LncRNAs Signature Associated with Chemoradiotherapy Response and Prognosis in Locally Advanced Rectal Cancer

Yiyi Zhang
First Affiliated Hospital of Fujian Medical University
Binjie Guan
Fujian Medical University Union Hospital
Yong Wu
First Affiliated Hospital of Fujian Medical University
Fan Du
First Affiliated Hospital of Nanchang University
Jinfu Zhuang
First Affiliated Hospital of Fujian Medical University
Yuanfeng Yang
First Affiliated Hospital of Fujian Medical University
Xing Liu
First Affiliated Hospital of Fujian Medical University
Guoxian Guan (fjxhggx@163.com)
First Affiliated Hospital of Fujian Medical University

Research

Keywords: rectal cancer, chemoradiotherapy, lncRNA, prognosis

Posted Date: May 7th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-452503/v1

Abstract

Background

Long non-coding RNAs (lncRNAs) are promising diagnostic and prognostic biomarkers in cancers. Neoadjuvant chemoradiotherapy (NCRT) is the standard of care for patients with locally advanced rectal cancer (LARC). However, studies are limited regarding lncRNAs associated NCRT response and prognosis of LARC patients. This study aimed to identify lncRNAs associated with NCRT response and prognosis in CRC patients, and to explore potential mechanisms.

Methods

LncRNA expression profiles from our previous gene chip data basing on the LASSO to identify a four-lncRNA signature that predicted NCRT response and prognosis and further validated in 138 colorectal cancer (CRC) patients and 36 LARC patients from our center. A Cox regression model was performed to identify prognostic risk factors. Moreover, we identified the function of the LINC00909 in vivo and in vitro in CRC cell lines.

Results

Four hub lncRNAs (DBET, LINC00909, FLJ33534, and HSD52) were screened by comparing the relative lncRNA expression of NCRT-responsive and non-responsive patients (AUC = 0.68, 0.73, 0.73, and 0.70, respectively, all $p < 0.05$). A competing endogenous RNA (ceRNA) network was constructed based on the four lncRNAs. Moreover, the four lncRNAs expression was identified by the external data in cancerous and adjacent non-cancerous tissues in CRC patients. The results demonstrated that the expression of the four lncRNAs was lower in the normal tissues than in the cancerous tissues (all $p < 0.05$), and the COX analysis demonstrated that the DBET, LINC00909 and FLJ33534 were assocaited with the DFS in CRC patients. The four lncRNAs were also identified in the LARC following NCRT patients, and the result demonstrated that LINC00909 and FLJ33534 had powerful ability to predict the NCRT response and prognosis (all $p < 0.05$). Basing on the multivariate COX analysis, we constructed a risk score and verified in the CRC and LARC patients in predicting NCRT response and prognosis. Moreover, The expression and prognosis of the DBET, LINC00909 and FLJ33534 in the CRC tissues were identified in the R2 platform and Oncomine database. Moreover, the over-expression LINC00909 cell lines demonstrated that over-expressed the LINC00909 increased the cell lines resistance to the 5-FU and radiotherapy in vivo and in vitro.

Conclusion

Our findings showed that DBET, LINC00909, and FLJ33534 could serve as novel biomarkers for prediction of NCRT response and prognosis in CRC patients. And LINC00909 could be a novel therapeutic targets in enhancing the NCRT response.

1 Background

Preoperative neoadjuvant chemoradiotherapy (NCRT) and radical surgery have become the standard of care for locally advanced rectal cancer (LARC) patients [1]. The benefits of this multimodality therapy have been well documented, including tumor downsizing and downstaging, increased radical resection rate, and reduce local recurrence[2–4]. However, rectal cancer patients could show heterogeneous treatment responses to NCRT. Approximately 15–45% of rectal cancers would develop resistance to NCRT and can be exposed to NCRT-related toxicities without oncological benefits, and even treatment failure[5]. Therefore, the identification of valid biomarkers for resistance to NCRT has become imperative.

Long non-coding RNAs (lncRNAs) are transcripts longer than 200 nucleotides (nt) in length and are lack of protein-coding ability[6]. LncRNAs play critical roles in many biological processes by affecting transcriptional modulation, splicing regulation, and post-transcriptional process[7–9]. LncRNAs are also involved in the process of proliferation, invasion, progression, and metastasis of cancers, including CRC[10–13]. Recently, lncRNAs have been reported to act as diagnostic and prognostic biomarkers for several cancers [14–19], including CRC. Several studies reported that lncRNAs also can act as effective biomarkers to chemotherapy resistance in mCRC patients[20, 21]. Recently, Li et.al[22] reported that several effective biomarkers, including mRNAs and lncRNAs, can effectively predict NCRT response in LARC patients. However, studies regarding lncRNAs associated with resistance to NCRT are limited. In addition, the incorporation of multiple lncRNAs' expression is needed for improving the prediction accuracy of NCRT response and prognosis of LARC patients.

In this context, this study aimed to screen lncRNAs relevant to NCRT response using our previous gene expression profile. Then, the lncRNAs were verified in internal and external datasets containing patient tissue samples, and a risk factor model to predict disease free survival was built based on the Cox regression analysis. Finally, we identified the function of the powerful lncRNA, LINC00909, in vivo and in vitro.

2 Materials And Methods

2.1 Data preprocessing

Totally 31 LARC patients receiving preoperative NCRT and radical surgery between March 2016 to December 2016 in Fujian Medical University Union Hospital, China were enrolled in this study. The inclusion criteria, exclusion criteria, treatment protocols, and follow-up protocols were described in our previous study[23]. The raw data can be obtained from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/, GSE145037) and used as an internal data set for screening effective

lncRNAs. Moreover, a total of 138 CRC patients without preoperative therapy between January 2017 and December 2017 were used for building the risk score model and validating the lncRNAs expression in cancer and adjuvant cancer tissues, named as the risk score training dataset, and the samples were collected after surgery. And a total of 58 LARC patients who received NCRT from 2017 to 2017 were included for external validation of predictive efficiency, named as the external validation dataset, and the samples were collected at diagnosis by the colonoscopy. All patients provided written informed consent. The workflow of this study was shown in Fig. 1.

2.2 Screening the hub lncRNAs, Gene Ontology (GO) enrichment, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses

The methods of screening the hub lncRNAs in the microarray were as follows: to distinguish the function of the genes in the microarray, we download the human annotation file from the Ensemble database[24]. Then annotate the function of each gene from the microarray. Genes can be classified into two categories, including non-coding RNAs and protein-coding RNAs. A total of 241 lncRNAs were found in the microarray. The differential expressions of the lncRNAs were screened out by $P < 0.05$. Finally, basing on the FDR and P value, we selected the most relevant four lncRNAs including, DBET, LINC00909, FLJ33534, and HSD52. Moreover, based on the differential expression protein-coding genes and the hypothesis that lncRNAs directly interact with mRNA and regulate the activity of mRNAs by acting as miRNA sponges, a lncRNA–miRNA–mRNA ceRNA competing endogenous RNA (ceRNA) network of the above lncRNAs was constructed. First, the differential expression lncRNAs and mRNAs were selected from the microarray. Then, the lncRNA–miRNA and miRNA–mRNA interactions were predicted. Based on the miRcode online tool (http://www.mircode.org), the MiRDB (http://www.mirdb.org/), miRTarBase (http://mirtarbase.mbc.nctu.edu.tw//), and Targetscan (http://www.targetscan.org//), the miRNAs negatively regulated by lncRNAs and mRNAs were selected to construct the ceRNA network. In addition, the differentially expressed mRNAs (DEmRNAs) from the ceRNA network was selected to perform the KEGG and GO analysis.

2.3 Real-time quantitative polymerase chain reaction (RT-qPCR)

Total RNA from patient tissues were isolated using TRIzol reagent (Invitrogen) according to the manufacturer's instruction. And 1μg total RNA was used for reverse transcription reaction using M-MLV Reverse Transcriptase Product (Promega). RT-qPCR was performed using an ABI 7500 real-time PCR system (Applied Biosystems; Thermo Fisher Scientific, Inc., Foster City, CA, USA). LncRNA levels were assessed by RT-qPCR with GAPDH used as an internal control. PCR amplification was performed by denaturation at 94˚C for 5 seconds, annealing and extension at 62˚C for 40 seconds for 40 cycles. The relative expression level of lncRNAs was calculated using the ΔCt method. In brief, the difference value between GAPDH Ct value and lncRNA Ct value was defined as the ΔCt value, and the high ΔCt value was

recognized as the relatively low expression of the lncRNA in each sample. All PCR amplification was performed in triplicate and repeated in three independent experiments. The RT-qPCR analysis was performed using primers in supplementary Table 1.

2.4 Internal and external validation for the hub lncRNAs

We first verified the hub lncRNAs expression between the NCRT-resistant and -sensitive groups in the microarray data. Then, we evaluated the hub lncRNAs expression between cancerous tissues and adjacent non-cancerous tissues in the external data. Additionally, the expressions of hub lncRNAs were analyzed in patients receiving NCRT. The receiver operating characteristic (ROC) curve was plotted and the area under the ROC curve (AUC) was calculated to evaluate the predictive ability of the hub genes.

2.5 Overexpression of the LINC00909 with the lentivirus

In order to overexpressed LINC0090, the sequences of the LINC00909 and control (CON) were as follows, LINC00909, F: CTTTTTTGTTAGACGGATCCGAAGGACTTCCGGTGGCTTCCAAGG; R: AAAGATATTTTATTACCGGTTTAC. Lentiviral virions were produced by co-transfection of HEK293T cells with 5 μg pLKO.1-puro vector and 5 μg packaging and envelope vectors using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. Lentivirus was harvested 48 h after transfection. SW620 and DLD cells were infected with lentivirus containing ov-LINC00909 or OV-CON for 24 h. Two days later, the virus-infected cells were selected by 4 μg/ml puromycin (Sigma-Aldrich; Merck KGaA, Darmstadt, Germany) for 48 h and subjected to required assays.

2.6 Colony formation assay

Colony formation assay was carried out as described previously[25]. Briefly, Cells were plated in6-well plates (500 cells per plate) cultured for 14 days for 24 h before the addition of 4Gy radiotherapy cultured for 14 days, fixed with 4% paraformaldehyde for 15 min, stained with 1% crystal violet for 10 min before counting the number of colonies. The number of colonies with diameters of more than 1.5 mm was counted.

2.7 Cell resistance to the 5-FU

Anchorage-dependent cell growth was evaluated by a CCK-8 Kit (Dojindo Laboratories, Japan) according to the manufacturer's instructions. Cells were plated in 96-well plates at $3x10^3$ cells per well. When cells reached 60% confluence, the medium was removed and replaced with fresh medium containing varying concentrations of 5-FU, and then incubated for 48 h. The optical density was detected at 450 nm using a microplate reader, and the cell viability was calculated.

2.8 Tumor xenografts in the Rat

Male athymic nude mice (15-20g, 6–8 weeks of age) were purchased from SHANGHAI SLAC LABORATORY ANIMAL CO. LTD (China). Care and treatment of all experimental mice were carried out in accordance with institutional guidelines (No. 2019-0023). Tumor xenografts were established by subcutaneous injection of a 100 ul cells (DLD ov-LINC00909 groups VS. DLD con-LIN00909 group;

SW620 ov-LINC00909 groups VS. SW620 con-LIN00909 group) suspension (1×10^7 /ml), in each foreleg of nude mice. Then we measured the long diameter as the tumor size each weeks and performed for 4 weeks.

2.9 Statistical analysis

All statistical analyses were performed using SPSS software (version 23, SPSS Inc, Chicago, IL) and R software (version 3.4.1). The optimal cut-off values for lncRNAs expression were determined by using the X-tile program (http://www.tissuearray.org/rimmlab/). Survival outcomes were assessed using the Kaplan-Meier method and the log-rank test. A Cox proportional hazards model was performed to identify risk factors for disease-free survival (DFS)[26]. LASSO Cox regression model was applied to determine the ideal coefficient for each prognostic feature and estimate the likelihood deviance[27–30]. The corresponding risk scores for the samples from validation datasets were calculated using a risk score system. Based on cut-off values determined by ROC analysis, patients were divided into high-risk and low-risk groups. The entire patient cohort was divided into two subgroups according to patient outcomes (dead or alive). Then, ROC curves were plotted based on the risk scores and survival status. The risk score was selected as the cut-off value when the AUC reached its maximum. Kaplan-Meier and Cox regression analyses were performed to compare DFS risk between high-risk and low-risk groups. The performance of the model was evaluated by time-dependent ROC analysis. Decision curve analysis (DCA) was performed to evaluate the clinical utility of the model for disease recurrence. DCA is a method for evaluation and comparison of the predictive value between different prediction models[31, 32]; therefore, this method was used to evaluate the clinical utility of the model for disease recurrence. The x-axis of the DCA represents the percentage of threshold probability, and the y-axis represents the net benefit of the predictive model. The net benefit was calculated according to the following formula: Net benefit = (true positives/n) − (false positives/n) * (pt/(1 − pt). $P < 0.05$ was considered statistically significant.

3 Results

3.1 Cluster analysis, GO enrichment and KEGG analysis

The gene microarray was used to examine gene expression profiles in primary tumor cells. A total of 18419 genes were detected, including 241 lncRNAs. Supervised hierarchical cluster analysis of lncRNA expression profiling data showed a clustering trend between the two groups (Fig. 2A, 2B). Moreover, a total of 16 differentially expressed lncRNAs (DElncRNAs) were found in the two groups, and the expression of DElncRNAs was higher in the NCRT-resistant group ($P < 0.05$).

GO enrichment analysis was performed to investigate the molecular mechanism of DElncRNAs involved in NCRT response for LARC patients. As shown in Fig. 2C, the top three significant GO terms were related to the positive regulation of transcription from RNA polymerase II promoter, negative regulation of transcription from RNA polymerase II promoter and positive regulation of transcription, DNA-templated. KEGG pathway analysis demonstrated that the top three KEGG pathways were pathways in cancer, MAPK signaling pathway, and neurotrophin signaling pathway (Fig. 2D). Moreover, we selected the top 6

lncRNAs to construct a ceRNA network basing on the differential mRNAs in the gene chip (Fig. 2E). Moreover, the LASSO analysis was performed to explore the significant predictors for disease. The result demonstrated that the DBET, LINC00909, FLJ33534, and HSD52 were the significant factors (Fig. 2F and G).

3.2 The four lncRNAs validation in internal data

To validate the four lncRNAs expression in the internal data, we examined the DBET, LINC00909, FLJ33534, and HSD52 expression in rectal cancer tissues between NCRT-resistant and -sensitive cases in our microarray datasets. The relative expression of four lncRNAs were significantly increased in NCRT-resistant tissues (3.75 ± 0.27 vs.4.03 ± 0.47, P = 0.04; 1.62 ± 0.15 vs. 1.83 ± 0.0.28, P = 0.01; 1.56 ± 0.20 vs. 1.77 ± 0.35, P = 0.03; 1.07 ± 0.17 vs. 1.29 ± 0.43, P = 0.04; Fig. 3A). ROC curves demonstrated that the two lncRNAs, LINC00909 (AUC = 0.73, P = 0.03) and FLJ33534 (AUC = 0.73, P = 0.03), could efficiently differentiate NCRT-resistant from NCRT-sensitive rectal cancer cases; (Fig. 3B and E). Moreover, DBET (AUC = 0.68, P = 0.09) and HSD52 (AUC = 0.70, P = 0.06) also displayed well predictive efficiency in NCRT response (Fig. 3C and D). GO functional and KEGG pathway analyses were performed to investigate the molecular mechanism of hub lncRNAs, DBET, LINC00909, FLJ33534, and HSD52. The top three significant GO terms were related to the transcriptional activator activity, RNA polymerase II transcription factor binding, basolateral plasma membrane and integral component of membrane. KEGG pathway analysis demonstrated that the top three KEGG pathways were related to vasopressin-regulated water reabsorption, RAS signaling pathway, and glioma signaling pathway (Fig. 3F). Moreover, the four lncRNAs relevant ceRNA network was constructed basing on the differential mRNAs in the gene chip (Fig. 3G).

Figure3 verified the three lncRNAs in the R2 platform and Oncomine database. High expression of the DBET, LINC00909, and FLJ33534 were associated with worse prognosis in the TCGA, Sveen, and Marisa datasets (all P < 0.05). Moreover, the DBET, LINC00909, and FLJ33534 were higher expression in the CRC cancer tissues compared with adjutant-cancerous tissues by eight datasets meat-analysis in Oncomine database.</fig>

3.3 Hub genes validation in the external without preoperative therapy data

To independently validate the hub genes, we analyzed the expression level of the hub genes in the cancerous tissues and adjacent non-cancerous tissues using qPCR (Fig. 4A and B). The results demonstrated that the expression of DBET, LINC00909, FLJ33534, and HSD52 was higher in the cancerous tissues than in the adjacent non-cancerous tissues (The ΔCT value, DBET, 9.26 ± 2.59 vs. 11.91 ± 2.16, P < 0.001; LINC00909, 8.83 ± 2.21 vs. 10.47 ± 1.74, P < 0.001; FLJ33534, 12.36 ± 1.83 vs. 13.61 ± 2.04, P < 0.001; HSD52, 12.45 ± 2.08 vs. 13.05 ± 2.42, P = 0.04).

A total of 138 patients were included in the validation set. The clinicopathological characteristics of CRC patients are summarized in Table 1. As seen in Supplementary Fig. 1, X-tile plots identified 7.4, 10.6, 11.8, and 6.7 as cut-off values for DBET, LINC00909, FLJ33534, and HSD52, respectively. Based on the cut-off

points, we divided the entire cohort into low and high subgroups in terms of DFS. Lower expression of DBET and LINC00909 were associated with a better DFS and overall survival (OS) in CRC patients (both $P < 0.01$, Fig. 4C, D, G, and H). The higher expression of the FLJ33534 was associated with a worse DFS ($P < 0.01$, Fig. 4E) and overall survival (OS) ($P = 0.06$, Fig. 4I) in CRC patients. The high expression of the HSD52 was associated with worse prognosis but there is no statistic difference DFS ($P = 0.12$, Fig. 4F) and overall survival (OS) ($P = 0.09$, Fig. 4J) in CRC patients.

Table 1
Cox regression analysis of predictive factors for overall survival in CRC patients (n = 138)

Variables	Univariate analysis			Multivariate analysis		
	HR	95% CI	P value	HR	95% CI	P value
Sex, male/female	1.112	0.429–2.882	0.828			
Age	0.966	0.930–1.004	0.081			
ASA	1.092	0.471–2.530	0.837			
Tumor size	1.143	0.938–1.394	0.186			
Pathological T stage	1.621	0.695–3.780	0.264			
Pathological N stage	1.679	0.950–2.969	0.074			
Pathological M stage	13.670	5.218–35.813	**< 0.001**	4.441	1.518–12.996	**0.006**
Emergency operation	4.098	0.541–31.038	0.172			
Postoperative hospital stay	1.028	0.980–1.078	0.259			
Tumor location			0.768			
Ascending colon	Reference	Reference				
Transverse colon	0.373	0.042–3.337	0.378			
Descending colon	0.669	0.069–5.548	0.669			
Sigmoid	0.528	0.118–2.359	0.403			
Rectal	1.022	0.308–3.39\	0.971			
Risk score	2.549	1.745–3.723	**< 0.001**	2.110	1.396–3.186	**< 0.001**
Nerval invasion	1.092	0.3563.350	0.878			
Vascular invasion	1.066	0.243–4.669	0.933			
Tumor differentiation	0.376	0.086–1.643	0.193			
Histopathology			0.236			
Ulcerating	Reference	Reference				
Infiltrating	0.378	0.123–1.160	0.089			
Expanding	7.435	0.087–32.254	0.987			

HR, hazard ratio; CI, confidential interval; ASA: American Society of Anesthesiologists;

3.4 Hub genes validation in the external LARC patients

We further explored the hub lncRNAs expression in the NCRT-resistant and -sensitive patients (male 25 and female 11) who received NCRT before surgery. As shown in the Fig. 5A, the expression of DBET, LINC00909, FLJ33534, and HSD52 were higher in the NCRT-resistance group than in the responsive group (The ΔCT value, DBET, 8.96 ± 2.41 vs. 0.47 ± 2.79, P = 0.07; LINC00909, 10.70 ± 1.52 vs. 12.50 ± 1.51, P < 0.01; FLJ33534, 7.97 ± 2.15 vs. 9.31 ± 2.34, P = 0.04; HSD52, 11.09 ± 1.91 vs. 11.98 ± 1.90, P = 0.18; Fig. 5A). Furthermore, we analyzed the predictive ability of each hub lncRNA in patients receiving NCRT before surgery. The hub gene with the biggest predictive power was LINC00909 (AUC = 0.82, P < 0.01, Fig. 5C). The predictive ability of other lncRNAs, such as DBET (AUC = 0.65, P = 0.07), FLJ33534 (AUC = 0.67, P = 0.04), and HSD52 (AUC = 0.66, P = 0.06) were as show in Fig. 5B, D, and E. Moreover, we analysis the relationship between four lncRNAs and prognosis in LARC patients. As shown in the Fig. 5F-M, the result demonstrated that the high expression of the DBET, LINC00909, FLJ33534, and HSD52 were associated with the worse DFS in LARC following NCRT patients (P = 0.02, P < 0.02, P = 0.02, and P = 0.06). However, we cannot find the similarly result in the OS (P = 0.77, P = 0.33, P = 0.06, and P = 0.71).

3.5 Construction of a risk factor model and validation

To explore the prognostic impact of the hub lncRNAs on DFS in CRC patients, we performed a Cox regression analysis. As shown in the supplementary Table 2, the result demonstrated that ΔCT value DBET (HR = 0.676, 95%CI: 0.573–0.796, P < 0.001), ΔCT value LINC00909 (HR = 0.681, 95%CI: 0.543–0.854, P < 0.001) and ΔCT value FLJ33534 (HR = 0.759, 95%CI: 0.636–0.906, P = 0.002) were significantly associated with DFS in CRC patients. Basing on the COX regression analysis, the hub lncRNAs risk score system was built as follows: risk score = (0.39) × (ΔCt value of DBET) + (0.38) × (ΔCt value of LINC00909) + (0.28) × (ΔCt value of FLJ33534). Using the cutoff value of 0.89 for risk scores generated from ROC curves, the patients were divided into high-risk and low-risk groups (Fig. 6A). Patients in the low-risk group had an improved DFS and OS than those in the high-risk group (both log-rank P < 0.001, Fig. 6B and C). Moreover, the risk score was also identified in the LARC patients, and the result demonstrated that the risk score can also predict the prognosis, DFS (P < 0.01, Fig. 6E) and OS (P = 0.08, Fig. 6D) in the LARC patients.

As depicted in Fig. 6F, Time-dependent AUC curves showed that the LINC00909 had the most powerful predictive ability among the hub lncRNAs. The Cox model showed a stronger predictive ability to predict DFS for CRC patients than any single hub lncRNA. To further explore the predictive ability of the risk score in predicting the NCRT response, ROC curve analysis was performed in the LARC patients. The results demonstrated that the risk score had better predictive power compared with any hub gene (AUC = 0.75, P = 0.01, Fig. 6G).

3.6 Association of risk score with patient characteristics patients and prognosis in CRC

All patients (n = 138) were equally divided into the low-risk score group (n = 68) and the high-risk score group (n = 68). A higher pathology M stage (P = 0.017) and nerval invasion (P = 0.040) were found in the high-risk group. No statistical differences were observed between two groups in terms of gender, age, American Society of Anaesthesiology (ASA) grade, tumor location, histopathology, tumor differentiation pathology T stage, pathology N stage, postoperative hospital stay (days), lymph nodes retrieved, metastatic lymph nodes, and tumor size, as shown in supplementary Table 3.

To further determine the prognostic factors in CRC patients, COX regression analysis was performed. On univariate analysis, higher pathological M stage (HR = 13.670, P < 0.001), and higher risk score (HR = 2.549, P < 0.001) were independently associated with OS in CRC patients. Multivariate Cox regression analysis demonstrated that higher pathological M stage (HR = 4.441, P = 0.006), and higher risk score (HR = 2.110, P < 0.001) remained significantly associated with OS, as demonstrated in Table 1.

On univariate analysis, higher pathological N stage (HR = 2.465, P < 0.001), vascular invasion (HR = 2.387, P = 0.040), higher pathological T stage (HR = 2.348, P = 0.008), and higher risk score (HR = 2.625, P < 0.001) were independently associated with DFS in CRC patients. Cox regression analysis demonstrated that higher risk score (HR = 1.224, P < 0.001) and higher pathological N stage (HR = 2.128, P = 0.001), remained significantly associated with increased risk of local recurrence, as demonstrated in Table 2.

Table 2
Cox regression analysis of predictive factors for disease-free survival in CRC patients (n = 138)

Variables	Univariate analysis			Multivariate analysis		
	HR	**95% CI**	P **value**	**HR**	**95% CI**	P **value**
Sex, male/female	1.527	0.785–2.970	0.212			
Age	0.988	0.961–1.017	0.414			
ASA	0.735	0.392–1.378	0.337			
Tumor size	1.075	0.928–1.246	0.337			
Pathological T stage	2.348	1.253–4.398	**0.008**	1.367	0.772–2.420	0.283
Pathological N stage	2.465	1.667–3.645	**< 0.001**	2.128	1.362–3.326	**0.001**
Emergency operation	2.749	0.375–20.154	0.320			
Postoperative hospital stay	1.006	0.963–1.051	0.790			
Tumor location			0.746			
Ascending colon	Reference	Reference				
Transverse colon	0.364	0.077–1.716	0.202			
Descending colon	0.573	0.122-2.700	0.482			
Sigmoid	0.775	0.299–2.008	0.599			
Rectal	1.312	0.571–3.013	0740			
Risk score	2.625	2.005–3.436	**< 0.001**	2.880	2.127–3.901	**< 0.001**
Nerval invasion	1.705	0.835–3.482	0.143			
Vascular invasion	2.387	1.039–5.484	**0.040**	2.406	0.963–6.012	0.060
Tumor differentiation	1.056	0.495–2.254	0.888			
Histopathology			0.057			
Ulcerating	Reference	Reference				
Infiltrating	0.588	0.286–1.207	0.148			
Expanding	6.001	0.791–45.523	0.083			

HR, hazard ratio; CI, confidential interval; ASA: American Society of Anesthesiologists;

3.7 The clinical application of the risk score

We further explored the association between the risk score and clinicopathological parameters. In the early pathology stage (stage 0-II), we found that the low-risk score group had better DFS and OS compared with the low-risk score group ($P < 0.01$, Fig. 7A and B). In the advanced pathology stage (stage III-IV), the low-risk group had a better prognosis compared with the high-risk group (all $P = 0.01$, Fig. 7C and D). DCA was used to evaluate the performance of the risk score. As shown in Fig. 7E, the risk score provided more benefit than either lncRNAs in the disease-free scheme or the disease recurrence scheme. The clinical impact curve (Fig. 7F) showed the prediction of risk stratification of 1,000 patients using a resampling bootstrap method. “Number high risk” indicated the number of patients classified as positive (high risk) by the risk score according to various threshold probabilities. “Number high risk with the event” was the true positive patient number according to various threshold probabilities.

3.8 The lncRNAs validate in the R2 platform and Oncomine database

Basing on the above result, the DBET, LINC00909 and FLJ33534 were further enrolled to verify in the external database. We verified the three lncRNAs in the R2 platform and Oncomine database in the CRC tissues. The result was shown in the supplemental Fig. 1, the high expression of the DBET, LINC00909 and FLJ33534 were associated with the worse prognosis in the three independent database in the R2 platform, which were similarly with our result. In the Oncomine database, basing the eight dates meta-analysis we can find the DBET, LINC00909 and FLJ33534 were high expression in the CRC tissues compared with adjutant-cancerous tissues. The above result supporting our result that the DBET, LINC00909 and FLJ33534 acted as the oncogene in the CRC patients.

3.9 Overexpression of the LINC00909 associated with the NCRT resistance in vivo and in vitro

Basing on the previous result, we found that the LINC00909 was the most powerful lncRNAs in predicting the NCRT response and prognosis in CRC patients. To further identified the function of the LINC00909 in CRC cell lines, we constructed two LINC00909 overexpression CRC cell lines. As shown in Fig. 8A, we successful constructed the LINC00909 overexpressed cell lines, DLD-over-LINC00909 and SW620-over-LINC00909 (all $P < 0.01$). Moreover, we detected the LINC00909 overexpression CRC cell lines resistance to the NCRT. The result demonstrated that the IC50 of the DLD-CON group was 112.80 ± 20.76 ug/ml to 5-FU, and the DLD-over group was 1104.74 ± 50.74 ug/ml to 5-FU; the IC50 of the SW620-CON group was 94.89 ± 9.887 ug/ml to 5-FU, and the SW620-over group was 845.62 ± 35.24 ug/ml to 5-FU (Fig. 8B and C). Moreover, we analysis the sensitive of the LINC00909 overexpression cell lines to the 4Gy radiotherapy(Fig. 8D and E). The result demonstrated that the colony number of DLD-CON group was 19.33 ± 10.12, and the colony number of DLD-over group was 62.33 ± 15.04 ($P < 0.01$); the colony number of DLD-CON group was 21.33 ± 11.37, and the colony number of DLD-over group was 101.67 ± 25.03 ($P <$

0.01). The tumor xenografts was performed to explored the LINC00909 function in vivo experiment. As shown in Fig. 9F-J, the tumor size of the two groups were verified the LINC00909 as the oncogene in vivo (all $P < 0.01$).

4 Discussion

In this study, four lncRNAs (DBET, LINC00909, FLJ33534, and HSD52) were identified, and DBET, LINC00909, FLJ33534 were validated as hub genes correlated with NCRT response and prognosis in CRC patients. A three-lncRNA based risk model was constructed to predict NCRT and prognosis of CRC patients. Moreover, we verified the function of the LINC00909 in CRC in vivo and in vitro experiment.

LncRNAs have been reported to act as potent biomarkers for diagnosis and prediction of the prognosis, progression in the CRC patients[33–35]. Li et.al revealed that several effective biomarkers, KRAS, PDPK1, PPP2R5C, PPP2R1B, and YES1 which including mRNAs and lncRNAs, can effectively predict the NCRT response in the LARC[22]. However, the predictive effect of lncRNAs and a predicting model basing on lncRNAs in NCRT resistance are still unclear. To explore the role of the lncRNAs in the LARC patients receiving NCRT, we re-analyzed the microarray and classified the genes basing on the coding function in the transcriptome. Then, basing on the LASSO analysis we selected the most effective four lncRNAs, DBET, LINC00909, FLJ33534, and HSD52, in our microarray data to predict the NCRT response and prognosis in CRC patients. And the DBET, LINC00909, and FLJ33534 have a well predictive power to predict the NCRT response in both internal and external data sets. The results indicated that the hub lncRNAs were effective biomarkers in predicting NCRT response and prognosis.

In the previous studies, the above hub lncRNAs had already been reported in several areas. LINC00909 had already reported in the human glioma, and LINC00909 could act as an oncogenic lncRNA in glioma tumorigenesis[36]. Moreover, in the Xu et. al study[37], the high expression of serum LINC00909 could serve as an effective diagnostic biomarker for CRC. HSD52 gene expression is associated with body mass index (BMI) in obese Korean women including overweight patients[38]. In addition, Ahmad et. al[39] demonstrated that FLJ33534 has an intronic variant, rs140133294, in association with BMI variance. Currently, no studies have investigated the biological functions of the DBET gene. To further explore the hub lncRNAs involved in the function and mechanism in NCRT patients, Pearson analysis and a ceRNA network were constructed to select relevant mRNAs[40]. The ceRNA network was constructed based on the hypothesis that lncRNAs directly interact with miRNAs and regulate the activity of mRNAs by acting as miRNA sponges[41]. Based on the associated mRNA, the GO enrichment and KEGG analysis were performed, the results demonstrated that the hub lncRNAs involve in the RAS signaling pathway and transcriptional activator activity, which had been reported in the previous studies[42–44]. Moreover, we successfully constructed the LINC00909 overexpression CRC cell lines to verified that overexpression LINC00909 could enhance the resistance to the NCRT in CRC. And the result of the in vivo and in vitro was according with our microarray and qPCR result.

The microarray as a useful tool to selected the hub lncRNAs. But there is no defining that the result from the microarray may contain several false-positive lncRNAs which resulting the lncRNAs cannot be verified in the external dataset. Thus, to further verify the hub lncRNAs screened by the microarray profiling, we examined the hub lncRNAs expression in the cancerous and adjacent non-cancerous tissues in 138 CRC patients. The results demonstrated that the three lncRNAs acted as the oncogenes in the CRC and high expression of the three lncRNAs was associated with a shorter DFS.

The risk factor model has been utilized for prognostication in several tumors, such as liver, lung, and colon cancers[10, 45, 46]. However, to our best knowledge, no study has focused on NCRT response and prognosis in CRC patients. In the present study, we successfully constructed a risk factor model based on a three-lncRNA signature that had a powerful ability in predicting the prognosis of CRC patients. Moreover, the time-dependent ROC curve demonstrated the risk factor model had the best AUC value than each of the single lncRNAs in predicting DFS in CRC patients. Additionally, to further explore the relationship between three lncRNAs and NCRT response in the CRC patients, we screened out 36 LARC patients who received the NCRT before surgery in an external data set. The results of ROC analysis revealed that the LINC00909, FLJ33534, and the risk factor score had a powerful predictive ability to predict NCRT response in LARC patients. Moreover, in both the internal and external data sets, the risk score model had a better predictive power than any single lncRNA. In summary, the risk factor model based on the four lncRNAs had a strong predictive ability in predicting prognosis and NCRT response in CRC patients.

In clinical practice, a patient's prognosis is usually influenced by a variety of clinical factors. DCA as a useful tool can assist the clinical decision. In the present study, we also found that the risk score basing on the four lncRNAs had a better benefit to estimate the patients' disease recurrence rate. Thus, we employed the risk score to analyze the patient's prognosis in CRC patients. The pathological TNM stage has been considered as the most useful factor to predict the prognosis of CRC patients in many studies[47, 48]. As we all known, ypCR-Stage II patients are associated with a better prognosis, such patients without risk factors are not recommended for postoperative chemotherapy according to the current NCCN guidelines[49]. However, the recurrence rates of the early-stage patients were 83.1–88.7% [50, 51]. Thus, screening out the above patients was an urgent task. In the present study, we analyzed the prognosis of the early-stage patients and the results demonstrated that a higher risk score was associated with a higher risk of disease recurrence in CRC patients. To sum up, the risk score model not only acted as an effective predictive biomarker in CRC patients but also can distinguish the ypCR-Stage II CRC patients who had disease recurrence risk.

There were some limitations to the current study. Firstly, the small gene chip sample size was a major limitation of our study. Due to the study design, we included LARC patients at diagnosis who did not receive any treatment before biopsy from a colonoscopy, which limited the sample size of our study. We will continue to expand our sample size in our future studies. Secondly, the involved pathways of hub lncRNAs were conducted by lncRNAs microarray profiling and bioinformatics methods, and they need to be further validated by *in vitro* and *in vivo* experimental studies in future research.

5 Conclusion

In conclusion, we identified and validated the three hub lncRNAs as new effective predictors for NCRT and a prognostic factor for CRC patients. Moreover, based on the three lncRNAs, we constructed a risk factor model that had a strong power to predict NCRT and prognosis in CRC patients. These results may help to discriminate CRC patients who are candidates for NCRT. Moreover, the risk score can distinguish the ypCR-Stage II patients CRC patients who had higher disease recurrence rate and the early-stage patients with a high risk score will be considered for postoperative chemotherapy. Moreover, we identified the function of the most powerful lncRNA, LINC00909, in resistance to the NCRT. Nevertheless, more insightful molecular mechanisms are warranted in future studies.

Declarations

Ethics approval and consent to participate

This study was carried out in accordance with the committee of Fujian Medical University Union Hospital with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by committee of the Fujian Medical University Union Hospital.

Consent for publication

Not applicable

Availability of data and material

I confirm that I have included a citation for available data in my references section and the data that support the findings of this study are available from the corresponding author (GXG) upon reasonable request.

Competing interests

None declared

Funding

This study was supported by the Science Foundation of the Fujian Province, (No. 2016J01602; 2019J0105), Special Financial Foundation of Fujian Provincial (No.2015-1297; 2020B1050), Young and middle-aged backbone training project in the health system of Fujian province (2016-ZQN-26), the Startup Fund for Scientific Research, Fujian Medical University (2017XQ1029, 2018QH2027), and the Professor Development Foundation of Fujian Medical University (No.JS11006).

Author Contributions

YYZ, BJG, YW, FD, XL and GXG designed the experiments, performed the experiments, analyzed the data, and wrote the paper. JFZ and YFY performed the experiments. All authors read and approved the final manuscript.

Acknowledgement

The authors would like to express our sincere thanks for sharing the data from The Gene Expression Omnibus (GEO) database.

References

1. Heald RJ, Husband EM, Ryall RD. The mesorectum in rectal cancer surgery--the clue to pelvic recurrence. Br J Surg. 1982;69:613-6.
2. Sauer R, Becker H, Hohenberger W, Rödel C, Wittekind C, Fietkau R, Martus P, Tschmelitsch J, Hager E, Hess CF, et al. Preoperative versus postoperative chemoradiotherapy for rectal cancer. N Engl J Med. 2004;351:1731-40.
3. Sauer R, Liersch T, Merkel S, Fietkau R, Hohenberger W, Hess C, Becker H, Raab HR, Villanueva MT, Witzigmann H, et al. Preoperative versus postoperative chemoradiotherapy for locally advanced rectal cancer: results of the German CAO/ARO/AIO-94 randomized phase III trial after a median follow-up of 11 years. J Clin Oncol. 2012;30:1926-33.
4. van Gijn W, Marijnen CA, Nagtegaal ID, Kranenbarg EM, Putter H, Wiggers T, Rutten HJ, Påhlman L, Glimelius B, van de Velde CJ. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer: 12-year follow-up of the multicentre, randomised controlled TME trial. Lancet Oncol. 2011;12:575-82.
5. Ha YJ, Tak KH, Kim CW, Roh SA, Choi EK, Cho DH, Kim JH, Kim SK, Kim SY, Kim YS, et al. PSMB8 as a Candidate Marker of Responsiveness to Preoperative Radiation Therapy in Rectal Cancer Patients. Int J Radiat Oncol Biol Phys. 2017;98:1164-73.
6. Ling H, Fabbri M, Calin GA. MicroRNAs and other non-coding RNAs as targets for anticancer drug development. Nat Rev Drug Discov. 2013;12:847-65.
7. Anderson DM, Anderson KM, Chang CL, Makarewich CA, Nelson BR, McAnally JR, Kasaragod P, Shelton JM, Liou J, Bassel-Duby R, et al. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. Cell. 2015;160:595-606.
8. Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. Nat Rev Genet. 2014;15:7-21.
9. Nelson BR, Makarewich CA, Anderson DM, Winders BR, Troupes CD, Wu F, Reese AL, McAnally JR, Chen X, Kavalali ET, et al. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. Science. 2016;351:271-5.
10. Dai W, Feng Y, Mo S, Xiang W, Li Q, Wang R, Xu Y, Cai G. Transcriptome profiling reveals an integrated mRNA-lncRNA signature with predictive value of early relapse in colon cancer. Carcinogenesis.

2018;39:1235-44.

11. Fernández-Barrena MG, Perugorria MJ, Banales JM. Novel lncRNA T-UCR as a potential downstream driver of the Wnt/β-catenin pathway in hepatobiliary carcinogenesis. Gut. 2017;66:1177-8.
12. Li W, Zhang Z, Liu X, Cheng X, Zhang Y, Han X, Zhang Y, Liu S, Yang J, Xu B, et al. The FOXN3-NEAT1-SIN3A repressor complex promotes progression of hormonally responsive breast cancer. J Clin Invest. 2017;127:3421-40.
13. Shi L, Hong X, Ba L, He X, Xiong Y, Ding Q, Yang S, Peng G. Long non-coding RNA ZNFX1-AS1 promotes the tumor progression and metastasis of colorectal cancer by acting as a competing endogenous RNA of miR-144 to regulate EZH2 expression. Cell Death Dis. 2019;10:150.
14. Ali MM, Akhade VS, Kosalai ST, Subhash S, Statello L, Meryet-Figuiere M, Abrahamsson J, Mondal T, Kanduri C. PAN-cancer analysis of S-phase enriched lncRNAs identifies oncogenic drivers and biomarkers. Nat Commun. 2018;9:883.
15. Casero D, Sandoval S, Seet CS, Scholes J, Zhu Y, Ha VL, Luong A, Parekh C, Crooks GM. Long non-coding RNA profiling of human lymphoid progenitor cells reveals transcriptional divergence of B cell and T cell lineages. Nat Immunol. 2015;16:1282-91.
16. Fu X, Ravindranath L, Tran N, Petrovics G, Srivastava S. Regulation of apoptosis by a prostate-specific and prostate cancer-associated noncoding gene, PCGEM1. DNA Cell Biol. 2006;25:135-41.
17. Jiang Z, Slater CM, Zhou Y, Devarajan K, Ruth KJ, Li Y, Cai KQ, Daly M, Chen X. LincIN, a novel NF90-binding long non-coding RNA, is overexpressed in advanced breast tumors and involved in metastasis. Breast Cancer Res. 2017;19:62.
18. Kurian L, Aguirre A, Sancho-Martinez I, Benner C, Hishida T, Nguyen TB, Reddy P, Nivet E, Krause MN, Nelles DA, et al. Identification of novel long noncoding RNAs underlying vertebrate cardiovascular development. Circulation. 2015;131:1278-90.
19. Sánchez Y, Huarte M. Long non-coding RNAs: challenges for diagnosis and therapies. Nucleic Acid Ther. 2013;23:15-20.
20. Li P, Zhang X, Wang H, Wang L, Liu T, Du L, Yang Y, Wang C. MALAT1 Is Associated with Poor Response to Oxaliplatin-Based Chemotherapy in Colorectal Cancer Patients and Promotes Chemoresistance through EZH2. Mol Cancer Ther. 2017;16:739-51.
21. Li Z, Chen Y, Ren WU, Hu S, Tan Z, Wang Y, Chen Y, Zhang J, Wu J, Li T, et al. Transcriptome Alterations in Liver Metastases of Colorectal Cancer After Acquired Resistance to Cetuximab. Cancer Genomics Proteomics. 2019;16:207-19.
22. Li N, Yu J, Luo A, Tang Y, Liu W, Wang S, Liu Y, Song Y, Fang H, Chen B, et al. LncRNA and mRNA signatures associated with neoadjuvant chemoradiotherapy downstaging effects in rectal cancer. J Cell Biochem. 2019;120:5207-17.
23. Zhang Y, Sun Y, Xu Z, Chi P, Lu X. Is neoadjuvant chemoradiotherapy always necessary for mid/high local advanced rectal cancer: A comparative analysis after propensity score matching. Eur J Surg Oncol. 2017;43:1440-6.

24. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, Billis K, Cummins C, Gall A, Girón CG, et al. Ensembl 2018. Nucleic Acids Res. 2018;46:D754-754D761.
25. Zhang Y, Sun L, Sun Y, Chen Y, Wang X, Xu M, Chi P, Xu Z, Lu X. Overexpressed CES2 has prognostic value in CRC and knockdown CES2 reverses L-OHP-resistance in CRC cells by inhibition of the PI3K signaling pathway. Exp Cell Res. 2020;389:111856.
26. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. 2010;33:1-22.
27. Agesen TH, Sveen A, Merok MA, Lind GE, Nesbakken A, Skotheim RI, Lothe RA. ColoGuideEx: a robust gene classifier specific for stage II colorectal cancer prognosis. Gut. 2012;61:1560-7.
28. Zhang JX, Song W, Chen ZH, Wei JH, Liao YJ, Lei J, Hu M, Chen GZ, Liao B, Lu J, et al. Prognostic and predictive value of a microRNA signature in stage II colon cancer: a microRNA expression analysis. Lancet Oncol. 2013;14:1295-306.
29. Jiang Y, Zhang Q, Hu Y, Li T, Yu J, Zhao L, Ye G, Deng H, Mou T, Cai S, et al. ImmunoScore Signature: A Prognostic and Predictive Tool in Gastric Cancer. Ann Surg. 2018;267:504-13.
30. Sveen A, Ågesen TH, Nesbakken A, Meling GI, Rognum TO, Liestøl K, Skotheim RI, Lothe RA. ColoGuidePro: a prognostic 7-gene expression signature for stage III colorectal cancer patients. Clin Cancer Res. 2012;18:6001-10.
31. Rousson V, Zumbrunn T. Decision curve analysis revisited: overall net benefit, relationships to ROC curve analysis, and application to case-control studies. BMC Med Inform Decis Mak. 2011;11:45.
32. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006;26:565-74.
33. Alidoust M, Hamzehzadeh L, Rivandi M, Pasdar A. Polymorphisms in non-coding RNAs and risk of colorectal cancer: A systematic review and meta-analysis. Crit Rev Oncol Hematol. 2018;132:100-10.
34. Pichler M, Rodriguez-Aguayo C, Nam SY, Dragomir MP, Bayraktar R, Anfossi S, Knutsen E, Ivan C, Fuentes-Mattei E, Lee SK, et al. Therapeutic potential of FLANC, a novel primate-specific long non-coding RNA in colorectal cancer. Gut. 2020.
35. Wei L, Wang X, Lv L, Zheng Y, Zhang N, Yang M. The emerging role of noncoding RNAs in colorectal cancer chemoresistance. Cell Oncol (Dordr). 2019;42:757-68.
36. Liu Z, Lu C, Hu H, Cai Z, Liang Q, Sun W, Jiang L, Hu G. LINC00909 promotes tumor progression in human glioma through regulation of miR-194/MUC1-C axis. Biomed Pharmacother. 2019;116:108965.
37. Xu W, Zhou G, Wang H, Liu Y, Chen B, Chen W, Lin C, Wu S, Gong A, Xu M. Circulating lncRNA SNHG11 as a novel biomarker for early diagnosis and prognosis of colorectal cancer. Int J Cancer. 2019.
38. Lee M, Kwon DY, Kim MS, Choi CR, Park MY, Kim AJ. Genome-wide association study for the interaction between BMR and BMI in obese Korean women including overweight. Nutr Res Pract. 2016;10:115-24.

39. Ahmad S, Zhao W, Renström F, Rasheed A, Zaidi M, Samuel M, Shah N, Mallick NH, Shungin D, Zaman KS, et al. A novel interaction between the FLJ33534 locus and smoking in obesity: a genome-wide study of 14 131 Pakistani adults. Int J Obes (Lond). 2016;40:186-90.
40. Kopp F, Mendell JT. Functional Classification and Experimental Dissection of Long Noncoding RNAs. Cell. 2018;172:393-407.
41. Fan CN, Ma L, Liu N. Systematic analysis of lncRNA-miRNA-mRNA competing endogenous RNA network identifies four-lncRNA signature as a prognostic biomarker for breast cancer. J Transl Med. 2018;16:264.
42. Bando H, Kagawa Y, Kato T, Akagi K, Denda T, Nishina T, Komatsu Y, Oki E, Kudo T, Kumamoto H, et al. Correction: A multicentre, prospective study of plasma circulating tumour DNA test for detecting RAS mutation in patients with metastatic colorectal cancer. Br J Cancer. 2020.
43. Samalin E, Mazard T, Assenat E, Rouyer M, De la Fouchardière C, Guimbaud R, Smith D, Portales F, Ychou M, Fiess C, et al. Triplet chemotherapy plus cetuximab as first-line treatment in RAS wild-type metastatic colorectal carcinoma patients. Ann Oncol. 2019;30 Suppl 4:iv26-26iv27.
44. Vera R, Mata E, González E, Juez I, Alonso V, Iranzo P, Martínez NP, López C, Cabrera JM, Safont MJ, et al. Is aflibercept an optimal treatment for wt RAS mCRC patients after progression to first line containing anti-EGFR. Int J Colorectal Dis. 2020.
45. Gu J, Zhang X, Miao R, Ma X, Xiang X, Fu Y, Liu C, Niu W, Qu K. A three-long non-coding RNA-expression-based risk score system can better predict both overall and recurrence-free survival in patients with small hepatocellular carcinoma. Aging (Albany NY). 2018;10:1627-39.
46. Liao M, Liu Q, Li B, Liao W, Xie W, Zhang Y. A group of long noncoding RNAs identified by data mining can predict the prognosis of lung adenocarcinoma. Cancer Sci. 2018;109:4033-44.
47. Liu Q, Luo D, Cai S, Li Q, Li X. Real-World Implications of Nonbiological Factors with Staging, Prognosis and Clinical Management in Colon Cancer. Cancers (Basel). 2018;10.
48. Zhou R, Zeng D, Zhang J, Sun H, Wu J, Li N, Liang L, Shi M, Bin J, Liao Y, et al. A robust panel based on tumour microenvironment genes for prognostic prediction and tailoring therapies in stage I-III colon cancer. EBioMedicine. 2019;42:420-30.
49. Benson AB, Venook AP, Cederquist L, Chan E, Chen YJ, Cooper HS, Deming D, Engstrom PF, Enzinger PC, Fichera A, et al. Colon Cancer, Version 1.2017, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw. 2017;15:370-98.
50. Ackermann CJ, Guller U, Jochum W, Schmied BM, Warschkow R. The prognostic value of signet ring cell histology in stage I/II colon cancer-a population-based, propensity score-matched analysis. Int J Colorectal Dis. 2018;33:1183-93.
51. Low YS, Blöcker C, McPherson JR, Tang SA, Cheng YY, Wong J, Chua C, Lim T, Tang CL, Chew MH, et al. A formalin-fixed paraffin-embedded (FFPE)-based prognostic signature to predict metastasis in clinically low risk stage I/II microsatellite stable colorectal cancer. Cancer Lett. 2017;403:13-20.

Figures

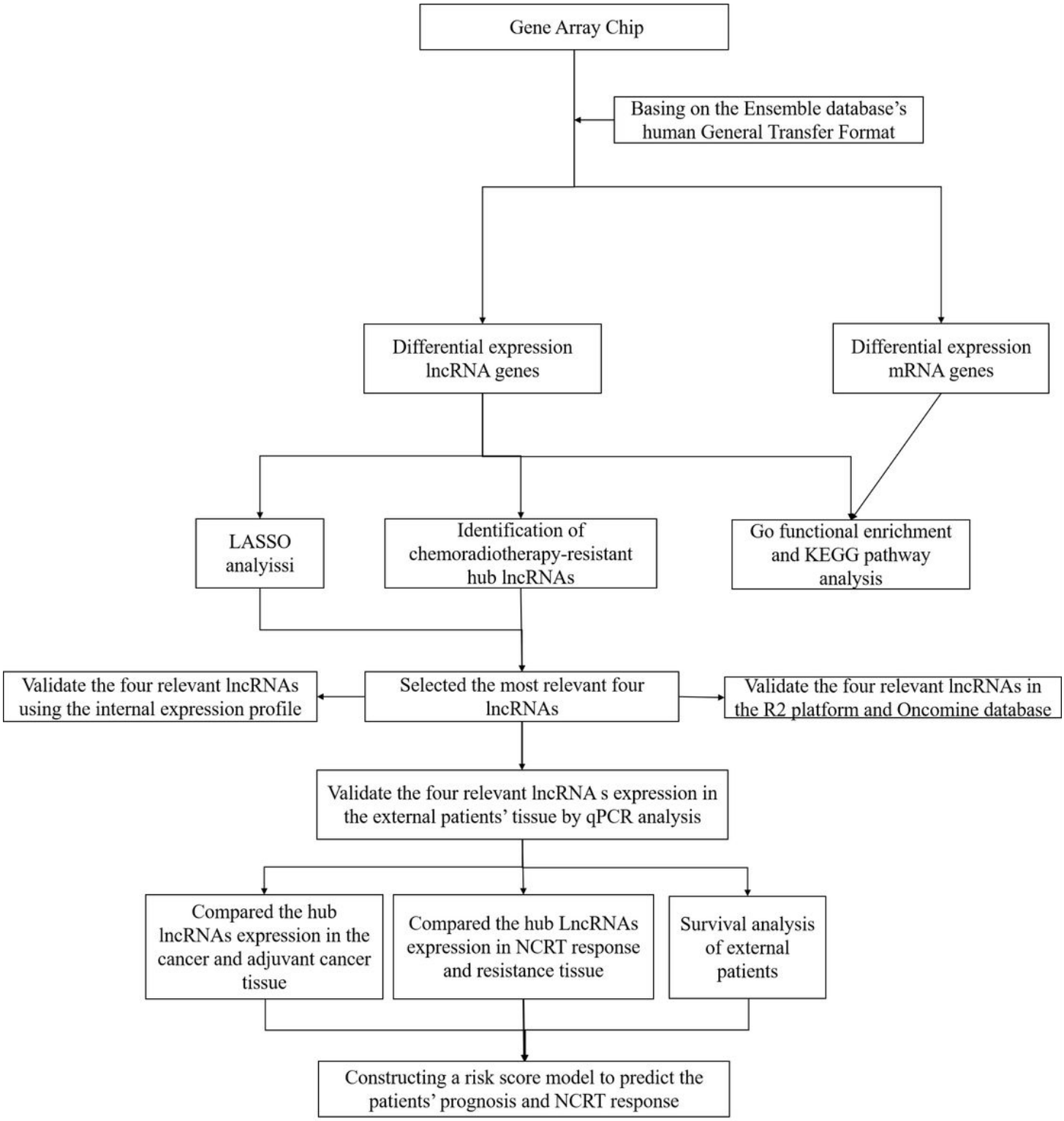

Figure 1

Work flow diagram of data preparation, processing, analysis, and validation in this study.

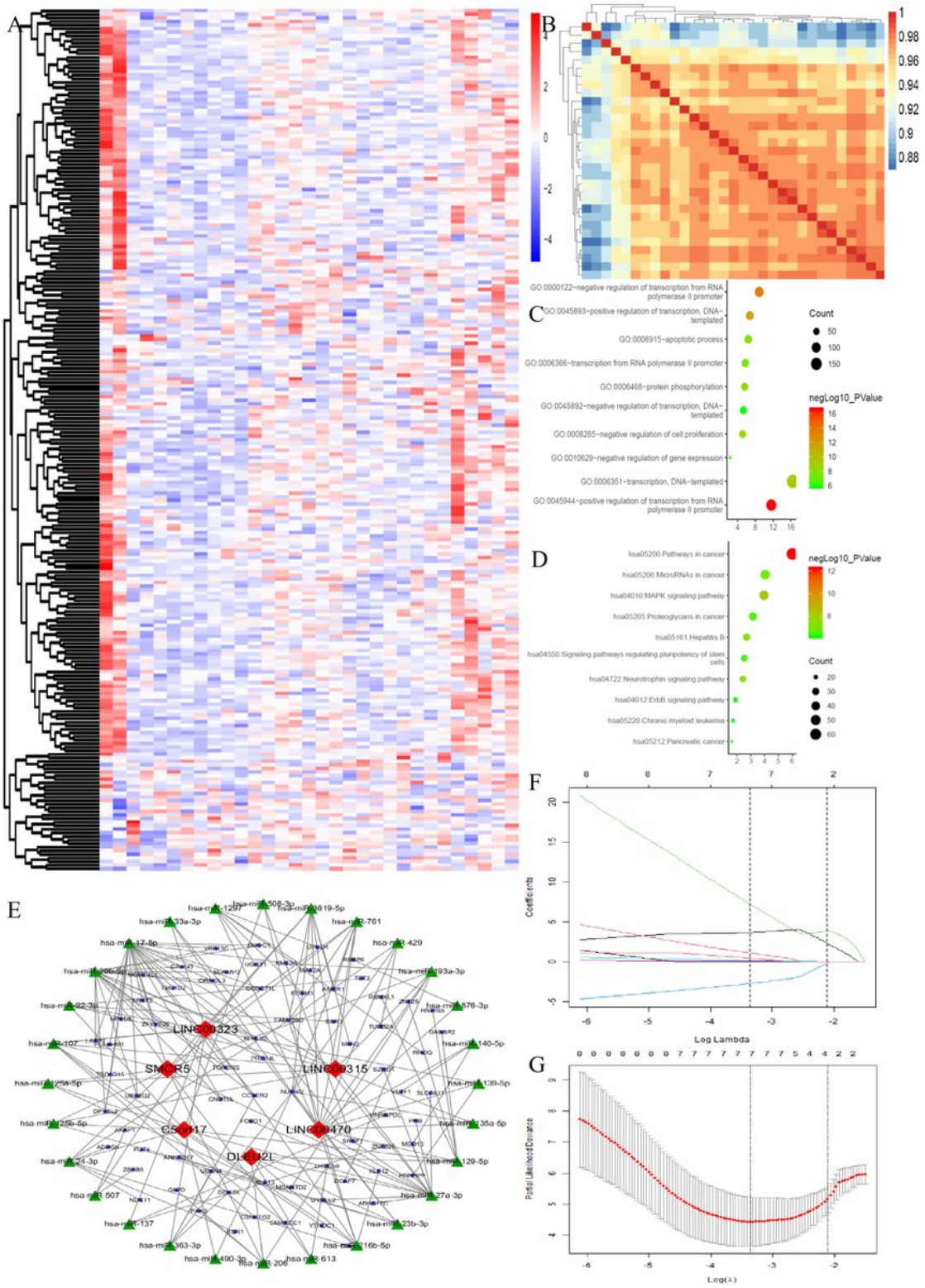

Figure 2

LncRNAs expression comparison between NCRT-resistant and NCRT-sensitivity groups. Gene Ontology (GO) functional and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of the differentially expressed lncRNAs. (A) The hierarchical clustering of all targets values of lncRNA expression among samples. (B) The heat map of the principal component analysis in each sample. (C) GO functional analysis of the top ten functional classifications. (D) KEGG pathway analysis of the top ten

significant pathways. (E) The ceRNA network was constructed basing on the differential expressed lncRNAs. (F)LASSO coefficient profiles of the 29 factors, (G) The AUC was estimated with cross-validation technique and the largest lambda value was chosen when the cross-validation error was within one standard error of the minimum.

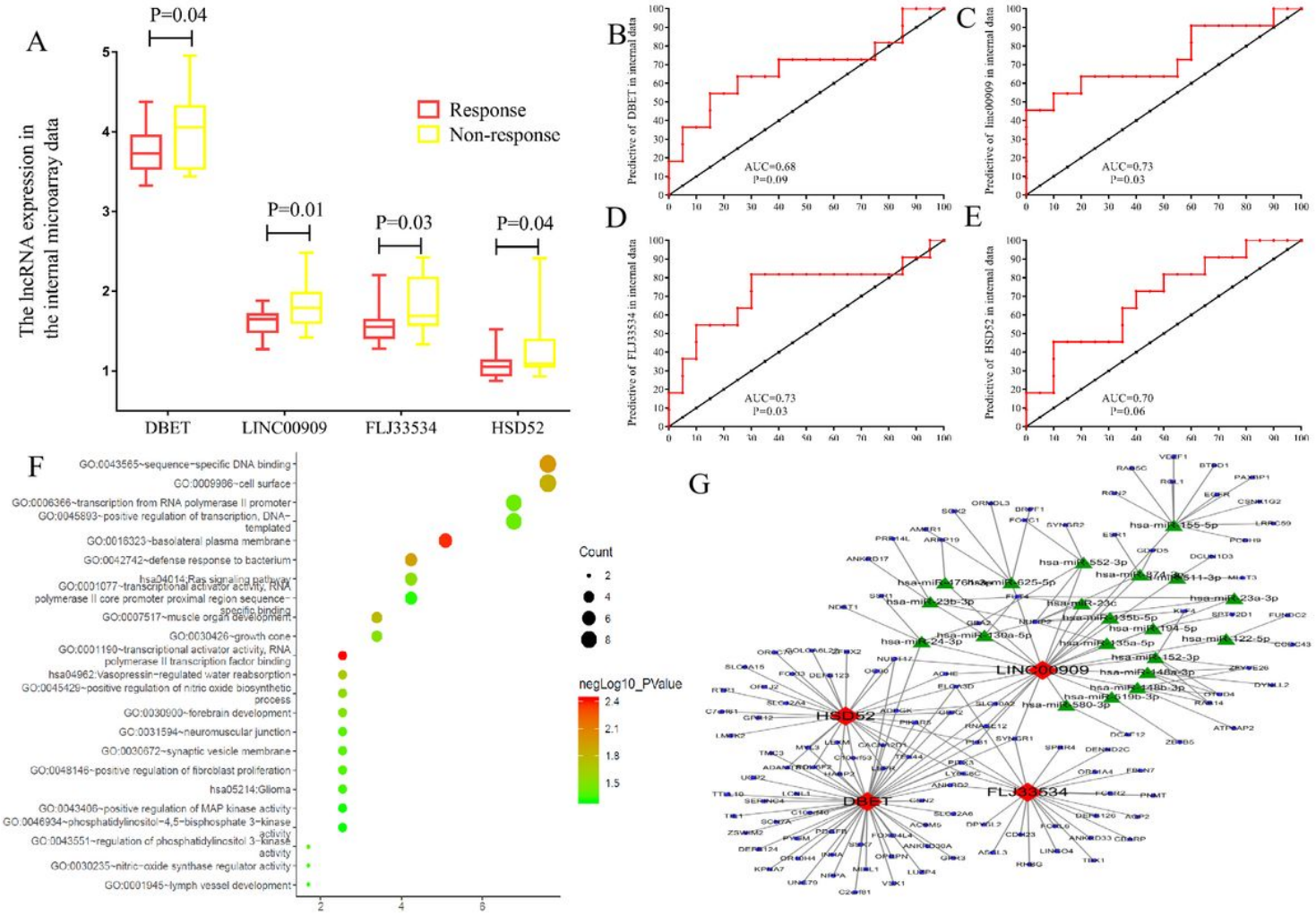

Figure 3

The hub lncRNAs validation and Gene Ontology (GO) functional and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. (A) In our data the hub lncRNAs expression (all P<0.01). (B-E) ROC curves and AUC analysis to evaluate the predictive efficiency of the hub LncRNAs in our data set. (F) GO functional analysis and KEGG pathway analysis (G) The network was constructed basing on the differential lncRNAs and mRNAs.

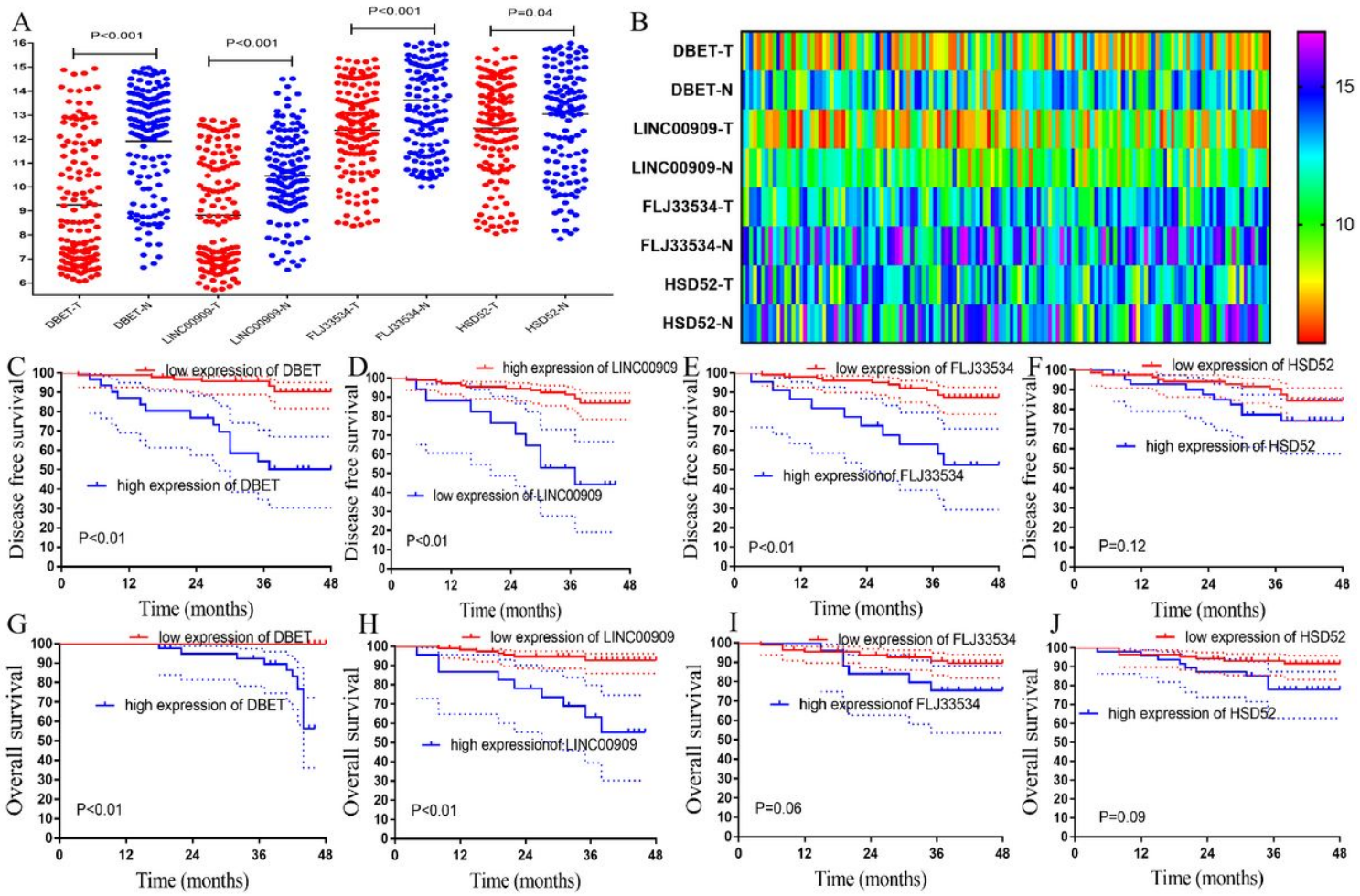

Figure 4

External validation of hub lncRNAs in cancer and adjuvant cancer tissue, and survival analysis (A)The hub lncRNAs ΔCt value in the cancerous tissue and adjacent non-cancer tissue in external CRC patients by qPCR (all P<0.01). (B) The heatmap of the ΔCt value. (C-F) The DFS analysis of the hub lncRNAs in the external dataset. (G-J) The OS analysis of the hub lncRNAs in the external dataset.

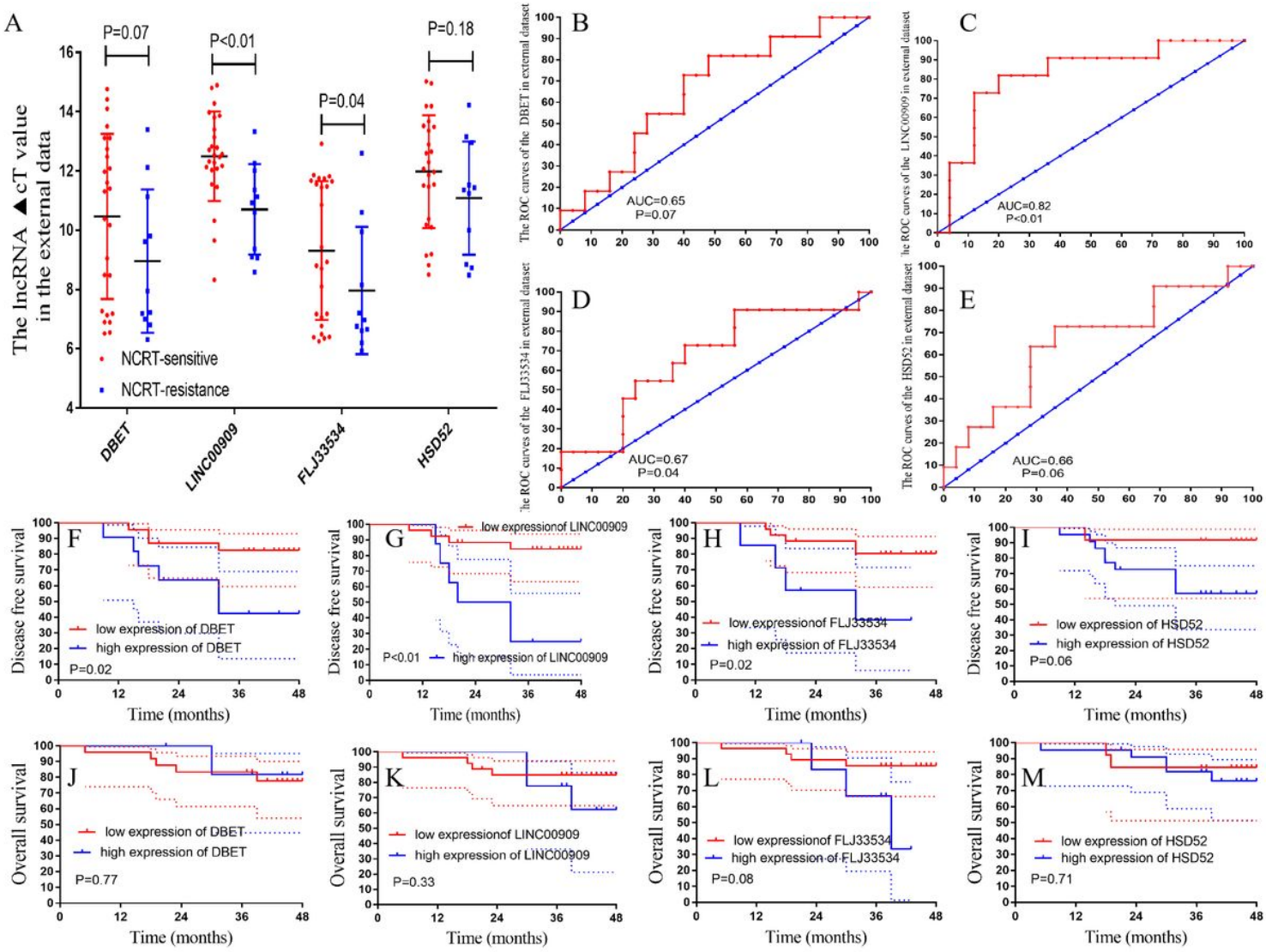

Figure 5

The hub lncRNAs validation in LARC patients following NCRT (A) In LARC patients following NCRT the hub lncRNAs expression. (B-E) ROC curves and AUC analysis to evaluate the predictive efficiency of the hub LncRNAs in LARC patients following NCRT. (F-I) The DFS analysis of the hub lncRNAs in LARC patients following NCRT. (J-M) The OS analysis of the hub lncRNAs in LARC patients following NCRT.

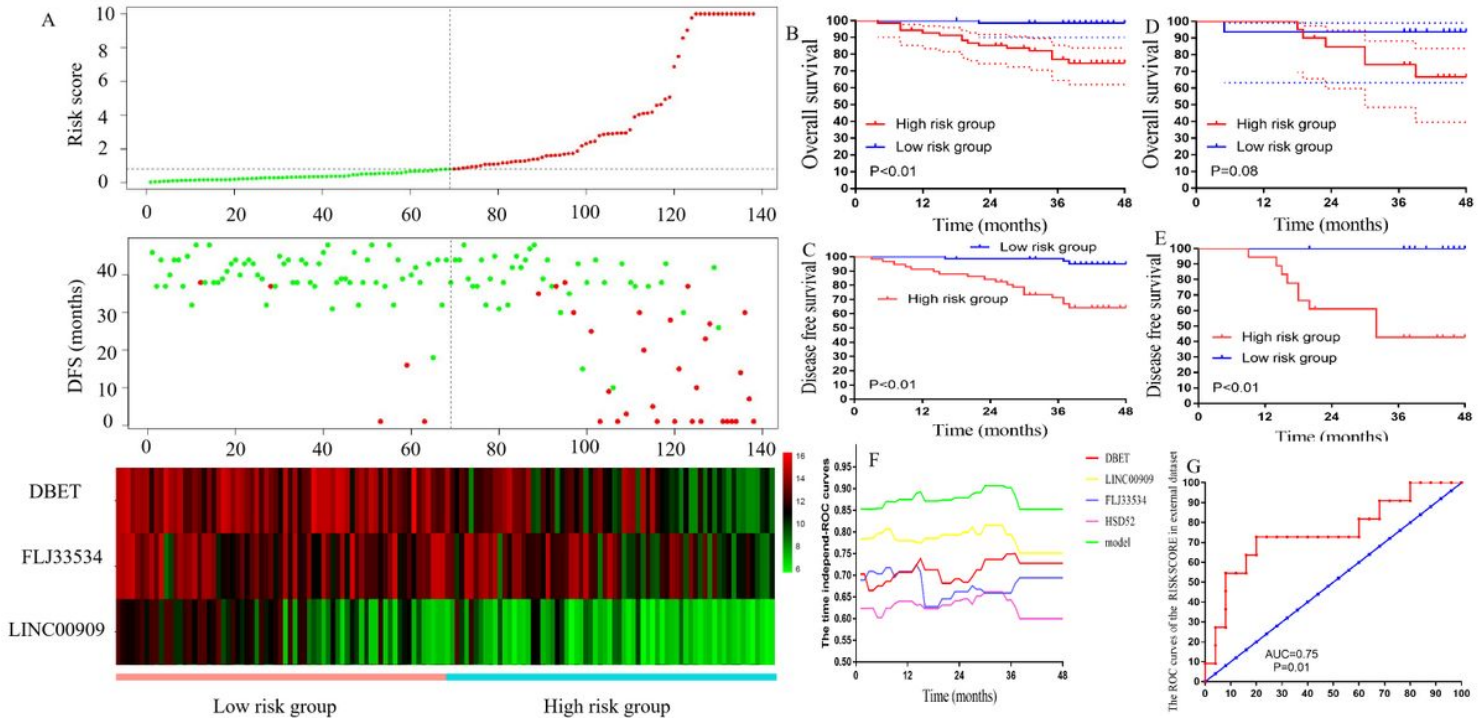

Figure 6

Rrisk factor model construction and verified in the prognosis and NCRT response. (A) The risk factor model of the hub lncRNAs in the external dataset. (Upper) LncRNA risk score distribution of 138 CRC patients. (Middle) Status of every patient in the external dataset (N=138). (Lower) Expression heatmap of the hub lncRNAs corresponding to each sample above. Red: high expression; Blue: low expression. (B) The OS analysis of the risk score in the CRC patients. (C) The DFS analysis of the risk score in the CRC patients. (D) The OS analysis of the risk score in the LARC following NCRT patients. (E) The DFS analysis of the risk score in the LARC following NCRT patients. (F)Time-dependent AUC curves of the hub lncRNAs and risk factor models for the prediction of DFS. (G) ROC curves and AUC analysis to evaluate the predictive efficiency of the risk score in LARC following NCRT patients.

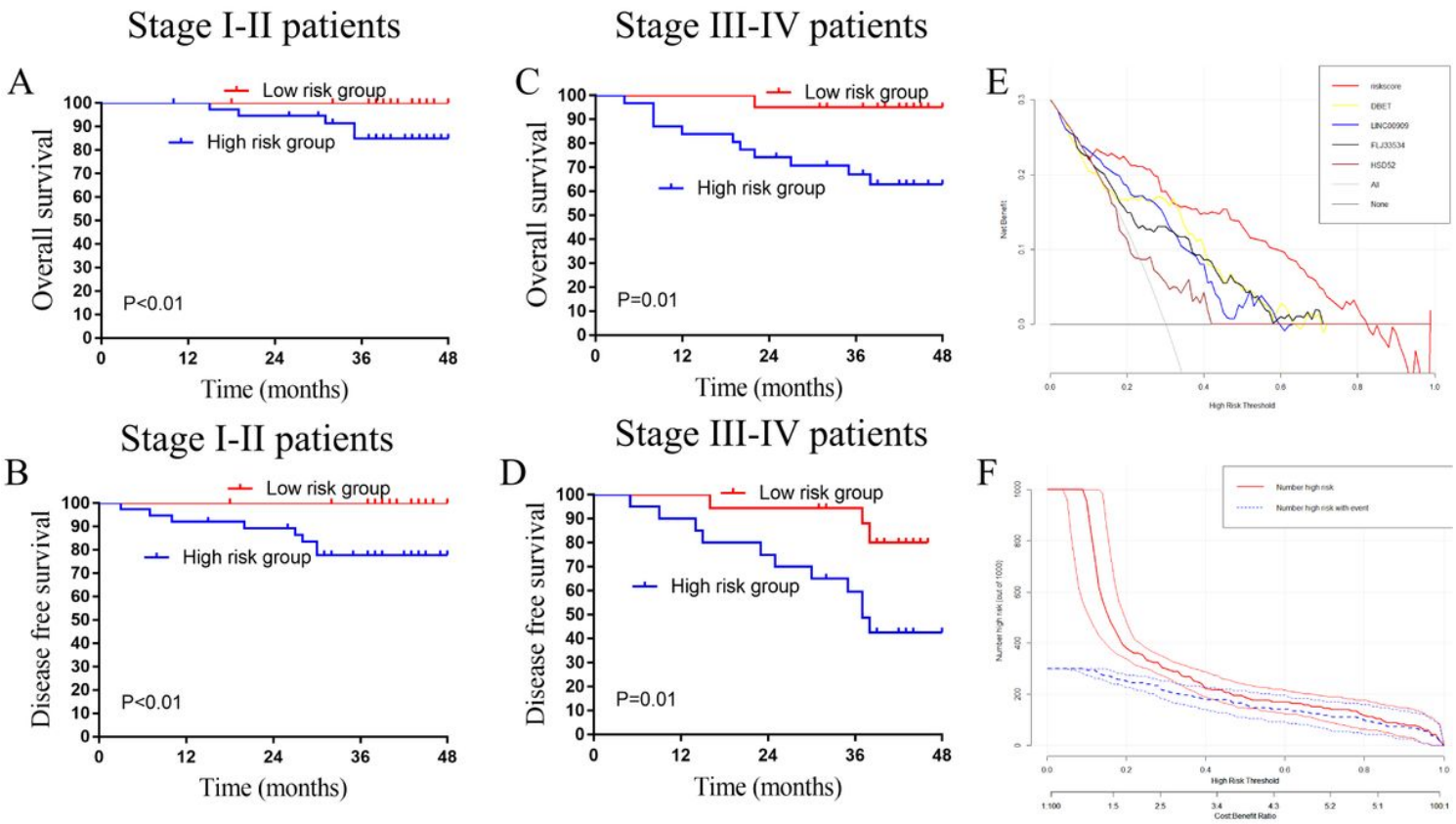

Figure 7

The application value of the lncRNAs in the colorectal cancer patients. (A) The OS analysis of the risk score in the Pathology stage I-II. (B) The DFS analysis of the risk score in the Pathology stage I-II. (C) The OS analysis of the risk score in the Pathology stage III-IV. (D) The DFS analysis of the risk score in the Pathology stage III-IV. (E) Decision curve analysis for disease recurrence. (F) Clinical impact curve for the risk score. Of 1,000 patients, the red solid line shows the total number of patients deemed to be at high risk for each risk threshold. The blue dashed line shows how many of those would be true positives.

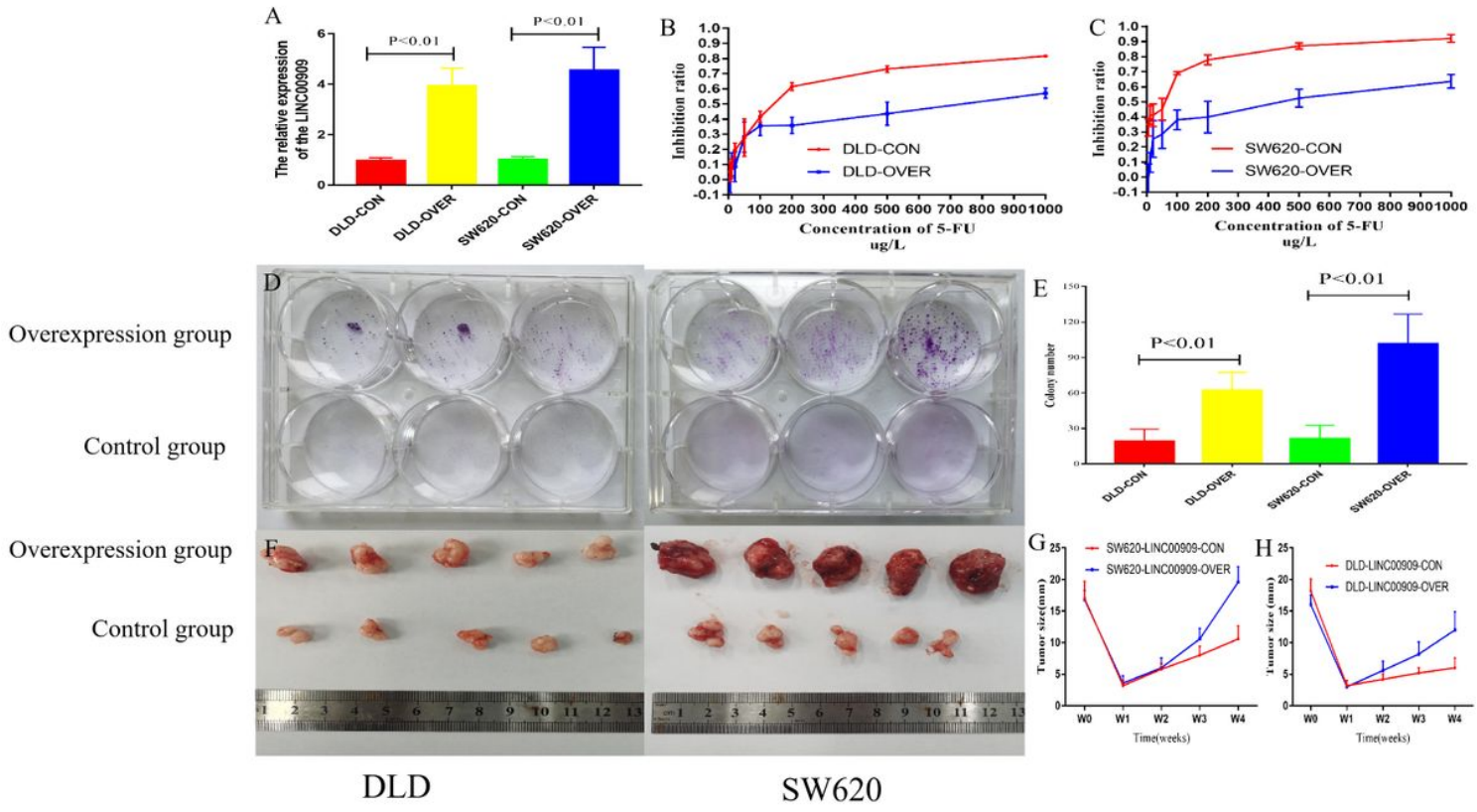

Figure 8

Overexpression of the LINC00909 enhanced the resistance to the NCRT in vivo and in vitro. The LINC00909 expression was increased in the LINC00909 overexpressed group compared with the control group ($P < 0.01$). (B and C) CCK-8 assays revealed that significantly enhanced the resistance to the 5-FU in overexpressed LINC00909 group compared with control group. (D and E) Colony formation assays demonstrated significantly increased cell number in overexpressed LINC00909 cell lines, compared to control cells. (F) Representative tumor images of control and overexpressed LINC00909 in SW620 and DLD cells. (G and H) Tumor growth curves of SW620 and DLD xenografts from the overexpressed LINC00909 and control groups.

Supplementary Files

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

- suppfigure1.tif
- suppfigure2.tif
- suppfigure3.tif
- suppletable1.doc
- suppletable2.doc
- suppletable3.doc