Int J Clin Exp Pathol 2019;12(10):3975-3981
www.ijcep.com /ISSN:1936-2625/IJCEP0101974

Original Article

Combined identification of LncRNA CCAT1 and SOX2OT in serum as an effective screening for non-small cell lung cancer

Zhipeng Wang[1,2], Jian Li[1,2], Yan Wu[3], Hongbin Wang[3]

[1]School of Medicine, Jiangsu University, Jiangsu, China; [2]Department of Pulmonary Medicine, Affiliated Hospital of Jiangsu University, Jiangsu, China; [3]College of Marine Life and Fisheries, Jiangsu Ocean University, Jiangsu, China

Received September 6, 2019; Accepted September 24, 2019; Epub October 1, 2019; Published October 15, 2019

Abstract: Background: Long non-coding RNAs (lncRNAs) CCAT1 and SOX2OT have been shown to play important regulatory roles in cancer biology. Tumor biomarkers need to be detectable in easily accessible body fluids, should be characterized by high specificity, sufficient sensitivity, and robustness against influencing factors. The aim of this study is to evaluate the clinical significance of serum CCAT1 and SOX2OT as a biomarker in the screening of NSCLC. Results: CCAT1 and SOX2OT were shown to be detectable in the cellular fraction of peripheral human blood, showing serum levels of CCAT1 and SOX2OT were significantly increased of cancer patients as compared to cancer-free controls. The ROC curves illustrated strong separation between the NSCLC patients and control group, with an AUC of 0.846 (95% CI 0.766-0.926; $P < 0.001$) for CCAT1 and 0.787 (95% CI: 0.691-0.883; $P < 0.001$) for SOX2OT. However, the combination of SOX2OT and CCAT1 yielded an AUC of 0.894 (95% CI: 0.825-0.963; $P < 0.001$), which was significantly improved as compared to CCAT1 or SOX2OT alone. Moreover, the interaction between lncRNAs and some functional proteins, such as TTF1, p63, Ck7, K-ras, EGFR, may contribute to the tumorigenesis. Conclusion: Our results demonstrated that increased serum CCAT1 and SOX2OT could be used as a predictive biomarker for NSCLC screening, and that combination of CCAT1 and SOX2OT had a higher positive diagnostic efficiency of NSCLC than CCAT1 or SOX2OT alone.

Keywords: Non-small cell lung cancer, long non-coding RNA, CCAT1, SOX2OT, tumor biomarker

Introduction

Lung cancer is one of the most fatal malignancies worldwide, which has a 5 year survival rate of about 15% [1]. Non-small cell lung cancer (NSCLC) accounts for 85% of lung cancers, whose ravages has violently hijacked the great burden of human welfare, health, and economics [2]. Although great progress has been made in the treatment of NSCLC in recent years, the overall survival time of patients with NSCLC has not improved significantly. One important reason is the lack of molecular biomarkers. Therefore, looking for the effective biomarkers of NSCLC is urgently needed.

LncRNAs are more than 200 nucleotides in length, longer than microRNAs. According to genomic organization, lncRNAs can be classified into five broad categories: sense, antisense, bidirectional, intronic, and intergenic [3]. Recently, several lncRNAs had been shown to serve as NSCLC biomarkers, such as PVT1 [4], MALAT1 [5], XSIT [6], and CCHE1 [7], may contribute to cancer development. Colon cancer-associated transcript 1, LncRNA CCAT1, which had been reported involving in colorectal cancer [8], cholangiocarcinoma [9], gastric cancer [10]. These results indicated that CCAT1 may be necessary for human cancer progression. In recent study, overexpression of CCAT1 was demonstrated to promote the migration and invasion of H358 lung adenocarcinoma cells; while downregulation of CCAT1 expression inhibited H1650 cell migration and invasion [11]. The lncRNA SOX2OT is transcribed in the same orientation like SOX2, which plays a significant role in regulation of tumor progression [12]. SOX2OT was found highly expressed in gastric cancer [13], hepatocellular carcinoma [14], and

colorectal cancer [15], suggesting that SOX2OT may serve as a biomarker for the diagnosis of these cancers. A study had shown that the lncRNA-SOX2OT regulated lung cancer cell proliferation and was a prognostic indicator of poor survival [16]. Although lncRNAs may play an important role in NSCLC, little is known whether lncRNA CCAT1 and SOX2OT in serum can serve as biomarkers in NSCLC.

In our current study, we examined the expression levels of CCAT1 and SOX2OT in serum, and their potential use as tumor markers for NSCLC detection were evaluated. We hypothesized that these NSCLC-related lncRNAs might be released into the circulation during NSCLC initiation and could be utilized to detect and screen NSCLC.

Materials and methods

Patients and specimens

A total of 48 patients with NSCLC were diagnosed as lung cancer patients by biopsy and CT results. Immunohistochemistry and electron microscopy were used to determine the histological classification of patients with non-small cell lung cancer, such as squamous cell carcinoma, adenocarcinoma, large cell carcinoma, and other subtypes. Blood specimens were collected from Affiliated Hospital of Jiangsu University between Jan 2017 and Feb 2019. All patients recruited in this study were not subjected to preoperative radiotherapy or chemotherapy and diagnosed with NSCLC based on histopathological evaluation. All clinical data including age, gender, clinical stage, smoking history, infiltration degree, and lymph node metastasis of these patients were recorded in a database. In addition, 48 people without NSCLC or other malignancies were recruited to act as healthy controls. Written informed consents were signed by all participators in advance.

RNA extraction and quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA was extracted from all tissue samples with TRIzol (Invitrogen, Carlsbad, CA, USA). Then the first chain of cDNA was synthesized using reverse transcriptase with High Capacity cDNA Reverse Transcription Kit (Takara, Dalian, China). Quantitative PCR was performed on the cDNA using specific primers (Takara, Dalian, China). RT-PCR reaction was performed in the Applied Biosystems 7500 Fast Real-Time PCR system (Applied Biosystems, Foster City, CA, USA). PCR with the following primers: CCAT1, Forward 5'-TTGCTCACCTTACTGCCTGA-3' and Reverse 5'-CTCATAAGGAGCGCACAACC-3'; SOX2OT, Forward 5'-AGAGGAGTTAGGCAAGTGGAGTGG-3' and Reverse 5'-GAAGCAGGTGCTCATCAGAGGAAC-3'; 5srRNA, Forward 5'-GTCTACGGCCATACCACCCTGAA-3' and Reverse 5'-AAGCCTACAGCACCCGGTATTCC-3'. 5srRNA as an internal control was used to normalize the data to determine the relative expression of the target genes. The reaction conditions were set according to the kit instructions. After completion of the reaction, the amplification curve and melting curve were analyzed. Gene expression values are represented using the $2^{-\Delta\Delta Ct}$ method. Data from two patients with non-small cell lung cancer were deleted because the patient's pathological index is incomplete. Data from three healthy samples were deleted because inflammation of other organs is detected during physical examination.

Statistical analysis

All statistical analyses were performed using GraphPad Prism 6.0 (GraphPad Software, LaJolla, CA, USA) software. Receiver operating characteristics (ROC) curve analysis was undertaken to determine the best cutoff value for NSCLC patients and control group levels to achieve 95% specificity. All quantified data was presented as mean ± SD. All values are two sided, $P < 0.001$ was regarded as statistical significant.

Bioinformatics analysis

In order to more deeply explore the relationship between clinical pathologic feature and lncRNAs in breast and lung cancers, computer analysis based on RNAProtein interaction prediction (RPISeq) program, which used both random forest (RF) and support vector machine (SVM) classifiers were utilized to predict the interaction between lncRNAs and breast cancer and lung cancer associated proteins. The interaction probabilities generated by RPISeq range from 0 to 1. If the predictions with probabilities were > 0.5, we considered them positive, indicating that the corresponding RNA and protein were likely to interact.

Table 1. Correlation clinicopathological factors and CCAT1 expression levels in NSCLC patients

Characteristics	Numbers	CCAT1-Low	CCAT1-High	*P* value
Gender				0.171
Male	33	16	17	
Female	15	8	7	
Age (years)				0.386
< 60	13	5	8	
≥ 60	35	17	18	
Tumor size (cm)				0.429
≤ 3 cm	16	6	10	
> 3 cm	32	13	19	
Histological classification				0.016
SCC (Squamous cell carcinoma)	17	5	12	
AD (adenocarcinoma or other)	31	11	20	
TNM stage				0.615
I and II	10	4	6	
III and IV	38	15	23	
Lymph node metastasis				0.293
Negative	14	5	9	
Positive	34	15	19	
History of smoking				0.007
Ever	22	8	14	
Never	26	10	16	

Results

Identification of lncRNA CCAT1 and SOX2OT implicated in NSCLC patients

Clinicopathological characteristics analyses were shown in **Tables 1** and **2**. The median age was 64.2 years (range 39-82). Approximate two thirds of patients were male (n = 33), and 54% of patients were never smokers (n = 26), and 65% of patients had nonsquamous histology (n = 31).

To explore whether these NSCLC-related lncRNAs could reach the circulation at levels sufficient to be detectable, real-time PCR analysis was performed to determine the expression level of CCAT1 and SOX2OT in 96 serum samples (48 NSCLC patients and 48 healthy donors). We found that the expression of CCAT1 and SOX2OT in NSCLC patients was conspicuously higher than that of the healthy specimens (P < 0.001, **Figure 1A** and **1B**). Next, we explored the correlation between the expression of CCAT1 and SOX2OT and the clinicpathological factors of patients with NSCLC. CCAT1 and SOX2OT expression, whereas patients with lymph node metastasis were associated with higher CCAT1 and SOX2OT, respectively (**Figure 2A**, P = 0.003) (**Figure 2B**, P = 0.009).

Evaluation of CCAT1 and SOX2OT in serum as predictive NSCLC-related biomarkers

To investigate the characteristics of CCAT1 and SOX2OT as potential biomarkers for NSCLC, ROC curves and the area under the ROC curves (AUC) were performed on data from all subjects, including 48 NSCLC patients and 48 healthy group. Using ROC analyses, for NSCLC patients and controls, 0.846 (95% CI: 0.766-0.926; P < 0.001), for CCAT1 and 0.787 (95% CI: 0.691-0.883; P < 0.001) for SOX2OT (**Figure 3A** and **3B**). Intriguingly, there is increasing evidence showing that combination several tumor markers could improve diagnostic accuracy [17]. We determined whether the combination of CCAT1 and SOX2OT could provide a more effective screening for NSCLC. The results indicated that combination of CCAT1 and SOX2OT yielded a 0.894 (95% CI: 0.825-0.963; P < 0.001), which was significantly improved as compared to CCAT1 (AUC = 0.846) or SOX2OT (AUC = 0.787) alone (**Figure 3C**).

Prediction of interaction between lncRNAs and cancer associated protein

RNA-protein interactions were proven to be very important in transcriptional and post-transcriptional regulation of gene expression. Identifying novel RNA protein interactions will add valuable information about the RNA-protein interaction networks. RPISeq, a bioinformatics method, can be used to predict RNA protein interactions by using sequence analysis [18]. (http://pridb.gdcb.iastate.edu/RPISeq/). RPISeq contained Random Forest (RF) and Support

Table 2. Correlation clinicopathological factors and SOX2OT expression levels in NSCLC patients

Characteristics	Numbers	SOX2OT-Low	SOX2OT-High	*P* value
Gender				0.825
Male	33	12	21	
Female	15	5	10	
Age (years)				0.544
< 60	13	6	7	
≥ 60	35	11	24	
Tumor size (cm)				0.737
≤ 3 cm	16	5	10	
> 3 cm	32	12	20	
Histological classification				0.028
SCC (Squamous cell carcinoma)	17	5	12	
AD (adenocarcinoma or other)	31	12	23	
TNM stage				0.169
I and II	10	2	6	
III and IV	38	15	25	
Lymph node metastasis				0.033
Negative	14	4	10	
Positive	34	13	21	
History of smoking				0.246
Ever	22	11	11	
Never	26	6	20	

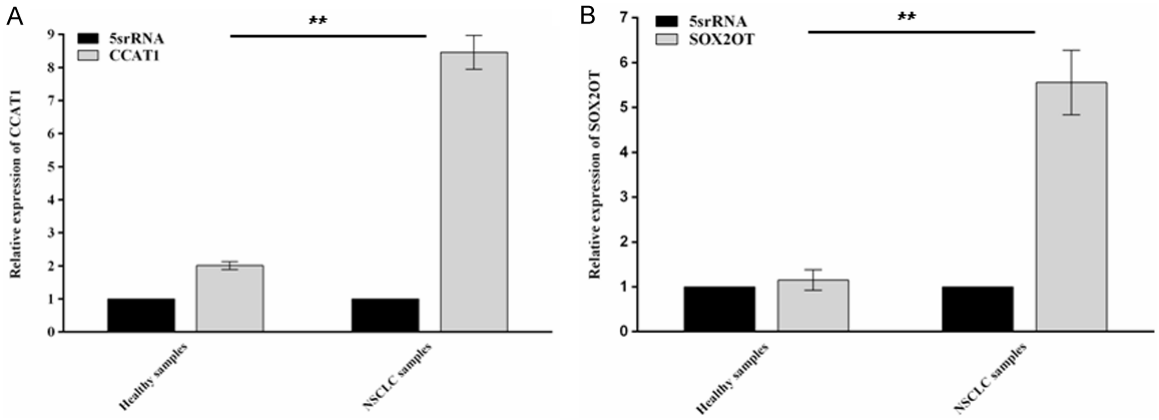

Figure 1. CCAT1 and SOX2OT were detectable in tumor tissues and serum. CCAT1 (A) and SOX2OT (B) expression was examined by real-time PCR and normalized to 5srRNA expression in 48 pairs of NSCLC patients compared with healthy samples. **, P < 0.01.

Vector Machine (SVM). In addition to clinical pathologic features, TTF1, p63, Ck7, K-ras, and EGFR are associated with the oncogenesis of lung cancer. The results in this study showed that lncRNA SOX2OT had great possibility of interaction with lung cancer associated proteins, while lncRNA CCAT1 were predicted to interact with functional protein except CK7. For example, EGFR was predicted to interact with CCAT1 by RF (score = 0.85) and SVM (score = 0.89). EGFR was predicted to interact with SOX2OT by RF (score = 0.8) and SVM (score = 0.82). Detail results were shown in **Figure 4**. It was suggested that interaction between lncRNAs and some functional proteins may contribute to the tumorigenesis.

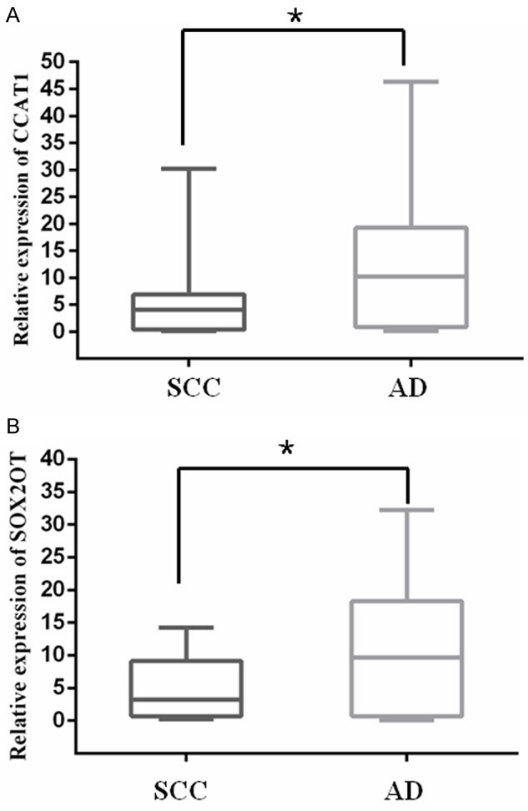

Figure 2. Comparison of lncRNAs CCAT1 (A) and SOX2OT (B) expression in NSCLC with lymph node metastasis than in those with negative lymph node metastasis. *, $P < 0.05$.

Discussion

NSCLC is commonly detected in late stages of the disease. Biomarkers have the potential to detect cancer at early stages, facilitating an earlier and therefore more curative therapy that ideally results in decreased mortality. In recent years, lncRNA differential expression profiles in various cancers have been shown by microarray studies, RNA-Seq, and quantitative reverse transcription PCR (qRT-PCR). It is therefore suggested that lncRNAs might be used as potential biomarkers for cancer diagnosis and prognosis. A great amount of lncRNAs have already been revealed to correlate with various disease processes, including carcinogenesis. lncRNA PVT1 were reported to be dysregulated and likely played important roles in a variety of cancers. PVT1 could be a novel biomarker for diagnosis and prognosis of non-small cell lung cancer [19]. Moreover, lncRNA H19 [20], CCHE1 [7], HIT [21] were associated with the process of NSCLC. In this study, we detected CCAT1 and SOX2OT levels in serum from NSCLC patients. The results of qRT-PCR showed that CCAT1 and SOX2OT levels were significantly upregulated in serum compared to control group respectively. Comparable results were achieved by Yin et al., showing a upregulation of CCAT1 in patients with NSCLC [11]. We used the ROC curve to analyze the diagnostic value of serum CCAT1 and SOX2OT. The results showed that the individual AUC of CCAT1 and SOX2OT for the diagnosis of NSCLC were about 0.846 and 0.787, respectively. Intriguingly, the combination of UCA and HIF1A-AS1 could provide a more effective screening for NSCLC. The results indicated that combination of CCAT1 and SOX2OT yielded an AUC of 0.894, which was significantly improved as compared to CCAT1 (AUC = 0.846) or SOX2OT (AUC = 0.787) alone. As we know, early discovery, early diagnosis, and early treatment could greatly increase the survival rate of cancer patients. Biomarkers in body fluid have a potential capacity to detect cancers in early stage [22]. Measurements obtained from tumor tissues and serums were strongly correlated for CCAT1 and SOX2OT. The results suggested that serum samples were acceptable for evaluation of NSCLC-related biomarkers.

Indeed, the interaction between lncRNAs and lung cancer associated proteins were suggested. EGRF is a potent mitogenic factor in a variety of cells, promoting epithelial cell proliferation. EGFR mutations were found in lung cancer. Alexander reported that patients with malignant pleural effusion of lung adenocarcinoma had increased frequency of EGFR and KRAS mutations [23]. TTF1 is a transcription termination code. The overexpression of TTF1 was associated with poor prognosis in patients with colorectal cancer [24]. P63 induced expression of many p53-target genes that were involved in cell-cycle arrest and apoptosis, so p63 is useful in distinguishing non-small cell lung carcinoma patients with or without EGFR/KRAS mutation [25]. In this study, we found RPISeq offered a convenient and inexpensive method for computational construction of RNA-protein interaction networks, and provided useful insights into the function of long non-coding RNAs.

Conclusions

In the present study, CCAT1 and SOX2OT could be detected in peripheral blood, showing dif-

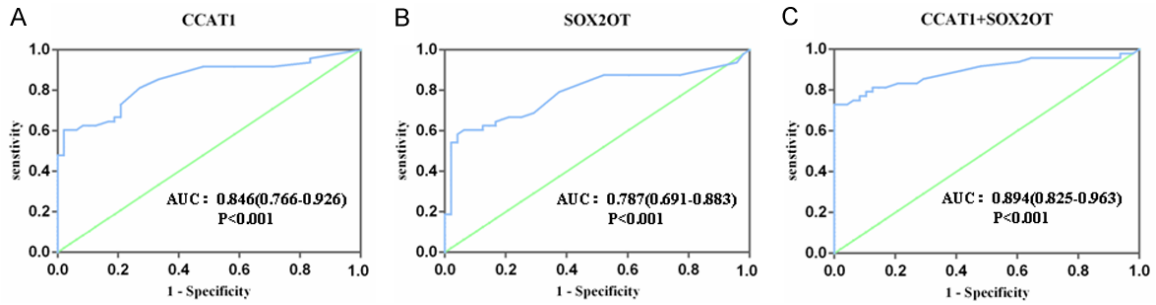

Figure 3. The ROC curve of CCAT1 and SOX2OT expression levels in serum for NSCLC diagnosis. ROC-CCAT1 (A) and ROC-SOX2OT (B) for detecting NSCLC from normal controls. ROC curves of a combination of CCAT1 and SOX2OT to discriminate NSCLC from normal controls (C).

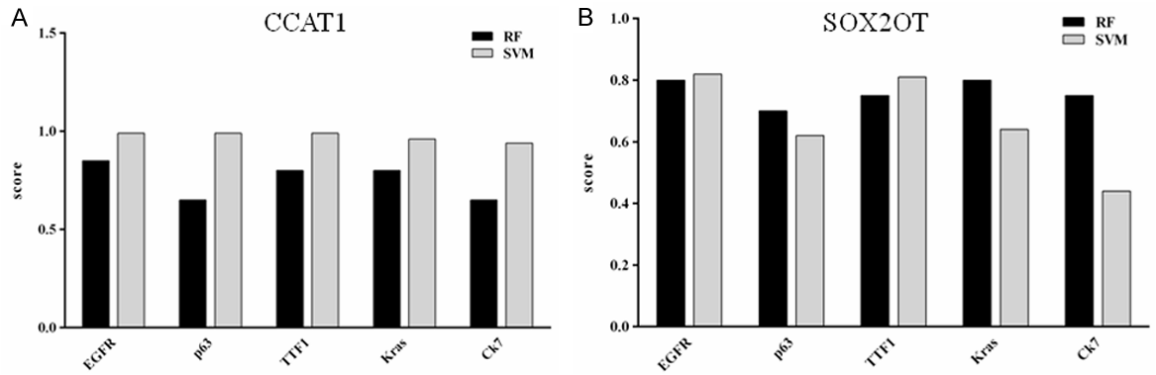

Figure 4. The scores of the interaction probability between lncRNAs and lung cancer associated protein predicted by RPISeq. A. The scores of the interaction probability between CCAT1 and lung cancer associated proteins. B. The scores of the interaction probability between SOX2OT and lung cancer associated proteins.

ferent expression levels between NSCLC patients and cancer-free controls. It was demonstrated that CCAT1 and SOX2O complied with key characteristics of diagnostic biomarkers, being minimally-invasive, exhibiting high specificity, and robustness. Moreover, the combination of CCAT1 and SOX2OT had a higher positive diagnostic rate of NSCLC than CCAT1 or SOX2OT alone.

Disclosure of conflict of interest

None.

Address correspondence to: Hongbin Wang, College of Marine Life and Fisheries, Jiangsu Ocean University, Jiangsu, China. E-mail: zstu_wy@163.com

References

[1] Tan DS, Yom SS, Tsao MS, Pass HI, Kelly K, Peled N, Yung RC, Wistuba II, Yatabe Y, Unger M, Mack PC, Wynes MW, Mitsudomi T, Weder W, Yankelevitz D, Herbst RS, Gandara DR, Carbone DP, Bunn PA Jr, Mok TS and Hirsch FR. The International Association for The Study of Lung Cancer Consensus statement on optimizing management of EGFR mutation-positive non-small cell lung cancer: status in 2016. J Thorac Oncol 2016; 11: 946-63.

[2] Molina JR, Yang P, Cassivi SD, Schild SE and Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. Mayo Clin Proc 2008; 83: 584-94.

[3] Khan S and Scholey JM. Assembly, functions, and evolution of archaella, flagella, and cilia. Curr Biol 2018; 28: R278-R292.

[4] Chen W, Zhu H, Yin L, Wang T, Wu J, Xu J, Tao H, Liu J and He X. lncRNA-PVT1 facilitates invasion through upregulation of MMP9 in nonsmall cell lung cancer cell. DNA Cell Biol 2017; 36: 787-793.

[5] Li S, Mei Z, Hu HB and Zhang X. The lncRNA MALAT1 contributes to non-small cell lung cancer development by modulating miR-124/STAT3 axis. J Cell Physiol 2018; 233: 6679-6688.

[6] Wang X, Zhang G, Cheng Z, Dai L, Jia L, Jing X, Wang H, Zhang R, Liu M, Jiang T, Yang Y and Yang M. Knockdown of lncRNA-XIST suppress-

es proliferation and TGF-β1-induced EMT in NSCLC through the notch-1 pathway by regulation of miR-137. Genet Test Mol Biomarkers 2018; 22: 333-342.
[7] Liao Y, Cheng S, Xiang J and Luo C. lncRNA CCHE1 increased proliferation, metastasis, and invasion of non-small lung cancer cells and predicted poor survival in non-small lung cancer patients. Eur Rev Med Pharmacol Sci 2018; 22: 1686-1692.
[8] Xiang JF, Yin QF, Chen T, Zhang Y, Zhang XO, Wu Z, Zhang S, Wang HB, Ge J, Lu X, Yang L and Chen LL. Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. Cell Res 2014; 24: 513-31.
[9] Jiang XM, Li ZL, Li JL, Zheng WY, Li XH, Cui YF and Sun DJ. LncRNA CCAT1 as the unfavorable prognostic biomarker for cholangiocarcinoma. Eur Rev Med Pharmacol Sci 2017; 21: 1242-1247.
[10] Zhu H, Zhao H, Zhang L, Xu J, Zhu C, Zhao H and Lv G. Dandelion root extract suppressed gastric cancer cells proliferation and migration through targeting lncRNA-CCAT1. Biomed Pharmacother 2017; 93: 1010-1017.
[11] Lin H, Cheng W, Yan H and Zhang X. Overexpression of the long noncoding RNA CCAT1 promotes metastasis via epithelial-to-mesenchymal transition in lung adenocarcinoma. Oncol Lett 2018; 16: 1809-1814.
[12] Tosetti V, Sassone J, Ferri ALM, Taiana M, Bedini G, Nava S, Brenna G, Di Resta C, Pareyson D, Di Giulio AM, Carelli S, Parati EA and Gorio A. Transcriptional role of androgen receptor in the expression of long non-coding RNA Sox-2OT in neurogenesis. PLoS One 2017; 12: e0180579.
[13] Qu F and Cao P. Long noncoding RNA SOX2OT contributes to gastric cancer progression by sponging miR-194-5p from AKT2. Exp Cell Res 2018; 369: 187-196.
[14] Shi XM and Teng F. Up-regulation of long noncoding RNA Sox2ot promotes hepatocellular carcinoma cell metastasis and correlates with poor prognosis. Int J Clin Exp Pathol 2015; 8: 4008-14.
[15] Liu S, Xu B and Yan D. Enhanced expression of long non-coding RNA Sox2ot promoted cell proliferation and motility in colorectal cancer. Minerva Med 2016; 107: 279-86.
[16] Hou Z, Zhao W, Zhou J, Shen L, Zhan P, Xu C, Chang C, Bi H, Zou J, Yao X, Huang R, Yu L and Yan J. A long noncoding RNA Sox2ot regulates lung cancer cell proliferation and is a prognostic indicator of poor survival. Int J Biochem Cell Biol 2014; 53: 380-8.
[17] 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.
[18] Muppirala UK, Honavar VG and Dobbs D. Predicting RNA-protein interactions using only sequence information. BMC Bioinformatics 2011; 12: 489.
[19] Cui D, Yu CH, Liu M, Xia QQ, Zhang YF and Jiang WL. Long non-coding RNA PVT1 as a novel biomarker for diagnosis and prognosis of non-small cell lung cancer. Tumour Biol 2016; 37: 4127-34.
[20] Zhang E, Li W, Yin D, De W, Zhu L, Sun S and Han L. Erratum to: c-Myc-regulated long non-coding RNA H19 indicates a poor prognosis and affects cell proliferation in non-small-cell lung cancer. Tumour Biol 2016; 37: 5653.
[21] Jia X, Wang Z, Qiu L, Yang Y, Wang Y, Chen Z, Liu Z and Yu L. Upregulation of Lnc RNA-HIT promotes migration and invasion of non-small cell lung cancer cells by association with ZEB 1. Cancer Med 2016; 5: 3555-3563.
[22] Li Q, Shao Y, Zhang X, Zheng T, Miao M, Qin L, Wang B, Ye G, Xiao B and Guo J. Plasma long noncoding RNA protected by exosomes as a potential stable biomarker for gastric cancer. Tumour Biol 2015; 36: 2007-12.
[23] Smits AJ, Kummer JA, Hinrichs JW, Herder GJ, Scheidel-Jacobse KC, Jiwa NM, Ruijter TE, Nooijen PT, Looijen-Salamon MG, Ligtenberg MJ, Thunnissen FB, Heideman DA, de Weger RA and Vink A. EGFR, and KRAS mutations in lung carcinomas in the Dutch population: increased EGFR mutation frequency in malignant pleural effusion of lung adenocarcinoma. Cell Oncol (Dordr) 2012; 35: 189-96.
[24] Ueda M, Iguchi T, Nambara S, Saito T, Komatsu H, Sakimura S, Hirata H, Uchi R, Takano Y, Shinden Y, Eguchi H, Masuda T, Sugimachi K, Yamamoto H, Doki Y, Mori M and Mimori K. Overexpression of transcription termination factor 1 is associated with a poor prognosis in patients with colorectal cancer. Ann Surg Oncol 2015; 22 Suppl 3: S1490-8.
[25] Thunnissen E, Boers E, Heideman DA, Grünberg K, Kuik DJ, Noorduin A, van Oosterhout M, Pronk D, Seldenrijk C, Sietsma H, Smit EF, van Suylen R, von der Thusen J, Vrugt B, Wiersma A, Witte BI and den Bakker M. Correlation of immunohistochemical staining p63 and TTF-1 with EGFR and K-ras mutational spectrum and diagnostic reproducibility in non small cell lung carcinoma. Virchows Arch 2012; 461: 629-38.