Hindawi Publishing Corporation
International Journal of Genomics
Volume 2013, Article ID 480534, 10 pages
http://dx.doi.org/10.1155/2013/480534

Research Article

Integrated Analysis of Long Noncoding RNA and Coding RNA Expression in Esophageal Squamous Cell Carcinoma

Wei Cao,[1] Wei Wu,[1,2] Fachun Shi,[3] Xiaobing Chen,[1] Lihua Wu,[1] Ke Yang,[1] Fu Tian,[1] Minghui Zhu,[1] Guoyong Chen,[1] WeiWei Wang,[1] Fred G. Biddle,[4] and Jianqin Gu[3]

[1] *Clinical Research Center, People's Hospital of Zhengzhou, 33 Yellow River Road, Zhengzhou, Henan 45003, China*
[2] *Department of Pathology and Experimental Medicine, University of Calgary, Calgary, AB, Canada T2N 4N1*
[3] *Science and Education Department, Health Bureau of Zhengzhou, China*
[4] *Departments of Medical Genetics and Biological Sciences, University of Calgary, Calgary, AB, Canada T2N 4N1*

Correspondence should be addressed to Wei Cao; caoweiyu@hotmail.com

Received 28 May 2013; Accepted 26 August 2013

Academic Editor: Soraya E. Gutierrez

Tumorigenesis is a complex dynamic biological process that includes multiple steps of genetic and epigenetic alterations, aberrant expression of noncoding RNA, and changes in the expression profiles of coding genes. We call the collection of those perturbations in genome space the "cancer initiatome." Long noncoding RNAs (lncRNAs) are pervasively transcribed in the genome and they have key regulatory functions in chromatin remodeling and gene expression. Spatiotemporal variation in the expression of lncRNAs has been observed in development and disease states, including cancer. A few dysregulated lncRNAs have been studied in cancers, but the role of lncRNAs in the cancer initiatome remains largely unknown, especially in esophageal squamous cell carcinoma (ESCC). We conducted a genome-wide screen of the expression of lncRNAs and coding RNAs from ESCC and matched adjacent nonneoplastic normal tissues. We identified differentially expressed lncRNAs and coding RNAs in ESCC relative to their matched normal tissue counterparts and validated the result using polymerase chain reaction analysis. Furthermore, we identified differentially expressed lncRNAs that are co-located and co-expressed with differentially expressed coding RNAs in ESCC and the results point to a potential interaction between lncRNAs and neighboring coding genes that affect ether lipid metabolism, and the interaction may contribute to the development of ESCC. These data provide compelling evidence for a potential novel genomic biomarker of esophageal squamous cell cancer.

1. Introduction

Esophageal squamous cell carcinoma (ESCC) is one of the most common types of cancer, and it ranks among the main causes of cancer deaths worldwide [1]. There are marked regional variation and exceptionally high incidence in certain areas of China. Despite advances in multidisciplinary treatment of ESCC, 5-year survival rate remains poor. The initiatome [2] of ESCC is a complex dynamic biological process in genome space, and it may include multiple steps of genetic and epigenetic alterations [3], aberrations in expression of noncoding RNA (e.g., microRNAs) [4], and changes in the expression profile of coding genes [5, 6]. In past decades, expression profiling of coding genes has defined important signaling pathways involved in tumorigenesis. The latest knowledge of actively transcribed long noncoding RNAs (lncRNAs) from high-throughput sequencing is revealing an even greater complexity about cancer genome regulatory networks.

LncRNAs are endogenous cellular RNA transcripts, ranging from 200 to 100,000 nucleotides in length, and they lack an open reading frame of significant length (less than 100 amino acids) [7]. LncRNAs are generally expressed at a lower level than protein-coding genes, and they display more tissue-specific and cell-specific expression patterns [8, 9]. LncRNAs were previously believed to be transcriptional noise, but now they have critical roles in development and differentiation as well as in the proliferation and progress

of disease, including cancer [10]. Mechanisms of action of transcribed lncRNAs are described as modifying chromatin architecture and regulating gene expression in a cis or trans manner. For example, H19 lncRNA cis-regulates *IGF2* gene expression at the same genomic locus, and HOTAIR lncRNA is transcribed on Chr 12, and it transregulates *HoxD* gene on Chr 2. Additionally, lncRNAs have also been reported to coordinate the regulation of neighboring coding genes through a "locus control" process [11], which mediates the localization of genes within nuclear regions to favor their transcription through the formation of domains of histone modification and intra- or interchromosomal loops [12]. Dysregulated lncRNAs have been identified with different screening methodologies in various types of cancer. For example, the cancer-related lncRNA, metastasis-associated lung adenocarcinoma transcript 1 (*MALAT-1*), was identified by subtractive hybridization during screening for early non-small cell lung cancer with metastasis [13]. Overexpression of *MALAT-1* is highly predictive of poor prognosis and shortened survival time in early stage lung cancer. Overexpression of *HOTAIR* lncRNA was found in several solid tumors [14–17] in association with cancer metastasis, and increased *HOTAIR* expression in breast cancer is transcriptionally induced by estradiol [18]. Prostate cancer associated lncRNA, PCGEM1 [19], and *PCAT-1* [20] appear to be prostate-specific regulators of cell apoptosis and proliferation. Recently, *AFAP1-AS1* lncRNA was reported to be overexpressed in esophageal adenocarcinoma [21].

The handful of dysregulated lncRNAs in different cancers suggests that lncRNAs are an enigmatic component of the whole transcriptome, which may participate in tumorigenesis, invasion, and metastasis. Efforts are being made to explore the "lncRNAome" of various cancers with advanced high-throughput RNA sequencing technologies [8, 9], and dynamic changes in lncRNA expression have been observed in cancer cells during different stages of cancer development and during treatment [22]. However, our understanding of the role of lncRNAs in cancer biology is still in an early stage, and a clearly defined, predictive set of biological functions for lncRNAs is lacking in cancer biology. Therefore, thorough searches and analyses of the interactions between lncRNA and coding genes may help to infer their potential biological roles.

In order to understand the role of lncRNAs in ESCC, we report a pilot study of the profiles of differentially expressed lncRNAs and coding RNAs from tumor and adjacent normal tissue of individual patients with ESCC. We assessed the whole transcriptomic landscape for potential interactions between lncRNAs and coding-gene expression. In particular, we evaluated the coding genes that are co-located and co-expressed with the differentially expressed lncRNAs during the genesis of ESCC.

2. Results and Discussion

2.1. Transcriptomic Landscape of ESCC. Our genome-wide gene expression profiling of both lncRNAs and coding genes from ESCC and adjacent nonneoplastic tissue was conducted to detect possible associations of lncRNAs with ESCC. We first asked whether these transcripts of 7,419 noncoding and 27,958 coding RNAs could distinguish ESCC from normal tissues. Figure 1(a) shows that the four ESCC samples are clustered together in one group and clearly separated from the samples of normal tissue. Next, we examined the whole transcriptomic pattern (lncRNAs + coding RNAs) from each sample and the landscapes of the whole transcriptome (represented by heatmaps in Figure 1(a)) of normal tissues differ from those of ESCC that exhibit more heterogeneous alterations. The overall changes from a respective normal to cancer state were also seen separately as a difference in expression profile of either the lncRNA or the coding RNA. These observations suggest that a potential dynamic interaction between lncRNAs and coding RNAs may be reshaping the landscape of the whole transcriptome during ESCC development.

To gain a detailed understanding of the biological themes of all RNA transcripts, we further identified those transcripts that are significantly and differentially expressed (DE) in ESCC tissue compared to matched normal tissue, based on the criteria described in the methods. There are 410 DE-lncRNAs and 1219 DE-mRNAs that represent about 5% of the transcripts in the respective microarrays (Supplementary Tables S1 and S2 available online at http://dx.doi.org/10.1155/2013/480534). DE-lncRAs distinguish a cancer cell from its normal cell state with three times fewer transcripts than DE-mRNAs (Figures 1(b) and 1(c)), suggesting that the DE-lncRNA profile is more informative and, potentially, a more faithful indicator of a specific cell state.

Enrichment analysis of DE-mRNAs demonstrated that the respective genes are involved in cancer-related pathways (Figure 1(d)). Since expression profiling of coding-RNA has been intensively studied in esophageal cancer, we validated 10 genes whose expression level in other studies [23–25] is significantly changed ($P < 0.05$) by at least 2-fold relative to normal tissues (Table 1).

2.2. Expression of lncRNAs in ESCC. LncRNAs are emerging as a novel class of noncoding RNAs that are pervasively transcribed in the genome, but there is limited functional knowledge about them. High-throughput screening of lncRNAs from ESCC has been poorly studied, except for a recent report of overexpressed lncRNA, AFAP1-AS1, in esophageal adenocarcinoma [21]. In our study, a total of 7,419 intergenic lncRNAs and other transcripts of uncertain coding potential were examined, and we identified 410 DE-lncRNAs in ESCC relative to adjacent normal esophageal tissues. We named the anonymous lncRNAs ESCC Associated Long noncoding RNAs (ESCCAL, Supplementary Table S1). Expression of HOTAIR lncRNA is increased in various cancers [14–17, 26], and it is also significantly increased in our analysis of ESCC (Figure 2(a)). In addition, we confirmed another two upregulated lncRNAs that are differentially expressed in ESCC and that we have named ESCCAL-1, and ESCCAL-5. The increased and differential expression of HOTAIR, ESCCAL-1 and ESCCAL-5 in ESCC tissue relative to adjacent nonneoplastic tissue was independently assessed with

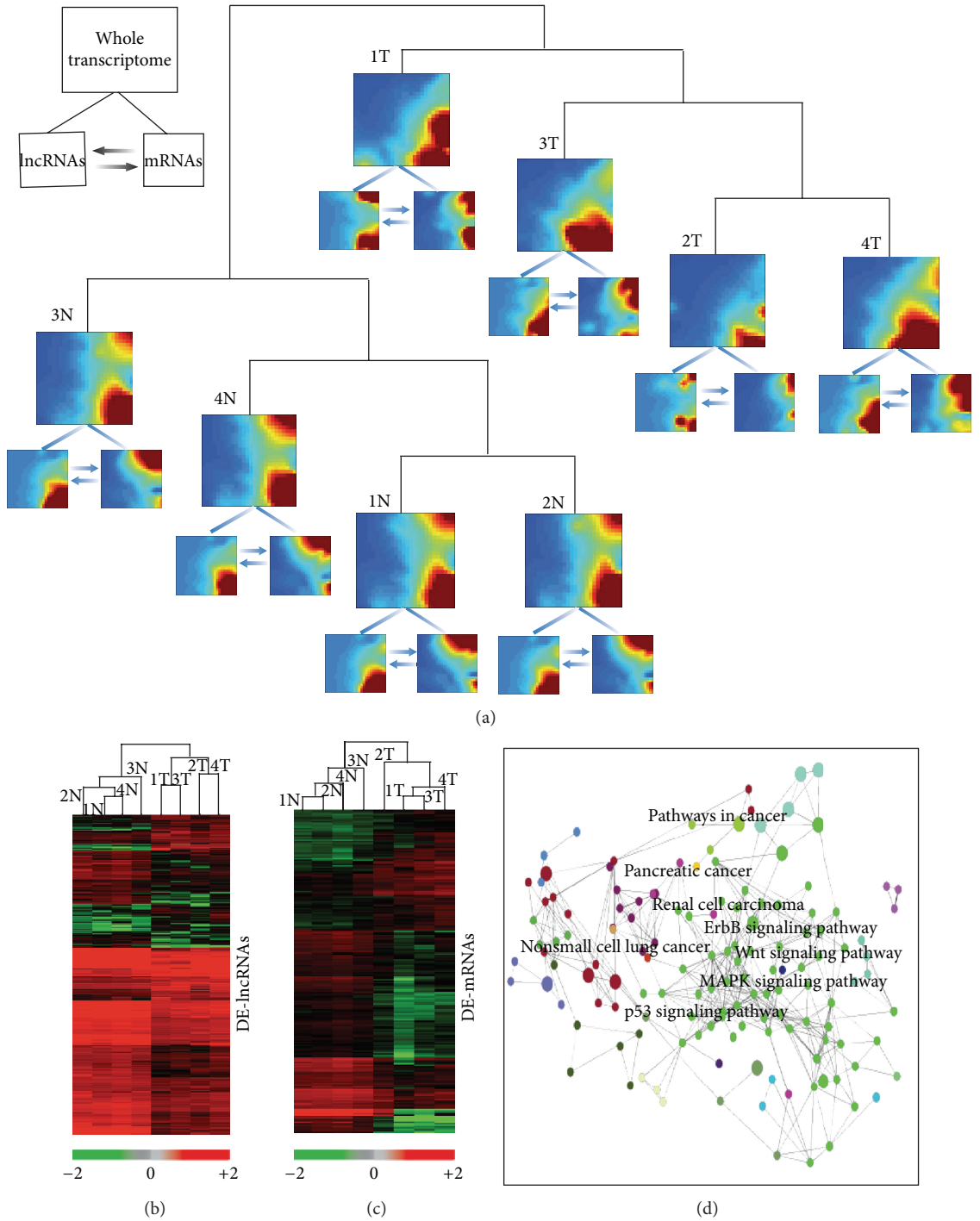

Figure 1: Transcriptomic landscape of esophageal squamous cell cancer (ESCC). (a) Whole transcriptome of tumor (T) and adjacent normal tissue (N) of four patients with ESCC were detected using a microarray with 7,419 long noncoding RNAs (lncRNAs) and 27,958 coding RNAs. Two main clusters (Ts and Ns) were generated using unsupervised clustering methods. Then, a self-organizing map (SOM) of either whole transcriptome (both lncRNAs and mRNAs) or lncRNAs or mRNA was produced from each sample (see legend in up-left corner of this figure, and the arrows are meant to indicate the potential interaction), using gene expression dynamic inspector (GEDI). Mosaic patterns are pseudocolored SOMs to show integrated biological entity in each sample. Red through blue color indicates high to low expression level. (b) and (c) Differentially expressed lncRNAs (DE-lncRNAs) and coding RNAs (DE-mRNAs) in ESCC. Hierarchical clustering analysis of 410 DE-lncRNAs (b) and 1219 DE-mRNAs (c) between ESCC tissue and adjacent normal tissue (fold change > or < 2-fold and $P < 0.05$). Red and green colors indicate high and low expression, respectively. In the heatmap, columns represent samples, and rows represent each gene. The scale of expression level is shown on the horizontal bar. (d) KEGG functional analysis of DE-mRNA networks in ESCC. The DE-mRNA genes are involved in cancer-related signaling functions, and a detailed list of significant GO terms is shown in Figure S1 and its associated legend in Supplementary Information.

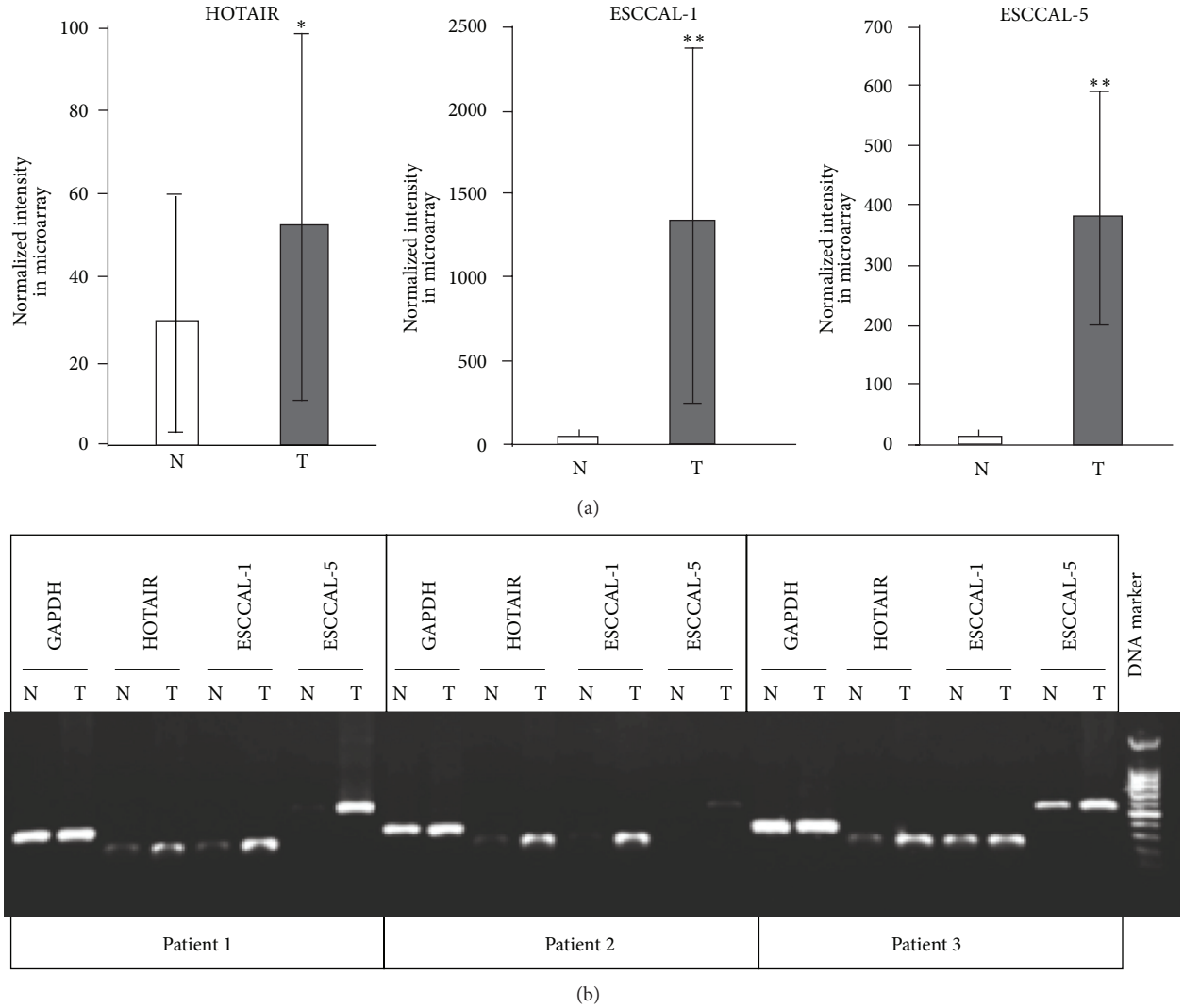

Figure 2: Long noncoding RNAs (lncRNAs) expression in esophageal squamous cell carcinoma (ESCC). (a) Three differentially expressed lncRNAs, HOTAIR, ESCCAL-1, and ESCCAL-5, from microarray detection. The average intensity of expression in normal tissues (N) and tumors (T) is plotted with their standard deviations. (b) Validation of HOTAIR, ESCCAL-1, and ESCCAL-5 with independent patient samples by PCR analysis. The amplicons were separated with 2% agarose gel. GAPDH was used as an internal control. Significance is $^{*}P < 0.05$, $^{**}P < 0.01$.

Table 1: Validation of selected differential expression of mRNAs in esophageal squamous cell carcinoma in independent studies.

Probe name	P value	FC	Regulation	Gene symbol	Genbank accession	Independent study	Reference
A_33_P3232692	0.005838984	8.301461	Up	IL24	NM_001185156		
A_24_P411121	0.00055	5.329484	Up	TNFRSF18	NM_148901		
A_23_P169097	$8.81E-05$	4.466178	Up	WISP1	NM_080838		
A_23_P304304	0.004822649	3.944957	Down	ARSF	NM_004042	Microarray	[20]
A_24_P56363	0.003573538	3.323955	Down	CAB39L	NM_030925		
A_23_P419760	0.001041661	32.70335	Down	CRISP3	NM_006061		
A_23_P413923	0.002921898	4.54899	Down	DMRTA1	NM_022160		
A_23_P56978	0.002093997	5.438183	Down	PTK6	NM_005975	RNA-seq	[21]
A_23_P115091	0.005171322	3.289834	Down	RAB25	NM_020387	Q-RT-PCR	[22]
A_33_P3258542	0.001039129	20.36035	Down	SPINK8	NM_001080525	Microarray	[20]

PCR methods in matched-pair tissue samples from three additional ESCC patients and the results are consistent with the microarray analysis (Figure 2(b)). Interestingly, except for HOTAIR, other previously reported lncRNAs (i.e., MALAT-1, PCAT-1 and AFAP1-AS1) are not differentially expressed in our analysis of ESCC. Therefore, the DE-lncRNAs that we have identified may be a unique property of ESCC, and we are currently using a population-based analysis to characterize these DE-lncRNAs as potential genomic biomarkers and regulatory elements in the dynamic process leading to ESCC.

2.3. LncRNAs Co-located and Co-expressed with Coding Genes in ESCC. LncRNAs have been reported to coordinate the regulation of neighboring coding genes through a "locus control" process [11]. We wondered whether such a "locus control" process could operate in ESCC development, and, therefore, we searched neighboring genes of the 410 DE-lncRNAs in the genome. The majority (98.8%) of the 410 DE-lncRNAs harbor neighboring coding genes whose genomic locations are within ~5 kb upstream and ~1 kb downstream of the lncRNA and may extend to 1000 kb in both directions (Figure 3(a)). Interrogation of 538 coding genes that are neighbors of these DE-lncRNAs (DE-lncRNAs co-located genes) revealed predicted functions in 9 common pathways such as the AP1 transcription factor network, integrin-linked kinase signaling, several signaling pathways in adherens junctions, and FOXO family signaling (Figure 3(b)).

We asked whether any DE-lncRNAs co-located genes are also differentially expressed in ESCC. Analysis of the DE-lncRNAs co-located genes with DE-mRNA data set identified 76 genomically co-located and differentially co-expressed genes (Figure 3(c) and Table 2). Strikingly, the co-located and co-expressed genes with DE-lncRNAs may be involved in ether lipid metabolism pathways by the participation of the *LPCAT1* gene encoding lysophosphatidylcholine acyltransferase1 and the *PLD1* gene encoding phospholipase D1 (Figure 3(c)). The lncRNA ESCCAL-337 (chr3:171506370-171528740) was downregulated in ESCC and located at 22,068 bp downstream of the *PLD1* gene, whose expression was also decreased in ESCC. In contrast, the lncRNA *ESCCAL-356* (chr5:1544500-1567142 reverse strand) was downregulated in ESCC and located at 21,250 bp upstream of *LPCAT1*, whose expression was upregulated in ESCC (Figure 3(c)). LPCAT1 modulates phospholipid composition by catalyzing lysophosphatidylcholine into phosphatidylcholine, and overexpression of LPCAT1 was reported to create favorable conditions for cancer cell proliferation [27, 28]. Therefore, at least two of the DE-lncRNAs have the potential to contribute to ESCC by a "locus control" process with neighboring coding genes.

In conclusion, we performed a genome-wide survey of the expression of lncRNAs and coding mRNAs from paired samples of primary neoplastic tissue and adjacent non-neoplastic normal tissue from four individuals. The overall transcriptomic landscape (both lncRNAs and mRNAs) is able to distinguish malignant from normal tissue in each person. We discovered a set of differentially expressed lncRNAs and their co-located and co-expressed coding mRNAs and demonstrated that lncRNAs may be involved in ether lipid metabolism in ESCC. Our study provides genomic support for a model of a "locus control" process in ESCC and a framework for further experimental study.

3. Materials and Methods

3.1. Specimens. Written informed consent was obtained from patients before surgery, and the study protocol was approved by the Institutional Review Board for the use of human subjects at Zhengzhou Hospital. Primary tumors and adjacent nonneoplastic tissues were obtained from patients with ESCC who underwent surgical treatment at Linxian Hospital in May 2012. All tissues were frozen in liquid nitrogen immediately after surgical resection. None of the patients had prior chemotherapy or radiotherapy, nor did they have any other serious diseases. All ESCC tissues were histopathologically diagnosed by at least two independent senior pathologists.

3.2. Microarray Hybridization. Total RNAs were extracted using Trizol reagent, following manufacturer's instructions (Invitrogen, Carlsbad, CA, USA). The quality of RNAs was measured with a 2100 Bioanalyzer (Agilent technology, USA). Input of 100 ng of total RNA was used to generate Cyanine-3 labeled cRNA, according to the Agilent One-Color Microarray-Based Gene Expression Analysis Low for Input Quick Amp Labeling kit (v6.0). Samples were hybridized on Agilent SurePrint G3 Human GE 8 × 60 K Microarray (Design ID 028004). Arrays were scanned with the Agilent DNA Microarray Scanner at a 3 μm scan resolution, and data were processed with Agilent Feature Extraction 11.0.1.1. 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 GSE45350 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE45350).

3.3. Validation by Polymerase Chain Reaction (PCR). PCR analysis was performed on additional matched ESCC and adjacent nonneoplastic tissues for selected lncRNAs. The primer sequences for PCR are as follows: HOTAIR, forward 5′-GGTAGAAAAAGCAACCACGAAGC-3′ and reverse 5′-ACATAAACCTCTGTCTGTGAGTGCC-3′; ESSCAL-1 (chr8:76121095-76189420 reverse strand), forward 5′-CCAGACAGCAGCAAAGCAAT-3′ and reverse 5′-GGAAGCAGCAAATGTGTCCAT-3′; ESSCAL-5 (chr2:216585154-216585719 forward strand), forward 5′-TACCAACATTGTCCACCGGG-3′ and reverse 5′-GCTGATGACAGTCCCTTGCT-3′. GAPDH was used as a control, forward 5′-CCGGGAAACTGTGGCGTGATGG-3′ and reverse 5′-AGGTGGAGGAGTGGGTGTCGCTGTT-3′. The thermocycle conditions are as follows: initial denaturation at 95°C for 10 minutes, followed by 94°C for 45 seconds, 65°C for 30 seconds, and 72°C for 1 minute for 15 cycles. Then, the annealing temperature was reduced by 0.5°C/cycle for the next 14 cycles, and the amplification was finished with another 24 cycles with the annealing temperature at 58°C.

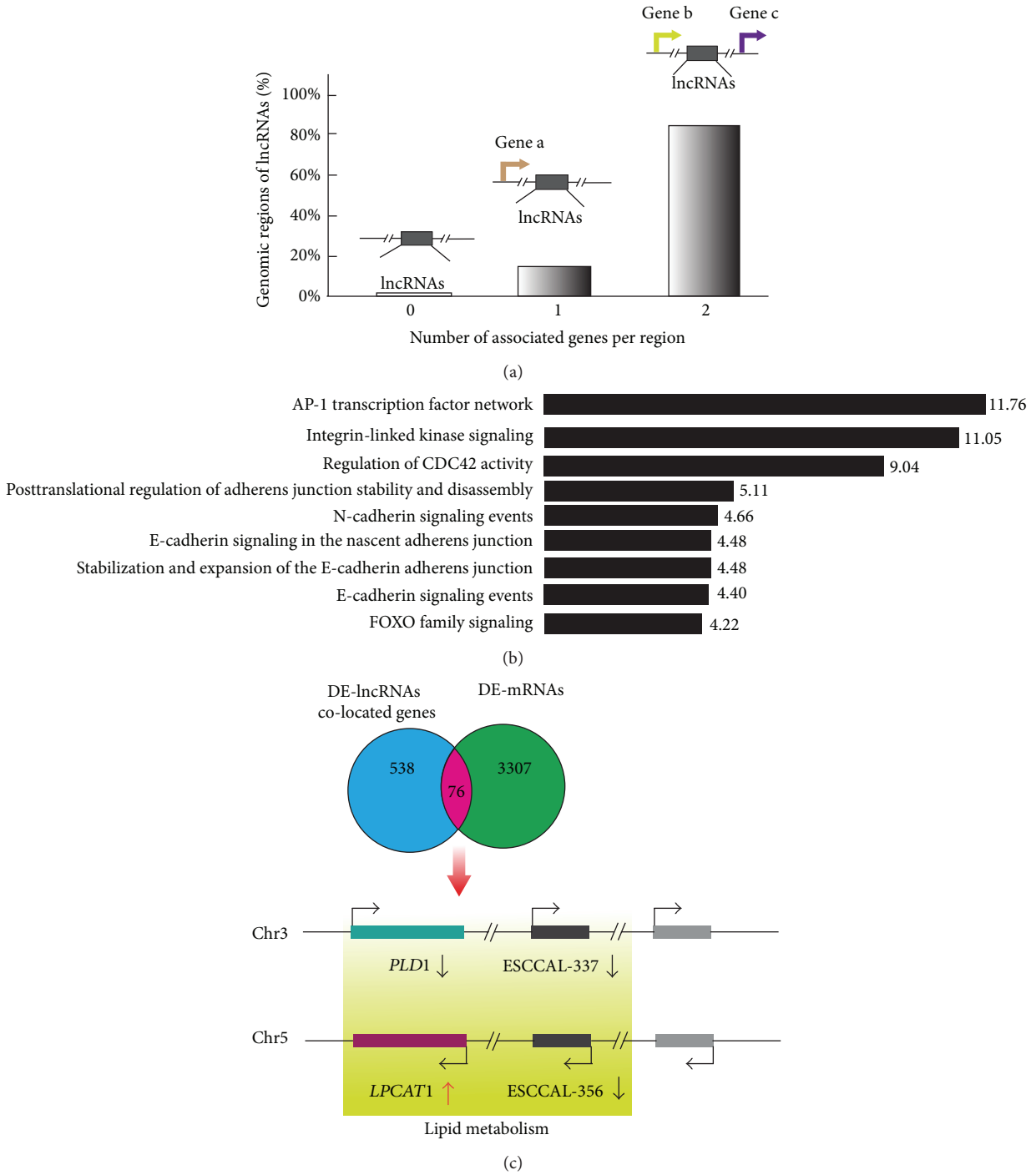

Figure 3: Identification of lncRNAs co-located and co-expressed neighboring genes in esophageal squamous cell carcinoma (ESCC). (a) Identification of neighboring genes of the DE-lncRNAs. The genomic coordinate information of 410 DE-lncRNAs was used to search neighboring genes whose genomic locations are within ~5 kb upstream and ~1 kb downstream of the lncRNA and may extend to 1000 kb in both directions using GREAT software (http://bejerano.stanford.edu/great/public/html/index.php). The percentage of DE-lncRNAs harboring zero, one, or two neighboring genes is presented. (b) Gene Ontology (GO) enrichment analysis of lncRNAs co-located genes. Identified gene enriched pathways/terms are listed on the left; the length of horizontal bars and the numbers on the right indicate the percentage of genes involved in each pathway/term. (c) LncRNAs co-located and co-expressed coding mRNAs. Overlap of 538 DE-lncRNA co-located genes with 3307 DE-mRNAs in microarrays identified 76 lncRNAs co-located and co-expressed coding mRNAs (list in Table 2). GO enrichment analysis suggests phospholipase D1 (*PLD1*) and lysophosphatidylcholine acyltransferase1 (*LPCAT1*) are involved in ether lipid metabolism pathway. Genomic location shows that *PLD1* is located at −22,068 bp upstream of ESCCAL-337 lncRNA on Chr 3 and *LPCAT1* is at −21,250 bp upstream of ESCCAL-356 lncRNA on Chr 5.

Table 2: List of identified co-located and co-expressed genes with differentially expressed lncRNAs in ESCC.

LncRNA name	Genomic coordinate	LncRNA expression	LncRNA associated genes	Gene expression	Gene symbol	Genbank accession
ESCCAL-177	chr1:205404138-205404079	Up	CDK18 (−69575), LEMD1 (−12895)	Up	LEMD1	NM_001001552
ESCCAL-348	chr1:222587441-222587382	Up	DUSP10 (−671896), HHIPL2 (+134032)	Up	DUSP10	NM_007207
ESCCAL-31	chr1:225237693-225237752	Up	DNAH14 (+120367), LBR (+378092)	Up	DNAH14	NM_001373
ESCCAL-159	chr1:225240300-225240359	Up	DNAH14 (+122974), LBR (+375485)	Up	DNAH14	NM_001373
ESCCAL-327	chr1:89887111-89887052	Up	LRRC8B (−103315), GBP6 (+57646)	Down	GBP6	NM_198460
XLOC_000915	chr1:91295528-91295469	Up	BARHL2 (−112705), ZNF644 (+192313)	Down	BARHL2	NM_020063
ESCCAL-65	chr11:2017146-2017205	Up	MRPL23 (+48674), IGF2 (+145165)	Up	IGF2	NM_000612
ESCCAL-19	chr12:66204406-66204347	Up	HMGA2 (−13863), MSRB3 (+531610)	Up	HMGA2	NM_003484
XLOC_011548	chr15:81953024-81952965	Up	TMC3 (−286577), MEX3B (+385366)	Up	MEX3B	NM_032246
ESCCAL-102	chr17:6766871-6766930	Up	ALOX12 (−132483), TEKT1 (−31841)	Down	ALOX12	NM_000697
ESCCAL-342	chr17:78302158-78302217	Up	ENDOV (−86779), RNF213 (+67521)	Up	RNF213	NM_020954
ESCCAL-342	chr17:78302158-78302217	Up	ENDOV (−86779), RNF213 (+67521)	Down	ENDOVNM	NM_001164638
ESCCAL-99	chr2:102034125-102034066	Up	CREG2 (−30131), RFX8 (+57069)	Up	CREG2	NM_153836
ESCCAL-5	chr2:216585455-216585514	Up	FN1 (−284694), MREG (+292861)	Down	MREG	NM_018000
ESCCAL-5	chr2:216585455-216585514	Up	FN1 (−284694), MREG (+292861)	Up	FN1	NM_054034
ESCCAL-288	chr2:27789661-27789602	Up	ZNF512 (−16261), GCKR (+69926)	Down	GCKR	NM_001486
ESCCAL-344	chr2:37327299-37327358	Up	CCDC75 (+15735), EIF2AK2 (+56861)	Up	EIF2AK2	NM_001135652
ESCCAL-10	chr22:48086469-48086528	Up	FAM19A5 (−798789), TBC1D22A (+927981)	Up	FAM19A5	NM_015381
ESCCAL-81	chr5:141710377-141710436	Up	SPRY4 (−5787), FGF1 (+355246)	Up	FGF1	NM_000800
ESCCAL-81	chr5:141710377-141710436	Up	SPRY4 (−5787), FGF1 (+355246)	Up	SPRY4	NM_030964
ESCCAL-204	chr6:126699769-126699828	Up	RSPO3 (−740249), CENPW (+38546)	Up	CENPW	NM_001012507
ESCCAL-106	chr6:21822910-21822969	Up	SOX4 (+228968), PRL (+480142)	Up	SOX4	NM_003107
ESCCAL-239	chr8:104258684-104258625	Up	FZD6 (−52006), BAALC (+105734)	Up	BAALC	NM_024812
ESCCAL-239	chr8:104258684-104258625	Up	FZD6 (−52006), BAALC (+105734)	Up	FZD6	NM_003506
ESCCAL-300	chr1:180918852-180918793	Down	STX6 (+73223), XPR1 (+317677)	Up	XPR1	NM_004736
XLOC_000515	chr1:200384539-200384598	Down	ZNF281 (−5403), KIF14 (+205293)	Up	KIF14	NM_014875
XLOC_000515	chr1:200384539-200384598	Down	ZNF281 (−5403), KIF14 (+205293)	Up	ZNF281	NM_012482
ESCCAL-77	chr1:201592869-201592810	Down	CSRP1 (−116873), NAV1 (−24610)	Up	NAV1	NM_020443
ESCCAL-76	chr1:90091283-90091342	Down	LRRC8C (−7331), LRRC8B (+100916)	Up	LRRC8C	NM_032270
ESCCAL-209	chr10:14549236-14549177	Down	FRMD4A (−176341), CDNF (+330776)	Down	CDNF	NM_001029954
ESCCAL-284	chr10:44848467-44848526	Down	HNRNPA3P1 (−562632), CXCL12 (+32048)	Down	CXCL12	NM_199168
ESCCAL-256	chr11:117671760-117671701	Down	DSCAML1 (−3755)	Down	DSCAML1	NM_020693
ESCCAL-254	chr11:126219961-126220020	Down	ST3GAL4 (−5549), DCPS (+46344)	Down	ST3GAL4	NM_006278
ESCCAL-307	chr11:14975924-14975983	Down	CYP2R1 (−62203), CALCA (+17878)	Down	CYP2R1	NM_024514
ESCCAL-105	chr11:17366661-17366720	Down	B7H6 (−6618), NUCB2 (+68405)	Down	NUCB2	NM_005013
ESCCAL-232	chr12:131245982-131245923	Down	RIMBP2 (−243491), STX2 (+77858)	Down	RIMBP2	NM_015347
ESCCAL-24	chr12:132824534-132824475	Down	GALNT9 (+81400), NOC4L (+195512)	Up	GALNT9	NM_021808

Table 2: Continued.

LncRNA name	Genomic coordinates	LncRNA expression	LncRNA associated genes	Gene expression	Gene symbol	Genbank accession
ESCCAL-122	chr12:2952238-2952297	Down	NRIP2 (−8047), FOXM1 (+34053)	Up	FOXM1	NM_202002
ESCCAL-338	chr12:32101291-32101232	Down	BICD1 (−158923), H3F3C (−156087)	Up	BICD1	NM_001714
ESCCAL-275	chr14:71180731-71180790	Down	TTC9 (+72257), MAP3K9 (+95127)	Down	MAP3K9	NM_033141
ESCCAL-275	chr14:71180731-71180790	Down	TTC9 (+72257), MAP3K9 (+95127)	Down	TTC9	NM_015351
ESCCAL-121	chr15:78088340-78088281	Down	LINGO1 (−163602), TBC1D2B (+281683)	Up	LINGO1	NM_032808
ESCCAL-231	chr16:88381397-88381338	Down	ZNF469 (−112511), BANP (+396330)	Up	ZNF469	NM_001127464
ESCCAL-74	chr2:47548478-47548537	Down	CALM2 (−144768), EPCAM (−47779)	Up	EPCAM	NM_002354
ESCCAL-152	chr20:56839403-56839344	Down	PMEPA1 (−554343), PPP4R1L (+45121)	Up	PMEPA1	NM_020182
ESCCAL-337	chr3:171506465-171506406	Down	TNIK (−328239), PLD1 (+22068)	Down	PLD1	NM_002662
ESCCAL-80	chr3:177935638-177935579	Down	KCNMB2 (−318615)	Down	KCNMB2	NM_181361
ESCCAL-90	chr3:195441846-195441787	Down	MUC20 (−5936), SDHAP2 (+56907)	Down	MUC20	NM_001098516
ESCCAL-59	chr3:64855048-64855107	Down	ADAMTS9 (−181713)	Up	ADAMTS9	NM_182920
ESCCAL-72	chr4:74922502-74922443	Down	CXCL3 (−17983), CXCL2 (+42524)	Up	CXCL3	NM_002090
ESCCAL-79	chr4:79626460-79626401	Down	BMP2K (−71101), ANXA3 (+153689)	Down	ANXA3	NM_005139
ESCCAL-253	chr4:8357097-8357038	Down	ACOX3 (+85384), HTRA3 (+85579)	Down	ACOX3	NM_003501
ESCCAL-179	chr4:8359416-8359357	Down	ACOX3 (+83065), HTRA3 (+87898)	Up	HTRA3	NM_053044
ESCCAL-11	chr4:84299039-84299098	Down	HPSE (−43035), HELQ (+77956)	Down	HPSE	NM_006665
ESCCAL-41	chr4:8512901-8512960	Down	GPR78 (−69286), METTL19 (+70399)	Down	GPR78	NM_080819
ESCCAL-353	chr5:1175322-1175263	Down	SLC12A7 (−63121), SLC6A19 (−26417)	Up	SLC12A7	NM_006598
ESCCAL-260	chr5:131808618-131808677	Down	IRF1 (+17817), SLC22A5 (+103247)	Up	IRF1	NM_002198
ESCCAL-132	chr5:134578804-134578863	Down	PITX1 (−208870), H2AFY (+156094)	Down	PITX1	NM_002653
ESCCAL-4	chr5:141732627-141732568	Down	SPRY4 (−27978), FGF1 (+333055)	Up	SPRY4	NM_030964
ESCCAL-356	chr5:1545355-1545296	Down	LPCAT1 (−21250), MRPL36 (+254630)	Up	LPCAT1	NM_024830
XLOC_004881	chr5:72570680-72570621	Down	TMEM174 (+101628), FOXD1 (+173701)	up	FOXD1	NM_004472
ESCCAL-36	chr6:106899513-106899572	Down	ATG5 (−125848), AIM1 (−59762)	Down	AIM1	NM_001624
XLOC_005849	chr6:138145112-138145053	Down	OLIG3 (−329552), TNFAIP3 (−43498)	Up	TNFAIP3	NM_006290
ESCCAL-262	chr6:2283830-2283771	Down	GMDS (−37933), MYLK4 (+467353)	Down	GMDS	NM_001500
ESCCAL-120	chr6:29716817-29716758	Down	LOC554223 (−42895), HLA-F (+25671)	Up	HLA-F	NM_018950
ESCCAL-257	chr6:29988410-29988469	Down	ZNRD1 (−40596), HLA-J (+14223)	Up	HLA-J	NR_024240
ESCCAL-73	chr6:36126663-36126722	Down	BRPF3 (−37857), MAPK13 (+28431)	Down	MAPK13	NM_002754
ESCCAL-97	chr6:40305634-40305575	Down	MOCS1 (−410150), LRFN2 (+249521)	Down	LRFN2	NM_020737
ESCCAL-333	chr6:72018004-72018063	Down	RIMS1 (−578616), OGFRL1 (+19557)	Down	OGFRL1	NM_024576
ESCCAL-115	chr7:12593405-12593464	Down	VWDE (−149583), SCIN (−16768)	Down	SCIN	NM_033128
ESCCAL-68	chr7:139487166-139487225	Down	TBXAS1 (−41756), HIPK2 (−9503)	Down	HIPK2	AF207702
ESCCAL-87	chr7:19958800-19958741	Down	TMEM196 (−146367), MACC1 (+298242)	Down	MACC1	NM_182762
ESCCAL-58	chr8:16355082-16355141	Down	MSR1 (−304812), FGF20 (+504562)	Up	MSR1	NM_002445
ESCCAL-130	chr8:37189082-37189141	Down	ZNF703 (−364189), KCNU1 (+547270)	Up	ZNF703	NM_025069
ESCCAL-199	chr8:38623612-38623671	Down	RNF5P1 (−164867), TACC1 (−21080)	Down	TACC1	BC041391
ESCCAL-8	chr9:34668311-34668252	Down	CCL27 (−5593), CCL19 (+22992)	Down	CCL19	NM_006274

Notes: Red-highlighted genes are whose expressions are significantly changed in Esophageal Squamous Cell Carcinoma (ESCC). Green color-highlighted rows: genes involved in lipid metabolism predicted with GO enrichment analysis. Yellow color-highlighted rows: The expression of lncRNA ESCCAL-5 was validated by PCR.

Final extension was at 72°C for 10 minutes. The amplicons were resolved in 2% agarose gel.

3.4. Bioinformatic Analysis. Intensity data were exported to GeneSpring 12.0 (Agilent Technologies, Santa Clara, CA, USA) for quantile normalization and the analysis of differentially expressed long noncoding RNAs and coding RNAs. Paired t-test analysis was used to obtain probe sets whose magnitude of change in expression of RNAs between ESCC tissue and adjacent normal esophageal tissue was either greater or less than 2.0 fold and P value < 0.05 (P values were corrected for multiple testing using the method of Benjamini-Hochberg). The normalized data containing 42544 probes were further analyzed using the R program. All control probes were removed. We then defined the coding ("NM_," "XM_") and noncoding ("lincRNA," "NR_," and "XR_") genes in the normalized data according to the definition of RefSeq accession format (http://www.ncbi.nlm.nih.gov/projects/RefSeq/key.html). Differentially expressed long noncoding RNAs (DE-lncRNAs) and coding RNAs (DE-mRNAs) were further identified. The landscapes of the whole transcriptome (lncRNAs + coding RNAs) or all lncRNAs or all coding RNAs were analyzed with gene expression dynamic inspector (GEDI).

3.5. Co-Location and Co-Expression Analysis between DE-lncRNAs and DE-mRNAs. Genomic coordinates of DE-lncRNAs were imported to GREAT software (http://bejerano.stanford.edu/great/public/html/index.php) for co-location analysis. Neighboring coding genes were then matched with DE-mRNAs to obtain a co-expression dataset. Three subgroups of genes (DE-lncRNA co-located genes, DE-mRNAs, and co-expressed genes) were used for gene expression network analysis using Cytoscape software (v2.8.3).

Abbreviations

AFAP1-AS1:	Actin filament-associated protein 1 antisense RNA
ESCC:	Esophageal squamous cell carcinoma
ESCCAL:	ESCC-associated lncRNA
lncRNA:	Long noncoding RNA
MALAT-1:	Metastasis-associated lung adenocarcinoma transcript 1
HOTAIR:	HOX antisense intergenic RNA
PCAT-1:	Prostate cancer associated noncoding RNA transcript 1
PCR:	Polymerase chain reaction.

Authors' Contribution

Wei Cao and Wei Wu contributed equally to this project.

Acknowledgments

The National Natural Science Foundation of China (no. 81171992) and the Zhengzhou Science and Technology Programme (121PPTGG494-8) supported this work. The authors thank Anne Haegert at The Laboratory for Advanced Genome Analysis at the Vancouver Prostate Centre, Vancouver, Canada, for expert technical support. They also thank GenomeSky Inc, Canada, for providing computational consultation.

References

[1] J. Ferlay, H. R. Shin, F. Bray, D. Forman, C. Mathers, and D. M. Parkin, "Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008," *International Journal of Cancer*, vol. 127, no. 12, pp. 2893–2917, 2010.

[2] W. Wu and J. A. Chan, "Understanding the role of long noncoding RNAs in the cancer genome," in *Next Generation Sequencing in Cancer Research-Decoding Cancer Genome*, W. Wu and H. Choudhry, Eds., pp. 199–215, Springer, New York, NY, USA, 2013.

[3] J. J. Hao, T. Gong, Y. Zhang et al., "Characterization of gene rearrangements resulted from genomic structural aberrations in human esophageal squamous cell carcinoma KYSE150 cells," *Gene*, vol. 513, no. 1, pp. 196–201, 2013.

[4] M. Kano, N. Seki, N. Kikkawa et al., "MiR-145, miR-133a and miR-133b: tumor-suppressive miRNAs target FSCN1 in esophageal squamous cell carcinoma," *International Journal of Cancer*, vol. 127, no. 12, pp. 2804–2814, 2010.

[5] H. Su, N. Hu, H. H. Yang et al., "Global gene expression profiling and validation in esophageal squamous cell carcinoma and its association with clinical phenotypes," *Clinical Cancer Research*, vol. 17, no. 9, pp. 2955–2966, 2011.

[6] D. M. Greenawalt, C. Duong, G. K. Smyth et al., "Gene expression profiling of esophageal cancer: comparative analysis of Barrett's esophagus, adenocarcinoma, and squamous cell carcinoma," *International Journal of Cancer*, vol. 120, no. 9, pp. 1914–1921, 2007.

[7] I. Dunham, A. Kundaje, S. F. Aldred et al., "An integrated encyclopedia of DNA elements in the human genome," *Nature*, vol. 489, no. 7414, pp. 57–74, 2012.

[8] E. A. Gibb, C. J. Brown, and W. L. Lam, "The functional role of long non-coding RNA in human carcinomas," *Molecular Cancer*, vol. 10, article 38, 2011.

[9] A. L. Brunner, A. H. Beck, B. Edris et al., "Transcriptional profiling of lncRNAs and novel transcribed regions across a diverse panel of archived human cancers," *Genome Biology*, vol. 13, no. 8, article R75, 2012.

[10] M. Guttman and J. L. Rinn, "Modular regulatory principles of large non-coding RNAs," *Nature*, vol. 482, no. 7385, pp. 339–346, 2012.

[11] K. C. Wang, Y. W. Yang, B. Liu et al., "A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression," *Nature*, vol. 472, no. 7341, pp. 120–126, 2011.

[12] A. Dean, "On a chromosome far, far away: LCRs and gene expression," *Trends in Genetics*, vol. 22, no. 1, pp. 38–45, 2006.

[13] P. Ji, S. Diederichs, W. Wang et al., "MALAT-1, a novel noncoding RNA, and thymosin β4 predict metastasis and survival in early-stage non-small cell lung cancer," *Oncogene*, vol. 22, no. 39, pp. 8031–8041, 2003.

[14] R. A. Gupta, N. Shah, K. C. Wang et al., "Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis," *Nature*, vol. 464, no. 7291, pp. 1071–1076, 2010.

[15] R. Kogo, T. Shimamura, K. Mimori et al., "Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers," *Cancer Research*, vol. 71, no. 20, pp. 6320–6326, 2011.

[16] K. Kim, I. Jutooru, G. Chadalapaka et al., "HOTAIR is a negative prognostic factor and exhibits pro-oncogenic activity in pancreatic cancer," *Oncogene*, vol. 32, pp. 1616–1625, 2013.

[17] D. Li, J. Feng, T. Wu et al., "Long intergenic noncoding RNA HOTAIR is overexpressed and regulates PTEN methylation in laryngeal squamous cell carcinoma," *The American Journal of Pathology*, vol. 182, no. 1, pp. 64–70, 2013.

[18] A. Bhan, I. Hussain, K. I. Ansari, S. Kasiri, A. Bashyal, and S. S. Mandal, "Antisense transcript long noncoding RNA (lncRNA) HOTAIR is transcriptionally induced by estradiol," *Journal of Molecular Biology*, vol. 425, no. 19, pp. 3707–3722, 2013.

[19] D. Cejka, D. Losert, and V. Wacheck, "Short interfering RNA (siRNA): tool or therapeutic?" *Clinical Science*, vol. 110, no. 1, pp. 47–58, 2006.

[20] J. R. Prensner and A. M. Chinnaiyan, "The emergence of lncRNAs in cancer biology," *Cancer Discovery*, vol. 1, pp. 391–407, 2011.

[21] W. Wu, T. D. Bhagat, X. Yang et al., "Hypomethylation of noncoding DNA regions and overexpression of the long noncoding RNA, AFAP1-AS1, in Barrett's esophagus and esophageal adenocarcinoma," *Gastroenterology*, vol. 144, no. 5, pp. 956–966, 2013.

[22] E. Özgür, U. Mert, M. Isin, M. Okutan, N. Dalay, and U. Gezer, "Differential expression of long non-coding RNAs during genotoxic stress-induced apoptosis in HeLa and MCF-7 cells," *Clinical and Experimental Medicine*, vol. 13, no. 2, pp. 119–126, 2013.

[23] J. D. Li, Q. C. Feng, and J. S. Li, "Differential gene expression profiling of oesophageal squamous cell carcinoma by dna microarray and bioinformatics analysis," *Journal of International Medical Research*, vol. 38, no. 6, pp. 1904–1912, 2010.

[24] S. Ma, J. Y. Bao, P. S. Kwan et al., "Identification of PTK6, via RNA sequencing analysis, as a suppressor of esophageal squamous cell carcinoma," *Gastroenterology*, vol. 143, pp. 675–686, 2012.

[25] M. Tong, K. W. Chan, J. Y. Bao et al., "Rab25 is a tumor suppressor gene with antiangiogenic and anti-invasive activities in esophageal squamous cell carcinoma," *Cancer Research*, vol. 72, no. 22, pp. 6024–6035, 2012.

[26] Y. J. Geng, S. L. Xie, Q. Li, J. Ma, and G. Y. Wang, "Large intervening non-coding RNA HOTAIR is associated with hepatocellular carcinoma progression," *Journal of International Medical Research*, vol. 39, no. 6, pp. 2119–2128, 2011.

[27] X. Zhou, T. J. Lawrence, Z. He, C. R. Pound, J. Mao, and S. A. Bigler, "The expression level of lysophosphatidylcholine acyltransferase 1 (LPCAT1) correlates to the progression of prostate cancer," *Experimental and Molecular Pathology*, vol. 92, no. 1, pp. 105–110, 2012.

[28] Y. Morita, T. Sakaguchi, K. Ikegami et al., "Lysophosphatidylcholine acyltransferase 1 altered phospholipid composition and regulated hepatoma progression," *Journal of Hepatology*, vol. 59, no. 2, pp. 292–299, 2013.