Functions and mechanisms of long noncoding RNAs in lung cancer

Zhenzi Peng
Chunfang Zhang
Chaojun Duan

Institute of Medical Sciences, Key Laboratory of Cancer Proteomics of Chinese Ministry of Health, Xiangya Hospital, Central South University, Changsha, People’s Republic of China

Abstract: Lung cancer is a heterogeneous disease, and there is a lack of adequate biomarkers for diagnosis. Long noncoding RNAs (lncRNAs) are emerging as an important set of molecules because of their roles in various key pathophysiological pathways, including cell growth, apoptosis, and metastasis. We review the current knowledge of the lncRNAs in lung cancer. In-depth analyses of lncRNAs in lung cancer have increased the number of potential effective biomarkers, thus providing options to increase the therapeutic benefit. In this review, we summarize the functions, mechanisms, and regulatory networks of lncRNAs in lung cancer, providing a basis for further research in this field.

Keywords: ncRNA, tumorigenesis, biomarker, network, proliferation, apoptosis

Introduction

Besides small-cell lung cancer, non-small-cell lung cancer (NSCLC) is any type of epithelial lung cancer and accounts for 85% of all lung cancers. The 5-year survival rate of this heterogeneous disease is 16.6%, and it has only improved slightly in the past few years.¹ NSCLC can be classified into discrete subclasses according to histological phenotypes, including squamous cell carcinoma (SCC), adenocarcinoma (ADC), and large-cell carcinoma. The most common type of NSCLC is ADC, which is less associated with smoking and chronic inflammation than SCC.² The clinical integration of ADC in terms of predictive biomarker signatures is characterized by positive immunostaining for cytokeratin 7 and TTF1; however, SCC is characterized by positivity of cytokeratin 5, cytokeratin 6 and/or SOX2, and p63.³,⁴ In addition to tissue biopsy for clinical diagnosis, specific gene mutations in tumors have highlighted their usefulness as diagnostic markers and molecular therapy targets. For instance, EGFR, ALK, and MET mutations are always found in ADC patients.³ However, DDR2, FGFR1, and FGFR2 mutations, as well as mutations in genes in the PI3K signaling pathway, are generally found more frequently in SCC.³

Characteristics of IncRNAs

Noncoding RNAs (ncRNAs) are a superclass of endogenous, non-protein-coding RNA transcripts, many of which have essential functions in various cellular processes. Based on their function, ncRNAs can be classified into two subclasses, namely housekeeper ncRNAs (tRNA, tRNA, snRNA, snoRNA) and the regulated ncRNAs. The regulated ncRNAs can be categorized by length as follows: any ncRNA <200 nucleotides (nt) in length is a short ncRNA (siRNA, miRNA, piRNA) and those >200 nt in length are long ncRNAs (lncRNAs). Recently, Iyer et al⁶ demonstrated that there are ~60,000 ncRNAs in the human genome, and >68% of these are IncRNAs, of which ~80%
are not annotated. Approximately 1% of lncRNAs harbor ultraconserved elements, and 7% of lncRNAs harbor disease-associated single-nucleotide polymorphisms.\(^6\)

An increasing number of lncRNAs have been identified as key regulators of a wide range of cellular processes, including dosage compensation, imprinting, transcription, mRNA splicing, translation, nuclear and cytoplasmic trafficking, and cellular localization. Notably, ectopic expression of lncRNAs is associated with a great variety of diseases.

However, compared with mRNAs, lncRNAs have their own peculiar characteristics. First, the mean length of lncRNAs is shorter than that of mRNAs, with an average of 592 nt compared to 2,453 nt for mRNAs. Second, lncRNAs have fewer exons, although they harbor standard canonical splice sites.\(^7\) Third, the methylation level of the transcription start site of lncRNAs is higher than that of mRNAs; therefore, the expression level of lncRNAs is significantly lower. In addition, lncRNAs show strict tissue specificity.\(^8\) Furthermore, lncRNAs tend to correlate with transposable elements, especially with endogenous retroviruses, compared with protein-coding genes.\(^9\) Moreover, lncRNA conservation includes more than introns or random intergenic regions, but they are less conserved than mRNAs.\(^10\) Despite the fact that the conservation of some lncRNAs at the sequence level is not high, they may have the same functions.\(^11,12\) Diederichs\(^11\) indicated that the conservation of lncRNAs should include four dimensions: sequence, structure, function, and expression from synthetic loci. Despite the differences between lncRNAs and mRNAs, they share certain common characteristics. Large-scale studies revealed that many identified lncRNAs are transcribed by RNA polymerase II, which is the same as that for mRNA. They also share the same posttranscriptional mRNA processing, including 5' - capping, splicing, and poly-adenylation at the 3' end. Certain transcription factors can bind to the promoter region of lncRNAs, such as c-myc, p53, and Sox2.

**Classification and functions of lncRNAs**

The most recent classification of lncRNAs is based on their location relative to that of target protein-coding genes. According to these criteria, lncRNAs can be classified as exonic, intronic, overlapping, or intergenic. Moreover, based on the transcriptional direction with respect to protein-coding genes, lncRNAs are divided into two groups, namely sense and antisense.\(^7\)

According to transcriptional modes, lncRNAs can be categorized as cis-acting and trans-acting lncRNAs. Cis-acting lncRNAs mediate gene expression based on their position in the vicinity of the target gene transcriptional site. However, trans-acting lncRNAs can control the expression of genes at any loci based on the recruitment of proteins to the target sites to participate in transcriptional regulation. lncRNAs hybridize with DNA or RNA molecules to form triple-stranded RNA–DNA structures that play essential roles in transcription.\(^12,13\)

lncRNAs can be classified into various discrete subclasses on the basis of their function as follows (Figure 1).\(^14\)

**Decoys**
lncRNA decoys exert biological functions by binding to proteins indirectly and playing a role in multiple processes of life.

**Scaffolds**
The second set of lncRNAs function as scaffolds, as lncRNAs act as platforms to bind relevant molecular complexes. Typically, \(HOTAIR\) acts as a bridging scaffold for the PRC2 and LSD1/CoREST/REST complex to suppress gene expression. PRC2 binds to an 89 bp fragment in the 5' end of \(HOTAIR\), and the LSD1/CoREST/REST complex binds to a 646 bp fragment in its 3' end.\(^15,16\) Genome-wide RNA immunoprecipitation analysis has shown that ~20% of the various types of lncRNAs are bound to PRC2.\(^17\) Another example of a molecular scaffold is \(Kcnqlot1\), which combines PRC2 and G9a to generate H3K27me3 and H3K9me3, leading to epigenetic silencing of genes.\(^18\) Indeed, several studies show that not only proteins but also lncRNAs play critical roles in bridging molecular components.

**Signal**
lncRNAs play critical roles in signal regulation and in the responses to various stimuli. lncRNAs are often expressed in a spatial- and temporal-specific pattern. Upon specific expression, they can modulate translation and integrate developmental cues.\(^19\) For instance, \(lincRNA-p21\), which binds to hnRNP-K, promotes the proper localization of hnRNP-K and results in the silencing of p53-regulated genes.\(^20\) Certain lncRNAs are emerging as signals of functionally significant biological events because of their roles in regulating transcriptional activity or pathways.

**Sponges**
\(PTEN1\) (pseudogene of PTEN) upregulates the tumor suppressor PTEN by attracting miRNAs to its 3' - untranslated region.\(^21\) The lncRNA \(GAPLINC\) acts as a “sponge” and modulates CD44 expression by attracting \(miR211-3p\).\(^22\) The lncRNA antisense ncRNA in the INK4 locus (\(ANRIL\)) can
function as a “sponge” to titrate miR-99a/miR-449a, thus activating CDK6 and inactivating p15INK4b/p16INK4a. Consequently, E2F1 is released inappropriately and contributes to gastric cancer cell proliferation. IncRNAs therefore function as sponges by interacting with miRNAs and suppressing their effects on target sites. Recent evidence highlights a classification of circRNAs as miRNA sponges that contribute to the downregulation of target genes. Compared to linear RNAs, the half-life of circRNAs is longer. It is plausible that the duration and the effect of circRNAs have more advantages.

Guide

Multiple studies indicate that nuclear-retained IncRNAs function in guiding chromatin modifiers to specific genomic loci. Typical is PRC2, which contains a histone methyltransferase (enhancer of zeste 2, EZH2) that inhibits gene expression via trimethylation of histone H3Lys27 (H3K27me3). In addition, chromatin conformation changes induced by nuclear-enriched IncRNAs can promote gene expression.

**Regulation modes of IncRNAs**

**IncRNAs involved in epigenetic regulation**

An example of the chromatin-modifying capabilities of IncRNAs is dosage compensation in mammals, which requires the preferential silencing of one parental allele. Xist gives rise to stable epigenetic silencing of large-scale genes in the X-chromosome by tethering PRC2 to the transcriptional site, inducing the formation of H3K27me3 to inactivate...
Another example is IncRNA p21, which can change DNA methylation levels by promoting histone methyltransferase and DNA methyltransferase binding to target sites, thus affecting the expression of reprogramming genes.17

**IncRNAs participate in transcriptional regulation**

Evidence to date indicates that IncRNAs, which are transcribed from enhancers or promoters, can act in cis pattern to control transcriptional efficiency. IncRNAs transcribed from enhancers can affect their activity or help recruit protein factors. Among IncRNAs in prostate cancer, two overexpressed IncRNAs bind to the androgen receptor, promoting androgen receptor binding to an enhancer.29

**IncRNAs affect posttranscriptional processing**

Most genes are transcribed via tissue-specific and cell-specific alternative-splicing patterns in humans.30 Noteworthy is that the majority of IncRNAs are often similarly expressed in a spatial- and temporal-specific manner. Accumulating evidence indicates that IncRNAs may regulate alternative splicing by cis-acting mechanisms or by recruiting regulatory splicing factors.31,32 Recently, Gonzalez et al showed that a conserved antisense IncRNA transcribed from the FGFR2 locus in humans regulates alternative splicing by recruiting the histone demethylase KDM2a and polycomb group proteins. These findings suggested that IncRNAs can regulate alternative splicing through the establishment of a splicing-specific chromatin signature.33

**Dysregulation and functional roles of IncRNAs in lung cancer**

Lung cancer is often associated with aberrant IncRNA transcriptomes, including onco-IncRNAs and tumor suppressor IncRNAs. Here, we discuss recent discoveries that implicate aberrant IncRNAs in lung cancer (Table 1). In addition, we provide a framework of systematically functionalized IncRNAs and integrate them with the protein-coding RNA dimension in complex networks (Figure 2).

**Onco-IncRNAs**

**MALAT1**

**Characteristics of MALAT1**

MALAT1, also known as NEAT2, is located on chromosome 6p24.3 in humans with a length of 8.7 kb.11,12 MALAT1 is a well-characterized IncRNA with markedly high expression in most NSCLC types. MALAT1 generates a primary noncoding transcript that is enriched in the nucleus.34 Furthermore, the MALAT1 gene shows strong sequence conservation from humans to zebra fish.35 MALAT1 transcription is initiated from multiple promoters and produces different MALAT1 transcript variants. However, most MALAT1 transcripts, which harbor two distinct nuclear speckle localization elements, are enriched in nuclear speckles, indicating essential role of MALAT1 in RNA metabolism. MALAT1 is modified at the 3′ end and forms a triple-helical structure, which may be the reason why MALAT1 appears as a very stable IncRNA.36,37 The half-life of MALAT1 ranges from 9 to 16.5 hours.38 After modification, MALAT1 generates a second small mascRNA that is localized to the cytoplasm.39 MALAT1 binds to the unmethylated PRC2 protein, therefore contributing to PRC2 preferential binding to H2AK5ac and H2AK13ac, marking transcriptional activation.40 At the transcriptional level, several studies have shown that MALAT1 regulates gene expression, such as that of growth control genes. MALAT1 affects the phosphorylation levels of serine/arginine (SR) splicing factors, such as B-MYB, leading to changes in gene expression.12 MALAT1 also regulates gene expression at the posttranscriptional level.41 For instance, MALAT1 affects a minor subset of transcripts in the process of alternative splicing, such as RNPS1, PRP6, and SON.42 In summary, MALAT1 can modulate gene expression both at the transcriptional and posttranscriptional level.

**MALAT1 in lung cancer**

MALAT1, which is found in abundance in various tissues and cell lines, is successfully knocked out by the zinc finger nuclease technique in A549 cells. MALAT1 has been shown to be a critical regulator of the metastasis phenotype in lung cancer cells.43 In A549 cells, inhibition of MALAT1 downregulates the expression of motility-associated genes, including HMMR, AIM1, SLC26A2, LAYN, CCT4, ROD1, CTHRC, and FHL1.44,45 Xenograft models have confirmed this effect. The prevailing view is that MALAT1 participates in the regulation of motility-related genes to enhance the motility of lung ADC cells.45 Recently, Shen et al demonstrated that MALAT1 overexpression promotes lung cancer brain metastasis by inducing epithelial-to-mesenchymal transition (EMT). Another study suggested that MALAT1 is involved in cell cycle regulation at the level of G2/M phase progression. MALAT1 interacts with nuclear hnRNP C and promotes hnRNP C translocation to the cytoplasm in the G2/M phase.47 In addition to its effect on cell cycle progression, MALAT1 affects the expression of the proapoptotic factor Bcl-2, which is specifically associated with prognosis in NSCLC.48 Moreover, a meta-analysis and various studies
have shown that high MALAT1 expression is related to poor overall survival in NSCLC.\textsuperscript{49}

HOTAIR

Characteristics of HOTAIR

HOTAIR is located on chromosome 12q13.13 in humans, and has a length of 2.1 kb. There are four gene clusters (HOXA, HOXB, HOXC, and HOXD) and 39 HOX genes in the genome.\textsuperscript{25} These clusters generate numerous IncRNAs that are often expressed in a spatial- and temporal-specific pattern.\textsuperscript{50} Although HOTAIR is located in the HOXC locus, it has been shown to repress gene expression in the HOXD locus. In addition, the secondary structure of HOTAIR contains four independently folding modules, two of which are evolutionarily conserved protein-binding domains.\textsuperscript{51} HOTAIR is enriched in the nucleus but is also expressed in the cytoplasm.\textsuperscript{52} HOTAIR has little sequence conservation in humans and mice, but the molecular mechanism is similar

<table>
<thead>
<tr>
<th>Type of IncRNA</th>
<th>Name</th>
<th>Location</th>
<th>Length (kb)</th>
<th>Expression level</th>
<th>Property</th>
<th>Intersection molecules and pathway</th>
<th>Cell processes and clinical features</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intergenic</td>
<td>MALAT1</td>
<td>6p24.3</td>
<td>8.7</td>
<td>Up</td>
<td>Oncogenic</td>
<td>Affects expression of Bcl-2 and metastasis-related genes</td>
<td>Promotes cell growth, metastasis, and EMT; inhibits apoptosis; poor survival</td>
<td>11,12, 34–49</td>
</tr>
<tr>
<td></td>
<td>HOTAIR</td>
<td>12q13.13</td>
<td>2.15</td>
<td>Up</td>
<td>Oncogenic</td>
<td>Induced by Col-1, HIF-1α. Promotes expression of gelatinases. Represses cell-adhesion related genes, p21, and HOXA5</td>
<td>Promotes cell growth, invasion, metastasis, EMT and relapse</td>
<td>25, 50–64</td>
</tr>
<tr>
<td></td>
<td>MEG3</td>
<td>14q32.3</td>
<td>1.6</td>
<td>Down</td>
<td>Tumor suppressive</td>
<td>Affects expression of p53</td>
<td>Inhibits proliferation. Induces apoptosis</td>
<td>88–91</td>
</tr>
<tr>
<td>Intrinsic</td>
<td>SOX2-OT</td>
<td>3p26.33</td>
<td>4.2</td>
<td>Up in SCC</td>
<td>Oncogenic</td>
<td>Inhibits the expression of EZH2</td>
<td>Promotes cell growth. Poor survival</td>
<td>65–68</td>
</tr>
<tr>
<td></td>
<td>SPRY4-IT</td>
<td>15q31</td>
<td>0.69</td>
<td>Down</td>
<td>Tumor suppressive</td>
<td>Silenced by EZH2. Inhibits the expression of EMT markers</td>
<td>Inhibits growth, invasion, and metastasis. Induces apoptosis</td>
<td>92–96</td>
</tr>
<tr>
<td></td>
<td>HNF1A-AS</td>
<td>12q24</td>
<td>2.46</td>
<td>Up in ADC</td>
<td>Oncogenic</td>
<td>Affects EMT-related genes via DNMT1</td>
<td>Promotes cell growth and metastasis</td>
<td>1</td>
</tr>
<tr>
<td>Overlapping</td>
<td>ANRIL</td>
<td>9p21.3</td>
<td>126</td>
<td>Up</td>
<td>Oncogenic</td>
<td>Represses KLF2 and p21 via PRC2</td>
<td>Promotes cell growth and metastasis</td>
<td>69–73</td>
</tr>
<tr>
<td></td>
<td>H19</td>
<td>11q15.5</td>
<td>2.3</td>
<td>Up</td>
<td>Oncogenic</td>
<td>Affects expression of miR675 and HGF/SF</td>
<td>Unknown</td>
<td>74–81</td>
</tr>
<tr>
<td></td>
<td>PANDAR</td>
<td>6p21.2</td>
<td>1.5</td>
<td>Down</td>
<td>Tumor suppressive</td>
<td>Affects expression of Bcl-2 via NF-YA</td>
<td>Inhibits proliferation. Induces apoptosis</td>
<td>97, 98</td>
</tr>
<tr>
<td>Exonic</td>
<td>CARLO-5</td>
<td>8q24.21</td>
<td>1.6</td>
<td>Up</td>
<td>Oncogenic</td>
<td>Suppresses p16, p21 and p27 and EMT markers</td>
<td>Affects cell cycle, proliferation, and invasion, EMT. Poor survival</td>
<td>82, 83</td>
</tr>
<tr>
<td></td>
<td>GASS</td>
<td>1q25</td>
<td>0.65</td>
<td>Down</td>
<td>Tumor suppressive</td>
<td>Affects expression of E2F1, p21, and p33. Suppresses miR-21 expression</td>
<td>Inhibits proliferation. Induces apoptosis</td>
<td>99–108</td>
</tr>
<tr>
<td></td>
<td>MVIH</td>
<td>10q22</td>
<td></td>
<td>Up</td>
<td>Oncogenic</td>
<td>Affects expression of MMP2/MMP9</td>
<td>Promotes proliferation and invasion. Poor survival</td>
<td>84</td>
</tr>
<tr>
<td>Unknown</td>
<td>TUG1</td>
<td>22q12.2</td>
<td>7.1</td>
<td>Down</td>
<td>Tumor suppressive</td>
<td>Induced by p53, represses HOXB7 via PRC2</td>
<td>Inhibits proliferation. Induces apoptosis</td>
<td>109–112</td>
</tr>
<tr>
<td></td>
<td>BANCR</td>
<td>9q21.11</td>
<td>0.69</td>
<td>Down</td>
<td>Tumor suppressive</td>
<td>Inhibits the expression of EMT markers</td>
<td>Inhibits migration, invasion, metastasis, EMT. Induces apoptosis</td>
<td>113–117</td>
</tr>
<tr>
<td></td>
<td>EVADR</td>
<td>6q13</td>
<td>0.39</td>
<td>Up in ADC</td>
<td>Oncogenic</td>
<td>Induced by MER48 ERV element</td>
<td>Unknown</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>PVT1</td>
<td>8q24.21</td>
<td>210</td>
<td>Up</td>
<td>Oncogenic</td>
<td>Induced by p53</td>
<td>Induces proliferation, invasion, and metastasis. Poor survival</td>
<td>85–86</td>
</tr>
</tbody>
</table>

Notes: Up, upregulated; down, downregulated.

Abbreviations: IncRNAs, long noncoding RNAs; EMT, epithelial-to-mesenchymal transition; SCC, squamous cell carcinoma; ADC, adenocarcinoma; MMP, matrix metalloproteinase.
Figure 2 Overview of the regulatory network of lncRNAs in lung cancer.
Notes: Blue frames represent onco-lncRNAs. Green frames represent tumor suppressor lncRNAs. Purple frames represent iFs pathway. 
Abbreviations: lncRNAs, long noncoding RNAs; miR, microRNA; MMPs, matrix metalloproteinases; EMT, epithelial-to-mesenchymal transition.
in the two species. Reciprocally, the conserved IncRNA sequence does not always possess the same function in other species as expected.

HOTAIR-mediated epigenetic gene silencing is dependent on its function as a bridge scaffold for PRC2 and LSD1/CoREST/REST. PRC2 binds to the 5’ end of HOTAIR and regulates chromosome occupancy by EZH2 (a subunit of PRC2), which leads to histone H3 lysine 27 trimethylation of the HOXD locus. Meanwhile, the LSD1/CoREST/REST complex binds to the 3’ end of HOTAIR.\(^5^3\) HOTAIR represses gene expression by increasing the occupation of H3K27me3 and decreasing the occupation of H3K4me3 on the promoter in a trans-acting manner.\(^5^4\) In addition to its role in the PRC2 and LSD1/CoREST/REST complexes, HOTAIR acts as a scaffold for E3 ubiquitin ligases and their corresponding substrates, namely E3 ubiquitin ligase Dzip3 and its substrate Ataxin-1, resulting in proteolysis.\(^5^5\) Furthermore, HOTAIR harbors a target site for miR-34a\(^6^6\) and miR-141 in its exon 6, both of which can attenuate the expression of HOTAIR. Exon 6 of HOTAIR contains a target site for let-7i that mediates the formation of a hetero-tetramer containing let-7i, Ago2, HOTAIR, and HuR.\(^5^7\)

**HOTAIR in lung cancer**

Ectopic expression of HOTAIR has been reported in a plethora of cancerous tissues. High HOTAIR levels are associated with invasion and metastases and linked to an advanced stage of disease and poor survival in patients with lung cancer.\(^5^8\)–\(^6^0\) EMT mediates the invasive phenotype of lung cancer cells. HOTAIR may promote the EMT process of lung epithelial cells by distinct mechanisms. One of these mechanisms is transcriptional repression of HOXA5 gene. HOXA5 is related to postnatal lung development. It is likely that HOTAIR and miR-196a act in common to repress the expression of HOXA5, therefore contributing to dedifferentiation during lung tumorigenesis.\(^6^1\) Moreover, HOTAIR facilitates the process of EMT by inhibiting the expression of cell adhesion-related genes in small-cell lung cancer epithelial cells.\(^5^9\) Meanwhile, tumor-promoting type I collagen, which is a potent inducer of EMT, can modulate the translational control of HOTAIR in NSCLC cells.\(^6^2\) In addition to the repression of EMT inhibitors and the promotion of EMT, HOTAIR also affects the expression of gelatinase, a matrix metalloproteinase (MMP) that plays a role in triggering invasion in lung cancer cells.\(^5^8\) HOTAIR can promote resistance to cisplatin via downregulation of p21 (WAF1/CIP1) protein levels in lung ADC cells.\(^6^3\) Furthermore, HOTAIR as a direct target of HIF-1α, promotes cell proliferation, migration, and invasion in hypoxic NSCLC. Evidence indicates that HIF-1α binds to HOTAIR via interaction with the upstream region of HOTAIR in NSCLC cells.\(^6^4\) The upregulation of HOTAIR has been shown to have a negative impact on lung cancer by regulating genes involved in invasion, metastasis, and poor survival.\(^5^9\)

**SOX2-OT**

SOX2-OT is located on chromosome 3p26.33 in humans and has a length of 4.2 kb. SOX2-OT is an intronic IncRNA that overlaps with the SOX2 gene, which is a major regulator of pluripotency.\(^6^5\) Recent studies have shown that SOX2 and SOX2-OT have similar expression patterns in lung SCC and ADC tissues.\(^6^6\) There are multiple transcription initiation sites in the human and mouse SOX2-OT locus, leading to many spliced variants.\(^6^7\) In addition, the human genomic region of SOX2-OT is characterized by multiple conserved transcription factor-binding sites. These sites have essential roles in the tumorigenesis process.\(^6^5\) SOX2-OT is expressed at higher levels in human primary lung cancer tissues than in adjacent non-tumor tissues. Typically, SOX2-OT exhibits significantly higher expression in SCC of the lung than in ADC.\(^6^8\) With respect to the mechanism of SOX2-OT, Hou et al showed that knockdown of SOX2-OT expression leads to cell cycle arrest at G2/M phase through the modulation of the expression of EZH2. Meanwhile, its expression level is significantly correlated with cell proliferation and colony formation ability in lung cancer cell lines.\(^6^9\) This study further indicates that high SOX2-OT expression predicts poor survival in lung cancer patients.

**HNF1A-ASI**

HNF1A-ASI is located on chromosome 12q24 and has a length of 2.46 kb. Overexpression of HNF1A-ASI has been reported in lung ADC tissues compared with the corresponding non-tumor tissues. In addition, elevated expression of HNF1A-ASI is linked to tumor–node–metastasis (TNM) stage, tumor size, and lymph node metastasis. HNF1A-ASI can regulate EMT-related protein expression via binding to DNMT1, therefore regulating cell growth and metastasis both in vitro and in vivo.\(^1\)

**ANRIL**

*ANRIL* is derived from the p15/CDKN2B-p16/CDKN2A-p14/ARF gene cluster, which maps to human chromosome 9p21.3 with a length of 126 kb. This gene family is associated with cutaneous malignant melanoma and neural system tumors. *ANRIL* includes several isoforms with
tissue-specific expression because it consists of 19 exons.69 Recently, elevated levels of ANRIL have been reported in NSCLC tissues, and its expression level is significantly correlated with poor prognosis. siRNA-mediated knockdown of ANRIL results in the inhibition of cell proliferation and the promotion of apoptosis both in vitro and in vivo.70 An ongoing study indicates that the subcellular localization of ANRIL is mostly in the cell nucleus. ANRIL is indicated as a “decoy”, and it represses KLF2 and p21 transcription by binding to PRC2 in NSCLC PC9 cells, which sheds light on the effect of ANRIL on NSCLC cell proliferation and apoptosis partly in trans.70 Generally, KLF2 as a tumor suppressor is significantly downregulated in various cancers, leading to inhibition of cell proliferation via KRAS.71 ANRIL can regulate the transcription of miR-99a and miR-449a by recruiting the PRC2 complex in gastric cancer.72 Recently, Ren et al73 demonstrated that high expression of miR-449a attenuates lung cancer cell proliferation, and the downregulation of miR-449a is correlated with a shorter disease-free survival of patients. It is plausible that ANRIL modulates the expression of miR-449a, thereby inhibiting the proliferation of lung cancer cells during lung tumorigenesis.

H19

Characteristics of H19
H19 is located on chromosome 11q15.5 in humans and has a length of 2.3 kb. H19 is a paternally imprinted gene that is spliced into five exons. The H19 gene locus is complex, harboring conserved miR-675 and antisense protein-encoding transcript (HOTS), which is a tumor suppressor.74 Another pro-tumorigenic antisense transcript, 91H, overlaps with the H19 gene locus.75 H19 and its nearby gene IGF2 show uniparental mono-allelic expression. There is an imprinting control region (ICR) between them. The ICR is unmethylated on maternal chromosomes, where it binds to the transcription factor CTCF and inhibits the enhancer from binding to the ICR. As a result, the enhancer binds to H19 and induces its expression. Conversely, on paternal chromosomes, the ICR is methylated and binds to the enhancer, resulting in H19 downregulation.76 H19 can act in both cis and trans patterns: for instance, H19 regulates the imprinting of IGF2 by silencing the expression of neighboring genes. An example of H19 acting in trans pattern is its role as a molecular sponge for miRNA let-7, which is involved in inducing EMT.77

H19 in lung cancer
Evidence to date indicates that H19 is associated with various tumorigenesis signaling pathways, including the p53 and HIF-1α, TGF-β, Bcr-Abl, Wnt/β-catenin, and HGF pathways.78 In lung cancer cells, H19 is induced by hypoxic stress via a p53-dependent manner. Knockdown of H19 expression in hypoxia has a suppressing effect on cancer cell proliferation, anchorage-independent growth, and colony formation.79,80 Furthermore, knockdown of H19 can reverse the tumorigenic and scattering effect of HGF/SF on A549 cells.79 The overexpression of H19 has a negative impact on lung cancer. Notably, the upregulated H19 is loss of imprinting independent in the airway epithelia of smokers in comparison with nonsmokers.81

CARLO-5
CARLO-5 is located on chromosome 8q24.21 in humans and has a length of 1.6 kb. CARLO-5 is significantly upregulated in NSCLC tissues. Overexpression of CARLO-5 in NSCLC tissues is significantly correlated with advanced TNM stage. The expression of p16, p21, and p27, which are G0/G1 arrest markers, decreases with the downregulation of CARLO-5.82 High level of CARLO-5 expression is a prognostic indicator of poor patient survival. Moreover, elevated expression of CARLO-5 is associated with increased proliferation and invasion ability, partially through the modulation of EMT.83

MVIH
MVIH is located on chromosome 10q22 in humans. Overexpression of MVIH has been reported in NSCLC tissues, and its expression level is significantly correlated with TNM stage and tumor size. High levels of MVIH expression are prognostic indicators of poor survival.84 siRNA-mediated knockdown of MVIH inhibits cell proliferation and invasion, partly via modulating the expression of MMP2 and MMP9. MMPs are involved in multiple biological processes, including remodeling of extracellular matrix, cell proliferation, differentiation, and metastasis.

PVT1
PVT1 is located on chromosome 8q24.21 in humans and has a length of 210 kb. The similar expression patterns of PVT1 and MYC gene might be explained by the shared genomic locus between them. A study identified a p53 transcription factor-binding site in the PVT1 promote region.85 Overexpression of PVT1 in NSCLC tissues is significantly correlated with TNM stage. In addition, patients with high levels of PVT1 expression show poor survival.86 Knockdown of PVT1 expression inhibits lung cancer cell proliferation, migration, and invasion.

EVADR
EVADR is located on chromosome 6q13 in humans and has a length of 0.39 kb. Recent evidence shows that EVADR is
overexpressed in ADC tissues, including lung ADC, and is correlated with decreased patient survival. Among nine MER48-associated IncRNAs, EVADR is the only one that is consistently expressed in ADC tissues. EVADR expression is regulated via an active promoter provided by the MER48 endogenous retrovirus element.

**Tumor suppressor IncRNAs**

**MEG3**
MEG3 is located on chromosome 14q32.3 in humans and has a length of 1.6 kb. Previous evidence indicates that MEG3 is a tumor suppressor because of its role in modulating angiogenesis. MEG3 can act in both p53-dependent and p53-independent manner during different processes. However, overexpression of MEG3 decreases NSCLC cell proliferation and induces apoptosis via the activation of p53. In addition, MEG3 expression is deregulated in NSCLC tissues, and a low expression level is significantly related with higher TNM stage, increased tumor size, and poor patient survival.

**SPRY4-IT1**
SPRY4-IT1 is located on chromosome 5q31 in humans and has a length of 0.69 kb. SPRY4-IT1 is upregulated in melanoma, esophageal SCC, and clear cell renal cell carcinoma, suggesting a common oncogenic role. Sun et al showed that SPRY4-IT1 is significantly downregulated in 94.2% of NSCLC cancerous tissues compared with normal tissues, which suggests an anti-oncogenic role. Ectopic expression of SPRY4-IT1 is associated with tumor size, advanced pathological stage, lymph node metastasis, and overall survival time in NSCLC patients. Reduced SPRY4-IT1 expression is an independent prognostic marker for NSCLC. EZH2, a methyltransferase and a catalytic subunit of PRC2, is overexpressed in NSCLC, and its downregulation prevents it from binding to the SPRY4-IT1 promoter region. This decreases the H3K27me3 modification, resulting in the inhibition of SPRY4-IT1 expression. SPRY4-IT1 has also been shown to promote NSCLC cell proliferation and metastasis by modulating the process of EMT.

**PANDAR**
PANDAR is located on chromosome 6p21.2 in humans and has a length of 1.5 kb. Although PANDAR is transcribed in antisense to CDKN1A, it is not a linked transcript of CDKN1A. PANDAR is a direct transcriptional target of p53 in NSCLC cells, and can modulate Bcl-2 expression by binding to NF-YA, thus affecting NSCLC cell apoptosis. PANDAR interacts with NF-YA (NF-YA is related to tumorigenesis) to decrease proapoptotic gene expression in a p53-dependent manner in normal human fetal lung fibroblasts. In NSCLC tissues, PANDAR interacts with NF-YA and is deregulated. PANDAR downregulation is associated with increased tumor size and advanced TNM stage. PANDAR expression is an independent prognostic predictor for NSCLC. Moreover, PANDAR, a transcriptional target of p53, affects NSCLC cell apoptosis partly by modulating Bcl-2 transcription through binding to NF-YA, thus affecting the proliferation of NSCLC cells in vitro and in vivo.

**GAS5**
Characteristics of GAS5
GAS5, comprising 12 exons, is located on chromosome 1q25 in humans and has a length of 0.65 kb. With respect to this locus, there are ten C/D box snoRNAs transcribed from its intronic regions. In addition, these snoRNAs have increasingly been linked to the functions of GAS5. GAS5 plays an essential role in cell apoptosis and growth. Evidence to date indicates that GAS5 modulates the activity of glucocorticoid-responsive genes. In this process, GAS5 represents a clear example of a decoy IncRNA, which competitively binds to the glucocorticoid receptor and prevents it from binding to glucocorticoid response elements.

**GAS5 in lung cancer**
Multiple studies have indicated that GAS5 acts as a tumor suppressor in various cancers, such as gastric cancer, hepatocellular cancer, and colorectal cancer, because of its role in the inhibition of proliferation and promotion of apoptosis. In NSCLC patient samples, decreased expression of GAS5 is linked to advanced TNM stage and increased tumor size. Increased expression of GAS5 deregulates E2F1 and drives the expression of p21 p53 in NSCLC cells, indicating that GAS5 has a regulatory effect on NSCLC cell proliferation. The activity of GAS5 can be regulated by miR-21, which has a putative binding site in GAS5. In turn, GAS5 suppresses miR-21 expression in a feedback loop between them. Dong et al. showed that overexpression of GAS5 in the lung ADC A549 cell line reverses gefitinib resistance, suggesting its tumor-suppressive function. Moreover, overexpression of GAS5 reverses the resistance to EGFR-tyrosine kinase inhibitors in ADC in vitro and in vivo. Generally, increased EGFR is related to poor prognosis in ADC.

TUG1
TUG1 is located on chromosome 22q12.2 in humans and has a length of 7.1 kb. The TUG1 gene displays a high level
of conservation in the human, mouse, rat, dog, and cow genomes.\textsuperscript{109} TUG1 is involved in photoreceptor development and is deregulated in many kinds of human cancers. In NSCLC, TUG1 possesses tumor suppressor features such as inhibition of cell proliferation and promotion of apoptosis.\textsuperscript{108} TUG1, which is induced by p53, is found binding to PRC2 and epigenetically regulates the expression of HOXB7.\textsuperscript{110} Thus, it is plausible that TUG1 modulates NSCLC cell growth via the AKT and MAPK signaling pathways because HOXB7 participates in these pathways. Moreover, patients with low level of TUG1 expression display a higher TNM stage, increased tumor size, and relatively poor overall survival.\textsuperscript{110} TUG1 has an oncogenic role in NSCLC, but it is a bona fide ncRNA in other cancer entities, including urothelial carcinoma of the bladder\textsuperscript{111} and osteosarcoma.\textsuperscript{112}

BANCR

BANCR is located on chromosome 9q21.11 and has a length of 0.69 kb. BANCR is upregulated in malignant melanoma, colorectal carcinoma, and papillary thyroid carcinoma tissues, suggesting a common oncogenic role.\textsuperscript{113\textendash}115 However, Sun et al showed that BANCR is significantly downregulated in NSCLC cancerous tissues compared with normal tissues.\textsuperscript{116} In addition, deregulated expression of BANCR is associated with increased tumor size, advanced pathological stage, lymph node metastasis, and poor survival in NSCLC patients. Reduced BANCR expression is an independent prognostic marker for NSCLC. Furthermore, knockdown of BANCR expression leads to the promotion of cell migration and invasion but inhibition of metastasis. It is plausible that downregulated BANCR promotes cell proliferation by downregulating p21 expression.\textsuperscript{117} Subsequent studies further indicate that BANCR has a critical role in EMT via modulation of E-cadherin, N-cadherin, and Vimentin expression. In sum, BANCR is proposed to modulate NSCLC cell-invasive and metastatic ability partially by modulating the EMT process.\textsuperscript{116}

Discussion

Lung cancer is responsible for the largest number of cancer-related deaths around the world. One of the main barriers to the success of lung cancer therapy is the lack of tumor biomarkers for early diagnosis. In previous studies, evidences have mainly focused on elucidation of lncRNAs in cellular and mouse models. As shown in Table 1, the expression of MALAT1, HOTAIR, SOX2-OT, HNF1A-AS1, ANRIL, MVIH, and PVT1 is associated positively with tumor size; reciprocally, the expression of MEG3, SPRY4-IT1, GAS5, and TUG1 is negatively correlated. Besides, the expression of MALAT1, HOTAIR, HNF1A-AS1, ANRIL, and PVT1 is associated positively with lymph node metastases, whereas the expression of SPRY4-IT1 and BANCR is negatively correlated. Indeed, to accurately and comprehensively understand the role of lncRNAs in human, large clinical cases of pathological characteristics as well as the prognosis are needed. The previous studies have indicated that overexpression of MALAT1, SOX2-OT, ANRIL, CARLO-5, MVIH, and PVT1 is a negative prognostic marker for patient survival. Indeed, the clinical integration of lncRNAs with respect to prognostic and predictive biomarker signatures will increase the therapeutic benefit. Here, we summarize the recent research and regulatory networks of lncRNAs (Figure 2) in lung cancer. Interestingly, the key role of PRC2 in modulating proliferation and cell cycle of lung cancer cells has been emerging. PRC2 indirectly functions as a double-edged sword through ROB1, ROCK1, and E2F1. First, MALAT1, ANAIL, SOX2-OT, and HOTAIR promote proliferation and cell cycle by regulating PRC2. TUG1 inhibits proliferation by regulating PRC2. Meanwhile, PRC2 is able to inhibit the expression of SPRY4-IT1. Several growth-related genes including p21 and p53 have been shown in the hinge of network. BANCR, GAS5, and MEG3 enhance activity of these tumor suppressor genes. CARLO-5 and HOTAIR have been proposed to highlight oncogenic feature by inhibiting the expression of p21 and p53. Generally, MMPs play a critical role in tumor cell growth and metastasis by altering the environments in which the cells grow.\textsuperscript{118} We have shown that HOTAIR and MVIH possibly regulate proliferation and metastatic ability of lung cancer cells. Noteworthy is that the altering expression of E-cadherin, Vimentin, and N-cadherin is a fundamental event in EMT.\textsuperscript{119} It is plausible that MALAT1, BANCR, SPRY4-IT1, and H19 act in concert with them to regulate the progress of EMT during lung tumorigenesis. Additionally, miRNAs as high-potential biomarkers have critical position in regulatory network of lung cancer; for example, miR-196a represses HOX5 expression, thereby promoting metastasis. It is very likely that lncRNAs harbor miRNA seed regions and enrich target RNA-binding motifs. Thus, miRNAs are proposed to modulate lung tumorigenesis by lncRNAs. Although an impressive number of studies in the last decade focused on the characteristics and functions of ncRNAs, research is still in its infancy and presents great challenges. First, as the sequence and structure of lncRNAs are only poorly conserved, the canonical knockdown and knockout methods may have no effect. In addition, ectopic expression of lncRNAs may not show obvious phenotypes...
as with protein-coding transcripts. Second, reference value is not always high between different researches, owing to different and multiple functions of IncRNAs in different tissues and cells. Therefore, elucidating the biological functions of IncRNAs is not easy. Third, the limited bioinformatical resources are another reason. Current IncRNA annotation is lacking compared with other RNA databases. Similarly, bioinformatic tools, such as IncRNA secondary structure prediction, remain to be developed. Unraveling the functions and regulatory mechanisms of IncRNAs in lung cancer might be a future breakthrough to improve our understanding of this network. The integration of miRNA and IncRNA signature profiling in lung cancer may be a useful tool for clinical applications.

Conclusion
IncRNAs are increasingly being recognized as critical molecules in various biological processes. In addition to the various types, there is a large number of IncRNAs, and they show numerous modes of interaction. Based on the location concerning the nearest protein-coding gene, IncRNAs can be classified into four subclasses, namely exonic, intronic, overlapping, and intergenic IncRNAs. According to their function, IncRNAs can be categorized as signal, decoy, sponge, guide, and scaffold molecules. It has become increasingly clear that IncRNAs are involved in tumorigenesis in many cancers. This network should serve as a guide for “navigating” through the IncRNAs research in the literature.

Acknowledgment
This work was supported by grants from the National Natural Science Foundation of China (Nos 30871189, 81171841, 81200366, 81372515, 81401901, and 81572281).

Disclosure
The authors report no conflicts of interest in this work.

References


