Int J Clin Exp Med 2015;8(4):4862-4882
www.ijcem.com /ISSN:1940-5901/IJCEM0005917

Original Article

Gene microarray analysis of lncRNA and mRNA expression profiles in patients with hypopharyngeal squamous cell carcinoma

Jieyu Zhou[1*], Wenming Li[1*], Tong Jin[1], Xuan Xiang[1], Maocai Li[1], Juan Wang[1], Guojun Li[2], Xinliang Pan[1], Dapeng Lei[1]

[1]Department of Otorhinolaryngology, Qilu Hospital, Shandong University; Key Laboratory of Otolaryngology, Ministry of Health, Jinan, Shandong 250012, P. R. China; [2]Department of Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [*]Equal contributors.

Received January 14, 2015; Accepted March 19, 2015; Epub April 15, 2015; Published April 30, 2015

Abstract: Background: Studies have shown that long noncoding RNAs (lncRNAs) are involved in the development and progression of many types of cancer. However, the mechanisms by which lncRNAs influence development and progression of hypopharyngeal squamous cell carcinoma (HSCC) are unclear. Method: We investigated differences in lncRNA and mRNA expression profiles between 3 pairs of HSCC tissues and adjacent nontumor tissues by microarray analysis. Results: In HSCC tissues, 1299 lncRNAs were significantly upregulated (n=669) or downregulated (n=630) compared to levels in adjacent nontumor tissues. Moreover, 1432 mRNAs were significantly upregulated (n=684) or downregulated (n=748) in HSCC tissues. We randomly selected 2 differentially expressed lncRNAs (AB209630, AB019562) and 2 differentially expressed mRNAs (SPP1, TJP2) for confirmation of microarray results using qRT-PCR. The qRT-PCR results matched well with the microarray data. The differentially expressed lncRNAs and mRNAs were distributed on each of the chromosomes, including the X and Y chromosomes. Pathway analysis indicated that the biological functions of differentially expressed mRNAs were related to 48 cellular pathways that may be associated with HSCC development. GO analysis revealed that 593 mRNAs involved in biological processes, 50 mRNAs involved in cellular components, and 46 mRNAs involved in molecular functions were upregulated in the carcinomas; 280 mRNAs involved in biological processes, 58 mRNAs involved in cellular components, and 71 mRNAs involved in molecular functions were downregulated in the carcinomas. In addition, 8 enhancer-like lncRNAs and 21 intergenic lncRNAs with their adjacent mRNA pairs were identified as coregulated transcripts. Conclusion: These findings provide insight into the mechanisms underlying HSCC tumorigenesis and will facilitate identification of new therapeutic targets and diagnostic biomarkers for this disease.

Keywords: Hypopharyngeal squamous cell carcinoma, lncRNA, mRNA, microarray, expression profile

Introduction

Hypopharyngeal squamous cell carcinoma (HSCC), a malignant neoplasm arising from the mucosa of the upper aerodigestive tract, is one of the most aggressive cancers in the head and neck area [1]. Even though surgical resection, radiation therapy, and neoadjuvant chemotherapy for HSCC are continuously improving, patients with HSCC remain exceedingly vulnerable to relapse and death [2]. The 5-year survival rate is only approximately 25% to 40% [2]. Studies to date have resulted in a large body of valuable experimental evidence regarding the cellular and molecular mechanisms of HSCC [3-6]. However, it is difficult to explain the germination of HSCC through a single molecule or gene, and we have not found specific markers for HSCC [7].

For the past several decades, cancer investigation has been focused mostly on protein-coding genes. In recent years, however, it has been well reported that the non-protein-coding portion of the human genome is also crucial for cancer biology [8]. Long noncoding RNAs (lncRNAs, >200 nucleotides) are defined as non-protein-coding RNAs distinct from housekeeping RNAs such as tRNAs, rRNAs, and snRNAs and independent from small RNAs such as microRNAs and piwiRNAs [9, 10]. LncRNAs play important roles in almost every

aspect of cell biology, including chromosome remodeling, transcription, and posttranscriptional processing [11-14]. Altered expression of lncRNAs is a feature of many types of cancer and has been shown to promote the development, invasion, and metastasis of tumors by a variety of mechanisms [15, 16]. Moreover, as mature lncRNA is the functional end product, the level of lncRNA expression correlates directly with the level of the active molecule [8]. Thus, the use of noncoding RNAs in diagnostics has intrinsic advantages over the use of protein-coding RNAs [8].

The expression and functional significance of lncRNAs in HSCC remain unclear. We hypothesized that lncRNAs, in combination with mRNAs, are involved in the germination and development of HSCC. To test this hypothesis and attempt to identify specific genes that may prove helpful for the diagnosis, treatment, and prevention of HSCC, we examined the genome-wide expression levels of both lncRNAs and mRNAs in HSCC tissues and paired adjacent nontumor tissues by microarray analysis. We identified numerous lncRNAs and mRNAs that were differentially regulated between HSCCs and paired nontumor tissues.

Materials and methods

Patients and tissue specimens

We retrospectively analyzed tissue samples and patient data from patients who had undergone surgical treatment for primary HSCC at Qilu Hospital of Shandong University, Jinan, China. All patients had a pathological diagnosis of HSCC before surgery. We retrospectively analyzed 23 tissue samples and patient data from patients who had undergone surgical treatment for primary HSCC between November 2012 and April 2013. All patients had a pathological diagnosis of HSCC before surgery. Primary tumor subsite, clinical stage, treatment, and vital status were abstracted from the medical records. Patients who had received neoadjuvant chemotherapy or radiation therapy before surgery were excluded from this study. Three primary HSCC samples and 3 paired adjacent nontumor tissue samples were used for global profiling of human lncRNA and mRNA expression using the Arraystar Human lncRNA Microarray (Arraystar, Rockville, MD, USA). Additionally, HSCC specimens and matched noncancerous mucosal epithelial tissues from 20 patients were obtained for confirmation of differential lncRNA and mRNA expression by qRT-PCR. The study protocol was approved by the institutional review board (IRB) of the Ethics Boards of Qilu Hospital (the permit number is 12040), and tissue specimen acquisition was carried out in accordance with institutional guidelines. All subjects signed written informed consent, and this consent procedure was approved by the IRB of the Ethics Boards of Qilu Hospital.

RNA extraction and microarray hybridization

Fresh tissue specimens had been stored immediately in liquid nitrogen for total RNA extraction. Total RNA was extracted from each sample using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions. RNA concentration was quantified with the NanoDrop ND-1000 (NanoDrop Technologies/Thermo Scientific, Wilmington, DE, USA), and RNA integrity was assessed by standard denaturing agarose gel electrophoresis. RNA from the 3 HSCC specimens and 3 paired nontumor tissue specimens was employed for microarray analysis. Sample preparation and microarray hybridization were performed according to the manufacturer's standard protocols with minor modifications. Briefly, total RNA was purified after removal of rRNA and tRNA (using an mRNA-ONLY Eukaryotic mRNA Isolation Kit, Epicentre, Madison, WI, USA). Then, each sample was amplified and transcribed into fluorescent cRNA along the entire length of the transcript, without 3' bias, utilizing a random priming method. The labeled cRNAs were hybridized onto the Human lncRNA Array v2.0 (8×60K, Arraystar). After the slides were washed, the arrays were scanned with an Agilent G2505C scanner (Agilent Technologies, Santa Clara, CA, USA).

Microarray data analysis

Agilent Feature Extraction software (v11.0.1.1) was used to analyze the acquired array images. Quantile normalization and subsequent data processing were performed using the GeneSpring GX v11.5.1 software package (Agilent Technologies). After quantile normalization of the raw data, lncRNAs and mRNAs that were flagged as Present or Marginal ("All Targets Value") in all 6 samples were chosen for further data analysis. Statistically significant differential expression of lncRNAs and mRNAs between

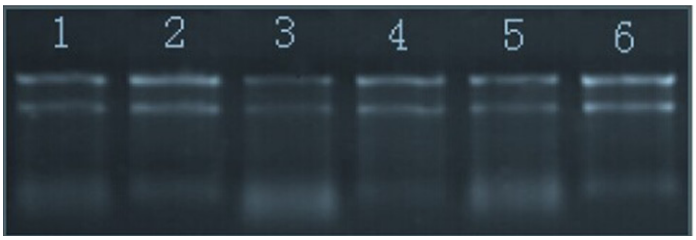


Figure 1. Analysis of RNA integrity and genomic DNA contamination through electrophoresis on a denaturing agarose gel. The 28 s and 18 s rRNA bands are clear and intact. Moreover, the 28 s band is twice as intense as the 18 s band.

Table 1. RNA quantification and quality assurance by NanoDrop ND-1000

Sample ID	OD260/280	OD260/230	Concentration, ng/μL	Volume, μL	Quantity, ng	Quality control pass or fail
PB	1.93	2.29	372.07	20	7441.40	Pass
TB	1.97	2.35	719.20	80	57536.00	Pass
P12068476	1.89	2.22	390.66	20	7813.20	Pass
T12068476	1.98	2.33	674.70	80	53976.00	Pass
P12071728	1.94	2.30	600.29	40	24011.60	Pass
T12071728	2.01	2.35	1031.40	80	82512.00	Pass

HSCC and paired nontumor tissue was identified through volcano plot filtering. Hierarchical clustering was performed to distinguish between the lncRNA and mRNA expression patterns among the samples. The differentially expressed mRNAs were submitted to the KEGG (Kyoto Encyclopedia of Genes and Genomes) database for pathway analysis and were then submitted to the GO (Gene Ontology) database for GO category analysis. LncRNAs with enhancer-like functions were identified using a GENCODE annotation [17] of the human genes [18]. Rinn lncRNA [19, 20] profiling and homeobox (HOX) cluster profiling [21] were analyzed based on papers published by the John Rinn laboratory. The differentially expressed lncRNAs, especially the enhancer-like lncRNAs and the Rinn intergenic lncRNAs (lincRNAs), were remapped on the genome and their nearby coding gene pairs (distance <300 kb) to identify for lncRNA-mRNA coexpression analysis.

Validation of the differentially expressed lncRNAs and mRNAs by quantitative real-time PCR (qRT-PCR)

qRT-PCR was used to validate the microarray data among 20 HSCC patients. Total RNA was reverse-transcribed to cDNA using PrimeScript Reverse Transcriptase (Takara, Dalian, China) following the manufacturer's protocol. qRT-PCR was performed using SYBR Green chemistry in the ABI 7900HT sequence detection machine (ABI Applied Biosystems, Foster City, CA, USA). The gene-specific primers used were as follows: AB019562, 5'-GGATGTCAGGTCTGCGAAACT-3' (sense), and 5'-GATAGTGTGGTTTATGGACTGAGGT-3' (antisense); AB209630, 5'-GGGCTATTGTCCCTAAGTTGAT-3' (sense), and 5'-TGTCTTGTAGAGCATAAGGAAACC-3' (antisense); SPP1, 5'-ACCTGCCAGCAACCGAAGT-3' (sense), and 5'-GGTGATGTCCTCGTCTGTAGCA-3' (antisense); TJP2, 5'-GCAGAGCGAACGAAGAGTATG-3' (sense), and 5'-ATGACGGGATGTTGATGAGG-3' (antisense); GAPDH, 5'-GGGAAACTGTGGCGTGAT-3' (sense), and 5'-GAGTGGGTGTCGCTGTTGA-3' (antisense). PCR was performed in a 10-μL reaction volume and consisted of an initial denaturation step at 95°C for 30 sec followed by amplification with 40 cycles at 95°C for 5 sec and 60°C for 30 sec. The threshold cycle (Ct) was defined as the cycle number at which the fluorescence passed a predetermined threshold. Both target and reference (GAPDH) genes were amplified in separate wells in triplicate. Gene expression was calculated using the comparative threshold cycle ($2^{-\Delta\Delta CT}$) method.

Statistical methods

SPSS 18.0 (SPSS Inc., Chicago, IL, USA) and GraphPad Prism 5.0 (GraphPad Software Inc., San Diego, CA, USA) statistical software were

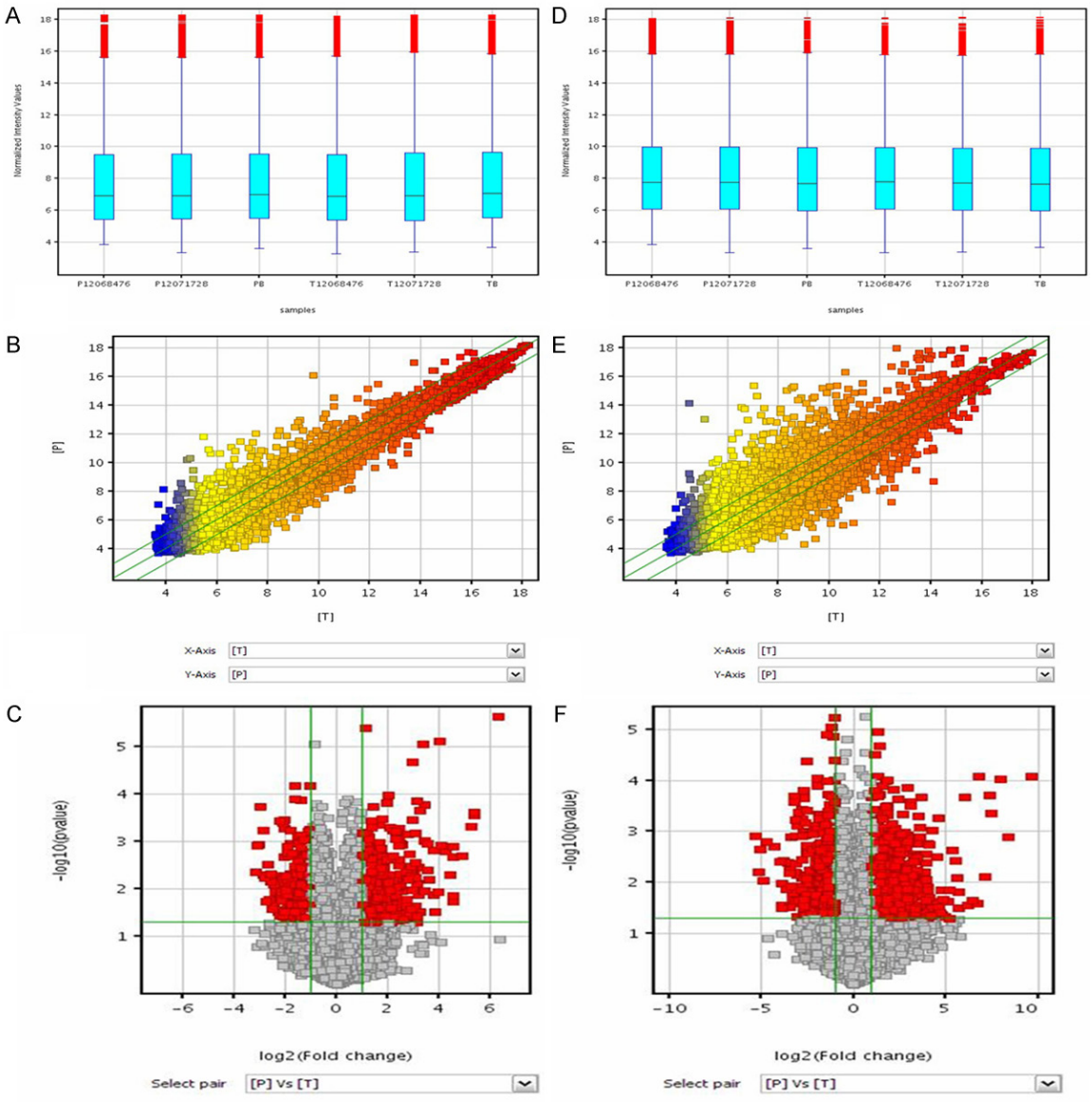


Figure 2. Differences in LncRNA and mRNA expression profiles between HSCC tissues and adjacent nontumor tissues. A and D: Box plots. All 6 samples in the dataset were normalized. For both lncRNA and mRNA, the distributions of log2 ratios were nearly the same in all tested samples. B and E: Scatter plots. The values plotted on the X and Y axes are the averaged normalized signal values in each group (log2 scaled). The green lines indicate the fold change. The middle line indicates a fold change of 1, or no difference in expression between HSCC and adjacent nontumor tissue. The values above the top green line and below the bottom green line indicate more than 2.0-fold difference between HSCC and nontumor tissue samples. C and F: Volcano plots. Volcano plots show the relationship between magnitude of expression difference and statistical significance. They also allow subsets of genes to be isolated on the basis of those values. The vertical green lines correspond to 2.0-fold upregulation and 2.0-fold downregulation of expression, and the horizontal green line indicates a *P* value of 0.05. Thus, the red points in the plot represent lncRNAs with statistically significant differential expression.

employed for the analysis. The statistical significance of the microarray results was analyzed by fold change and Student's *t*-test. The false discovery rate was calculated to correct the *P* value. The threshold value used to screen differentially expressed lncRNAs and mRNAs was a fold change of ≥2.0. The Wilcoxon matched pairs test was used to compare the RNA expression levels in tumors versus adjacent nontumor tissues. In all analyses, a 2-sided *P* value <0.05 was considered statistically significant.

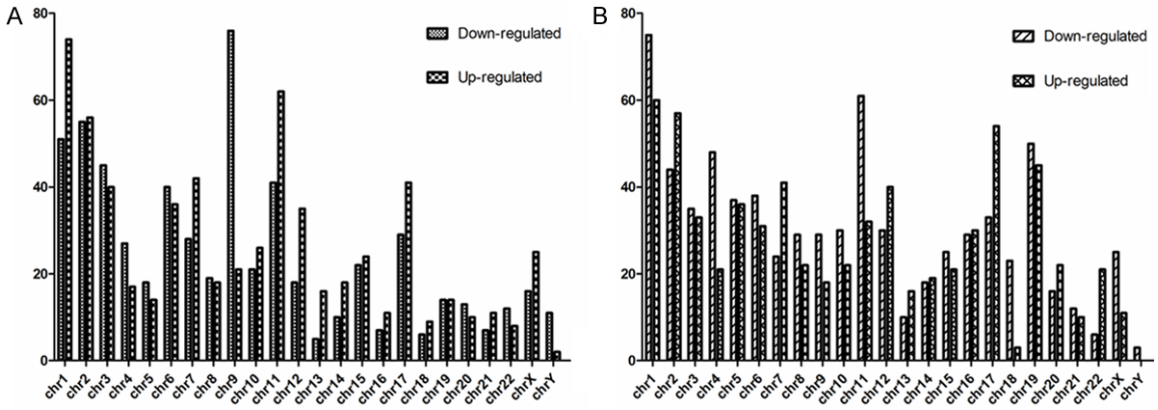


Figure 3. Chromosomal distribution of differentially expressed lncRNAs and mRNAs. Most of the differentially expressed lncRNAs were found on chromosomes 1, 2, 11, 9, and 3. Most of the differentially expressed mRNAs were found on chromosomes 1, 2, 19, 11, and 17.

Results

Quality of the sample RNAs

The integrity of RNAs was assessed by electrophoresis on a denaturing agarose gel. Intact total RNA run on a denaturing gel will have sharp 28S and 18S rRNA bands, and the 28S rRNA band should be approximately twice as intense as the 18S rRNA band. This 2:1 intensity ratio was observed for the RNA in this study (**Figure 1**), indicating that the RNA was intact. The concentration of RNAs (OD260), protein contamination of RNAs (ratio OD260/OD280), and organic compound contamination of RNAs (ratio OD260/OD230) were measured with the NanoDrop ND-1000. All samples had OD260/OD280 ratios of total RNA higher than 1.8, indicating adequate RNA concentration (**Table 1**).

Overview of lncRNA and mRNA profiles

Box plots: Box plots were used to compare the distributions of the intensities of the samples. After normalization, for both lncRNA and mRNA, the distributions of log2 ratios were nearly the same in all tested samples (**Figure 2A**, **2D**).

Scatter plots: Scatter plots were used to visualize differences in lncRNA and mRNA expression between the HSCC and nontumor tissue samples. The values plotted on the X and Y axes are the averaged normalized signal values of groups of samples (log2 scaled). The green lines are fold change lines (the default fold change value given is 2.0). The values above the top green line and below the bottom green line indicate more than 2.0-fold difference between HSCC and nontumor samples (**Figure 2B**, **2E**).

Volcano plot filtering: Volcano plot filtering was used to identify lncRNAs and mRNAs with statistically significant differences in expression between HSCC and nontumor samples (fold change ≥2.0, *P* value cut-off 0.05) (**Figure 2C**, **2F**). The microarray data showed that 1299 lncRNAs were significantly differentially expressed; of those, 669 were upregulated and 630 were downregulated in the carcinomas compared to the adjacent nontumor tissues. In addition, 1432 mRNAs were differentially expressed; of those, 684 were upregulated and 748 were downregulated in the carcinomas compared to the adjacent nontumor tissues. The differentially expressed lncRNAs and mRNAs were distributed on each of the chromosomes, including the X and Y chromosomes (**Figure 3**). Most of the differentially expressed lncRNAs were found on chromosomes 1, 2, 11, 9, and 3. Most of the differentially expressed mRNAs were found on chromosomes 1, 2, 19, 11, and 17. Volcano plot filtering was also used to identify the 10 most upregulated and downregulated lncRNAs (**Table 2**) and mRNAs (**Table 3**) in HSCC tissues.

Heat map and hierarchical clustering: Hierarchical clustering allowed us to hypothesize the relationships between HSCCs and adjacent nontumor tissues with regard to lncRNA (**Figure 4A**) and mRNA (**Figure 4B**) expression patterns.

Table 2. Ten most upregulated and downregulated lncRNAs in the carcinomas compared to the adjacent nontumor tissues by volcano plot

Probe name	FC Absolute	Regulation	Seqname	Gene symbol	Source	Chromosome	Relationship	Associated gene name	Associated protein name
ASHG19A3A021323	8.940928	Up	ENST00000478252	RP13-503K1.1	Ensembl	3	Intronic antisense	NSUN3	Putative methyltransferase NSUN3
ASHG19A3A022918	8.39297	Up	AL359062		misc_RNA	3	Intergenic		
ASHG19A3A042290	8.289279	Up	ENST00000433897	RP11-397A15.4	Ensembl	1	Intergenic		
ASHG19A3A045578	7.7143497	Up	AK002107		misc_RNA	1	Exon sense-overlapping	RAB3B	Ras-related protein Rab-3B
ASHG19A3A040116	7.2975035	Up	ENST00000447818	RP11-325E14.3	Ensembl	X	Intergenic		
ASHG19A3A034842	7.117814	Up	ENST00000419422	RP11-132A1.4	Ensembl	7	Intergenic		
ASHG19A3A030085	6.985488	Up	ENST00000399980	RP11-325M4.1	Ensembl	6	Intergenic		
ASHG19A3A049457	6.731179	Up	AK021444		NRED	13	Exon sense-overlapping	POSTN	Periostin isoform 4
ASHG19A3A049457	6.731179	Up	AK021444		NRED	13	Exon sense-overlapping	POSTN	Periostin isoform 3
ASHG19A3A049457	6.731179	Up	AK021444		NRED	13	Exon sense-overlapping	POSTN	Periostin isoform 2
ASHG19A3A007356	80.72981	Down	NR_024602	CLCA4	RefSeq_NR	1	Exon sense-overlapping	CLCA4	Calcium-activated chloride channel regulator 4
ASHG19A3A049241	42.46477	Down	ENST00000453176	RP11-38M15.6	Ensembl	13	Intergenic		
ASHG19A3A018626	42.141758	Down	NR_026755	C21orf15	RefSeq_NR	21	Intergenic		
ASHG19A3A030827	38.225327	Down	uc003qvy.1	AL832737	UCSC_knowngene	6	Intergenic		
ASHG19A3A010514	29.642584	Down	NR_026756	LOC284233	RefSeq_NR	18	Intergenic		
ASHG19A3A018630	23.840233	Down	ENST00000451663	C21orf81	Ensembl	21	Intergenic		
ASHG19A3A007578	23.742765	Down	NR_027763	GYS1	RefSeq_NR	19	Bidirectional	RUVBL2	ruvB-like 2
ASHG19A3A007578	23.742765	Down	NR_027763	GYS1	RefSeq_NR	19	Exon sense-overlapping	GYS1	“Glycogen [starch] synthase, muscle isoform 2”
ASHG19A3A007578	23.742765	Down	NR_027763	GYS1	RefSeq_NR	19	Exon sense-overlapping	GYS1	“Glycogen [starch] synthase, muscle isoform 1”
ASHG19A3A051379	23.452465	Down	BC008699		misc_RNA	14	Intergenic		

Table 3. Ten most upregulated and downregulated mRNAs in the carcinomas compared to the adjacent nontumor tissues by volcano plot

Probe name	FC Absolute	Regulation	Seqname	Gene symbol	Chromosome	Protein accession	Product
ASHG19A3A020207	41.58339	Up	NM_005940	MMP11	22	NP_005931	Matrix metallopeptidase 11 (stromelysin 3)
ASHG19A3A050427	35.78423	Up	NM_007129	ZIC2	13	NP_009060	Zic family member 2 (odd-paired homolog, Drosophila)
ASHG19A3A049790	34.567513	Up	NM_033132	ZIC5	13	NP_149123	Zic family member 5 (odd-paired homolog, Drosophila)
ASHG19A3A030137	30.225765	Up	NM_001126063	KHDC1L	6	NP_001119535	KH homology domain containing 1-like
ASHG19A3A017482	26.772041	Up	NM_001898	CST1	20	NP_001889	Cystatin SN
ASHG19A3A025453	19.727985	Up	NM_004181	UCHL1	4	NP_004172	Ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase)
ASHG19A3A019550	15.542832	Up	NM_006115	PRAME	22	NP_996839	Preferentially expressed antigen in melanoma
ASHG19A3A045632	15.119103	Up	NM_002421	MMP1	11	NP_002412	Matrix metallopeptidase 1 (interstitial collagenase)
ASHG19A3A045635	14.78575	Up	NM_002427	MMP13	11	NP_002418	Matrix metallopeptidase 13 (collagenase 3)
CUST_2_PI426249100	14.759951	Up	NM_017410_Exon2+	NM_017410	12		
ASHG19A3A031264	798.7413	Down	NM_001010909	MUC21	6	NP_001010909	Mucin 21, cell surface associated
ASHG19A3A005265	319.68326	Down	NM_022438	MAL	2	NP_071885	Mal, T-cell differentiation protein
ASHG19A3A025361	242.24474	Down	NM_012128	CLCA4	1	NP_036260	Chloride channel accessory 4
ASHG19A3A005267	176.59032	Down	NM_022440	MAL	2	NP_071885	Mal, T-cell differentiation protein
ASHG19A3A024262	162.92448	Down	NM_182502	TMPRSS11B	4	NP_872308	Transmembrane protease, serine 11B
ASHG19A3A008290	138.02931	Down	NM_019016	KRT24	17	NP_061889	Keratin 24
ASHG19A3A004016	106.28888	Down	NM_002371	MAL	2	NP_071885	Mal, T-cell differentiation protein
ASHG19A3A010362	103.64228	Down	NM_002974	SERPINB4	18	NP_002965	Serpin peptidase inhibitor, clade B (ovalbumin), member 4
ASHG19A3A010363	103.04679	Down	NM_006919	SERPINB3	18	NP_008850	Serpin peptidase inhibitor, clade B (ovalbumin), member 3
ASHG19A3A053404	85.175095	Down	NM_031948	PRSS27	16	NP_114154	Protease, serine 27

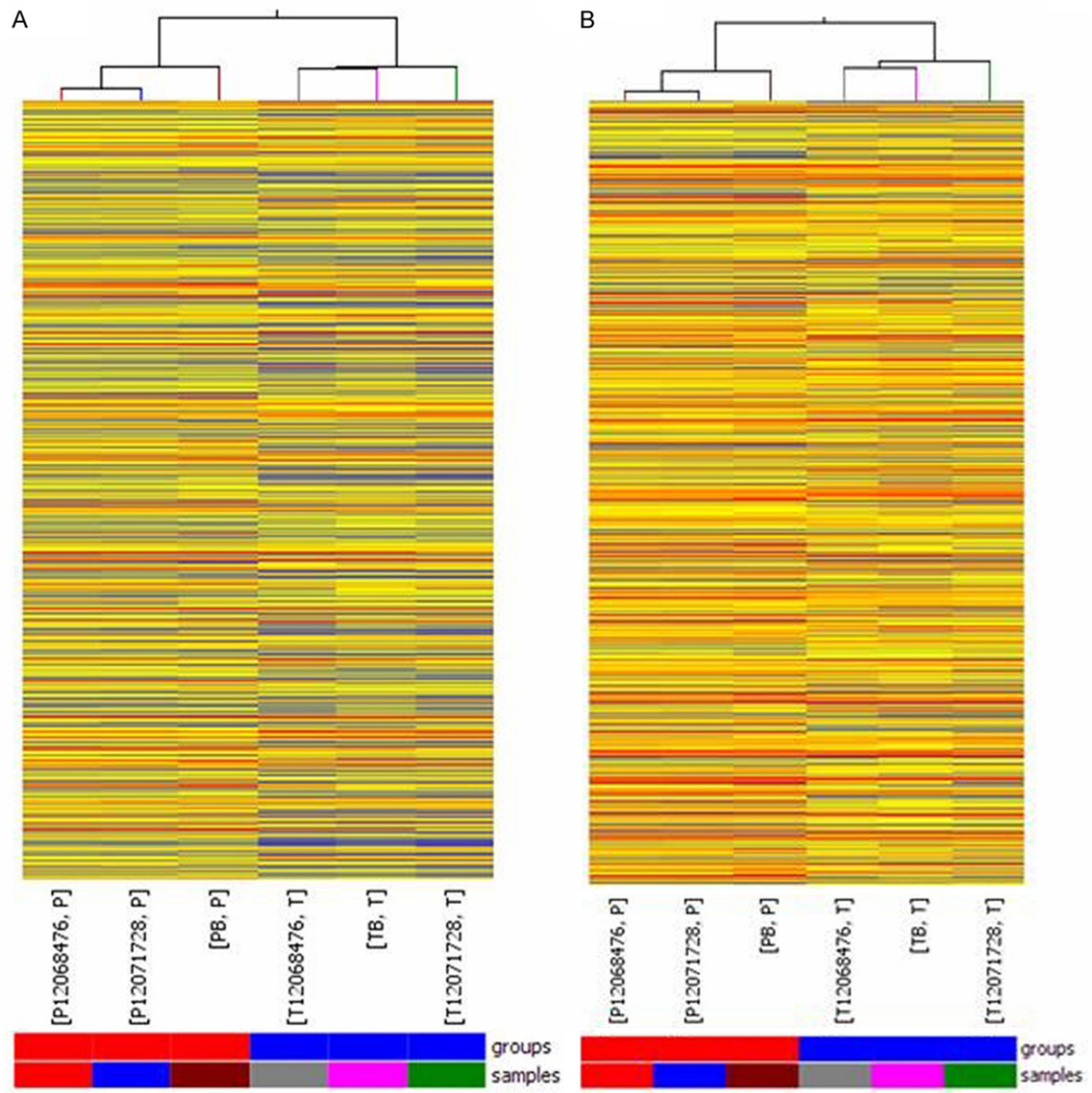


Figure 4. Heat map and hierarchical clustering of differences in lncRNA (A) and mRNA (B) expression profiles between HSCC tissues and adjacent nontumor tissues. Heat map and hierarchical clustering is one of the most widely used clustering methods for analyzing lncRNA and mRNA expression data. Cluster analysis arranges samples into groups based on their expression levels, which allowed us to hypothesize the relationships between HSCC tissues and adjacent nontumor tissues. In the dendrogram, red indicates high relative expression, and blue indicates low relative expression.

Confirmation of differential expression by qRT-PCR in a cohort of 20 patients with HSCC

To confirm the reliability and validity of the microarray data, we randomly selected 2 differentially expressed lncRNAs (AB209630, AB019562) and 2 differentially expressed mRNAs (SPP1, TJP2) and analyzed their expression in 20 HSCC samples and paired adjacent nontumor tissue samples with qRT-PCR. The relative expression levels of target RNA were given as ratios of RNA transcript level to GAPDH transcript level in the same RNA sample. AB209630 expression was significantly lower in carcinomas than in adjacent nontumor tissues (P<0.0001, 2.23-fold, **Figure 5A**). AB019562 expression was significantly higher in carcinomas than in adjacent nontumor tissues (P=0.0004, 7.83-fold, **Figure 5B**). SPP1 expression was significantly higher in carcinomas than in adjacent nontumor tissues (P=0.0001, 10.56-fold, **Figure 5C**). TJP2 expression was signifi-

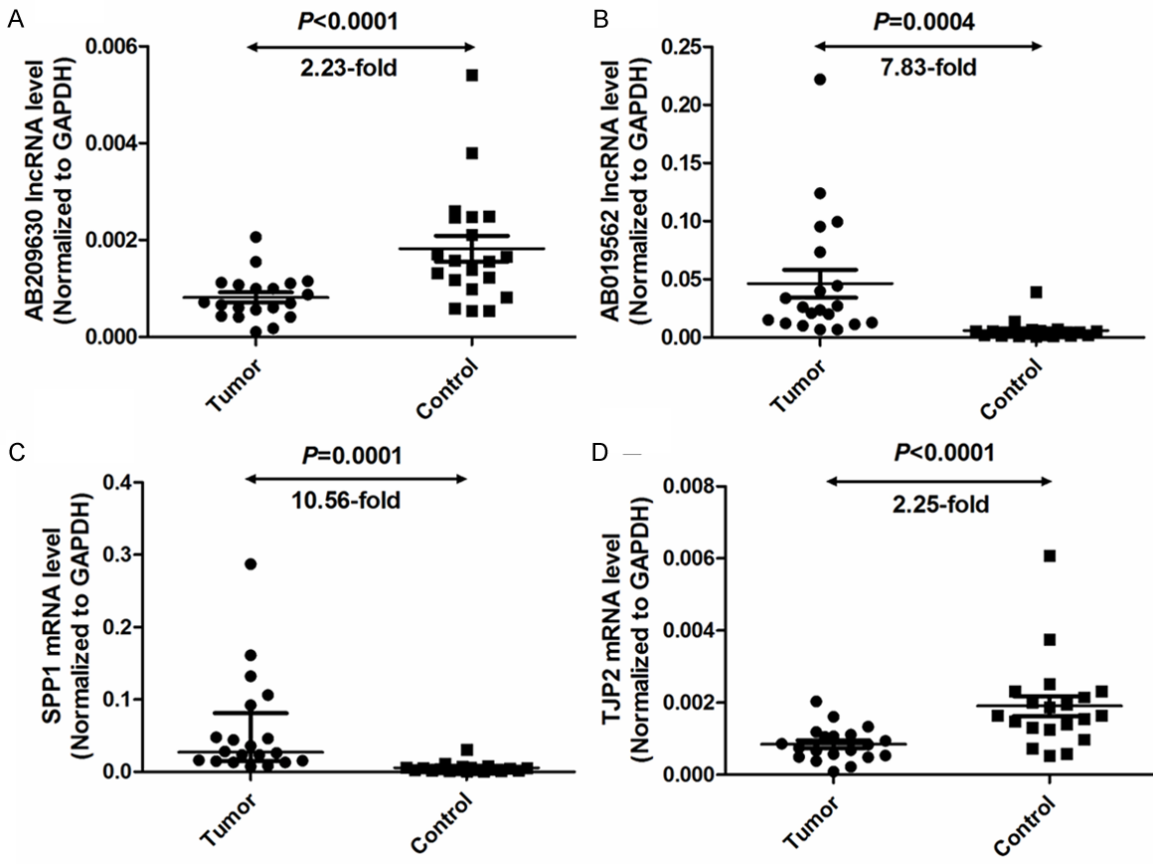


Figure 5. Quantitative determination of lncRNA and mRNA by means of qRT-PCR for lncRNAs AB209630 (A) and AB019562 (B) and mRNAs SPP1 (C) and TJP2 (D) in HSCC tissues or adjacent nontumor tissues. Target RNA relative expression levels were given as ratios of RNA transcript level to GAPDH transcript level in the same RNA sample. Scatter plots were shown with medians with interquartile range. Statistical analyses were performed with the Wilcoxon matched paired test (2-tailed).

cantly lower in carcinomas than in adjacent nontumor tissues (P<0.0001, 2.25-fold, **Figure 5D**). The expression levels of these 4 genes were consistent with the microarray results; thus, these qRT-PCR results confirmed the reliability of the microarray data.

Bioinformatics analysis

Pathway analysis: Pathway analysis based on the latest KEGG database (http://www.genome.jp/kegg/) allows users to determine the biological pathways with significant enrichment of differentially expressed mRNAs. The P value denotes the significance of the pathway, with lower P values indicating greater significance (the P value cut-off to determine significance is 0.05). From the whole pathway analysis, we identified 48 pathways with significant differences in gene expression between HSCC and adjacent nontumor tissues; of those, 29 pathways were upregulated and 19 pathways were downregulated in the carcinomas compared to the adjacent nontumor tissues (**Table 4**). The top 10 upregulated pathways included ECM-receptor interaction, focal adhesion, and melanogenesis signaling (**Figure 6A**). The top 10 downregulated pathways included tight junction, African trypanosomiasis, and peroxisome signaling (**Figure 6B**).

GO analysis: GO analysis is a functional analysis associating differentially expressed mRNAs with GO categories. The GO categories are derived from Gene Ontology (www.geneontology.org), which comprises 3 structured networks of defined terms that describe gene product attributes. The P value denotes the significance of GO Term enrichment in the differentially expressed mRNA list. The lower the P value, the more significant the GO Term (the P value cut-off is 0.05). GO analysis showed that 593 genes

Table 4. Results of KEGG pathway analysis of mRNAs differentially expressed in carcinomas vs adjacent nontumor tissues

Pathway list (selection counts)	Enrichment score	Regulation	Genes
ECM-receptor interaction (14)	6.271957	Up	COL1A1, COL1A2, COL4A2, COL5A2, COL5A3, COL6A1, COL6A2, COL6A3, FN1, ITGA1, ITGA5, LAMC2, SPP1, THBS2
Focal adhesion (21)	5.696962	Up	COL1A1//COL1A2//COL4A2//COL5A2//COL5A3//COL6A1//COL6A2//COL6A3//CTNNB1//FLT1//FN1//ITGA1//ITGA5//LAMC2//MYL9//PDGFB//PRKCA//ROCK2//SPP1//THBS2//VEGFA
Melanogenesis (12)	3.979435	Up	ADCY1//ADCY6//CAMK2D//CREB3L1//CTNNB1//FZD1//FZD2//FZD7//FZD8//FZD9//PRKACA//PRKCA
Amoebiasis (12)	3.69807	Up	ADCY1//COL1A1//COL1A2//COL4A2//COL5A2//COL5A3//FN1//GNA11//HSPB1//LAMC2//PRKACA//PRKCA
Wnt signaling pathway (14)	3.388262	Up	CAMK2D//CTNNB1//FZD1//FZD2//FZD7//FZD8//FZD9//NFATC1//NKD2//PORCN//PPP2CB//PRKACA//PRKCA//ROCK2
Dilated cardiomyopathy (10)	3.090323	Up	ADCY1//ADCY6//ITGA1//ITGA5//PRKACA//SGCB//SGCD//TNNI3//TPM1//TPM2
Protein digestion and absorption (9)	2.899216	Up	COL1A1//COL1A2//COL4A2//COL5A2//COL5A3//COL6A1//COL6A2//COL6A3//COL9A2
Basal cell carcinoma (7)	2.701192	Up	CTNNB1//FZD1//FZD2//FZD7//FZD8//FZD9//GLI3
Vascular smooth muscle contraction (10)	2.33321	Up	ADCY1//ADCY6//ARHGEF1//GNA11//KCNMA1//MYL6B//MYL9//PRKACA//PRKCA//ROCK2
Chagas disease (American trypanosomiasis) (9)	2.158455	Up	ADCY1//CALR//CCL3//CCL3L1//CCL3L3//GNA11//MAPK12//PPP2CB//SERPINE1
Pathogenic Escherichia coli infection (6)	2.051097	Up	CTNNB1//EZR//PRKCA//ROCK2//TUBB4//YWHAZ
Chemokine signaling pathway (13)	2.040568	Up	ADCY1//ADCY6//CCL11//CCL3//CCL3L1//CCL3L3//CCL4//CCL4L1//CXCR2//GNB4//PRKACA//ROCK2//XCL1
Pyruvate metabolism (5)	1.980173	Up	ACSS1//AKR1B1//ALDH1B1//PC//PKM2
Cholinergic synapse (9)	1.954947	Up	ADCY1//ADCY6//CAMK2D//CHRNA3//CREB3L1//GNA11//GNB4//PRKACA//PRKCA
Retrograde endocannabinoid signaling (8)	1.73084	Up	ADCY1//ADCY6//GABRA1//GABRA5//GNB4//MAPK12//PRKACA//PRKCA
Hypertrophic cardiomyopathy (7)	1.724737	Up	ITGA1//ITGA5//SGCB//SGCD//TNNI3//TPM1//TPM2
Salmonella infection (7)	1.648289	Up	CCL3//CCL3L1//CCL3L3//CCL4//CCL4L1//MAPK12//ROCK2
Gap junction (7)	1.575761	Up	ADCY1//ADCY6//GNA11//PDGFB//PRKACA//PRKCA//TUBB4
GABAergic synapse (7)	1.575761	Up	ADCY1//ADCY6//GABRA1//GABRA5//GNB4//PRKACA//PRKCA
Salivary secretion (7)	1.575761	Up	ADCY1//ADCY6//ATP2B4//CST1//KCNMA1//PRKACA//PRKCA
Bile secretion (6)	1.544129	Up	ABCC4//ADCY1//ADCY6//AQP9//PRKACA//SLC27A5
Oocyte meiosis (8)	1.514502	Up	ADCY1//ADCY6//ANAPC11//CAMK2D//MAPK12//PPP2CB//PRKACA//YWHAZ
Morphine addiction (7)	1.506877	Up	ADCY1//ADCY6//GABRA1//GABRA5//GNB4//PRKACA//PRKCA
Gastric acid secretion (6)	1.467153	Up	ADCY1//ADCY6//CAMK2D//EZR//PRKACA//PRKCA
Leukocyte transendothelial migration (8)	1.436409	Up	CTNNB1//EZR//MAPK12//MMP9//MYL9//PRKCA//ROCK2//SIPA1
Arginine and proline metabolism (5)	1.415061	Up	ALDH1B1//GAMT//NAGS//P4HA1//PYCR1
Pathways in cancer (17)	1.386872	Up	BIRC5//COL4A2//CTNNB1//FGFR1//FN1//FZD1//FZD2//FZD7//FZD8//FZD9//GLI3//LAMC2//MMP1//MMP9//PDGFB//PRKCA//VEGFA
Regulation of actin cytoskeleton (12)	1.336989	Up	ABI2//ARHGEF1//EZR//FGFR1//FN1//ITGA1//ITGA5//MYL9//PDGFB//RDX//ROCK2//SSH1
GnRH signaling pathway (7)	1.319729	Up	ADCY1//ADCY6//CAMK2D//GNA11//MAPK12//PRKACA//PRKCA

Tight junction (13)	2.943522	Down	CGN//CLDN3//CLDN7//EPB41L1//HRAS//LLGL2//MAGI3//MYL12B//MYL5//PPP2R1B//PRKCH//TJP1//TJP2
African trypanosomiasis (6)	2.66711	Down	HBA1//HBA2//HBB//HPR//IL18//SELE
Peroxisome (9)	2.65428	Down	ACAA1//ACOX1//ACOX3//CROT//EPHX2//FAR1//HMGCL//HSD17B4//MLYCD
Drug metabolism-other enzymes (7)	2.597677	Down	CES2//CYP3A4//CYP3A43//CYP3A5//CYP3A7//DPYD//TK2
Drug metabolism-cytochrome P450 (8)	2.311757	Down	ALDH3B2//CYP3A4//CYP3A43//CYP3A5//CYP3A7//FMO4//MAOA//MGST2
Malaria (6)	1.964824	Down	HBA1//HBA2//HBB//IL18//KLRB1//SELE
Nicotinate and nicotinamide metabolism (4)	1.839552	Down	C9ORF95//NAPRT1//NT5C3//PNP
Steroid hormone biosynthesis (6)	1.775109	Down	CYP3A4//CYP3A43//CYP3A5//CYP3A7//CYP7B1//SRD5A3
Histidine metabolism (4)	1.727958	Down	ALDH3B2//AMDHD1//FTCD//MAOA
Linoleic acid metabolism (4)	1.675968	Down	CYP3A4//CYP3A43//CYP3A5//CYP3A7
Steroid biosynthesis (3)	1.633233	Down	DHCR24//NSDHL//TM7SF2
Phenylalanine metabolism (3)	1.565825	Down	ALDH3B2//GOT1//MAOA
Influenza A (12)	1.558672	Down	CREBBP//FDPS//IL18//MAPK8//NXF1//PRSS2//PRSS3//RAF1//RNASEL//TLR3//TLR7//TRIM25
Metabolism of xenobiotics by cytochrome P450 (7)	1.550808	Down	ALDH3B2//CBR3//CYP3A4//CYP3A43//CYP3A5//CYP3A7//MGST2
Glycolysi/gluconeogenesis (6)	1.488156	Down	ALDH3B2//BPGM//DLAT//LDHC//PGAM2//PGM2
Alpha-Linolenic acid metabolism (3)	1.38884	Down	ACAA1//ACOX1//ACOX3
Terpenoid backbone biosynthesis (3)	1.38884	Down	FDPS//HMGCR//IDI1
Biosynthesis of unsaturated fatty acids (3)	1.38884	Down	ACAA1//ACOX1//ACOX3
Neurotrophin signaling pathway (9)	1.341531	Down	GAB1//HRAS//MAP3K5//MAPK7//MAPK8//NTRK1//RAF1//RPS6KA1//SORT1

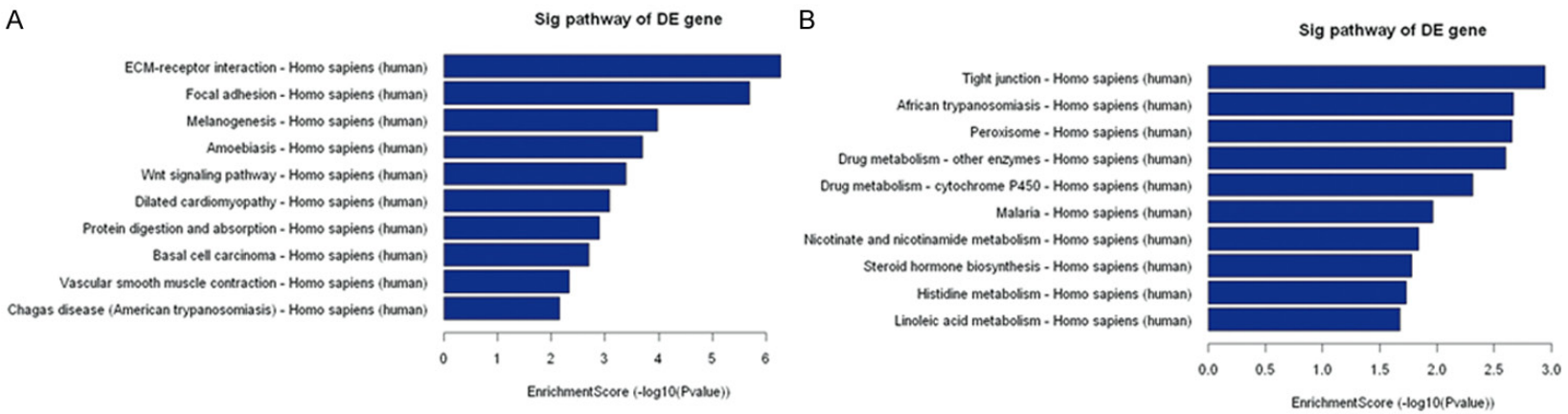


Figure 6. Pathway analysis report. A: The top 10 pathways that were upregulated in the carcinomas compared to the adjacent nontumor tissues. B: The top 10 pathways that were downregulated in the carcinomas compared to the adjacent nontumor tissues.

Table 5. Ten most upregulated and downregulated GO classifications for carcinomas versus adjacent nontumor tissues

Biological process classification	No. of genes	Cellular component classification	No. of genes	Molecular function classification	No. of genes	Regulation
Cellular process	436	Cytoplasm	296	Protein binding	247	Up
Biological regulation	309	Cytoplasmic part	237	Receptor binding	54	Up
Primary metabolic process	296	Extracellular region	102	Structural molecule activity	35	Up
Regulation of biological process	290	Membrane-enclosed lumen	96	Carbohydrate binding	21	Up
Regulation of cellular process	276	Organelle lumen	94	Pattern binding	16	Up
Macromolecule metabolic process	247	Cytosol	86	Polysaccharide binding	16	Up
Response to stimulus	241	Plasma membrane part	79	Glycosaminoglycan binding	15	Up
Multicellular organismal process	225	Extracellular region part	70	Chromatin binding	15	Up
Developmental process	192	Endoplasmic reticulum	57	G-protein coupled receptor binding	14	Up
Cellular response to stimulus	189	Cell projection	47	Protein complex binding	14	Up
Cellular response to stimulus	183	Cell part	508	Binding	426	Down
Cell communication	179	Cell	508	Protein binding	255	Down
Signaling	176	Cytoplasm	341	Catalytic activity	215	Down
Signal transduction	160	Membrane	286	Transferase activity	76	Down
Positive regulation of biological process	129	Cytoplasmic part	255	Oxidoreductase activity	44	Down
Positive regulation of cellular process	121	Cell periphery	185	Enzyme regulator activity	44	Down
Response to chemical stimulus	110	Plasma membrane	182	Calcium ion binding	38	Down
Small molecule metabolic process	108	Extracellular region	106	Protein dimerization activity	36	Down
Regulation of response to stimulus	93	Plasma membrane part	101	Identical protein binding	35	Down
Catabolic process	84	Cytosol	90	Protein homodimerization activity	27	Down

involved in biological processes, 50 genes involved in cellular components, and 46 genes involved in molecular functions were upregulated in the carcinomas compared to the adjacent nontumor tissues; 280 genes involved in biological processes, 58 genes involved in cellular components, and 71 genes involved in molecular functions were downregulated in the carcinomas compared to the adjacent nontumor tissues (data not shown). However, the top 10 upregulated and downregulated GO functions of Counts Enrichment Terms are shown in **Table 5**. The top 10 upregulated GO functions of Fold Enrichment Terms and Score Enrichment Terms are shown in **Figure 7**. The top 10 downregulated GO functions of Fold Enrichment Terms and Score Enrichment Terms are shown in **Figure 8**.

LncRNA classification and subgroup analysis

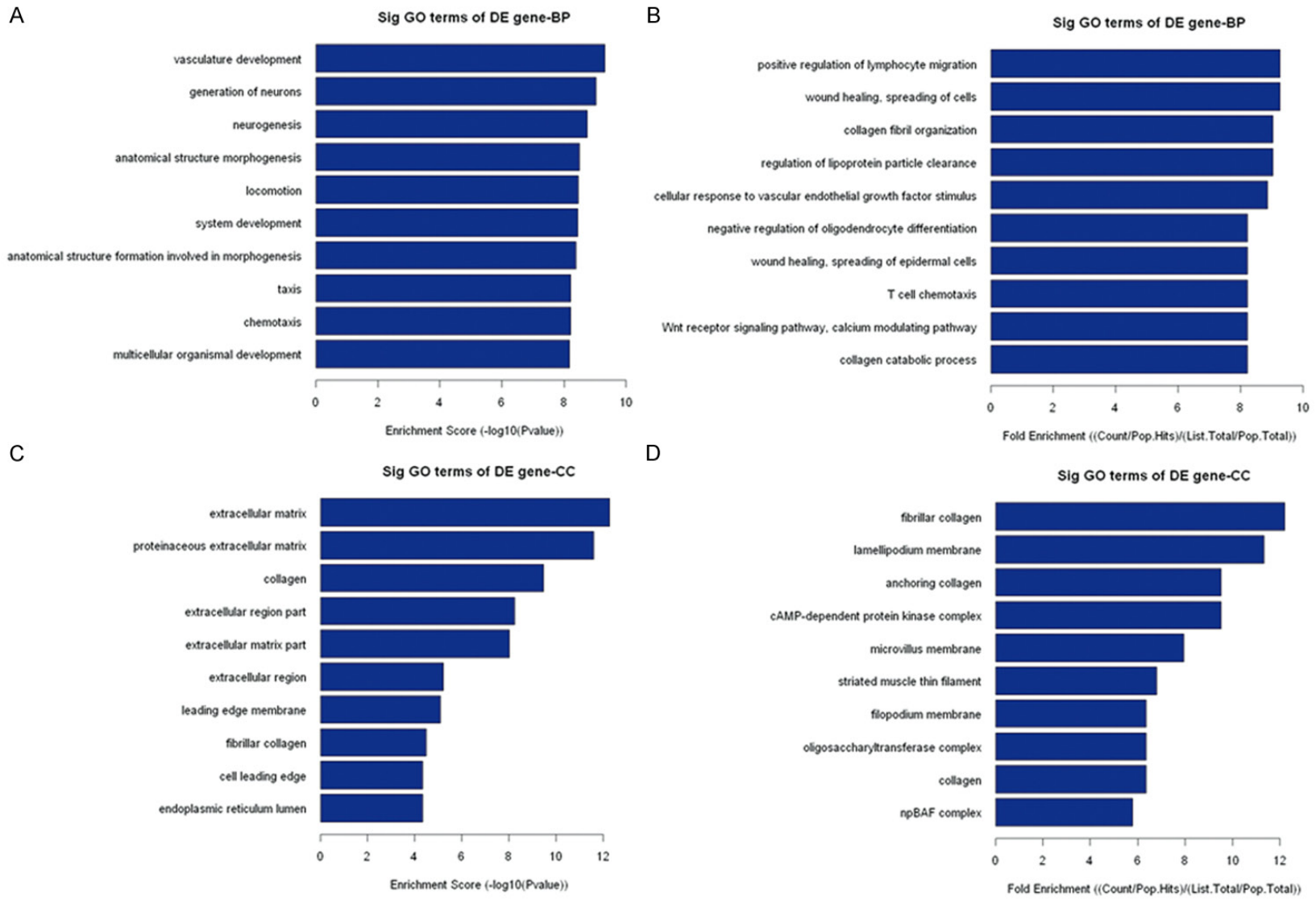
Enhancer lncRNA profiling: LncRNAs with enhancer-like function are identified using GENCODE annotation [17] of the human genes [18]. The consideration of selection of lncRNAs with enhancer-like function exclude transcripts mapping to the exons and introns of annotated protein-coding genes, the natural antisense transcripts, overlapping the protein coding genes and all known transcripts. We identified the profiling data for all probes for lncRNAs with enhancer-like function, eight differentially expressed enhancer-like lncRNAs were detected near 15 coding genes that were differentially expressed (distance <300 kb); 7 of the lncRNA-mRNA pairs were regulated in the up-up direction, 7 pairs were regulated in the down-down direction, and 1 pair was regulated in the down-up direction as shown in **Table 6**.

Rinn lincRNA profiling: All lincRNAs based on John Rinn's papers [19, 20] are identified. As shown in **Table 7**, a total of 21 differentially expressed lincRNAs had 27 adjacent coding gene pairs; 5 of the lincRNA-mRNA pairs were regulated in the up-up direction, 10 pairs were regulated in the down-down direction, 9 pairs were regulated in the up-down direction, and 3 pairs were regulated in the down-up direction.

HOX cluster profiling: Rinn et al characterized the transcriptional landscape of the 4 human Hox loci and identified 407 discrete transcribed regions in the four Hox loci [21]. Among the 407 targeted discrete transcribed regions, 166 coding transcripts and 299 noncoding transcripts were detected after we used the profiling data of all the probes targeting the 4 HOX loci.

Discussion

Although the molecular mechanism of HSCC has been extensively investigated [3-6], the


A
Sig GO terms of DE gene-BP
vasculature development
generation of neurons
neurogenesis
anatomical structure morphogenesis
locomotion
system development
anatomical structure formation involved in morphogenesis
taxis
chemotaxis
multicellular organismal development
0
2
4
6
8
10
Enrichment Score (-log10(Pvalue))
B
Sig GO terms of DE gene-BP
positive regulation of lymphocyte migration
wound healing, spreading of cells
collagen fibril organization
regulation of lipoprotein particle clearance
cellular response to vascular endothelial growth factor stimulus
negative regulation of oligodendrocyte differentiation
wound healing, spreading of epidermal cells
T cell chemotaxis
Wnt receptor signaling pathway, calcium modulating pathway
collagen catabolic process
0
2
4
6
8
10
Fold Enrichment ((Count/Pop.Hits)/(List.Total/Pop.Total))
C
Sig GO terms of DE gene-CC
extracellular matrix
proteinaceous extracellular matrix
collagen
extracellular region part
extracellular matrix part
extracellular region
leading edge membrane
fibrillar collagen
cell leading edge
endoplasmic reticulum lumen
0
2
4
6
8
10
12
Enrichment Score (-log10(Pvalue))
D
Sig GO terms of DE gene-CC
fibrillar collagen
lamellipodium membrane
anchoring collagen
cAMP-dependent protein kinase complex
microvillus membrane
striated muscle thin filament
filopodium membrane
oligosaccharyltransferase complex
collagen
npBAF complex
0
2
4
6
8
10
12
Fold Enrichment ((Count/Pop.Hits)/(List.Total/Pop.Total))

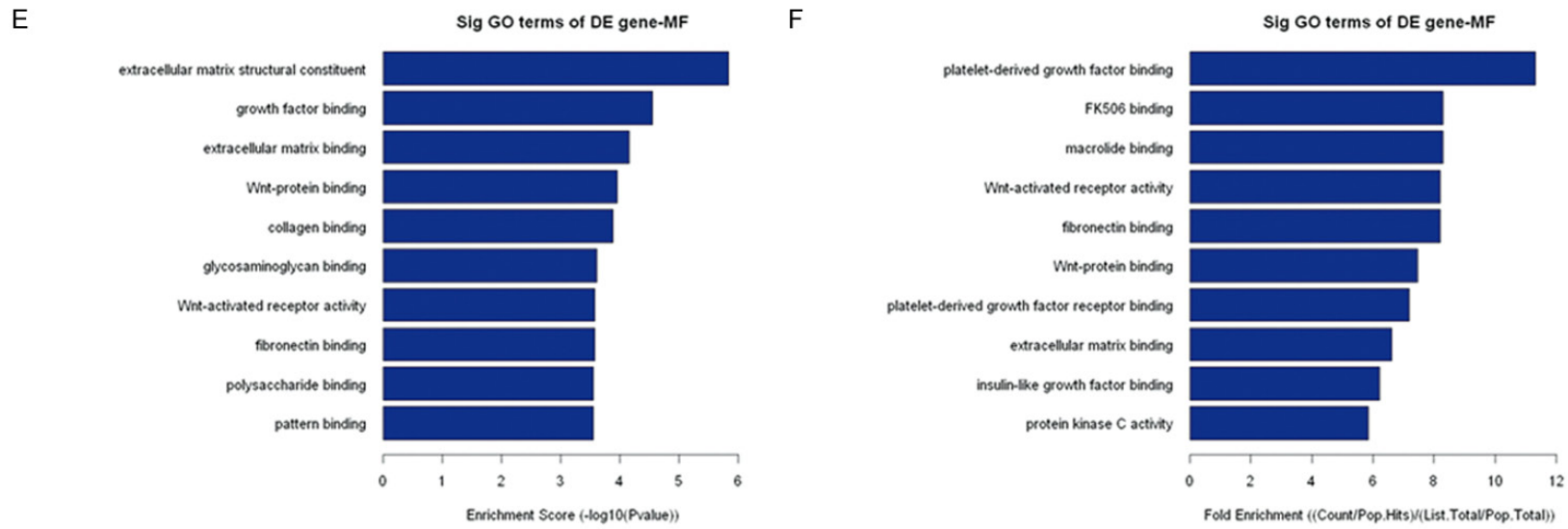
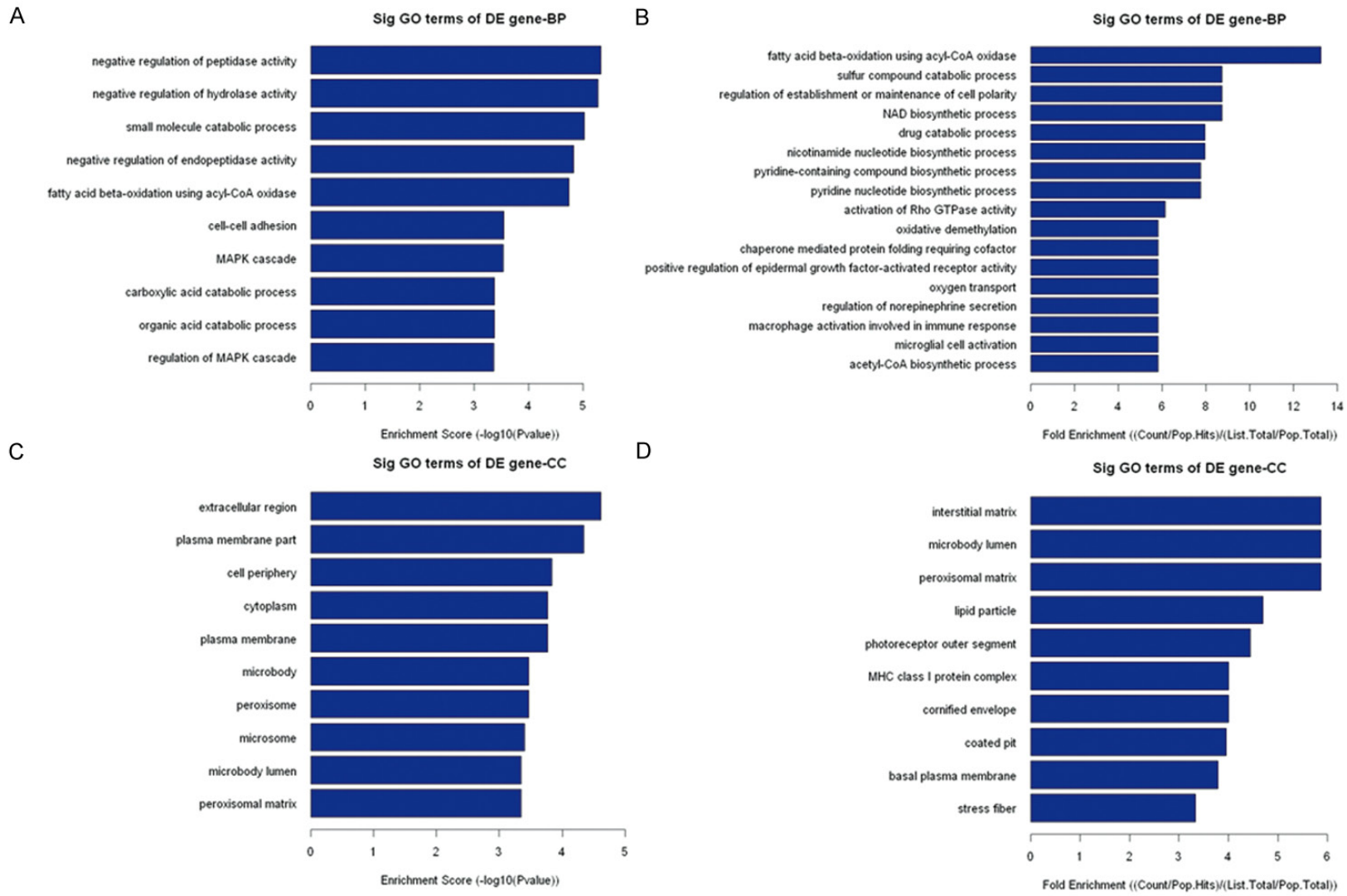


Figure 7. The top 10 upregulated GO functions of Fold Enrichment Terms and Score Enrichment Terms in the carcinomas compared to the adjacent nontumor tissues.


A
Sig GO terms of DE gene-BP
negative regulation of peptidase activity
negative regulation of hydrolase activity
small molecule catabolic process
negative regulation of endopeptidase activity
fatty acid beta-oxidation using acyl-CoA oxidase
cell-cell adhesion
MAPK cascade
carboxylic acid catabolic process
organic acid catabolic process
regulation of MAPK cascade
0 1 2 3 4 5
Enrichment Score (-log10(Pvalue))
B
Sig GO terms of DE gene-BP
fatty acid beta-oxidation using acyl-CoA oxidase
sulfur compound catabolic process
regulation of establishment or maintenance of cell polarity
NAD biosynthetic process
drug catabolic process
nicotinamide nucleotide biosynthetic process
pyridine-containing compound biosynthetic process
pyridine nucleotide biosynthetic process
activation of Rho GTPase activity
oxidative demethylation
chaperone mediated protein folding requiring cofactor
positive regulation of epidermal growth factor-activated receptor activity
oxygen transport
regulation of norepinephrine secretion
macrophage activation involved in immune response
microglial cell activation
acetyl-CoA biosynthetic process
0 2 4 6 8 10 12 14
Fold Enrichment ((Count/Pop.Hits)/(List.Total/Pop.Total))
C
Sig GO terms of DE gene-CC
extracellular region
plasma membrane part
cell periphery
cytoplasm
plasma membrane
microbody
peroxisome
microsome
microbody lumen
peroxisomal matrix
0 1 2 3 4 5
Enrichment Score (-log10(Pvalue))
D
Sig GO terms of DE gene-CC
interstitial matrix
microbody lumen
peroxisomal matrix
lipid particle
photoreceptor outer segment
MHC class I protein complex
cornified envelope
coated pit
basal plasma membrane
stress fiber
0 1 2 3 4 5 6
Fold Enrichment ((Count/Pop.Hits)/(List.Total/Pop.Total))

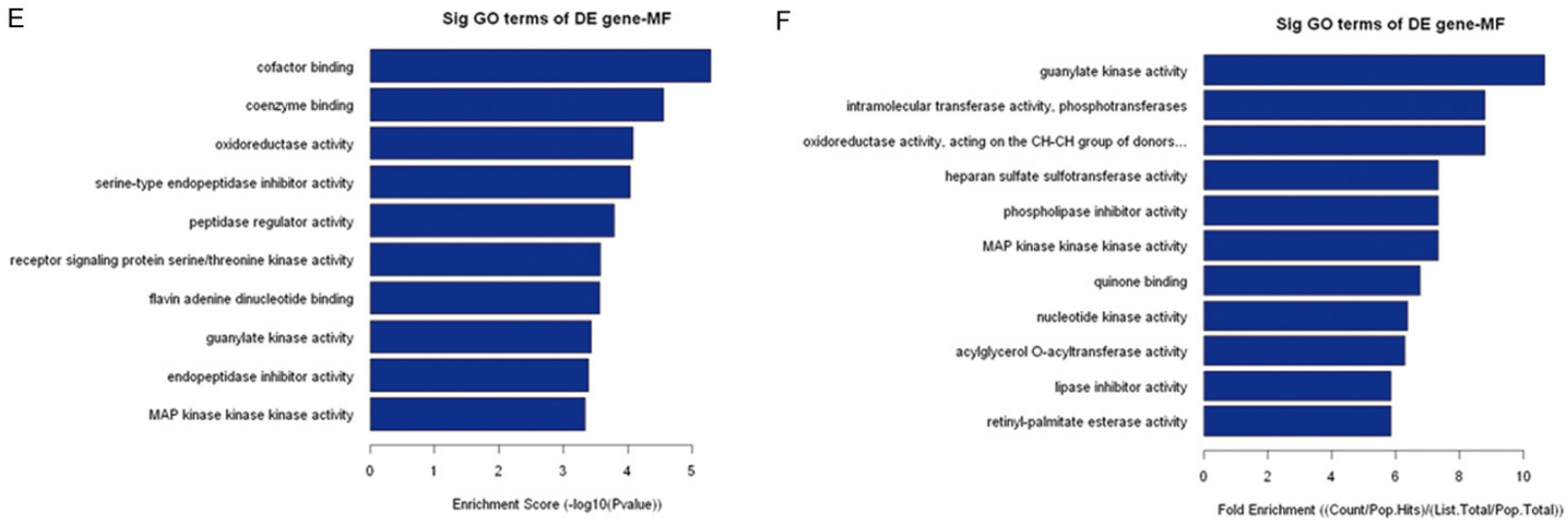


Figure 8. The top 10 downregulated GO functions of Fold Enrichment Terms and Score Enrichment Terms in the carcinomas compared to the adjacent nontumor tissues.

Table 6. Differentially expressed enhancer-like lncRNAs and their nearby coding gene pairs that were differentially expressed in carcinomas compared to adjacent nontumor tissues (distance <300 kb)

lncRNAs				mRNAs				Direction (LncRNA-mRNA)
Seqname	Gene symbol	*P* Value	Fold Change	Nearby gene	Nearby gene symbol	*P* Value	Fold change	
NR_015410	FLJ22536	0.0132885	3.6317198	NM_003107	SOX4	0.010653352	2.9999847	Up-up
Uc002wkq.2	BC069037	0.0349635	2.4205506	NM_175841	SMOX	0.015637238	2.8002222	Up-up
Uc002wkq.2	BC069037	0.0349635	2.4205506	NM_001134338	RNF24	0.016596192	2.018034	Up-up
Uc002wkq.2	BC069037	0.0349635	2.4205506	NM_175839	SMOX	0.026806494	3.3381217	Up-up
Uc002wkq.2	BC069037	0.0349635	2.4205506	NM_175840	SMOX	0.021685325	2.0501275	Up-up
ENST00000417473	AC099344.4	0.0040785	2.2474518	NM_004850	ROCK2	0.04852472	2.2684875	Up-up
ENST00000419064	RP11-13P5.2	0.0049485	3.0325863	NM_032532	FNDC1	0.016545452	3.0720844	Up-up
Uc003nsj.1	AK094433	0.0389439	4.254517	NM_001010909	MUC21	8.35326E-05	798.7413	Down-down
Uc003nsj.1	AK094433	0.0389439	4.254517	NM_001954	DDR1	0.000148092	4.3601513	Down-down
ENST00000433357	RP11-255A11.21	0.004358	6.7445364	NM_002771	PRSS3	0.02605703	3.0595872	Down-down
ENST00000416894	RP11-54A4.8	0.0294946	14.243775	NM_025008	ADAMTSL4	0.045489967	3.7833204	Down-down
ENST00000416894	RP11-54A4.8	0.0294946	14.243775	NM_004425	ECM1	0.000824244	11.762663	Down-down
ENST00000416894	RP11-54A4.8	0.0294946	14.243775	NM_022664	ECM1	0.003080209	23.130184	Down-down
ENST00000416894	RP11-54A4.8	0.0294946	14.243775	NM_019032	ADAMTSL4	0.012949571	7.710182	Down-down
HIT000395572	RP11-54A4.8	0.0307008	2.790272	NM_203459	CAMSAP1L1	0.007201749	2.2794487	Down-up

exact pathogenesis of this disease is still unclear. Until recently, lncRNAs had been considered as simply transcriptional noise [22]. However, recent studies showed that lncRNAs can regulate not only basal transcription but also posttranscriptional processes, including pre-mRNA processing, splicing, transport, translation, and siRNA-directed gene regulation [23]. Some lncRNAs can directly bind proteins and regulate protein function [24]. Furthermore, lncRNAs are involved in epigenetic modifications, including DNA methylation [25] and histone modification [26]. Several association studies have recognized that lncRNAs may function in various aspects of cell biology and have identified a large number of lncRNAs that are differentially expressed in disease states, including oncogenesis [27]. It has been reported that the expression of lncRNAs differs significantly between normal tissue and tumor tissue [28, 29]. Thus, lncRNAs are emerging as new players in the cancer paradigm, having regulatory functions in both oncogenic and tumor-suppressive pathways [30-33]. Dysregulation of lncRNAs, such as PCGEM1, ANRIL, DD3, HOTAIR, XIST, HULC, MALAT1, and Neat2, has been regarded as an important feature of several human cancers, including prostate cancer [34-36], breast cancer [37, 38], colorectal cancer [39], laryngeal cancer [40], male testicular cancer [41], non-small cell lung cancer [42], hepatocellular cancer [43], and colorectal cancer [44]. However, the expression and functional significance of lncRNAs in HSCC tumorigenesis have not been characterized.

In the present study, in order to reveal the molecular mechanisms underlying HSCC, we performed a genomewide screen of differences in mRNA and lncRNA expression profiles between HSCC tissues and matched adjacent nontumor mucosal epithelial tissues. We found that 1299 lncRNAs and 1432 mRNAs were differentially expressed in carcinomas compared with the adjacent nontumor tissues, indicating that many lncRNAs and mRNAs were significantly upregulated or downregulated in HSCC. To confirm the reliability and validity of the microarray results, we used qRT-PCR to validate the expression patterns of lncRNAs AB209630 and AB019562 and mRNAs SPP1 and TJP2 in 20 HSCC patients. The qRT-PCR results matched well with the microarray data. These differentially expressed genes were subsequently organized into hierarchical categories based on heat map and hierarchical clustering. We also found that differentially expressed lncRNAs and mRNAs were distributed on each of the chromosomes. This result proved that all of the chromosomes, including the X and Y chromosomes, can display different quantities and degrees of abnormalities in HSCC tumorigenesis. Furthermore, pathway analysis revealed 48 pathways that may play key roles in the different core epigenetic mechanisms of HSCC, including ECM-receptor inter-

Table 7. Differentially expressed lincRNAs and their nearby coding gene pairs that were differentially expressed in carcinomas compared to adjacent nontumor tissues (distance <300 kb)

lncRNAs				mRNAs				Direction (LncRNA-mRNA)
Seqname	Gene symbol	*P* Value	Fold change	Nearby gene	Nearby gene symbol	*P* Value	Fold change	
ENST00000508827	AL355916.1	0.0066856	2.3141046	NM_003082	SNAPC1	0.0434955	2.556988	Up-up
NR_027005	C6orf147	0.0386217	3.5584037	NM_138441	MB21D1	0.0095622	2.0235877	Up-up
AF085351		0.0053471	2.3140755	NM_001135651	EIF2AK2	0.0018426	5.02935	Up-up
uc002ilc.1	CR602880	0.0232957	2.092423	NM_001012511	GOSR2	0.0391195	2.0630145	Up-up
ENST00000417473	AC099344.4	0.0040785	2.2474518	NM_004850	ROCK2	0.0485247	2.2684875	Up-up
chr11:8215149-8227449+	LincRNA-LM01-2	0.0050277	2.042942	NM_024557	RIC3	0.0196345	2.4191647	Up-down
AX721103		0.004123	2.2437043	NM_003447	ZNF165	0.0192913	2.0826676	Up-down
AK131566	LincRNA-NR4A1	0.0026101	2.054437	NM_005556	KRT7	0.0031214	4.3845367	Up-down
ENST00000366365	C17orf86	0.0100169	2.6049378	NM_001113494	SEPT9	0.0001254	2.989717	Up-down
NR_003013	SCARNA16	0.0210531	2.6479173	NM_001113494	SEPT9	0.0001254	2.989717	Up-down
ENST00000508827	AL355916.1	0.0066856	2.3141046	NM_006255	PRKCH	0.0159917	2.4145923	Up-down
ENST00000403367	RP1-72A23.1	0.0396599	2.1803179	NM_006813	PNRC1	0.0491512	2.110001	Up-down
Uc001nnk.1	AB231722	0.0121469	2.0080948	NM_080661	GLYATL1	0.0046781	2.2372477	Up-down
Chr2:192293450-192304436+	LincRNA-OBFC2A-4	0.0483022	2.4108155	NM_001031716	OBFC2A	0.0350843	2.0829911	Up-down
NR_003587	MY015B	0.043617	3.2647016	NM_001031803	LLGL2	0.0237679	3.1176605	Down-down
BF724558	LincRNA-MCL1	0.0042217	3.1531017	NM_019032	ADAMTSL4	0.0129496	7.710182	Down-down
AI683742	LincRNA-ATG5-4	0.0193867	2.8032198	NM_001624	AIM1	0.0052337	5.3022714	Down-down
AK125137	LincRNA-AIM1-2	0.0023624	5.41028	NM_001624	AIM1	0.0052337	5.3022714	Down-down
BF724558	LincRNA-MCL1	0.0042217	3.1531017	NM_004425	ECM1	0.0008242	11.762663	Down-down
BF724558	LincRNA-MCL1	0.0042217	3.1531017	NM_025008	ADAMTSL4	0.04549	3.7833204	Down-down
BF724558	LincRNA-MCL1	0.0042217	3.1531017	NM_022664	ECM1	0.0030802	23.130184	Down-down
NR_003587	MY015B	0.043617	3.2647016	NM_001015002	LLGL2	0.0445101	2.3903623	Down-down
NR_003587	MY015B	0.043617	3.2647016	NM_004524	LLGL2	0.0020565	3.233063	Down-down
ENST00000427394	RP11-497D6.4	0.0058733	2.3810697	NM_001127715	STXBP5	0.0001852	2.5615492	Down-down
AI218855	LincRNA-RPS14-2	0.0179113	6.2049127	NM_001012301	ARSI	0.0224058	4.1109605	Down-up
Exon397+	LincRNA-FAM107B	0.0129884	2.944849	NM_031453	FAM107B	0.0228996	2.1127982	Down-up
HIT000395572		0.0307008	2.790272	NM_203459	CAMSAP1L1	0.0072017	2.2794487	Down-up

action, focal adhesion, melanogenesis signaling, tight junction, African trypanosomiasis, and peroxisome signaling. GO analysis revealed that 593 mRNAs involved in biological processes, 50 mRNAs involved in cellular components, and 46 mRNAs involved in molecular functions were upregulated in the carcinomas and 280 mRNAs involved in biological processes, 58 mRNAs involved in cellular components, and 71 mRNAs involved in molecular functions were downregulated in the carcinomas.

LncRNAs are known to function via a variety of mechanisms. However, a common and important function of lncRNAs is to alter the expression of nearby coding genes by affecting transcription [19, 45, 46] or playing a direct enhancer-like role [18, 47]. To gain insight into the function of lncRNAs in lncRNA-mRNA coexpression, we further identified nearby coding genes (<300 kb) that may be regulated by enhancer-like lncRNAs and Rinn lincRNAs, which may be used for predicting target genes of lncRNAs. The expression profiles included 8 differentially expressed enhancer-like lncRNAs with nearby coding genes that were differentially expressed. For example, the lncRNA-mRNA pairs FLJ22536-SOX4, BC069037-SMOX, and BC069037-RNF24 were regulated in the up-up direction; however, the pairs AK094433-MUC21, AK094433-DDR1, and RP11-255A11.21-PRSS3 were regulated in the down-down direction. A total of 21 differentially expressed lincRNAs had adjacent coding gene pairs that were differentially expressed. For example, the lincRNA-mRNA pairs AL355916.1-SNAPC1 and C6orf147-MB21D1 were regulated in the up-up direction; lincRNA-LMO1-2-RIC3 and lincRNA-NR4A1-KRT7 were regulated in the up-down direction; MYO15B-LLGL2 and lincRNA-MCL1-ADAMTSL4 were regulated in the down-down direction; and lincRNA-RPS14-2-ARSI and lincRNA-FAM107B-FAM107B were regulated in the down-up direction. Rinn et al characterized the transcriptional landscape of the 4 human Hox loci and identified a total of 407 discrete transcribed regions [21]. It was reported that lncRNAs in the Hox loci became systematically dysregulated during some cancer progression [15]. Among the 407 targeted discrete transcribed regions, 166 coding transcripts and 299 noncoding transcripts were detected in this study. Although the function of most aberrantly expressed lncRNAs is yet unknown, the information from this study may be useful for further studies on the mechanisms of HSCC tumorigenesis.

In conclusion, we demonstrated for the first time the expression profiles of human lncRNAs and mRNAs in patients with HSCC by microarray analysis. We identified 1299 lncRNAs and 1432 mRNAs that were differentially expressed in carcinomas compared to the adjacent nontumor tissues. It is likely that these deregulated lncRNAs and mRNAs play key roles in the development of HSCC. In addition, we identified potential regulatory mechanisms with bioinformatics analyses. Such information will be used to investigate the functions of these lncRNAs and mRNAs in the occurrence and development of HSCC and will facilitate identification of new therapeutic targets and diagnostic biomarkers for this disease.

Acknowledgements

This work was supported by the Taishan Scholars Program (No. tshw20130950), Shandong Province, and the Department of Science & Technology of Shandong Province (No. ZR2013HM107 and ZR2014HM005), and Science Foundation of Qilu Hospital Of Shandong University; and the Fundamental Research Funds Of Shandong University (No. 2014QLKY05).

Disclosure of conflict of interest

None.

Abbreviations

HSCC, hypopharyngeal squamous cell carcinoma; lincRNA, intergenic long noncoding RNA; lncRNA, long noncoding RNA; qRT-PCR, quantitative real-time polymerase chain reaction; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology; HOX, homeobox.

Address correspondence to: Dr. Dapeng Lei, Department of Otorhinolaryngology, Qilu Hospital, Shandong University; Key Laboratory of Otolaryngology, Ministry of Health, 107 West Wenhua Road, Jinan 250012, Shandong, P. R. China. Tel: 531-82166781; Fax: 531-82167544; E-mail: leidapeng@sdu.edu.cn

References

[1] Hall SF, Groome PA, Irish J, O'Sullivan B. The natural history of patients with squamous cell carcinoma of the hypopharynx. Laryngoscope 2008; 118: 1362-1371.

[2] Wycliffe ND, Grover RS, Kim PD, Simental A Jr. Hypopharyngeal cancer. Top Magn Reson Imaging 2007; 18: 243-258.

[3] Liu J, Lei DP, Jin T, Zhao XN, Li G. Altered expression of miR-21 and PTEN in human laryngeal and hypopharyngeal squamous cell carcinomas. Asian Pac J Cancer Prev 2011; 12: 2653-2657.

[4] Yu XM, Liu Y, Jin T, Liu J, Wang J. The expression of SIRT1 and DBC1 in laryngeal and hypopharyngeal carcinomas. PLoS One 2013; 8: e66975.

[5] Wang J, Pan XL, Ding LJ, Liu DY, Da-Peng L. Aberrant expression of Beclin-1 and LC3 correlates with poor prognosis of human hypopharyngeal squamous cell carcinoma. PLoS One 2013; 8: e69038.

[6] Zhao X, Ma C, Cai X, Lei D, Liu D. RNA interference of caveolin-1 via lentiviral vector inhibits growth of hypopharyngeal squamous cell carcinoma FaDu cells In Vitro and In Vivo. Asian Pac J Cancer Prev 2011; 12: 397-401.

[7] Pattje WJ, Schuuring E, Mastik MF, Slagter-Menkema L, Schrijvers ML. The phosphatase and tensin homologue deleted on chromosome 10 mediates radiosensitivity in head and neck cancer. Br J Cancer 2010; 102: 1778-1785.

[8] Cheetham SW, Gruhl F, Mattick JS, Dinger ME. Long noncoding RNAs and the genetics of cancer. Br J Cancer 2013; 108: 2419-2425.

[9] Sanchez Y, Huarte M. Long non-coding RNAs: challenges for diagnosis and therapies. Nucleic Acid Ther 2013; 23: 15-20.

[10] Ernst C, Morton CC. Identification and function of long non-coding RNA. Front Cell Neurosci 2013; 7: 168.

[11] Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. BMC Genomics 2007; 8: 39.

[12] Bolton EM, Tuzova AV, Walsh AL, Lynch T, Perry AS. Noncoding RNAs in prostate cancer: the long and the short of it. Clin Cancer Res 2014; 20: 35-43.

[13] Jia W, Chen W, Kang J. The functions of microRNAs and long non-coding RNAs in embryonic and induced pluripotent stem cells. Genomics Proteomics Bioinformatics 2013; 11: 275-283.

[14] Shi X, Sun M, Liu H, Yao Y, Song Y. Long noncoding RNAs: a new frontier in the study of human diseases. Cancer Lett 2013; 339: 159-166.

[15] Gupta RA, Shah N, Wang KC, Kim J, Horlings HM. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 2010; 464: 1071-1076.

[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: 135-141.

[17] Harrow J, Denoeud F, Frankish A, Reymond A, Chen CK. GENCODE: producing a reference annotation for ENCODE. Genome Biol 2006; 7 Suppl 1: S4, 1-9.

[18] Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A. Long noncoding RNAs with enhancer-like function in human cells. Cell 2010; 143: 46-58.

[19] Khalil AM, Guttman M, Huarte M, Garber M, Raj A. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A 2009; 106: 11667-11672.

[20] Guttman M, Amit I, Garber M, French C, Lin MF. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 2009; 458: 223-227.

[21] Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 2007; 129: 1311-1323.

[22] Zhu S, Zhang XO, Yang L. Panning for Long Noncoding RNAs. Biomolecules 2013; 3: 226-241.

[23] Yoon JH, Abdelmohsen K, Gorospe M. Posttranscriptional gene regulation by long noncoding RNA. J Mol Biol 2013; 425: 3723-3730.

[24] Arun G, Akhade VS, Donakonda S, Rao MR. mrhl RNA, a long noncoding RNA, negatively regulates Wnt signaling through its protein partner Ddx5/p68 in mouse spermatogonial cells. Mol Cell Biol 2012; 32: 3140-3152.

[25] Mohammad F, Pandey GK, Mondal T, Enroth S, Redrup L. Long noncoding RNA-mediated maintenance of DNA methylation and transcriptional gene silencing. Development 2012; 139: 2792-2803.

[26] Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell 2011; 44: 667-678.

[27] Maruyama R, Suzuki H. Long noncoding RNA involvement in cancer. BMB Rep 2012; 45: 604-611.

[28] Gibb EA, Vucic EA, Enfield KS, Stewart GL, Lonergan KM. Human cancer long non-coding

RNA transcriptomes. PLoS One 2011; 6: e25915.
[29] Zhu Y, Yu M, Li Z, Kong C, Bi J. ncRAN, a newly identified long noncoding RNA, enhances human bladder tumor growth, invasion, and survival. Urology 2011; 77: 510, e511-515.
[30] Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell 2010; 142: 409-419.
[31] Hung T, Wang Y, Lin MF, Koegel AK, Kotake Y. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nat Genet 2011; 43: 621-629.
[32] Gibb EA, Brown CJ, Lam WL. The functional role of long non-coding RNA in human carcinomas. Mol Cancer 2011; 10: 38.
[33] Huarte M, Rinn JL. Large non-coding RNAs: missing links in cancer? Hum Mol Genet 2010; 19: R152-161.
[34] Srikantan V, Zou Z, Petrovics G, Xu L, Augustus M. PCGEM1, a prostate-specific gene, is overexpressed in prostate cancer. Proc Natl Acad Sci U S A 2000; 97: 12216-12221.
[35] Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. Mol Cell 2010; 38: 662-674.
[36] Bussemakers MJ, van Bokhoven A, Verhaegh GW, Smit FP, Karthaus HF. DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. Cancer Res 1999; 59: 5975-5979.
[37] Richardson AL, Wang ZC, De Nicolo A, Lu X, Brown M. X chromosomal abnormalities in basal-like human breast cancer. Cancer Cell 2006; 9: 121-132.
[38] Chisholm KM, Wan Y, Li R, Montgomery KD, Chang HY. Detection of long non-coding RNA in archival tissue: correlation with polycomb protein expression in primary and metastatic breast carcinoma. PLoS One 2012; 7: e47998.
[39] Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S. Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. Cancer Res 2011; 71: 6320-6326.
[40] Li D, Feng J, Wu T, Wang Y, Sun Y. Long intergenic noncoding RNA HOTAIR is overexpressed and regulates PTEN methylation in laryngeal squamous cell carcinoma. Am J Pathol 2013; 182: 64-70.
[41] Kawakami T, Okamoto K, Ogawa O, Okada Y. XIST unmethylated DNA fragments in male-derived plasma as a tumour marker for testicular cancer. Lancet 2004; 363: 40-42.
[42] Schmidt LH, Spieker T, Koschmieder S, Schaffers S, Humberg J. The long noncoding MALAT-1 RNA indicates a poor prognosis in non-small cell lung cancer and induces migration and tumor growth. J Thorac Oncol 2011; 6: 1984-1992.
[43] Panzitt K, Tschernatsch MM, Guelly C, Moustafa T, Stradner M. Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. Gastroenterology 2007; 132: 330-342.
[44] Matouk IJ, Abbasi I, Hochberg A, Galun E, Dweik H. Highly upregulated in liver cancer noncoding RNA is overexpressed in hepatic colorectal metastasis. Eur J Gastroenterol Hepatol 2009; 21: 688-692.
[45] Mattick JS, Gagen MJ. The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. Mol Biol Evol 2001; 18: 1611-1630.
[46] Popadin K, Gutierrez-Arcelus M, Dermitzakis ET, Antonarakis SE. Genetic and epigenetic regulation of human lincRNA gene expression. Am J Hum Genet 2013; 93: 1015-1026.
[47] Mattick JS. Linc-ing Long noncoding RNAs and enhancer function. Dev Cell 2010; 19: 485-486.