENCES

- Aran, D., Hu, Z., and Butte, A. J. (2017). xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 18:220. doi: 10.1186/s13059-017-1349-1
- Atianand, M. K., Caffrey, D. R., and Fitzgerald, K. A. (2017). Immunobiology of Long Noncoding RNAs. *Annu. Rev. Immunol.* 35, 177–198. doi: 10.1146/annurev-immunol-041015-055459
- Balachandran, V. P., Gonen, M., Smith, J. J., and DeMatteo, R. P. (2015). Nomograms in oncology: more than meets the eye. *Lancet. Oncol.* 16, e173–e180. doi: 10.1016/S1470-2045(14)71116-7
- Baldus, S. E., Mönig, S. P., Huxel, S., Landsberg, S., Hanisch, F. G., Engelmann, K., et al. (2004). MUC1 and nuclear beta-catenin are coexpressed at the invasion front of colorectal carcinomas and are both correlated with tumor prognosis. *Clin. Cancer Res.* 10, 2790–2796. doi: 10.1158/1078-0432.ccr-03-0163
- Bhandari, A., Woodhouse, M., and Gupta, S. (2017). Colorectal cancer is a leading cause of cancer incidence and mortality among adults younger than 50 years in the USA: a SEER-based analysis with comparison to other young-onset cancers. *J. Investig. Med.* 65, 311–315. doi: 10.1136/jim-2016-000229
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Cancer Genome Atlas Network (2012). Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487:330. doi: 10.1038/nature11252
- Cao, S., Zheng, J., Liu, X., Liu, Y., Ruan, X., Ma, J., et al. (2019). FXR1 promotes the malignant biological behavior of glioma cells via stabilizing MIR17HG. *J. Exp. Clin. Cancer Res.* 38, 1–22. doi: 10.1186/s13046-018-0991-0
- Chen, Y. G., Satpathy, A. T., and Chang, H. Y. (2017). Gene regulation in the immune system by long noncoding RNAs. *Nat. Immunol.* 18, 962–972. doi: 10.1038/ni.3771
- Crespo, M., Vilar, E., Tsai, S. Y., Chang, K., Amin, S., Srinivasan, T., et al. (2018). Corrigendum: colonic organoids derived from human induced pluripotent stem cells for modeling colorectal cancer and drug testing. *Nat. Med.* 24:526. doi: 10.1038/nm0418-526a
- Dai, G. P., Wang, L. P., Wen, Y. Q., Ren, X. Q., and Zuo, S. G. (2020). Identification of key genes for predicting colorectal cancer prognosis by integrated bioinformatics analysis. *Oncol. Lett.* 19, 388–398. doi: 10.3892/ol.2019.11068
- Doonan, B. B., Schaafsma, E., Pinto, J. T., Wu, J. M., and Hsieh, T. C. (2017). Application of open-access databases to determine functional connectivity between resveratrol-binding protein QR2 and colorectal carcinoma. *In Vitro Cell. Dev. Biol. Anim.* 53, 575–578. doi: 10.1007/s11626-017-0174-x
- Finotello, F., Mayer, C., Plattner, C., Laschober, G., Rieder, D., Hackl, H., et al. (2019). Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome Med.* 11:34. doi: 10.1186/s13073-019-0638-6
- He, S., Hu, D., Feng, H., Xue, Y., Jin, J., and Wang, X. (2020). Efficacy of immunotherapy with PD-1 inhibitor in colorectal cancer: a meta-analysis. *J. Comp. Eff. Res.* 9, 1285–1292. doi: 10.2217/cer-2020-0040
- Iasonos, A., Schrag, D., Raj, G. V., and Panageas, K. S. (2008). How to build and interpret a nomogram for cancer prognosis. *J. Clin. Oncol.* 26, 1364–1370. doi: 10.1200/JCO.2007.12.9791
- Ishida, Y., Agata, Y., Shibahara, K., and Honjo, T. (1992). Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *EMBO J.* 11, 3887–3895.
- Iyer, M. K., Niknafs, Y. S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., et al. (2015). The landscape of long noncoding RNAs in the human transcriptome. *Nat. Genet.* 47, 199–208. doi: 10.1038/ng.3192
- Joyce, J. A., and Pollard, J. W. (2009). Microenvironmental regulation of metastasis. *Nat. Rev. Cancer* 9, 239–252. doi: 10.1038/nrc2618
- Kouo, T., Huang, L., Pucsek, A. B., Cao, M., Solt, S., Armstrong, T., et al. (2015). Galectin-3 shapes antitumor immune responses by suppressing CD8+ T cells via LAG-3 and inhibiting expansion of plasmacytoid dendritic cells. *Immunol. Res.* 3, 412–423. doi: 10.1158/2326-6066.CIR-14-0150
- Lee, J. S., Adler, L., Karathia, H., Carmel, N., Rabinovich, S., Auslander, N., et al. (2018). Urea cycle dysregulation generates clinically relevant genomic and biochemical signatures. *Cell* 174, 1559–1570.e22. doi: 10.1016/j.cell.2018.07.019
- Li, H., Xiong, H. G., Xiao, Y., Yang, Q. C., Yang, S. C., Tang, H. C., et al. (2020). Long non-coding RNA LINC02195 as a regulator of MHC I molecules and favorable prognostic marker for head and neck squamous cell carcinoma. *Front. Oncol.* 10:615. doi: 10.3389/fonc.2020.00615
- Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* 77:e108. doi: 10.1158/0008-5472.CAN-17-0307
- Li, Z., Wang, D., and Yin, H. (2020). A seven immune-related lncRNA signature predicts the survival of patients with colon adenocarcinoma. *Am. J. Transl. Res.* 12, 7060–7078.
- Lin, Y., Pan, X., Chen, Z., Lin, S., and Chen, S. (2020). Identification of an immune-related nine-lncRNA signature predictive of overall survival in colon cancer. *Front. Genet.* 11:318. doi: 10.3389/fgene.2020.00318
- Liu, Z., Lu, T., Li, J., Wang, L., Xu, K., Dang, Q., et al. (2021). Development and clinical validation of a novel six-gene signature for accurately predicting the recurrence risk of patients with stage II/III colorectal cancer. *Cancer. Cell. Int.* 21:359. doi: 10.1186/s12935-021-02070-z
- Lv, Y., Lin, S. Y., Hu, F. F., Ye, Z., Zhang, Q., Wang, Y., et al. (2020). Landscape of cancer diagnostic biomarkers from specifically expressed genes. *Brief. Bioinform.* 21, 2175–2184. doi: 10.1093/bib/bbz131
- Ma, J. Y., Liu, S. H., Chen, J., and Liu, Q. (2021). Metabolism-related long non-coding RNAs (lncRNAs) as potential biomarkers for predicting risk of recurrence in breast cancer patients. *Bioengineered* 12, 3726–3736. doi: 10.1080/21655979.2021.1953216
- Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12, 453–457. doi: 10.1038/nmeth.3337
- Qi, B., Liu, H., Zhou, Q., Ji, L., Shi, X., Wei, Y., et al. (2021). An immune-related lncRNA signature for the prognosis of pancreatic adenocarcinoma. *Aging* 13, 18806–18826. doi: 10.18632/aging.203323
- Quail, D. F., and Joyce, J. A. (2013). Microenvironmental regulation of tumor progression and metastasis. *Nat. Med.* 19, 1423–1437. doi: 10.1038/nm.3394
- Russo, A. G., Andreano, A., Sartore-Bianchi, A., Mauri, G., Decarli, A., and Siena, S. (2019). Increased incidence of colon cancer among individuals younger than 50 years: a 17 years analysis from the cancer registry of the municipality of Milan, Italy. *Cancer Epidemiol.* 60, 134–140. doi: 10.1016/j.canep.2019.03.015
- Schmitt, A. M., and Chang, H. Y. (2016). Long noncoding RNAs in cancer pathways. *Cancer Cell* 29, 452–463. doi: 10.1016/j.ccell.2016.03.010
- Statello, L., Guo, C. J., Chen, L. L., and Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell. Biol.* 22, 96–118. doi: 10.1038/s41580-020-00315-9
- Tammaing, M., Hiltermann, T., Schuurung, E., Timens, W., Fehrmann, R. S., and Groen, H. J. (2020). Immune microenvironment composition in non-small cell lung cancer and its association with survival. *Clin. Transl. Immunol.* 9:e1142. doi: 10.1002/cti2.1142
- Toor, S. M., Murshed, K., Al-Dhaheri, M., Khawar, M., Abu Nada, M., and Elkord, E. (2019). Immune checkpoints in circulating and tumor-infiltrating CD4(+) T cell subsets in colorectal cancer patients. *Front. Immunol.* 10:2936. doi: 10.3389/fimmu.2019.02936
- Van Veldhoven, C. M., Khan, A. E., Teucher, B., Rohrmann, S., Raaschou-Nielsen, O., Tjønneland, A., et al. (2011). Physical activity and lymphoid neoplasms in the European Prospective Investigation into Cancer and nutrition (EPIC). *Eur. J. Cancer* 47, 748–760. doi: 10.1016/j.ejca.2010.11.010
- Wang, M., Zhao, J., Zhang, L., Wei, F., Lian, Y., Wu, Y., et al. (2017). Role of tumor microenvironment in tumorigenesis. *J. Cancer* 8, 761–773. doi: 10.7150/jca.17648
- Wang, R., Zheng, Z., Mao, S., Zhang, W., Liu, J., Li, C., et al. (2021). Construction and validation of a novel eight-gene risk signature to predict the progression and prognosis of bladder cancer. *Front. Oncol.* 11:632459. doi: 10.3389/fonc.2021.632459
- Wang, X., Dai, C., Ye, M., Wang, J., Lin, W., and Li, R. (2021). Prognostic value of an autophagy-related long-noncoding-RNA signature for endometrial cancer. *Aging* 13, 5104–5119. doi: 10.18632/aging.202431
- Wang, Z., Zhang, J., Liu, Y., Zhao, R., Zhou, X., and Wang, H. (2020). An integrated autophagy-related long noncoding RNA signature as a prognostic biomarker for human endometrial cancer: a bioinformatics-based approach. *Biomed. Res. Int.* 2020:5717498. doi: 10.1155/2020/5717498

- Wanitsuan, W., Kannurn, S., Boonpipattanapong, T., Sangthong, R., and Sangkhathat, S. (2008). Overall expression of beta-catenin outperforms its nuclear accumulation in predicting outcomes of colorectal cancers. *World. J. Gastroenterol.* 14, 6052–6059. doi: 10.3748/wjg.14.6052
- Wu, Q., Li, Q., Zhu, W., Zhang, X., and Li, H. (2021). Identification of autophagy-related long non-coding RNA prognostic signature for breast cancer. *J. Cell. Mol. Med.* 25, 4088–4098. doi: 10.1111/jcmm.16378
- Xiao, Y., and Freeman, G. J. (2015). The microsatellite instable subset of colorectal cancer is a particularly good candidate for checkpoint blockade immunotherapy. *Cancer Discov.* 5, 16–18. doi: 10.1158/2159-8290.CD-14-1397
- Xing, Y., Zhao, Z., Zhu, Y., Zhao, L., Zhu, A., and Piao, D. (2018). Comprehensive analysis of differential expression profiles of mRNAs and lincRNAs and identification of a 14-lincRNA prognostic signature for patients with colon adenocarcinoma. *Oncol. Rep.* 39, 2365–2375. doi: 10.3892/or.2018.6324
- Xu, J., Meng, Q., Li, X., Yang, H., Xu, J., Gao, N., et al. (2019). Long Noncoding RNA MIR17HG Promotes Colorectal Cancer Progression via miR-17-5p. *Cancer Res.* 79, 4882–4895. doi: 10.1158/0008-5472.CAN-18-3880
- Zhang, F., Wang, X., Hu, H., Yang, Y., Wang, J., Tang, Y., et al. (2021). A hypoxia related long non-coding RNA signature could accurately predict survival outcomes in patients with bladder cancer. *Bioengineered* 12, 3802–3823. doi: 10.1080/21655979.2021.1948781
- Zhou, P., Lu, Y., Zhang, Y., and Wang, L. (2021). Construction of an immune-related six-lincRNA signature to predict the outcomes, immune cell infiltration, and immunotherapy response in patients with hepatocellular carcinoma. *Front. Oncol.* 11:661758. doi: 10.3389/fonc.2021.661758

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Wang, Chen, Wang, Xu, Dai, Bai, Chen, Yang and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.