Am J Cancer Res 2015;5(7):2258-2265
www.ajcr.us /ISSN:2156-6976/ajcr0010845

Original Article

Circulating lncRNAs associated with occurrence of colorectal cancer progression

Jian Shi[1,2*], Xiaohua Li[3*], Fan Zhang[2], Changxi Zhang[2], Qinghai Guan[2], Xuefeng Cao[2], Wentao Zhu[2], Xingyuan Zhang[2], Yu Cheng[2], Kun Ou[2], Qiangpu Chen[2], Sanyuan Hu[1]

[1]Department of Hepatobiliary Surgery, Binzhou Medical University Hospital, 661 Second Huanghe Road, Binzhou 256603, China; [2]Department of General Surgery, Qilu Hospital of Shandong University, 107 West Culture Road, Jinan, Shandong 250012, China; [3]Department of Pediatrics, Binzhou Medical University Hospital, 661 Second Huanghe Road, Binzhou 256603, China

Received May 30, 2015; Accepted June 11, 2015; Epub June 15, 2015; Published July 1, 2015

Abstract: Screening for the potential biomarker of colorectal cancer (CRC) is necessary to improve the early detection. The aim of this study was to investigate the potential of circulating cell-free long non-coding RNAs (lncRNA) as biomarkers of CRC. In this study, we applied an lncRNA microarray to screen the potential biomarker for CRC with a multi-stage validation and risk score formula detection. We discovered three lncRNA, XLOC_006844, LOC152578 and XLOC_000303, which were up-regulated in CRC comparing with the cancer-free controls with the merged area under curve (AUC) in training set and validation set of 0.919 and 0.975. The three lncRNAs might be the potential biomarker for the tumorigenesis prediction of CRC in the future.

Keywords: CRC, lncRNA, high throughput, risk score function, biomarker

Introduction

Colorectal cancer (CRC) is the third leading cause of cancer deaths [1, 2]. CRC cases are treated through surgical resection as it is the only curative technique that is available till date [3]. Nevertheless, colorectal cancer is associated with high mortality as approximately 30% of patients are diagnosed when this cancer has reached an advanced stage [4, 5]. Screening at risk populations for CRC has significantly improved the outcome for patients, for instance diagnosis while the disease remains localised to the colon dramatically improves patient survival, and removal of early lesions such as adenomatous polyps may prevent disease formation [6, 7]. Although cheap and non-invasive approaches has been explored, this test is vulnerable to false positive and negative results due to incorrect sample storage, or confounding medical complaints [8]. The other examinations involve more costly and invasive procedures which although allow direct access to colorectal lesions also suffer from low patient acceptance and procedural risks such as perforation of the colon [9]. Therefore, we need to discover novel biomarker as this would ultimately help us in improving the early diagnosis of patients suffering CRC and may improve the outcomes.

The long non-coding RNAs (lncRNAs) have been identified as biomarkers for predicting occurrence and metastasis, and in the diagnosis of multiple diseases [10-12]. LncRNAs were characterized with the relatively stable style with their secondary structure in body fluids, thus, the detection of lncRNAs in human plasma or urine was possible for researchers [13, 14]. Although numerous studies have investigated small RNAs such as microRNAs (miRNAs) as potential biomarkers for the diagnosis or metastasis prediction for CRC [15-17], the diagnostic utility of circulating lncRNAs in CRC was rarely reported. In this study, we are approaching to investigate the potential use of circulating lncRNAs in plasma as biomarkers for CRC. By using lncRNAs microarray to characterize the

Table 1. Clinicopathological features of surgical colorectal carcinoma (CRC) and cancer-free control samples

	CRC	Control	*P* value
N	290	250	
Age Mean (SE) year	58.11 (0.79)	58.18 (0.12)	0.17[a]
Sex (male/female)	250/40	225/25	0.18[b]
Location			
Proximal colon	140		
Distal colon	150		
Differentiation grade			
Well	97		
Moderate	102		
Poorly	91		
Tumor Size (cm)			
≤ 30 cm^3	154		
> 30 cm^3	136		
Tumor Number			
Solitary	161		
Multiple	129		
TNM stage			
I-II	133		
III-IV	157		
Metastasis			
Yes	176		
No	114		

[a]Student t-test. [b]Chi-square test.

genome-wide lncRNAs expression profile in plasma from CRC patients by comparing with the cancer-free controls, we sought to identify a panel of plasma lncRNAs that might serve as a novel biomarker for diagnosis of CRC.

Materials and methods

Samples

The study totally enrolled 290 patients who had been diagnosed as CRC in department of General Surgery in Qilu Hospital of Shandong University and Hepatobiliary Surgery in Binzhou Medical University Hospital between 2011 and 2013. The study was approved by the Institutional Ethics Committee of the Qilu Hospital of Shandong University and Binzhou Medical University. All research was performed in compliance with government policies and the Helsinki Declaration. Experiments were undertaken with the understanding and written consent of each subject.

Real time polymerase chain reaction

The plasma was separated by centrifugation at 800 g for 10 min at room temperature. The supernatant plasma was recovered and stored at -80°C until analysis. We extracted total RNA from 600 μL plasma by Trizol reagent according to the protocol of manufactory (Invitrogen, CA, USA) as described previously. We carried out RT-qPCR assay with a commercial kit (TAKARA, Japan) as reported.

Microarray analysis of lncRNAs

Total RNA from each sample was quantified by the NanoDrop ND-1000 (Thermo, CA, USA). For microarray analysis, Agilent Array platform was employed. The sample preparation and microarray hybridization were performed based on the manufacturer's standard protocols with minor modifications. Then, each sample was amplified and transcribed into fluorescent cRNA along the entire length of the transcripts without 3' bias utilizing a random priming method. The labeled cRNAs were hybridized onto the Human LncRNA Array v3.0 (8 × 60K, Arraystar). After having washed the slides, the arrays were scanned by the Agilent Scanner G2505C (Agilent, CA, USA).

Statistical analysis

The paired student-test was used to compare differences in plasma lncRNA expression between CRC groups and healthy control groups. Chi-square tests and the student's t test analysis of variance were used to evaluate statistical differences in demographic and clinical characteristics. Risk score analysis was performed to investigate the effectiveness of lncRNA in CRC predicting. The upper 95% reference interval of each lncRNA value in controls was set as the threshold to code the expression level of the corresponding lncRNA for each sample as 0 and 1 in the training set. A risk score function (RSF) to predict CRC was defined according to a linear combination of the expression level for each lncRNA. For example, the RSF for sample i using information from three lncRNAs was: rsfi = $\sum$3j-1Wj.sij. In the above equation, sij is the risk score for lncRNA j on sample i, and Wj is the weight of the risk score of lncRNA j. The risk score of three lncRNAs was calculated using the weight by the regression coefficient that

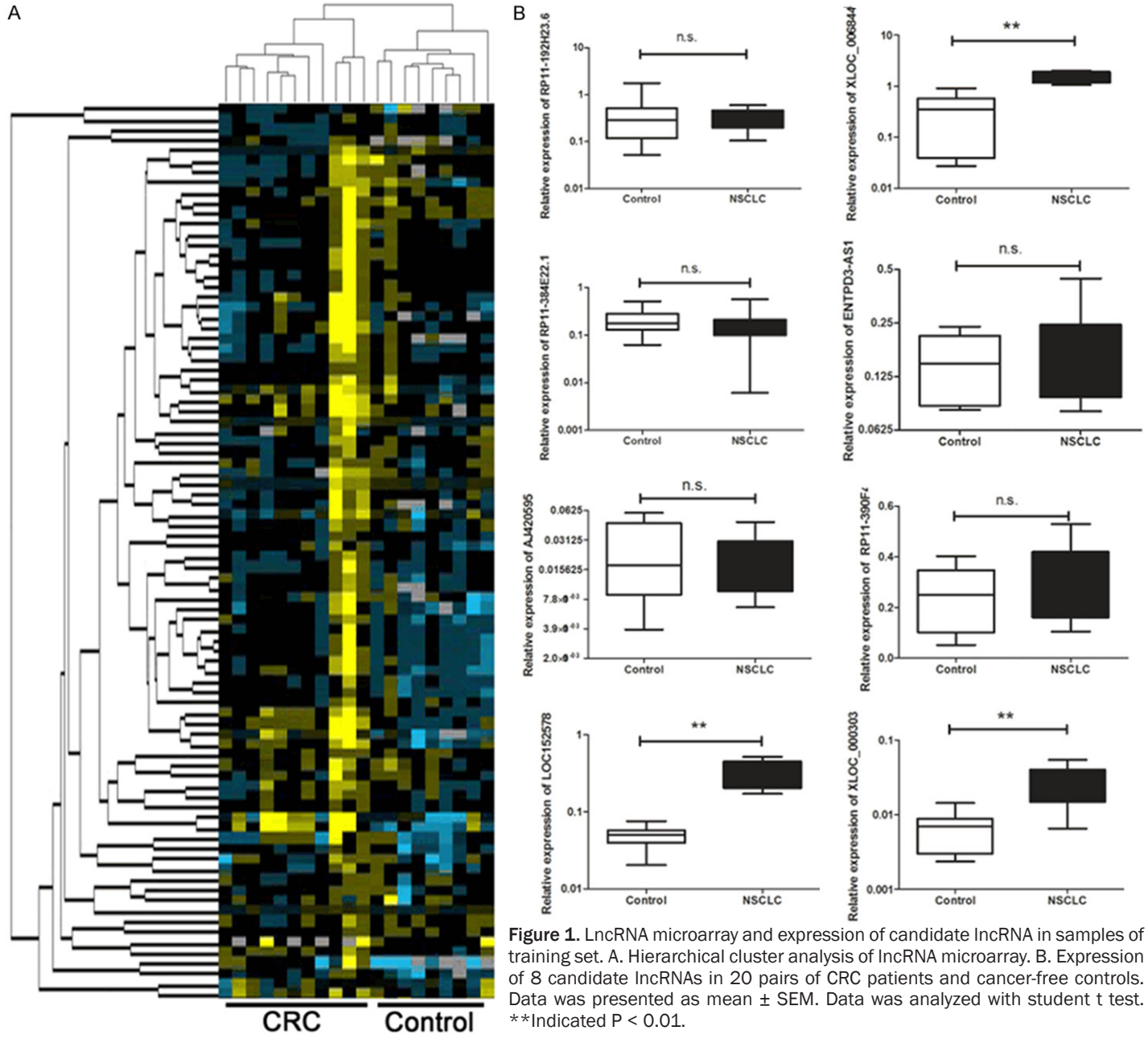
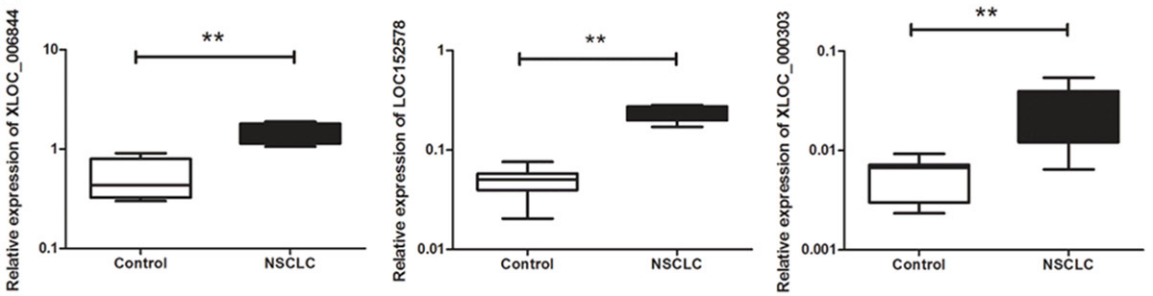
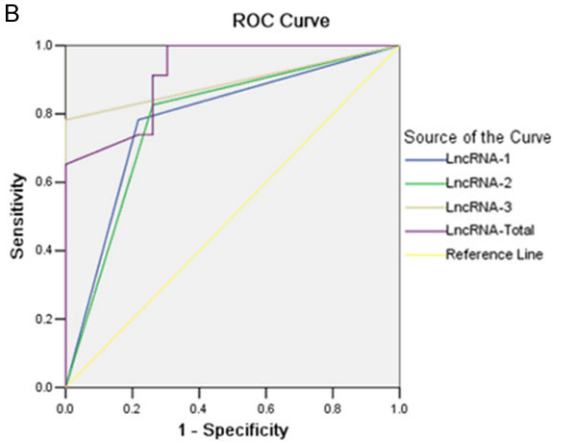


Figure 1. LncRNA microarray and expression of candidate lncRNA in samples of training set. A. Hierarchical cluster analysis of lncRNA microarray. B. Expression of 8 candidate lncRNAs in 20 pairs of CRC patients and cancer-free controls. Data was presented as mean ± SEM. Data was analyzed with student t test. **Indicated $P < 0.01$.

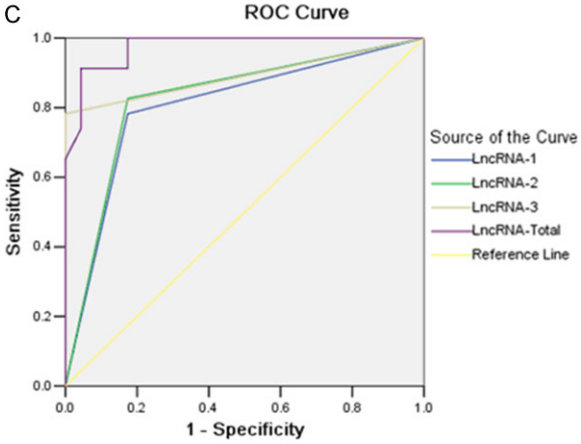
A



B



C



D

Area Under the Curve

Test Result Variable(s)	Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
LncRNA-1	.783	.071	.001	.643	.922
LncRNA-2	.783	.071	.001	.643	.922
LncRNA-3	.891	.054	.000	.786	.996
LncRNA-Total	.919	.038	.000	.844	.993

The test result variable(s): LncRNA-1, LncRNA-2, LncRNA-3, LncRNA-Total has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

E

Area Under the Curve

Test Result Variable(s)	Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
LncRNA-1	.804	.068	.000	.671	.938
LncRNA-2	.826	.065	.000	.698	.954
LncRNA-3	.891	.054	.000	.786	.996
LncRNA-Total	.975	.018	.000	.940	1.011

The test result variable(s): LncRNA-1, LncRNA-2, LncRNA-3, LncRNA-Total has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

Figure 2. Validation of the three lncRNA by RT-PCR and ROC analysis. A. Total 220 plasma from CRC patients and 180 healthy controls were used in RT-qPCR analysis. Data was presented as mean ± SEM and was analysis with student t test. **Indicated P < 0.01. B, D. ROC curve analysis was conducted for discrimination between CRC patients and control group in training set. C. ROC curve for the three-lncRNA signature to separate 220 CRC cases from 108 controls in the validation set.

was derived from the univariate logistic regression analysis of each lncRNAs. Statistical analysis was performed using STATA 9.2, and presented with GraphPad Prism 5.0 software.

Table 2. Risk score analysis of in CRC and cancer-free control plasma samples

Score	0-5.828	5.828-10.556	PPV[a]	NPV[b]
Training set			1.00	0.95
HCC	1	19		
Control	20	0		
Validation set			0.80	0.84
HCC	45	175		
Control	151	29		

[a]PPV, positive predictive value. [b]NPV, negative predictive value.

Results were considered statistically significant at $P < 0.05$.

Results

Screening of candidate lncRNA by microarray

Human LncRNA Array v3.0 was used to detect the lncRNA derived from plasma of 11 patients with CRC and the 9 cancer-free controls in this study. The clinicopathological relevance analysis of total 540 patients was summarized in **Table 1**. All 290 patients enrolled in this study were clinically and pathologically diagnosed with CRC. There were no significant differences in the distribution of age and sex between the cancer patients and the cancer-free controls.

A total of 39 lncRNA transcripts were specifically de-regulated (25 lncRNA transcripts up-regulated and 14 lncRNA transcripts down-regulated with the cutoff 4/025) in patients with CRC compared with cancer-free controls. Next, Filtering of all the 39 deregulated transcripts for high signal intensity (≥ 10) and at least 4-fold deregulation yielded 8 lncRNA candidates. All of the 8 candidate lncRNAs were confirmed to be consistently amplified in all individual samples (**Figure 1A**).

Training set and validation set analysis

We further examined these differentially expressed lncRNAs by RT-qPCR in a training sample set including 20 cases and 20 controls (including the same samples used in microarray assay). This phase generated a panel of three lncRNAs (XLOC_006844, LOC152578 and XLOC_000303) that were significantly up-regulated in CRC samples (**Figure 1B**).

To validate the accuracy and specificity of these three lncRNA as a CRC potential signature, we also examined their expression levels through a larger individual samples (220 cases and 180 controls). As shown in **Figure 2A**, the expression of three lncRNAs in plasma of CRC were all significantly higher than those in controls, which was consist with the results in training set.

To assess the diagnostic value of the three lncRNAs profiling system, we used a risk score formula to calculate the risk score function for cases and control samples. First, the risk score of each plasma sample in the training set was calculated, as the basis of their risk scores and a set cut off, plasma samples were then divided into a high-risk group, representing the possible CRC group, and a low-risk group, representing the predicted controls. At the optimal cutoff value (Value = 5.828) with the value of sensitivity + specificity considered to be maximal, the positive predictive value and negative predictive value was 95% and 100% in the training set, respectively. Similarly, when the same cut-off value was applied to calculate the risk score of samples in the larger validation sets, the positive predictive value and negative predictive value was 80% and 84% (**Table 2**). We also used the the ROC curves analysis to explorer the diagnostic sensitivity and specificity of the three-lncRNAs signature for CRC. Single As we presented in **Figure 2B**, **2D**, the areas under the curve (AUC) of XLOC_006844, LOC152578 and XLOC_000303 and the merged factor were 0.783, 0.783, 0.891, and 0.919, respectively. In the validation set the AUC of which were 0.804, 0.826, 0.891 and 0.975, respectively (**Figure 2C**, **2E**).

Double-blind test and stable expression validation

The double-blind fashion was further applied in 100 plasma samples (50 CRC patients and 50 cancer-free controls) to validate the accuracy of the three plasma lncRNA. After analyzing the expression levels of the three lncRNAs in these samples, a clear separation of CRC cases from controls was observed, with the accuracy rate of the three-lncRNAs profile as a CRC biomarker being 85%. We next amplified the three lncRNAs in three healthy controls at room temperature for 12 h, 24 h or subjecting it to up to 3 cycles of freezing and thawing. As presented in **Figure 3**, all of the process had minimal effects on the concentrations of the three

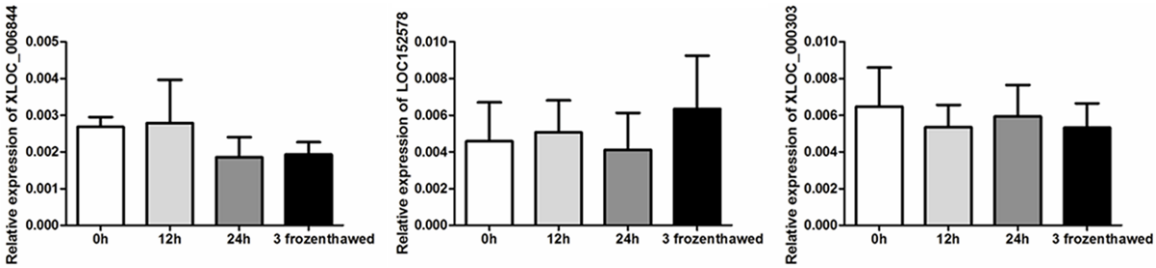


Figure 3. Stability detection of lncRNAs in human plasma. Human plasmas obtained from three healthy controls were incubated at room temperature for 12 h, 24 h or subjecting it to up to 3 cycles of freezing and thawing. Data was presented as mean ± SEM.

lncRNAs, demonstrating that these miRNAs are sufficiently stable in human plasma.

Discussion

Screening for the early detection of CRC is important to improve patient survival and facilitate cancer prevention through the detection and removal of polyps. The development of high throughput microarray and secondary generation sequencing have discovered some new biomarkers for CRC such as a miRNA panel which have been reported to be considerable clinical value in diagnosing early-stage CRC and to identify novel therapeutic targets [18, 19]. Not only the miRNA, but also the lncRNA have been annotated acting as a biomarker in predicting the feature of tumor [20, 21]. For example, genomic mapping studies have identified lncRNA as a biomarker for predicting HCC [21].

The aim of this study therefore was to investigate the potential of circulating cell-free lncRNAs as biomarkers of CRC. In our study, we performed a case-control study through the high throughput lncRNA microarray, and discovered three novel lncRNA, XLOC_006844, LOC152578 and XLOC_000303, acting as the biomarker for predicting the occurrence of CRC. The risk score analysis including a multistage validation was employed to evaluate the association between CRC and the lncRNAs expression levels. It has been widely reported that cancer-specific miRNAs are detectable in blood, sputum, urine, and other biological fluids of cancer patients. Therefore, miRNAs have been investigated as potential biomarkers for cancer diagnosis and prognosis. Likewise, lncRNAs have demonstrated utility as fluid-based markers of specific cancers. Serum and plasma harbor clinical discriminatory proteomic and transcriptomic biomarkers which can be easily assessed for clinical use.

In conclusion, we have identified three lncRNAs differentially expressed in plasma of the CRC patients compared with those of healthy controls. These findings indicate for the first time that the lncRNA in plasma might be used as a novel and a rapid diagnostic biomarker for CRC.

Acknowledgements

This study was financed by Shandong Province Science and Technology Development Project of China (2009GG10002047).

Disclosure of conflict of interest

None.

Address correspondence to: Qiangpu Chen, Department of General Surgery, Qilu Hospital of Shandong University, 107 West Culture Road, Jinan, Shandong 250012, China. Tel: +86 543 3258597; E-mail: dr_cqp@163.com; Sanyuan Hu, Department of Hepatobiliary Surgery, Binzhou Medical University Hospital, 661 Second Huanghe Road, Binzhou 256603, China. Tel: +86 18678861999; E-mail: hu_sanyuan@sina.com

References

[1] Weren RD, Ligtenberg MJ, Kets CM, de Voer RM, Verwiel ET, Spruijt L, van Zelst-Stams WA, Jongmans MC, Gilissen C, Hehir-Kwa JY, Hoischen A, Shendure J, Boyle EA, Kamping EJ, Nagtegaal ID, Tops BB, Nagengast FM, Geurts van Kessel A, van Krieken JH, Kuiper RP and Hoogerbrugge N. A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. Nat Genet 2015; 47: 668-71.

[2] Tu HC, Schwitalla S, Qian Z, LaPier GS, Yermalovich A, Ku YC, Chen SC, Viswanathan SR, Zhu H, Nishihara R, Inamura K, Kim SA, Morikawa T, Mima K, Sukawa Y, Yang J, Meredith G, Fuchs CS, Ogino S and Daley GQ.

LIN28 cooperates with WNT signaling to drive invasive intestinal and colorectal adenocarcinoma in mice and humans. Genes Dev 2015; 29: 1074-86.

[3] Medico E, Russo M, Picco G, Cancelliere C, Valtorta E, Corti G, Buscarino M, Isella C, Lamba S, Martinoglio B, Veronese S, Siena S, Sartore-Bianchi A, Beccuti M, Mottolese M, Linnebacher M, Cordero F, Di Nicolantonio F and Bardelli A. The molecular landscape of colorectal cancer cell lines unveils clinically actionable kinase targets. Nat Commun 2015; 6: 7002.

[4] Gallego CJ, Shirts BH, Bennette CS, Guzauskas G, Amendola LM, Horike-Pyne M, Hisama FM, Pritchard CC, Grady WM, Burke W, Jarvik GP and Veenstra DL. Next-Generation Sequencing Panels for the Diagnosis of Colorectal Cancer and Polyposis Syndromes: A Cost-Effectiveness Analysis. J Clin Oncol 2015; 33: 2084-91.

[5] Azvolinsky A. Colorectal cancer: to stack or sequence therapy? J Natl Cancer Inst 2015; 107.

[6] Bazzocco S, Andretta E, Rodrigues P, Garrido M, Alazzouzi H, Chionh F, Carton-Garcia F, Macaya I, Nieto R, Sanchez A, Schwartz S Jr, Dopeso H, Bilic J, Mariadason JM and Arango D. Highly Expressed Genes in Rapidly Proliferating Tumor Cells as New Targets for Colorectal Cancer Treatment. Clin Cancer Res 2015; [Epub ahead of print].

[7] Aravalli RN and Steer CJ. Circulating microRNAs: novel biomarkers for early detection of colorectal cancer. Transl Res 2015; [Epub ahead of print].

[8] Mooiweer E, van der Meulen-de Jong AE, Ponsioen CY, van der Woude CJ, van Bodegraven AA, Jansen JM, Mahmmod N, Kremer W, Siersema PD and Oldenburg B. Incidence of interval colorectal cancer among inflammatory bowel disease patients undergoing regular colonoscopic surveillance. Clin Gastroenterol Hepatol 2015; [Epub ahead of print].

[9] Romaguera D, Ward H, Wark PA, Vergnaud AC, Peeters PH, van Gils CH, Ferrari P, Fedirko V, Jenab M, Boutron-Ruault MC, Dossus L, Dartois L, Hansen CP, Dahm CC, Buckland G, Sanchez MJ, Dorronsoro M, Navarro C, Barricarte A, Key TJ, Trichopoulou A, Tsironis C, Lagiou P, Masala G, Pala V, Tumino R, Vineis P, Panico S, Bueno-de-Mesquita HB, Siersema PD, Ohlsson B, Jirstrom K, Wennberg M, Nilsson LM, Weiderpass E, Kuhn T, Katzke V, Khaw KT, Wareham NJ, Tjonneland A, Boeing H, Quiros JR, Gunter MJ, Riboli E and Norat T. Pre-diagnostic concordance with the WCRF/AICR guidelines and survival in European colorectal cancer patients: a cohort study. BMC Med 2015; 13: 107.

[10] Tang J, Zhuo H, Zhang X, Jiang R, Ji J, Deng L, Qian X, Zhang F and Sun B. A novel biomarker Linc00974 interacting with KRT19 promotes proliferation and metastasis in hepatocellular carcinoma. Cell Death Dis 2014; 5: e1549.

[11] Dong S, Qu X, Li W, Zhong X, Li P, Yang S, Chen X, Shao M and Zhang L. The long non-coding RNA, GAS5, enhances gefitinib-induced cell death in innate EGFR tyrosine kinase inhibitor-resistant lung adenocarcinoma cells with wide-type EGFR via downregulation of the IGF-1R expression. J Hematol Oncol 2015; 8: 43.

[12] Kurian L, Aguirre A, Sancho-Martinez I, Benner C, Hishida T, Nguyen TB, Reddy P, Nivet E, Krause MN, Nelles DA, Esteban CR, Campistol JM, Yeo GW and Izpisua Belmonte JC. Identification of Novel Long Noncoding RNAs Underlying Vertebrate Cardiovascular Development. Circulation 2015; 131: 1278-1290.

[13] Lorenzen JM, Schauerte C, Kielstein JT, Hubner A, Martino F, Fiedler J, Gupta SK, Faulhaber-Walter R, Kumarswamy R, Hafer C, Haller H, Fliser D and Thum T. Circulating long noncoding RNATapSaki is a predictor of mortality in critically ill patients with acute kidney injury. Clin Chem 2015; 61: 191-201.

[14] Tong YS, Wang XW, Zhou XL, Liu ZH, Yang TX, Shi WH, Xie HW, Lv J, Wu QQ and Cao XF. Identification of the long non-coding RNA POU3F3 in plasma as a novel biomarker for diagnosis of esophageal squamous cell carcinoma. Mol Cancer 2015; 14: 3.

[15] Ng EK, Chong WW, Jin H, Lam EK, Shin VY, Yu J, Poon TC, Ng SS and Sung JJ. Differential expression of microRNAs in plasma of patients with colorectal cancer: a potential marker for colorectal cancer screening. Gut 2009; 58: 1375-1381.

[16] Kanaan Z, Roberts H, Eichenberger MR, Billeter A, Ocheretner G, Pan J, Rai SN, Jorden J, Williford A and Galandiuk S. A plasma microRNA panel for detection of colorectal adenomas: a step toward more precise screening for colorectal cancer. Ann Surg 2013; 258: 400-408.

[17] Kanaan Z, Rai SN, Eichenberger MR, Roberts H, Keskey B, Pan J and Galandiuk S. Plasma miR-21: a potential diagnostic marker of colorectal cancer. Ann Surg 2012; 256: 544-551.

[18] Perilli L, Vicentini C, Agostini M, Pizzini S, Pizzi M, D'Angelo E, Bortoluzzi S, Mandruzzato S, Mammano E, Rugge M, Nitti D, Scarpa A, Fassan M and Zanovello P. Circulating miR-182 is a biomarker of colorectal adenocarcinoma progression. Oncotarget 2014; 5: 6611-6619.

[19] Giraldez MD, Lozano JJ, Ramirez G, Hijona E, Bujanda L, Castells A and Gironella M. Circulating microRNAs as biomarkers of colorectal cancer: results from a genome-wide profiling and validation study. Clin Gastroenterol Hepatol 2013; 11: 681-688 e683.

[20] Uchida S and Dimmeler S. Long noncoding RNAs in cardiovascular diseases. Circ Res 2015; 116: 737-750.

[21] Tang J, Jiang R, Deng L, Zhang X, Wang K and Sun B. Circulation long non-coding RNAs act as biomarkers for predicting tumorigenesis and metastasis in hepatocellular carcinoma. Oncotarget 2015; 6: 4505-4515.