RESEARCH ARTICLE

Long Non-coding RNAs are Differentially Expressed in Hepatocellular Carcinoma Cell Lines with Differing Metastatic Potential

Ting-Ting Fang, Xiao-Jing Sun, Jie Chen, Yan Zhao, Rui-Xia Sun, Ning Ren*, Bin-Bin Liu*

Abstract

Background: Metastasis is a major reason for poor prognosis in patients with cancer, including hepatocellular carcinoma (HCC). A salient feature is the ability of cancer cells to colonize different organs. Long non-coding RNAs (lncRNAs) play important roles in numerous cellular processes, including metastasis. Materials and Methods: In this study, the lncRNA expression profiles of two HCC cell lines, one with high potential for metastasis to the lung (HCCLM3) and the other to lymph nodes (HCCLYM-H2) were assessed using the Arraystar Human LncRNA Array v2.0, which contains 33,045 lncRNAs and 30,215 mRNAs. Coding-non-coding gene co-expression (CNC) networks were constructed and gene set enrichment analysis (GSEA) was performed to identify lncRNAs with potential functions in organ-specific metastasis. Levels of two representative lncRNAs and one representative mRNA, RP5-1014O16.1, lincRNA-TSPAN8 and TSPAN8, were further detected in HCC cell lines with differing metastasis potential by qRT-PCR. Results: Using microarray data, we identified 1,482 lncRNAs and 1,629 mRNAs that were differentially expressed (≥1.5 fold-change) between the two HCC cell lines. The most upregulated lncRNAs in H2 were RP11-672F9.1, RP5-1014O16.1, and RP11-501G6.1, while the most downregulated ones were lincRNA-TSPAN8, lincRNA-CALCA, C14orf132, NCRNA00173, and CR613944. The most upregulated mRNAs in H2 were C15orf48, PSG2, and PSG8, while the most downregulated ones were CALCB, CD81, CD24, TSPAN8, and SOST. Among them, lincRNA-TSPAN8 and TSPAN8 were found highly expressed in high lung metastatic potential HCC cells, while lowly expressed in no or low lung metastatic potential HCC cells. RP5-1014O16.1 was highly expressed in high lymphatic metastatic potential HCC cell lines, while lowly expressed in no lymphatic metastatic potential HCC cell lines. Conclusions: We provide the first detailed description of lncRNA expression profiles related to organ-specific metastasis in HCC. We demonstrated that a large number of lncRNAs may play important roles in driving HCC cells to metastasize to different sites; these lncRNAs may provide novel molecular biomarkers and offer a new basis for combating metastasis in HCC cases.

Keywords: Hepatocellular carcinoma - organ-specific metastasis - long non-coding RNA - RP5-1014O16.1 - expression

Asian Pac J Cancer Prev, **15 (23)**, 10513-10524

Introduction

Hepatocellular carcinoma (HCC) is the third most common cause of cancer-related deaths worldwide (Jemal et al., 2011; Llovet et al., 2012). Prognosis for the HCC patient is dismal due to poor therapeutic response and a high probability of relapse after treatment (Hong et al., 2003). Metastasis is not only a sign of deterioration but also the main obstacle to ameliorating HCC's poor prognosis (Genda et al., 1999; Poon and Wong, 2000; Ye et al., 2003). Therefore, exploring the molecular mechanisms under lying metastasis is important for developing effective intervention methods and improving patient prognosis.

Metastasis is a complex process involving a series of sequential and interrelated steps, often termed the invasion-metastasis cascade (Talmadge and Fidler, 2010). A salient feature of metastasis is the ability of cancer cells to colonize specific organ sites (Hanahan and Weinberg, 2011). For instance, HCC cells metastasize to the lung more frequently than the lymph nodes. It is well recognized that the probability of metastatic seeding and growth is determined by both intrinsic genetic properties of the cancer cells involved and characteristics of the stromal microenvironment that surrounds them (Lu and Kang, 2007; Nguyen et al., 2009). Functional genomic analysis of HCC cell lines with different organ-specific metastatic potential may provide new opportunities for identifying diagnostic biomarkers and developing therapies to combat HCC metastasis.

Large-scale genomic studies have demonstrated that the mammalian genome encodes many thousands

*Key Laboratory of Carcinogenesis and Cancer Invasion, Fudan University, Ministry of Education, Liver Cancer Institute, Zhongshan Hospital, Xuhui, Shanghai, P.R. China *For correspondence: ren.ning@zs-hospital.sh.cn, liu.binbin@zs-hospital.sh.cn*

of non-coding transcripts, both short (<200 nucleotides (nt) in length; sRNAs) and long (>200 nt; lncRNAs) (Kapranov et al., 2002; Bertone et al., 2004; Guttman et al., 2009; Cabili et al., 2011). miRNAs (microRNAs), the best-studied class of sRNAs, regulate their mRNA targets post-transcriptionally (Bartel, 2009). LncRNA transcripts were first revealed after large-scale sequencing of cDNA libraries in mouse (Okazaki et al., 2002). While only a small fraction of lncRNAs have been characterized in detail, it is clear that mRNA-like lncRNAs may act through a range of molecular mechanisms to function in epigenetic pathways and infrastructure developmental processes. LncRNAs have been reported to interact with chromatin modifications (Guttman et al., 2011; Guttman and Rinn, 2012), serve as precursors for the generation of sRNAs (Fejes-Toth et al., 2009; Guttman and Rinn, 2012), and play regulatory roles in the expression and activity of genes and the localization of the proteins they encode (Wilusz et al., 2009). In addition, lncRNAs function in numerous cellular processes, ranging from embryonic stem cell pluripotency to cell cycle regulation to disease, including cancer (Ponting et al., 2009; Wapinski and Chang, 2011). Recently, hundreds of lncRNAs have been discovered, as those "dark matters of the genome" are selectively over- or underexpressed in different tumors, but the cellular mechanisms of lncRNAs in tumors are still poorly understood (Qiu et al., 2013). Deregulation of lncRNAs is associated with the occurrence of various tumors and has potential significance for cancer diagnosis. Some lncRNAs are already considered biomarkers for specific cancers and outcomes. Examples include lncRNA DD3 for prostate cancer (Tinzl and Horvath, 2004; Wang et al., 2014), long intergenic non-coding RNA (lincRNA) HOTAIR for primary breast cancer (Gupta et al., 2010; Zhang et al., 2014), and serum lncRNA HULC for HCC (Panzitt et al., 2007; Xie et al., 2013). The expression profile of lncRNAs in HCC and their biological function in metastasis remain poorly understood. Better understanding of the roles that lncRNAs play will advance our understanding of cell regulatory and disease mechanisms.

In the present study, we compared the expression profiles of HCC cell lines with a similar genetic background but different potential for lung or lymph node metastasis. We found that expression signatures comprising lncRNAs and protein-coding mRNAs were significantly associated with organ-specific HCC metastasis. Our results indicate that lncRNA expression profiles may represent new molecular biomarkers for HCC metastasis.

Materials and Methods

HCC Cell lines and animals

Nine HCC cell lines (Hep3B, HepG2, SMMC-7721, MHCC-97L, MHCC-97H, HCCLM3, HCCLM6, HCCLYM-H, HCCLYM-H2) used in this study. MHCC-97L, MHCC-97H, HCCLM3, HCCLM6 were all established by the authors' institute. Hep3B, HepG2, SMMC-7721 were obtained from the Cell Bank of Shanghai Institutes of Biological Sciences, Chinese Academy of Sciences. The HCCLM3 (LM3) and HCCLM6 cell lines were derived from the same parental cell line MHCC-97H. LM3 metastasizes to the lung, while HCCLM6 can metastasize to multiple organs in a mouse model (Li et al., 2001; Li et al., 2003). Hep3B was cultured in MEM (GibocloBRL, USA) supplemented with 10% fetal bovine serum (GibocloBRL, USA), and other cells were cultured in high glucose DMEM (GibocloBRL, USA) supplemented with 10% fetal bovine serum (GibocloBRL, USA). By subcloning HCCLM6, we established another cell line, HCCLYM-H (unpublished data), which frequently metastasizes to the lymph nodes. We used limited dilution method (Dexter et al., 1978) and in vivo consecutive selection (Sun et al., 1996) to optimize the HCCLYM-H cell line. HCCLYM-H cells were seeded into the footpads of nude mice (BALB/C-nu/nu, male, 4 weeks old, 18-20 g) to assess tumor growth. Athymic nude mice were obtained from Shanghai Institute of Material Medicine and maintained in a pathogen-free environment. Animal care and experimental protocols were performed in accordance with the guidelines established by the Shanghai Medical Experimental Animal Care Commission. Ethical approval was obtained from the Zhongshan Hospital Research Ethics Committee. In addition, liver orthotopic transplantation was performed in nude mice, and 5 weeks later, metastasis to the lungs and lymph nodes were detected by pathological examination. In this way, we established a cell line named HCCLYM-H2 (H2), which showed stable and high metastatic potential specific to the lymph nodes.

RNA extraction and RNA quality control

Total RNA was extracted from HCC cell lines LM3 and H2 using Trizol (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. RNA clean up including a DNase I digestion step performed with RNeasy spin columns (Qiagen, Germany). RNA quantity and quality were assessed using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, USA). RNA integrity was measured by the relative abundance of 28S/18S ribosomal subunits, verified through denaturing agarose gel electrophoresis.

Microarray experiment and data analysis

Gene expression measurements were performed using the Arraystar Human LncRNA Microarray v2.0 (8660 K, Arraystar, Rockville, MD) which is designed for the global profiling of human LncRNAs and protein-coding transcripts. 33,045 LncRNAs and 30,215 coding transcripts can be detected by the second-generation LncRNA microarray. The LncRNAs are carefully collected from the most authoritative databases such as RefSeq, UCSC Knowngenes, Ensembl and many related literatures. We measured three samples for each cell line, for a total of six samples. Sample labeling and array hybridization were performed according to the Agilent One-Color Microarray-Based Gene Expression Analysis protocol (Agilent Technologies, Santa Clara, CA). The hybridized arrays were washed and fixed, then processed slides were scanned using the DNA Microarray Scanner (Agilent Technologies). Agilent Feature Extraction software (version 11.0.1.1) was used to analyze array

images. Quantile normalization and subsequent data processing were performed using the GeneSpring GX v11.5.1 software package (Agilent Technologies, Santa Clara, CA). LncRNAs and mRNAs in which at least one out of the six samples had flags in Present or Marginal for the "All Targets Value" was chosen for further data analysis. We identified lncRNAs and mRNAs that were differentially expressed between the two groups using volcano plotfiltering. P-values were calculated using the paired t-test. The threshold set for up and downregulated genes was a fold change≥1.5 and a P-value≤0.05. The microarray data discussed in this article have been deposited in National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) and are accessible through (GEO) Series accession number GSE61015 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE61015). Hierarchical clustering was performed using Agilent GeneSpring GX software (version 11.5.1).

Construction of the co-expression network

Coding non-coding gene co-expression (CNC) networks were constructed according to the normalized signal intensity of differentially expressed genes. We determined the network adjacency between two genes, X and Y, which was defined as the power of the Pearson correlation between the corresponding gene expression profiles, Kx and Ky. In this way, we obtained the gene adjacency matrix, M (X, Y) (Prieto et al., 2008). The adjacency matrix, M (X, Y), was visualized as a graph, and the topological properties of this graph were examined. To make a visual representation, only the strongest correlations (≥0.99) were drawn. In co-expression networks, each gene corresponds to a node and two genes are connected by an edge. Within the network, a degree is defined as the number of directly linked neighbors; this is the most important measure of the centrality of a gene within a network and determines its relative importance (Barabasi, 2004). We chose 132 lncRNAs in H2 and 173 in LM3 which with the highest level of co-expression to construct the H2 and the LM3 co-expression network respectively. And ten lncRNAs (CR613944, KLHL23, TPRXL, AX747582, AX746887, NCRNA00173, BC058547, RP11-672F9.1, SMEK3P, AX747284) which have the highest level of co-expression were chose to construct the subnetwork. CR613944, KLHL23, TPRXL, AX747582, AX746887 and NCRNA00173 are up-regulated in LM3 compared with H2. BC058547, RP11-672F9.1, SMEK3P and AX747284 are up-regulated in H2 compared with LM3.

Gene set enrichment analyses

Functional associations were computed using Gene Set Enrichment Analysis (GSEA), performed with Gene Spring GX software (version 11.5.1). Correlations between the expression levels of each lncRNA locus and selected protein-coding loci were calculated, similar to the method described by Guttman et al (Guttman et al., 2009). For each lncRNA locus, a list of protein-coding loci ranked based on correlation was constructed and subjected to GSEA (Mootha et al., 2003; Subramanian et al., 2005). Gene sets were filtered using a false-discovery rate (FDR) threshold of 0.05. An association matrix between lncRNA loci and gene ontology (GO) terms was constructed, using a FDR threshold of 0.01. Rows (lncRNA loci) and columns (GO terms) were clustered (k-means, 10 clusters), resulting in distinct subsets of lncRNAs associated with functional GO terms. To determine the enrichment level of positively associated GO terms for each cluster with respect to other clusters, positively correlated GO terms were ranked according to a binomial test.

Quantitative real-time PCR

RNA was isolated with RNA-iso Plus reagent (Takara, Japan) then reverse transcribed using PrimeScript® RT Master mix (Takara, Japan). Quantitative real-time PCR (qRT-PCR) was performed using a SYBR® Premix Ex TaqTMII Kit (Takara, Japan) according to the manufacturer's instructions, and the reactions were run in a Eppendorf master cycle repreal plex Real-Time PCR System (Eppendorf, Germany) for 40 cycles at 95°C for 10 sec, 60°C for 20 sec. Primer sequences are included below: CR613944, sense: 5'TCTCACCAAATCCCTTCAACAA3', antisense: 5'GCATGGGTTCAAATCCCAGTT3'; BC058547, sense: 5'GCTGGCCACAGTCACGTCTA3', antisense: 5'GGAAAACCCCTTTATCACTTTG3'; RP5-1014O16.1, sense: 5'TCCCTCGGCCTCTCTTATGA3', antisense: 5'CATGAAGATCACACCCAGGACTAT3'; NCRNA00173, sense: 5'CCAGGTTTTCCCGAATGTATCT3', antisense: 5'CCACGATCGAGGGAAGGA3'; lincRNA-CALCA, sense: 5'CCAGAAGAGAGCCTGCAACAC3', antisense: 5'CTTCACCATGCCCCCTGAT3'; lincRNA-TSPAN8, sense: 5'TCATCATGATTCTGGGCTTCCT3', antisense: 5'GAAGCAAGCCTATGAAAAACAACA3'; mRNA-TSPAN8, sense: 5'CATCTCTCATTGACTTATCTGGTAGC3', antisense: 5'ACGTCCCCCTAAG GTTTGGT3'; mRNA-CALCB, sense: 5'TCTGTTGTTTTTCATAGGCTTGCT3', antisense: 5'ACTTAGATTTGAAAACAGCTCCTAGGA3'; GAPDH, sense: 5'TGACTTCAACAGCGACACCCA3', antisense: 5'CACCCTGTTGCTGTAGCCAAA3'. The level of each transcript was normalized by the level of GAPDH and represented as fold change using the 2-ΔCt method.

Results

Establishment of lymph node-specific metastasis in the HCCLYM-H cell line

We established the HCCLYM-H2 cell line, which has high lymphatic metastatic potential (100%) and low lung metastatic potential (20%). Figure 1A shows the size of subcutaneous/orthotopic implantation tumors. Figure 1B shows lymph nodes metastases.

LncRNA and mRNA expression profiles in both cell lines

Hierarchical clustering was used to identify lncRNAs and protein-coding mRNAs that were differentially expressed between H2 and LM3 (Figure 3D and E). The expression profiles were shown by calculating log fold

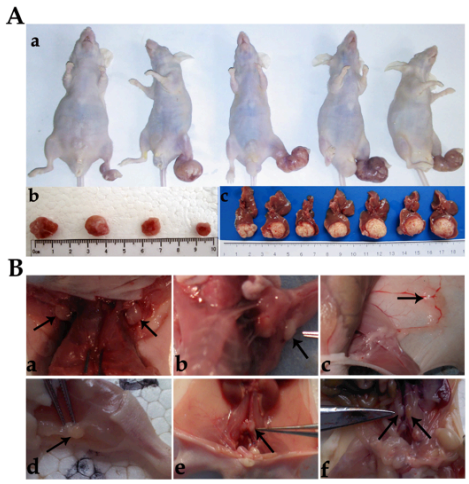

Figure 1. Evaluation of the HCCLYM-H2 cell line's Ability to Initiate Tumor Growth and Metastasize to the Lymph Node in a Mouse Model. A) Tumor size in nude mice with subcutaneous (a & b) and orthotopic implantations(c) of HCCLYM-H2 cells. **B)** HCC metastases in regional lymph nodes: (a) subclavian, (b) axillary, (c) inguinal, (d) popliteal, (e) common iliac, and (f) para-aortic

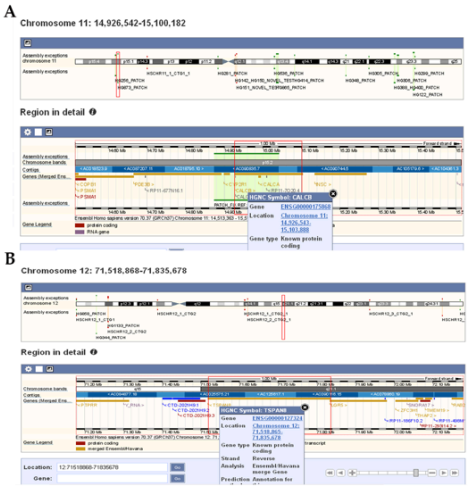

Figure 2. The Expression of lincRNAs CALCA and TSPAN8 is Associated with the Expression of the Protein-coding Genes CALCB and TSPAN8/LGR, Respectively. A) lincRNA CALCA is located on human chromosome 11:14,926,542-15,100,182 next to the CALCB mRNA, which is located 1 bp upstreamon chromosome 11:14,926,543-15,103,888. There is some overlap between the two genes. **B)** lincRNA TSPAN8 is located on chromosome 12: 71,518,868-71,835,678 and is associated with the mRNAs for TSPAN8 and LGR5, located on chromosome12: 71,518,865-71,835,678 and chromosome 12: 71,833,550-71,980,090, respectively. The TSPAN8 and LGR5 coding genes are 3bp and 2.2 kbp, respectively, from lincRNA TSPAN8

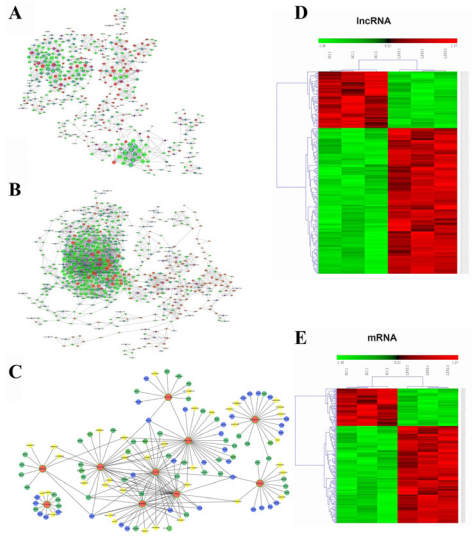

Figure 3. A) The H2 Co-expression Network Consists of lncRNAs (with a blue circle) and mRNAs (without a Blue Circle). Red denotes upregulation and green denotes downregulation in H2. In this network, 132 lncRNAs and 306 mRNAs were linked by 1,462 edges. **B)** In the LM3 co-expression network, 173 lncRNAs and 398 mRNAs were linked by 3,115 edges. **C)** The subnetwork. This subnetwork consists of 10 lncRNAs (red) and co-expressed lncRNAs (yellow) and mRNAs (blue and green). Green denotes mRNAs related to tumors. **D & E)** Changes in expression profiles (rows) across six samples, three from H2 and three from LM3 (columns). Shown are heatmaps of normalized expression values for 1,483 lncRNA loci and 1,629 mRNA loci. The sum of expression across all stages per locus is set to one. The two expression patterns identified correspond to the two HCC cell lines (H2 and LM3)

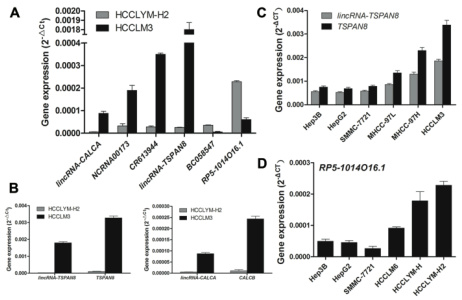

Figure 4. Differential Expression of RNAs. A) Differential expression of lincRNA-CALCA, NCRNA00173, CR613944, lincRNA-TSPAN8, BC058547, and RP5-1014O16.1 in the two HCC cell lines (p<0.01). **B)** Differential expression of lincRNAs CALCA and TSPAN8 and their corresponding mRNAs CALCB and TSPAN8, in the two HCC cell lines (p<0.01). **C)** Differential expression of lincRNA-TSPAN8 and its corresponding mRNA TSPAN8 in the different lung metastatic potentials HCC cell lines (p<0.05). **D)** Differential expression of RP5-1014O16.1 in the different lymphatic metastatic potentials HCC cell lines (p<0.05)

change H2/LM3. A total of 1, 629 mRNAs and 1, 483 lncRNAs exhibited significant differential expression (P≤0.05, ≥1.5-fold change) between the two cell lines; these included antisense lncRNAs, bidirectional lncRNAs, transcripts with overlap to RefSeq exons, and transcripts mapping to intronic and intergenic regions. More differentially expressed lncRNAs were detected in LM3 than H2. Three hundred and six lncRNAs were upregulated in H2, and 1, 177 lncRNAs were downregulated (Table 1). The numbers of expressed intronic and intergenic lncRNAs in H2 (42 and 151, respectively) were comparable to those in LM3 (209 and 545, respectively). RP11-672F9.1 was the lncRNA that showed the greatest increase in expression in H2 (log2 fold change = 9.543524). Long intergenic non-coding RNAs (lincRNAs) TSPAN8 (log2 fold change = 65.7) and CALCA (log2 fold change = 57.2) showed the greatest decrease in expression of all lncRNAs (Figure 2). Interestingly, the fold changes in expression of TSPAN8

Table 1. LncRNAs Differentially Expressed in H2 *vs* LM3 Cell Lines

Seqname	Gene Name	P-value	Fold Change	Chr Length (bp)	Chr	Regulation
ENST00000450980	RP11-672F9.1	0.0000479	9.54	425	chr10	up
ENST00000436499	RP5-1014O16.1	0.00125	7.49	275	chrX	up
ENST00000433664	RP11-501G6.1	0.00126	7.28	817	chr13	up
G43501		0.00873	7.23	143	chr7	up
BC021684		0.0051	5.71	1227	chr2	up
ENST00000394662	RP4-644F6.3	0.0205	5	1740	chr1	up
ENST00000427677	AC024038.1	0.0178	4.71	1064	chrY	up
uc001tvk.1	AK096932	0.000243	4.62	2088	chr12	up
NR_002784	SMEK3P	0.00141	4.61	3133	chrX	up
ENST00000398175	AC090377.1	0.0149	4.47	1785	chr18	up
ENST00000415536	AC003092.1	0.00195	4.19	639	chr7	up
uc002zzz.1	AX747284	0.000795	4.17	1167	chr22	up
ENST00000413041	AC140481.1	0.0413	4.13	594	chr2	up
ENST00000438347	RP11-449I17.5	0.00933	3.99	580	chr10	up
CR615581	lincRNA-ANGPT4	0.0353	3.92	1325	chr20	up
NR_027416	LOC100272146	0.0114	3.9	3216	chr17	up
uc003igi.1	CR604304	0.000363	3.79	1536	chr4	up
ENST00000466692	RP11-1398P2.1	0.00471	3.72	579	chr4	up
ENST00000493219	RP11-80H8.5	0.00384	3.63	1254	chr3	up
ENST00000442417	AC093818.1	0.028	3.62	950	chr2	up
uc004cff.2	BC058547	0.00558	3.56	864	chr9	up
ENST00000313544	OR5J1P	0.0244	3.54	939	chr11	up
BC070168	lincRNA-TSPAN8	0.0000661	65.7	1543	chr12	down
AK054728	lincRNA-CALCA	0.00000194	57.2	2258	chr11	down
NR_023938	C14orf132	0.0000955	30.6	7504	chr14	down
ENST00000510597	RP11-722M1.1	0.00000275	28.8	776	chr4	down
AK095203		0.0000227	19.4	1776	chr12	down
uc001szi.3	CR613944	0.00247	19.1	3386	chr12	down
ENST00000489452	NCRNA00173	0.000128	18.6	543	chr12	down
NR_027345	NCRNA00173	0.00297	15.3	1597	chr12	down
uc003ycp.2	BC038578	0.00158	14.1	1516	chr8	down
ENST00000503677	RP11-91J3.3	0.0000321	12.4	536	chr4	down
BX952962	lincRNA-OBFC2A-2	0.0142	11	428	chr2	down
uc001svy.1	BC047427	0.00335	10.1	2558	chr12	down
uc004bdv.2	AL390170	0.0233	9.59	1685	chr9	down
AF088004		0.0029	8.42	637	chr3	down
AL713738		0.000873	8.38	1301	chr5	down
ENST00000460407	RP11-38P22.2	0.0221	8.16	2050	chr3	down
HIT000098379		0.00345	7.52	227	chr7	down
uc002ufj.3	KLHL23	0.00195	7.3	1863	chr2	down
uc001gfa.1	AX746887	0.00104	7.25	1149	chr1	down
AK090694		0.00849	7.23	2488	chr5	down
NR_027378	LOC643763	0.000377	7.15	7056	chr8	down
uc003szc.1	AK093987	0.0124	7.04	2285	chr7	down
AK127243		0.00321	6.99	3948	chr4	down
BC013657		0.0304	6.53	969	chr5	down
NR_002191	PPP1R2P9	0.00503	6.23	894	chrX	down
AK123638		0.000859	6.15	2113	chr7	down
ENST00000477702	NCRNA00173	0.00665	6.05	314	chr12	down
NR_021490	FLJ42709	0.000431	5.89	2791	chr5	down
NR_023388	PRINS	0.00635	5.74	2199	chr10	down
ENST00000485347	RP11-889D3.1	0.00000537	5.71	429	chr3	down
nc-HOXA13-99+	nc-HOXA13-99	0.00386	5.65	148	chr7	down
AF070632		0.00599	5.57	1446	chr1	down
AL050204		0.00991	5.57	1655	chr11	down
ENST00000413244	CTA-929C8.8	0.00946	5.55	571	chr22	down
uc002odz.1	AX747582	0.00052	5.55	1724	chr19	down
CR601061		0.0171	5.52	1793	chr21	down
AF297014		0.00681	5.49	1496	chr1	down
ENST00000417339	RP11-56H2.1	0.00138	5.48	486	chrX	down
AK027294		0.00848	5.45	1673	chr8	down
AK098142		0.000893	5.35	2364	chr11	down
ENST00000432535	RP11-456A18.1	0.0199	5.27	477	chr10	down
ENST00000496674	RP11-314M24.1	0.000421	5.27	339	chr3	down
AF452715		0.000127	5.27	1270	chr19	down
CR599788		0.0000451	5.18	1536	chr17	down
ENST00000429725	C1orf99	0.00362	5.17	2393	chr1	down
uc001trz.2	AJ276555	0.000296	5.08	363	chr12	down
BC024158		0.00575	5.08	1816	chrX	down

*LncRNAs are selected from the 306 upregulated in H2 according to greatest fold change and 1177 downregulated in H2 according to greatest fold change. "up" means that the gene is upregulated in H2 compared with LM3, and "down" means that the gene is downregulated in H2 compared with LM3

and CALCB mRNAs (48.6 and 74.1, respectively) were as high as those for the lincRNAs, according to the microarray data.

Of 1, 629 differentially expressed mRNAs, 717 mRNAs were upregulated in H2 while 912 were downregulated (Table 2). GO analysis showed that differentially expressed mRNAs were related to cell adhesion, cell migration, apoptosis, and regulation of transcription. Pathway analysis revealed that the functions of the differentially expressed mRNAs included primary substance metabolism, tumorigenesis, and inflammatory factors associated with signaling pathways.

LncRNA classification and subgroup analysis

Our data contained subgroups of lncRNAs, such as Rinn lincRNAs (Guttman et al., 2009; Khalil et al., 2009), enhancer-like lncRNAs, and HOX lncRNAs. Microarray results suggested that 2, 163 lincRNA transcripts could be detected in both cell lines, and 43 of them were differentially expressed. Among the differentially expressed lincRNA transcripts, 15 were upregulated and 28 were downregulated in H2 (Table 3). The coding genes near the differentially expressed lincRNAs are also shown in Table 3. Differentially expressed enhancer-like lncRNAs and nearby coding genes (distance < 300 kb) are shown in Table 4. These 12 differentially expressed enhancer-like lncRNAs included 1 lncRNA upregulated in H2 and 11 downregulated and regulated nearby coding mRNAs in cis. Moreover, 407 HOX lncRNAs from four HOX loci were detected.

Co-expression network

Microarray-based CNC networks were used to cluster lncRNAs and mRNAs into phenotypically relevant co-expression modules. The structure of the CNC networks of lung metastatic and lymph node metastatic cell lines was significantly different (Figure 3A and B), indicating the inter-regulation of lncRNAs and mRNAs varied in these cell lines with different metastatic potential. In the H2 CNC network, 132 lncRNAs and 306 mRNAs were linked by 1, 462 edges. Nearly 680 edges (46.51%) connected mRNAs, 647 (44.25%) connected mRNAs and lncRNAs, and 135 (9.23%) linked pairs of lncRNAs (Figure 3A). In the LM3 CNC network, 173 lncRNAs and 398 mRNAs were linked by 3, 115 edges. In this CNC, 1, 481 edges (47.54%) connected mRNAs, 1, 293 edges (41.51%) connected mRNAs and lncRNAs, and 341 (10.95%) linked pairs of lncRNAs (Figure 3B). The CNC networks indicated that a single mRNA could be linked to 1-10 lncRNAs, and the same was true of lncRNAs. A subnetwork was constructed from the ten lncRNAs (CR613944, KLHL23, TPRXL, AX747582, AX746887, NCRNA00173, BC058547, RP11-672F9.1, SMEK3P, AX747284) with the highest level of co-expression (Figure 3C). In the subnetwork, the ten lncRNAs are linked with their co-expreesion genes including lncRNAs and mRNAs. Some mRNAs are related to tumors. Four intergenic lncRNAs, TPRXL, AX747582, AX746887, and SMEK3P, were all co-expressed with TFF1 mRNA. The gene for lncRNA RP11-672F9.1, upregulated in H2, was co-expressed with 10 mRNAs. The gene for lncRNA CR613944, downregulated in H2, was co-expressed with 22 mRNAs. Among the co-expressed genes, RASAL1 was also located on chromosome 12, relatively close to CR613944. Two lincRNAs, CR613944 and AX747582, were downregulated in H2 and co-expressed with WISP2 mRNA.

Gene set enrichment analysis

To define the functions of the differentially expressed lncRNAs and their co-expressed mRNAs, we performed GSEA. When analyzing both lncRNA and mRNA expression levels, the six samples (three from H2 and three from LM3) could be grouped into two broad classes (Figure 3D and E). These co-expression patterns corresponded to the two HCC cell lines. These results suggest that the cell lines'different lncRNA expression profiles may be related to their distinct patterns of organ-specific metastasis.

The verification to the microarray data

To further validate microarray data, we selected six lncRNAs (CR613944, BC058547, RP5-1014O16.1, NCRNA00173, lincRNA-CALCA, lincRNA-TSPAN8) and two mRNAs (TSPAN8, CALCB) in the two HCC cell lines using qRT-PCR. The results demonstrated that RP5-1014O16.1 and BC058547 were upregulated in H2 and lincRNA-TSPAN8, lincRNA-CALCA, CR613944, and NCRNA00173 were downregulated (p<0.01 for each lncRNA, Figure 4A). The expression of lincRNAs CALCA and TSPAN8 were concordant with that of the corresponding mRNAs (CALCB and TSPAN8, respectively) in the two cell lines (Figure 4B). Data obtained from qRT-PCR and the microarray was consistent.

Expression of lincRNA-TSPAN8 and TSPAN8 were evaluated in HCC cells with different lung metastatic potiential by qRT-PCR

Since lincRNA-TSPAN8 was found the most upregulated lncRNA in LM3 (Table 1), to evaluate the biological functions of lincRNA-TSPAN8, we first examined the expression of the gene in a variety of cell lines with different lung metastatic potential, including Hep3B, HepG2, SMMC-7721, MHCC-97L, MHCC-97H and LM3 by qRT-PCR. As shown in Figure 4C, comparing with no or low lung metastatic potiential cell lines (Hep3B, HepG2, SMMC-7721 and MHCC-97L), lincRNA-TSPAN8 was found highly expressed in high lung metastatic potiential cell lines (MHCC-97H and LM3). And the trend of the corresponding mRNA TSPAN8 expression in these HCC cells was the same as lincRNA-TSPAN8 (Figure 4C).

Expression of RP5-1014O16.1 was evaluated in HCC cells with different lymphatic metastatic potential by qRT-PCR

RP5-1014O16.1 was the second most upregulated lncRNA in H2 (Table 1). It is a new-found lncRNA, which is located on chrX: 149758021-149758295. To obtain the expression information about RP5-1014O16.1 in HCC, we detected its level in several different lymphatic metastatic potential HCC cell lines. We found that comparing with

Table 2. mRNAs Differentially Expressed in H2 *vs* LM3 Cell Lines

Seqname	Gene Name	P-value	Fold Change	Chr Length (bp)	Chr	Regulation
NM_032413	C15orf48	0.00195	30.8	929	chr15	up
NM_031246	PSG2	0.000198	24.8	1533	chr19	up
NM_001130167	PSG8	0.000216	22.4	2031	chr19	up
NM_197955	C15orf48	0.000814	18.7	815	chr15	up
NM_002784	PSG9	0.000276	12.7	1706	chr19	up
NM_001078	VCAM1	0.00305	12.4	3119	chr1	up
NM_001130014	PSG5	0.000508	12.4	1601	chr19	up
NM_002780	PSG4	0.000126	10	2059	chr19	up
NM_145006	SUSD3	0.00375	8.68	1205	chr9	up
NM_021016	PSG3	0.000202	8.67	1922	chr19	up
NM_002571	PAEP	0.00154	8.2	828	chr9	up
NM_002781	PSG5	0.00429	8.2	1669	chr19	up
NM_003175	XCL2	0.000349	8.2	566	chr1	up
NM_001128850	RRAD	0.00181	8.14	1476	chr16	up
NM_002922	RGS1	0.00753	8.11	1403	chr1	up
NM_182707	PSG8	0.000125	7.94	1441	chr19	up
NM_001145155	NR2F2	0.00781	7.72	3869	chr15	up
NM_203287	PSG11	0.000983	7.65	1243	chr19	up
NM_001898	CST1	0.00117	7.43	782	chr20	up
NM_002783	PSG7	0.00212	7.41	2046	chr19	up
NM_001150	ANPEP	0.0141	7.28	3740	chr15	up
NM_006498	LGALS2	0.00158	7.15	543	chr22	up
NM_001322	CST2	0.0044	7	694	chr20	up
NM_001012301	ARSI	0.000318	6.86	3225	chr5	up
NM_030754	SAA2	0.0262	6.79	545	chr11	up
NM_001113410	PSG11	0.00243	6.78	1175	chr19	up
NM_199161	SAA1	0.00478	6.41	531	chr11	up
NM_001136017	CCND3	0.000224	6.37	2104	chr6	up
NM_002658	PLAU	0.00387	6.27	2395	chr10	up
NM_130759	GIMAP1	0.0245	6.21	1248	chr7	up
NM_000728	CALCB	0.000152	74.1	1031	chr11	down
NM_004356	CD81	0.000119	62	1497	chr11	down
NM_013230	CD24	0.000752	52. 6	2194	chrY	down
NM_004616	TSPAN8	0.00000354	48.6	1159	chr12	down
NM_025237	SOST	0.000231	38.1	2322	chr17	down
NM_000597	IGFBP2	0.0000244	36.7	1439	chr2	down
NM_001102470	ADH6	0.00291	36.3	2803	chr4	down
NM_002364	MAGEB2	4.95E-09	35.7	1628	chrX	down
NM_001400	S1PR1	0.0000179	34	3050	chr1	down
NM_000669	ADH1C	0.0000435	27.8	1497	chr4	down
NM_006744	RBP4	0.0000247	26.9	941	chr10	down
NM_153488	MAGEA2B	0.000633	24.2	1629	chrX	down
NM_214462	DACT2	0.000321	22.4	2942	chr6	down
NM_024993	LRRTM4	0.00000302	19.6	3623	chr2	down
NM_012072	CD93	0.000245	16	6701	chr20	down
NM_004496	FOXA1	0.000399	15.4	3124	chr14	down
NM_001080848	CSAG2	0.0015	13.6	847	chrX	down
NM_002333	LRP3	0.00321	13	3714	chr19	down
NM_020855	ZNF492	1.91E-08	13	4245	chr19	down
NM_014505	KCNMB4	0.00139	12.8	1631	chr12	down
NM_018962	DSCR6	0.0000805	12.2	2260	chr21	down
NM_005288	GPR12	0.0107	11.6	4863	chr13	down
NM_003385	VSNL1	0.00839	11.5	2014	chr2	down
NM_015696	GPX7	0.000962	11.4	1246	chr1	down
NM_002196	INSM1	0.00113	10.6	2838	chr20	down
NM_016593	CYP39A1	0.00133	10.4	2288	chr6	down
NM_019066	MAGEL2	0.00141	10.2	4298	chr15	down
NM_001032278	MMP28	0.000728	10.1	1042	chr17	down
NM_004744	LRAT	0.0159	9.91	4909	chr4	down
NM_031474	NRIP2	0.00157	9.85	2787	chr12	down
NM_020361	CPA6	0.00155	9.72	1907	chr8	down
NM_018397	CHDH	0.00105	9.6	3642	chr3	down
NM_021969	NR0B2	0.0000964	9.47	1277	chr1	down
NM_018476	BEX1	0.0312	9.43	862	chrX	down
NM_003225	TFF1	0.000276	9.29	508	chr21	down
NM_000277	PAH	0.000664	9.03	2680	chr12	down
NM_005654	NR2F1	0.000124	8.7	3210	chr5	down
NM_153832	GPR161	0.000298	8.63	2733	chr1	down
NM_004497	FOXA3	0.000836	8.61	2046	chr19	down
NM_005139	ANXA3	0.00841	8.57	1634	chr4	down
NM_005213	CSTA	0.0256	8.52	838	chr3	down
NM_032947	C5orf62	0.0247	8.38	1993	chr5	down
NM_000522_Exon2-	NM_000522	0.00295	8.31	1563	chr7	down
NM_001295	CCR1	0.000882	8.29	2690	chr3	down
NM_000522	HOXA13	0.002	8.2	2514	chr7	down
NM_017709	FAM46C	0.00486	8.17	5720	chr1	down
NM_206852	RTN1	0.0136	8.13	1710	chr14	down
NM_001134745	LRRTM4	0.000165	7.98	3167	chr2	down
NM_173833	SCARA5	0.000462	7.69	3643	chr8	down
NM_174936	PCSK9	0.000645	7.58	3636	chr1	down
NM_001166243	FHIT	0.0000052	7.28	1122	chr3	down
NM_212559	XKRX	0.0024	7.11	2854	chrX	down
NM_005284	GPR6	0.0478	7.1	1645	chr6	down
NM_016084	RASD1	0.0000871	7.1	1758	chr17	down
NM_020437	ASPHD2	0.0166	7.1	3313	chr22	down

*mRNAs are selected from the 717 upregulated in H2 according to greatest fold change and the 912 downregulated in H2 according to greatest fold change. "up" means that the gene is upregulated in H2 compared with LM3, and "down" means that the gene is downregulated in H2 compared with LM3

Table 3. LincRNAs Differentially Expressed in H2 *vs* LM3 Cell Lines

Gene Name	Fold Change	Regulation	Nearby Gene	Fold Change	Regulation
AL358933.1	1.81	down	HIST1H2AL	2.03	down
RP11-314A15.2	2.02	up	VSIG4	4.99	down
lincRNA-LOC100506581-1	2.76	down	FOXL1	1.65	down
CR933665	1.6	up	TATDN1	1.93	up
BC025370	1.68	down	SERP2	2.47	down
lincRNA-RPS14-2	1.78	up	ARSI	6.86	up
lincRNA-CALCA	57.2	down	CALCB	74.1	down
LOC100130872-SPON2	1.71	down	IDUA	1.51	up
lincRNA-PGS1	2.87	up	SOCS3	1.62	up
lincRNA-PROX1-3	1.89	down	RPS6KC1	2.78	up
AC079767.4	3.15	up	PLEKHM3	1.55	down
RP11-792D21.1	1.53	down	ANXA3	8.57	down
lincRNA-CENPW-2	2.02	down	HINT3	1.56	up
lincRNA-TSPAN8	65.7	down	TSPAN8	48.6	down
AC002066.1	2.2	up	TFEC	3.09	up
lincRNA-SP100	2.57	up	SP140L	2.2	up
AC079767.4	3.15	up	CREB1	1.83	up
lincRNA-WISP1	1.81	down	PHF20L1	1.83	down
RP1-72A23.3	1.72	up	RNGTT	1.6	up
AL358933.1	1.81	down	HIST1H4J	1.7	down
lincRNA-KCNA2-1	2.47	down	CD53	1.5	up
lincRNA-BUB1B	1.53	down	C15orf23	2.64	up
lincRNA-BUB1B	1.53	down	C15orf23	3.89	up
AL358933.1	1.81	down	HIST1H2AM	2.34	down
lincRNA-ZNF532	1.85	down	MALT1	2.93	up
RP11-366O17.2	2.64	down	DNAJC15	1.85	down
lincRNA-RFC2	1.88	down	CLIP2	2.64	up
lincRNA-CDON-2	1.59	down	DDX25	6.68	down
lincRNA-CDON-1	1.56	down	DDX25	6.68	down
lincRNA-ADCY1-1	2.4	down	RAMP3	2.05	down
lincRNA-NKX2-5-1	1.79	down	STC2	2.08	up
AC002066.1	2.2	up	TFEC	3.32	up
lincRNA-ZNF365-2	4.87	down	ZNF365	4.32	down
RP4-612B18.1	1.68	down	METTL13	2.15	up
CR590356	1.76	up	RIPK2	2.26	up
lincRNA-INHBB-4	1.86	up	INHBB	4.76	down
AK026965	2.17	down	DDX46	1.62	down
lincRNA-KIAA1199	1.99	up	MESDC1	1.67	up
BC038570	2.5	down	PRR15	2.75	down
LOC100272228	1.52	up	HSFX2	1.56	down
CR933665	1.6	up	MTSS1	1.74	down
lincRNA-ZNF532	1.85	down	GRP	1.75	up
lincRNA-TSPAN8	65.7	down	LGR5	3.77	down

*"down" means that the gene is downregulated in H2 compared with LM3, and "up" means that the gene is upregulated in H2 compared with LM3

Table 4. Differentially Expressed Enhancer-Like lncRNAs and their Nearby mRNAs

Gene Name	Fold Change	Regulation	Nearby mRNA	Fold Change	Regulation
BC025370	1.68	down	SERP2	2.467	down
ANTXRL	1.86	down	ANXA8L2	1.94	up
AMZ2P1	1.75	up	LRRC37A3	2.01	up
LOC100289019	2.47	down	LCN2	3.42	up
RP4-673D20.3	2.72	down	SIRPA	2.07	down
RP5-1065P14.2	3.18	down	SPHAR	2.17	down
AC004009.3	4.09	down	HOXA13	8.2	down
RP11-439L18.1	1.73	down	AIG1	1.57	up
RP3-425P12.1	2.45	down	GMNN	1.73	up
RP11-112J3.16	3.63	down	RUSC2	2.61	down
RP3-331H24.4	1.76	down	OGFRL1	3.56	up
RP11-366O17.2	2.64	down	DNAJC15	1.85	down

*"down" means that the gene is downregulated in H2 compared with LM3, and "up" means that the gene is upregulated in H2 compared with LM3

no lymphatic metastatic potential HCC cell lines (Hep3B, HepG2, SMMC-7721), RP5-1014O16.1 was highly expressed in high lymphatic metastatic potential HCC cell lines (HCCLM6, HCCLYM-H, H2) (Figure 4D).

Discussion

Recent studies have shown that many thousands of lncRNAs are encoded in the human genome. They serve as transcriptional and post-transcriptional regulators and as guides for chromatin-modifying complexes, and many affect various cellular and developmental pathways (Mercer et al., 2009; Wilusz et al., 2009; Taft et al., 2010). It is not surprising that the dysregulation of lncRNAs appears to be a significant feature of many complex human diseases, especially cancer (Ren et al., 2013). For example, the expression of lncRNA BC200 is upregulated invarious human tumors, such as breast, lung, parotid gland, ovary, cervix, and tongue cancers, but it is undetectable incorresponding normal tissue (Chen et al., 1997; Iacoangeli et al., 2004). Similarly, H19, which is located within a cluster of imprinted genes on human chromosome 11 in p15.5, is expressed in many types of cancers, such as gastric, breast, liver, and esophageal, at significantly higher levels than in corresponding normal tissue (Hibi et al., 1996; Berteaux et al., 2008; He et al., 2014; Zhang et al., 2014). Other lncRNAs are expressed in normal tissues and act as tumor suppressor genes. One example, MEG3, is expressed in normal tissues of the adrenal gland, pancreas, ovary, brain tissue, and pituitary; it is rarely expressed in pituitary tumors or human cancer cell lines. Moreover, ectopic expression of this gene inhibits the growth of human cancer cell lines, including HeLa, MCF-7 and H4, indicating that MEG3 may represent a novel tumor suppressor (Zhang et al., 2003). The loss of GAS5 expression have been found in many types of tumors including melanoma, prostate cancers and breast but the concrete mechanism still needs further research (Smedley et al., 2000; Nupponen and Carpten, 2001; Mourtada-Maarabouni et al., 2009). In addition, the expression of GAS5 in Renal Cell Carcinoma specimens was obviously lower than that in adjacent normal tissues, and GAS5 can arrest cell cycling, induce cell apoptosis, also suppress cell migration and invasion (Qiao et al., 2013). These results suggest that GAS5 may have potential value to become a tumor marker for several tumors. . Previous studies showed that TUG1 over-express in bladder cancer and is connected to some characteristics of tumor cells, such as proliferation , apoptosis and so on (Khalil et al., 2009). Additionally, TUG1 can inhibit cancer cell proliferation and promote apoptosis in osteosarcoma (Zhang et al., 2013). TUG1 may act as a new diagnostic marker and therapeutic target of these tumors.

It is believed that lncRNAs play an important regulatory role in cancer progression. Some research suggests that lncRNAs are associated with cancer metastasis and prognosis. According to Gupta et al. (2010), an lncRNA called HOTAIR can be overexpressed nearly two-thousand-fold in breast cancer metastases. High HOTAIR expression levels are significantly associated with breast tumor metastasis and a low survival rate. HOTAIR regulates metastatic progression by recruiting the PRC2 complex to specific, genome-wide target genes; this, in turn, leads to H3K27 methylation and epigenetic silencing of metastasis suppressor genes such as JAM2 and PCDH1 (Simon and Kingston, 2009). It has been known for some time that another lncRNA, MLATA1, is associated with lung cancer metastasis and poor prognosis (Ji et al., 2003). Recent studies show that MLATA1 is also overexpressed in breast cancer, prostate tumors, rectal carcinoma, HCC, and cervical cancer (Lin et al., 2007; Guo et al., 2010). In addition, a number of reports have demonstrated that MLATA1 overexpression is linked to cancer metastasis (Yamada et al., 2006). Recent studies have also revealed that lncRNAs exhibit different patterns of expression in different types of tumors. Some lncRNAs are very sensitive and specific markers of tumors; one example includes DD3 (also known as PCA3) in prostate tumors. Because its expression appears to be restricted to the prostate and it is highly overexpressed in prostate cancer cells, DD3 is a promising marker for the early diagnosis of prostate cancer (Bussemakers et al., 1999; de Kok et al., 2002). Similarly, OCC1 encodes two non-coding regulatory RNAs and its expression is restricted to human colon carcinoma cells (Pibouin et al., 2002). Finally, the HOST genes are rarely expressed in normal tissues or non-ovarian cancers, but they are frequently expressed in ovarian cancer-derived cell lines and primary tumors. Therefore, the HOST genes have been proposed to be specific biomarkers and their study may lead to novel strategies for ovarian cancer diagnosis and therapy (Rangel et al., 2003). Multiple studies have addressed the ectopic expression of lncRNAs in HCC. For example, the highly upregulated in liver cancer (HULC) gene encodes an mRNA-like non-coding RNA (ncRNA) that is highly upregulated in HCC tissue (Panzitt et al., 2007; Matouk et al., 2009). HULC may downregulate miR-372 and induce phosphorylation of the cAMP responsive element binding protein1 (CREB1) in liver cancer (Wang et al., 2010). In addition, Yang et, al's studies (Yang et al., 2011) indicate that the lncRNA HEIH promotes tumor progression. The expression level of lncRNA-HEIH in hepatitis Bvirus (HBV)-related HCC is associated with recurrence and is an independent prognostic factor for survival. Measuring lncRNA-HEIH levels may help predict HCC patient prognosis. According to another study, levels of the lncRNA uc.338 are elevated in HCC cells and may be a promising marker for this type of cancer. Aberrant expression of its transcript, TUC338, is found in transformed hepatocytes, and its functional role in modulating growth may make it a desirable therapeutic target for selected HCC cases (Braconi et al., 2011). Collectively, these studies lead us to propose that lncRNAs may serve as key regulatory hubs in HCC progression. Until now, however, there have been few reports on the role of lncRNAs in the organ-specific metastasis of HCC.

In this study, we established the HCC cell line H2, which has high potential to metastasize to the lymph nodes. It and cell line LM3, which has high potential to metastasize to the lung, are similar in terms of genetic background, but very different in terms of metastatic potential. By comparing their expression profiles and validating a significant portion of differentially expressed lncRNAs using qRT-PCR, we were able to identify a small number of lncRNAs related to organ-specific metastasis in HCC. There have been several reports of aberrant lncRNA expression in various types of human cancers, including

HCC (Zhu et al., 2012). However, to the best of our knowledge, this is the first time that lncRNA expression profiles have been studied with regard to organ-specific metastasis in HCC.

Based on microarray data, we detected thousands of expressed lncRNAs in both cell lines, thousands of which were differentially expressed. We found that 306 lncRNAs were upregulated and 1177 downregulated in the H2 cell line, and the function of a large number of these RNAs is unknown. They may be involved in the occurrence, progression, and organ-specific metastasis of HCC. Hence, our work contributes new potential biomarkers for HCC metastasis. It also highlights the importance of investigating the biological relevance of lncRNAs to fully understand the molecular basis of organ-specific HCC metastasis. Furthermore, this study may help identify more effective therapeutic targets and facilitate the development of new personalized therapeutic strategies.

We constructed CNC networks to identify the mRNAs associated with 10 lncRNAs then performed GSEA to define their functions. The ten lncRNAs have the highest co-expression. The results revealed that four intergenic lncRNAs, TPRXL, AX747582, AX746887, and SMEK3P, were all co-expressed with TFF1 mRNA, which is associated with gastrointestinal tumors (Uchino et al., 2000). The gene for lncRNA RP11-672F9.1, upregulated in H2, was co-expressed with 10 mRNAs. One of these mRNAs, H2AFY2, is located near lncRNA RP11-672F9.1, which is reported to predict lung cancer recurrence (Sporn et al., 2009). The gene for lncRNA CR613944, downregulated in H2, was co-expressed with 22 mRNAs, 12 of which were related to cell adhesion, cell migration, and various types of cancer. Among the co-expressed genes, RASAL1 was also located on chromosome 12, relatively close to CR613944. This gene has been reported to contribute to colon tumor progression and gastric tumorigenesis (Ohta et al., 2009; Seto et al., 2011). Two lincRNAs, CR613944 and AX747582, were downregulated in H2 and co-expressed with WISP2 mRNA, which is translated into an important regulator involved in tumor cell invasion and metastasis (Fritah et al., 2008). Whether the progression and organ-specific metastasis of HCC are regulated by these lncRNAs warrants further study.

It has been shown that the transcription of lncRNAs can affect the expression of nearby coding genes (Da Sacco et al., 2012). LncRNAs can recruit chromatin-modifying enzymes to regulate the expression of genes, either in cis (near the site of lncRNA production) or in trans (when the genes involved are distant) through a phenomenon called transvection (Da Sacco et al., 2012). In our study, the lncRNAs that displayed the greatest decrease in expression in H2 were the lincRNAs TSPAN8 and CALCA; the nearby genes encoding TSPAN8 and CALCB mRNA were also differentially expressed in the H2 and LM3 cell lines and the fold change in their expression was equally high (48.58408 and 74.100555, respectively). These results indicate that the TSPAN8 and CALCA lncRNAs may increase the expression of nearby coding transcripts in cis. The possibility that these lncRNAs contribute to the different metastatic potential of the cell lines via cis regulatory functions requires further investigation.

The subgroup of lncRNAs contains Rinn lincRNAs, enhancer-like lncRNAs, and HOX lncRNAs. The subgroup analysis of lncRNAs consists of Rinn lincRNAs profiling, enhancer-like lncRNAs profiling, and HOX lncRNAs profiling, lincRNAs nearby coding gene data and enhancer LncRNAs nearby coding gene data. It may also help clarify the relationship between lncRNAs and the organ-specific metastasis of HCC. A large number of HOX lncRNAs, clustered around four chromosomal loci, termed HOXA through HOXD, are essential for specifying the positional identities of cells (Rinn et al., 2007). This subgroup of lncRNAs is expressed in a temporal and spatial pattern and could play a significant role in HOX regulation (Kmita and Duboule, 2003; Lemons and McGinnis, 2006). The HOTAIR lincRNA is one of the best-studied HOX lincRNAs, initially discovered as a repressor of the HOXD genes that is expressed from the HOXC lincRNA locus. HOTAIR lincRNA is a potential biomarker for metastasis and high probability of death in breast cancer patients (Gupta et al., 2010). Recently, another set of lncRNAs were identified as gene expression enhancers in multiple human cell lines, such as fibroblasts, keratinocytes, and HeLa cells. It was reported that the depletion of some of these lncRNAs led to decreased expression of neighboring protein-coding genes, such as the master regulator of hematopoiesis, SCL, Snai1, and Snai2 (Orom et al., 2010) (Orom et al., 2010) (Orom et al., 2010) (Orom et al., 2010).

Six lncRNAs and two mRNAs were selected for further validation by qRT-PCR, and the results were consistent with the microarray data. To obtain more information about these genes expression in HCC, we investigated their levels in some different metastatic potentials HCC cell lines. We found that the expression level of RP5-1014O16.1 in high lymphatic metastatic potentials HCC cell lines is evidently higher than that in no lymphatic metastatic potentials HCC cell lines. The data indicate that RP5-1014O16.1 is possible promote the lymphatic metastasis of HCC. RP5-1014O16.1 is a new-found lncRNA, which is located on chrX:149758021-149758295. The mechanisms of RP5-1014O16.1 promoting tumor metastasis need further research. In addition, the expression levels of lincRNA-TSPAN8 and the corresponding mRNA TSPAN8 in high lung metastatic potentials HCC cell lines were found to be evidently higher than that in low or no lung metastatic potentials HCC cell lines. Protein Tspan8 is encoded by TSPAN8, also known as TM4SF3 (transmembrane 4 superfamily 3), CO-029, D6.1A (in rats), a member of tetraspanin family, has been reported as a cancer associated gene in many types of tumors, like tumors of the gastrointestinal tract-gastric, colorectal, pancreatic and liver tumors (Kanetaka et al., 2001; Zoller, 2006). Tspan8 overexpression has been reported to be correlated with intrahepatic spread of HCC (Kanetaka et al., 2003). The tumor growth-promoting and metastasis-promoting activity of the Tspan8 is possible due to its capacity to induce angiogenesis and cancer cell motility (Gesierich et al., 2006; Yue et al., 2013). Taken together, these results show that lincRNA-TSPAN8

and TSPAN8 might play an important role in the lung metastasis of HCC. We speculate that lincRNA-TSPAN8 promotes the lymphatic metastasis of HCC possible via regulating TSPAN8 expression. The concrete mechanisms need further study.

In summary, the present work compares the expression profiles of two HCC cell lines with different metastatic potential. The results may contribute to more efficient diagnosis, aid in the identification of therapeutic targets, and facilitate the development of new, personalized therapeutic strategies. Future studies may reveal whether the progression and organ-specific metastasis of HCC is regulated by these lncRNAs or by other targets.

Acknowledgements

This research project was supported by grants from the National Basic Reseach Program of China (no. 2013CB910501).

References

Barabasi AL (2004). Network biology: Understanding the cell's functional organization. *Nat Rev Genet*, **5**, 101-15.

Bartel DP (2009). MicroRNAs: Target Recognition and Regulatory Functions. *Cell*, **136**, 215-33.

Berteaux N, Aptel N, Cathala G, et al (2008). A novel H19 antisense rna overexpressed in breast cancer contributes to paternal IGF2 expression. *Mol Cell Biol*, **28**, 6731-45.

Bertone P, Stolc V, Royce TE, et al (2004). Global identification of human transcribed sequences with genome tiling arrays. *Science*, **306**, 2242-6.

Braconi C, Valeri N, Kogure T, et al (2011). Expression and functional role of a transcribed noncoding RNA with an ultraconserved element in hepatocellular carcinoma. *Proc Nat Acad Sci USA*, **108**, 786-91.

Bussemakers M, Verhaegh GW, Karthaus H, et al (1999). DD3: A new prostate-specific gene, highly overexpressed in prostate cancer. *Cancer Res*, **59**.

Cabili MN, Trapnell C, Goff L, et al (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev*, **25**, 1915-27.

Chen W, Bocker W, Brosius J, et al (1997). Expression of neural BC200 RNA in human tumours. *J Pathol*, **183**, 345-51.

Da Sacco L, Baldassarre A, Masotti A (2012). Bioinformatics Tools and Novel Challenges in Long Non-Coding RNAs (lncRNAs). Functional Analysis. *Int J Mol Sci*, **13**, 97-114.

de Kok JB, Verhaegh GW, Roelofs RW, et al (2002). DD3 (PCA3)., a very sensitive and specific marker to detect prostate tumors. *Cancer Res*, **62**, 2695-8.

Dexter DL, Kowalski HM, Blazar BA, et al (1978). Heterogeneity of tumor cells from a single mouse mammary tumor. *Cancer Res*, **38**, 3174-81.

Fejes-Toth K, Sotirova V, Sachidanandam R, et al (2009). Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs. *Nature*, **457**, 1028-32.

Fritah A, Saucier C, De Wever O, et al (2008). Role of WISP-2/CCN5 in the maintenance of a differentiated and noninvasive phenotype in human breast cancer cells. *Mol Cell Biol*, **28**, 1114-23.

Genda T, Ichida T, Kojiro M, et al (1999). Cell motility mediated by Rho and Rho-associated protein kinase plays a critical role in intrahepatic metastasis of human hepatocellular carcinoma. *Hepatology*, **30**, 1027-36.

Gesierich S, Berezovskiy I, Ryschich E, et al (2006). Systemic induction of the angiogenesis switch by the tetraspanin D6.1A/CO-029. *Cancer Res*, **66**, 7083-94.

Guo FJ, Li YL, Liu Y, et al (2010). Inhibition of metastasis-associated lung adenocarcinoma transcript 1 in CaSki human cervical cancer cells suppresses cell proliferation and invasion. *Acta Biochim Biophys Sin*, **42**, 224-9.

Gupta RA, Shah N, Wang KC, et al (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*, **464**, 1071-6.

Guttman M, Amit I, Garber M, et al (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*, **458**, 223-7.

Guttman M, Donaghey J, Carey BW, et al (2011). lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature*, **477**, 295-300.

Guttman M, Rinn JL (2012). Modular regulatory principles of large non-coding RNAs. *Nature*, **482**, 339-46.

Hanahan D, Weinberg RA (2011). Hallmarks of cancer: the next generation. *Cell*, **144**, 646-74.

He Y, Meng XM, Huang C, et al (2014). Long noncoding RNAs: Novel insights into hepatocelluar carcinoma. *Cancer Lett*, **344**, 20-7.

Hibi K, Nakamura H, Hirai A, et al (1996). Loss of H19 imprinting in esophageal cancer. *Cancer Res*, **56**, 480-2.

Hong SS, Kim TK, Sung KB, et al (2003). Extrahepatic spread of hepatocellular carcinoma: a pictorial review. *Eur Radiol*, **13**, 874-82.

Iacoangeli A, Lin Y, Morley EJ, et al (2004). BC200 RNA in invasive and preinvasive breast cancer. *Carcinogenesis*, **25**, 2125-33.

Jemal A, Bray F, Center MM, et al (2011). Global cancer statistics. *CA-Cancer J Clin*, **61**, 69-90.

Ji P, Wang WB, Metzger R, et al (2003). MALAT-1, a novel noncoding RNA, and thymosin beta 4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene*, **22**, 8031-41.

Kanetaka K, Sakamoto M, Yamamoto Y, et al (2001). Overexpression of tetraspanin CO-029 in hepatocellular carcinoma. *J Hepatol*, **35**, 637-42.

Kanetaka K, Sakamoto M, Yamamoto Y, et al (2003). Possible involvement of tetraspanin CO-029 in hematogenous intrahepatic metastasis of liver cancer cells. *J Gastroenterol Hepatol*, **18**, 1309-14.

Kapranov P, Cawley SE, Drenkow J, et al (2002). Large-scale transcriptional activity in chromosomes 21 and 22. *Science*, **296**, 916-9.

Khalil AM, Guttman M, Huarte M, et al (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A*, **106**, 11667-72.

Kmita M, Duboule D (2003). Organizing axes in time and space; 25 years of colinear tinkering. *Science*, **301**, 331-3.

Lemons D, McGinnis W (2006). Genomic evolution of Hox gene clusters. *Science*, **313**, 1918-22.

Li Y, Tang Y, Ye L, et al (2003). Establishment of a hepatocellular carcinoma cell line with unique metastatic characteristics through *in vivo* selection and screening for metastasis-related genes through cDNA microarray. *J Cancer Res Clin Oncol*, **129**, 43-51.

Li Y, Tang ZY, Ye SL, et al (2001). Establishment of cell clones with different metastatic potential from the metastatic hepatocellular carcinoma cell line MHCC97. *World J Gastroenterol*, **7**, 630-6.

Lin R, Maeda S, Liu C, et al (2007). A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. *Oncogene*, **26**, 851-8.

Llovet JM, Lencioni R, Gaile PR, et al (2012). EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol*, **56**, 908-43.

Lu X, Kang Y (2007). Organotropism of breast cancer metastasis. *J Mammary Gland Biol Neoplasia*, **12**, 153-62.

Matouk IJ, Abbasi I, Hochberg A, et al (2009). Highly upregulated in liver cancer noncoding RNA is overexpressed in hepatic colorectal metastasis. *Eur J Gastroenterol Hepatol*, **21**, 688-92.

Mercer TR, Dinger ME, Mattick JS (2009). Long non-coding RNAs: insights into functions. *Nat Rev Genet*, **10**, 155-9.

Mootha VK, Lindgren CM, Eriksson KF, et al (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*, **34**, 267-73.

Mourtada-Maarabouni M, Pickard MR, Hedge VL, et al (2009). GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. *Oncogene*, **28**, 195-208.

Nguyen DX, Bos PD, Massague J (2009). Metastasis: from dissemination to organ-specific colonization. *Nat Rev Cancer*, **9**, 274-84.

Nupponen NN, Carpten JD (2001). Prostate cancer susceptibility genes: many studies, many results, no answers. *Cancer Metastasis Rev*, **20**, 155-64.

Ohta M, Seto M, Miyabayashi K, et al (2009). Decreased expression of the RAS-GTPase activating protein RASAL1 is associated with colorectal tumor progression. *Gastroenterology*, **136**, 206-16.

Okazaki Y, Furuno M, Kasukawa T, et al (2002). Analysis of the mouse transcriptome based on functional annotation of 60, 770 full-length cDNAs. *Nature*, **420**, 563-73.

Orom UA, Derrien T, Beringer M, et al (2010). Long Noncoding RNAs with Enhancer-like Function in Human Cells. **Cell**, **143**, 46-58.

Panzitt K, Tschernatsch MM, Guelly C, et al (2007). Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. *Gastroenterology*, **132**, 330-42.

Pibouin L, Villaudy J, Ferbus D, et al (2002). Cloning of the mRNA of overexpression in colon carcinoma-1: a sequence overexpressed in a subset of colon carcinomas. *Cancer Genet Cytogenet*, **133**, 55-60.

Ponting CP, Oliver PL, Reik W (2009). Evolution and Functions of Long Noncoding RNAs. *Cell*, **136**, 629-41.

Poon R, Wong J (2000). Risk factors, prevention, and management of postoperative recurrence after resection of hepatocellular carcinoma. *Ann Surg*, **232**.

Prieto C, Risueno A, Fontanillo C, et al (2008). Human gene coexpression landscape: confident network derived from tissue transcriptomic profiles. *PLoS Onee*, **3**.

Qiao HP, Gao WS, Huo JX, et al (2013). Long non-coding RNA GAS5 functions as a tumor suppressor in renal cell carcinoma. *Asian Pac J Cancer Prev*, **14**, 1077-82.

Qiu MT, Hu JW, Yin R, et al (2013). Long noncoding RNA: an emerging paradigm of cancer research. *Tumour Biol*, **34**, 613-20.

Rangel L, Wernyj RP, Cho KR (2003). Characterization of novel human ovarian cancer-specific transcripts (HOSTs). identified by serial analysis of gene expression. *Oncogene*, **22**, 7225-32.

Ren S, Wang F, Shen J, et al (2013). Long non-coding RNA metastasis associated in lung adenocarcinoma transcript 1 derived miniRNA as a novel plasma-based biomarker for diagnosing prostate cancer. *Eur J Cancer*, **49**, 2949-59.

Rinn JL, Wang JK, Xu X, et al (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by Noncoding RNAs. *Cell*, **129**, 1311-23.

Seto M, Ohta M, Ikenoue T, et al (2011). Reduced expression of RAS protein activator like-1 in gastric cancer. *Int J Cancer*, **128**, 1293-302.

Simon JA, Kingston RE (2009). Mechanisms of Polycomb gene silencing: knowns and unknowns. *Nat Rev Mol Cell Biol*, **10**, 697-708.

Smedley D, Sidhar S, Birdsall S, et al (2000). Characterization of chromosome 1 abnormalities in malignant melanomas. *Genes Chromosomes Cancer*, **28**, 121-5.

Sporn JC, Kustatscher G, Hothorn T, et al (2009). Histone macroH2A isoforms predict the risk of lung cancer recurrence. *Oncogene*, **28**, 3423-8.

Subramanian A, Tamayo P, Mootha VK, et al (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*, **102**, 15545-50.

Sun FX, Tang ZY, Lui KD, et al (1996). Establishment of a metastatic model of human hepatocellular carcinoma in nude mice via orthotopic implantation of histologically intact tissues. *Int J Cancer*, **66**, 239-43.

Taft RJ, Pang KC, Mercer TR, et al (2010). Non-coding RNAs: regulators of disease. *J Pathol*, **220**, 126-39.

Talmadge JE, Fidler IJ (2010). AACR Centrspective. *Cancer Res*, **70**, 5649-69.

Tinzl M, Horvath S (2004). DD3 (PCA3). RNA analysis in urine - A new perspective for detecting prostate cancer. *Eur Urol*, **46**, 182-7.

Uchino H, Kataoka H, Itoh H, et al (2000). Overexpression of intestinal trefoil factor in human colon carcinoma cells reduces cellular growth *in vitro* and *in vivo*. *Gastroenterol*, **118**, 60-9.

Wang JY, Liu XF, Wu HC, et al (2010). CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. *Nucleic Acids Res*, **38**, 5366-83.

Wang Y, Liu XJ, Yao XD (2014). Function of PCA3 in prostate tissue and clinical research progress on developing a PCA3 score. *Chin J Cancer Res*, **26**, 493-500.

Wapinski O, Chang HY (2011). Long noncoding RNAs and human disease. *Trends Cell Biol*, **21**, 354-61.

Wilusz JE, Sunwoo H, Spector DL (2009). Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev*, **23**, 1494-504.

Xie H, Ma H, Zhou D (2013). Plasma HULC as a promising novel biomarker for the detection of hepatocellular carcinoma. *Biomed Res Int*, **2013**, 136106.

Yamada K, Kano J, Tsunoda H, et al (2006). Phenotypic characterization of endometrial stromal sarcoma of the uterus. *Cancer Sci*, **97**, 106-12.

Yang F, Zhang L, Huo XS, et al (2011). Long noncoding RNA high expression in hepatocellular carcinoma facilitates tumor growth through enhancer of zeste homolog 2 in humans. *Hepatology*, **54**, 1679-89.

Ye QH, Forgues M, Kim JW, et al (2003). Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. *Nat Med*, **9**, 416-23.

Yue S, Mu W, Zoller M (2013). Tspan8 and CD151 promote metastasis by distinct mechanisms. *Eur J Cancer*, **49**, 2934-48.

Zhang EB, Han L, Yin DD, et al (2014). c-Myc-induced, long, noncoding H19 affects cell proliferation and predicts a poor prognosis in patients with gastric cancer. *Med Oncol*, **31**, 914.

Zhang H, Cai K, Wang J, et al (2014). MiR-7, Inhibited indirectly by lincRNA HOTAIR, directly inhibits SETDB1 and reverses the EMT of breast cancer stem cells by downregulating the STAT3 pathway. *Stem Cells*, **32**, 2858-68.