We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900 Open access books available
116,000 International authors and editors
120M Downloads

154 Countries delivered to
TOP 1% Our authors are among the most cited scientists
12.2% Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Chapter 11

Long Non-Coding RNA in Non-Small Cell Lung Cancers

Zule Cheng and Hongju Mao

Abstract

Non-small cell lung cancer (NSCLC) accounts for nearly 80% of diagnosed lung cancers. Due to the predominantly late diagnosis of NSCLC and drug resistance in the targeted therapy approaches, the 5-year overall survival rate is still less than 19%. Thus, novel diagnosis and treatment approaches are needed. Many efforts have been made to achieve great progress in understanding the genomic landscape of NSCLC and the molecular mechanisms involved in tumorigenesis. Long non-coding RNAs (lncRNAs) are transcripts longer than 200 nucleotides with little or no protein-coding potential. They are encoded across the genome and are involved in a wide range of cellular and biological processes. Dysregulation of lncRNAs is associated with a number of cancer-related processes, including epigenetic regulation, microRNA silencing, and DNA damage. Furthermore, lncRNAs have been reported to have the potential as biomarker for diagnosis and prognosis, as well as the therapy targets. Here in this chapter, we review some well-characterized lncRNAs associated with NSCLCs and the potential of lncRNAs as biomarkers in the diagnosis and prognosis of NSCLCs.

Keywords: NSCLC, lncRNA, diagnosis, prognosis, epigenetic

1. Introduction

Lung cancer is the leading cause of cancer-related deaths worldwide. According to the estimation of National Cancer Institute, USA, the estimated new cases for lung cancer in 2016 will be 224,390, accounting 13.3% of all new cancer cases, and the estimated deaths of lung cancer in 2016 will be 158,080 accounting 26.5% of all new cancer cases [1]. Non-small cell lung cancer (NSCLC) is the most common type of lung cancer which accounts for nearly 80–85% of diagnosed lung cancers [2]. NSCLC can be further histologically classified into three major subtypes: lung adenocarcinoma (ADC), lung squamous cell carcinoma (SCC), and large cell carcinoma. Much attention has been paid to the clinical diagnosis and treatment of NSCLC, however, the 5-year overall survival rate of NSCLC is still less than 19% in these
days [1]. This may attribute to the advanced stage of the disease at the time of diagnosis for many patients. The predominantly late diagnosis of NSCLC has limited the therapy options. The low-dose computed tomography (CT) scan can detect NSCLC early and has become the dominant detection approach. However, the high cost and the risk of false positive has overshadowed the benefits of swift diagnosis [3]. Thus, it is important to develop novel early detection approach with high sensitivity and specificity.

Biomarker is a powerful approach for cancer detection and treatment. It is defined as an indicator of biologic processes, pathogenic processes, or pharmacologic responses to therapeutic interventions. Traditional protein biomarkers such as CEA, SSC, CY211, and CA125 are classic tumour biomarkers commonly used in the diagnoses of NSCLC patients [4]. However, the current lack of diagnostic sensitivity and specificity has limited their usefulness in early detection of NSCLC. The occurrence of NSCLC always comes with the genetic changes. A thorough understanding of the genetic aberrations that contribute to NSCLC would assist in identifying biomarkers that could aid in earlier diagnoses and serve as drug targets, thus increasing treatment efficacy. Considerable efforts have been made to achieve great progress in understanding genomic landscape of NSCLC and the molecular mechanisms involved in tumorigenesis, several cancer-related genes such as TP53, EGFR, and KRAS, have been identified which play a vital role in cancer-related pathways [5–7]. Identification and characterization of specific driver mutations has transformed the diagnosis and treatment of NSCLC.

With the development of sequencing technology and bioinformatics databases, researchers have identified that more than 90% of genome is transcribed; of these transcripts, most are non-coding RNAs with little or no protein-coding potentials [8]. The enormous number and complex kinds of non-coding RNAs have drawn peoples’ attention to their roles in biological processes. MicroRNA (miRNA) is a well-studied small non-coding RNA of 18–25 nucleotides [9]. The functions of miRNAs can be summarized as mediating gene silencing by interfering with translational process or inducing mRNA degradation [10, 11]. miRNAs can be classified into oncomiRNAs and tumour suppressor miRNAs in relation to their function in carcinogenic processes; meanwhile, some of them show both oncogenic and suppressive activities under different situations [12]. Another advantage for miRNAs is their high stability and easy detection in tissue and blood [13, 14]. Several studies have reported the deregulation of various miRNAs in NSCLC [13, 15]. Screening studies have uncovered the potential of miRNAs as biomarkers in the diagnosis and prognosis of NSCLC [16, 17]. As the researches for novel biomarkers and therapy targets go further, another class of non-coding RNA molecules, long non-coding RNAs (lncRNAs) with longer length and more complex biological functions have drawn people’s attentions and become a new star in the RNA world.

2. Long non-coding RNAs in non-small cell lung cancer

2.1. Long non-coding RNAs

The development of microarray and high-throughput sequencing technologies have enabled us to explore the RNA world. As the understanding of the heterogeneous RNA molecules
Long Non-Coding RNA in Non-Small Cell Lung Cancers
http://dx.doi.org/10.5772/66487

goes deeper, the functional RNA molecules gain increasing attention again, among these, IncRNAs play a major role in the centre stage. Long non-coding RNAs can be loosely defined as a class of non-coding RNA, which are longer than 200 nucleotides. Different from the small non-coding RNAs, although IncRNAs have little or no protein coding potentials, several common features are still shared with mRNAs. Most of the IncRNAs are transcribed by RNA polymerase II, subsequent post-transcriptional processing including alternative splicing, 5′-capping, and polyadenylation are prevalently found in many IncRNAs [18]. Like mRNAs, the expression of IncRNAs is also under the regulation of transcriptional and epigenetic factors. Active or repressive histone marks that indicate the transcription status can also be found around the transcription start site of the IncRNAs [19]. On the other hand, IncRNAs have their own characters. LncRNAs have shorter median transcript length (2453 nucleotides for mRNAs and 592 nucleotides for IncRNAs) and less median exons number (8 exons for mRNAs and 3 exons for IncRNAs) than mRNAs [18]. Most IncRNAs are located in the nucleus, as most of them are functioned as regulation factors. The expression level of IncRNAs is always lower in cells than mRNAs, but with higher tissue specificities [18]. On the epigenetic level, the transcription start sites of IncRNAs have a higher density of DNA methylation compare with the mRNAs, however, this high methylation density is independent of their expression status [19].

The various features associated with mRNAs imply the complex origin and functions of IncRNAs. Study on the origin of IncRNAs are relatively scant, several hypotheses of the emergence of IncRNAs have been put forward. Some IncRNAs, such as Xist IncRNA, are believed to originate by undergoing a metamorphosis from erstwhile protein-coding gene while incorporating transposable sequence [20]. Other studies report the IncRNA can also originate from chromosome’s rearrangement, duplication of a non-coding gene by retrotransposition, neighbouring repeat, or transposable elements insertion [21, 22]. Along with the complex originations, IncRNAs also have heterogeneous groups with multiple classifications. The most prevalent classification method is based on the genomic location and context. According to this method, IncRNAs can be defined as: (1) sense IncRNAs; (2) antisense IncRNAs that are transcribed from the sense or antisense strands, respectively, overlapping one or more exons of protein-coding gene on the same or opposite strand; (3) bidirectional IncRNAs, whose transcription and neighbouring coding transcript on the opposite strand is initiated in close genomic proximity; (4) intronic IncRNA that are transcribed entirely from introns of protein-coding genes; (5) intergenic IncRNA that lies within the genomic interval between two genes [20, 23, 24]. Functions of IncRNAs may differ from each category. In general, through molecular mechanisms like signalling, decoying, guiding, and scaffolding [25], IncRNAs are widely involved in gene progresses like chromatin modification, transcriptional regulation, and post-transcriptional regulation [26–29]. As the exploration of IncRNAs goes further, growing evidences demonstrate that IncRNAs play important roles in various cellular processes [25, 30, 31]. Studies have identified the aberrant expression of IncRNAs in various cancers [32, 33] including non-small cell lung cancer [34]. The deregulation of some specific functional IncRNAs is proved to be important drivers implicated in tumour initialization and malignant transformation. Thus, IncRNAs have the potential as cancer biomarkers in the diagnosis and prognosis, as well as the targets for cancer therapy.
2.2. Functional lncRNAs in NSCLC

According to the current version of GENCODE (encyclopaedia of genes and gene variants), 15,767 long non-coding RNA genes encoding 27,692 long non-coding loci RNA have been identified based on manual curation, computational analysis, and experimental validation [35]. Along with it is the plethora of deregulated lncRNAs that are found in plenty of high-throughput lncRNA screen works. However, compared with the numerous screened lncRNAs, only few lncRNAs are well characterized and validated, the roles of most deregulated lncRNAs in diseases still remained unknown, and the data on the mechanism are scarce. In this section, some well characterized lncRNAs with reported deregulation and associated pathophysiological functions in NSCLC are reviewed.

2.2.1. HOX transcript antisense RNA (HOTAIR)

HOX transcript antisense RNA (HOTAIR) is a 2158 bps long antisense lncRNA transcribed from human HOXC locus in chromosome 12q13 [36]. As one of the most well-studied lncRNA implicated in cancer, HOTAIR is mainly involved in the epigenetic regulation as a molecular scaffold. HOTAIR can interact with polycomb repressive complex 2 (PRC2) and lysine-specific demethylase 1 (LSD1) in its 5′- and 3′-domain, respectively, and recruits PRC2 and LSD1 to the HOXD locus located on chromosome 2, inducing H3K27 methylation and H3K4 demethylation, thus silences a gene cluster involved in metastasis suppression.

HOTAIR was first reported to be highly overexpressed in primary breast cancer and metastatic breast cancer tissues. The high expression level of HOTAIR in breast cancer was closely associated with the metastasis [37]. A further study in the breast carcinoma cells showed that the enforced overexpression of HOTAIR led to the methylation of H3K27 [37]. Other researches also reported the up-regulated HOTAIR as a negative prognostic predictor in hepatocellular carcinoma [38], colorectal carcinoma [39], pancreatic cancer [40], oesophageal carcinoma [41], lung cancer [42], and gastric cancer [43, 44]. In gastric cancer, HOTAIR was also reported as an endogenous sponge of miR-331-3p, thus abolishing repression of target gene HER2 [44].

In non-small cell lung cancer, Liu et al. analysed the HOTAIR expression level in 42 NSCLC tissues and 4 NSCLC cell lines and reported the high expression of HOTAIR in both NSCLC samples and cell lines compared with corresponding normal counterparts. Results showed that high expression level of HOTAIR was correlated with advanced disease stage, metastasis, and short disease free interval. Furthermore, knockdown of HOTAIR decreased the migration and invasion of NSCLC cells in vitro and impeded cell metastasis in vivo, without altering cell vitality, which suggests a potential therapeutic role of lncRNA targeted therapies [42]. Nakagawa et al. examined the expression of HOTAIR in 77 NSCLCs and 6 brain metastases. They confirmed the negative prognostic effects of HOTAIR high expression and pointed out the overexpression of HOTAIR enhanced the invasion of NSCLC cells [45]. In lung adenocarcinoma (ADC) cells, the upregulation of HOTAIR contributed to the cisplatin resistance of ADC cells through the regulation of p21 expression, and a silence of HOTAIR resulted in the increase of chemosensitivity, which further led to the inhibition of cell proliferation, induction of G0/G1 cell-cycle arrest and enhancement of apoptosis [46]. Another study of lung ADC found that HOTAIR was an important mediator for the ratio of FOXA1 and FOXA2, which
might infect the migration and invasion [47]. A tumour micro-environment study performed in a type I collagen (Col-1) supplemented three-dimensional organotypic culture model found that the expression of HOTAIR could be upregulated by the tumour-promoting Col-1 in lung cancer cells. This finding provided a deeper insight into the mechanism of HOTAIR regulation [48].

In summary, various studies have worked on illuminating the mechanism of HOTAIR deregulation and function in NSCLCs. HOTAIR is widely involved in the chromatin modifications, and can also interact with various molecules like miRNAs and proteins. Although the facts that HOTAIR promotes cancer progress and drug resistance in NSCLC cells have been revealed, there are still many unclarified details in the mechanisms. Hence, more analyses on HOTAIR regulation and modes of action are needed.

2.2.2. Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1)

Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), also known as nuclear-enriched abundant transcript 2 (NEAT2), is an 8 kbs nuclear IncRNA expressed in chromosome 11q13 [49]. The mature MALAT1 transcript is generated through the procession by RNase P and RNase Z from the primary transcript [50]. MALAT1 is located in the nuclear speckles, and is mainly involved in alternative splicing process [51, 52]. MALAT1 exists widely and conservatively in lung, pancreas, and other healthy organs, the abundant amount of MALAT1 in these organs suggests significant functions for MALAT1 [49].

As one of the earliest identified cancer-associated IncRNAs, MALAT1 was firstly regarded as a high-risk predictor for metastasis in early stage NSCLC patients [49]. Since then, accumulating evidences confirmed the negative prognostic factor of MALAT1 in various cancers. Overexpression of MALAT1 was identified in pancreatic cancers and colorectal cancers, the high expression level was correlated with clinical progression and poor prognosis [53, 54]. Upregulation of MALAT1 was reported to promote the proliferation and metastasis of osteosarcoma and gallbladder cancer via different pathways such as PI3K/AKT and ERK/MAPK pathways [55, 56]. In oesophageal squamous cell carcinoma, silencing the MALAT1 expression resulted in the inhabitation of proliferation, migration, and invasion [57, 58]. Other studies also reported the deregulation of MALAT1 in bladder cancer and renal cancers [59, 60].

High expression of MALAT1 was identified in both early stage lung adenocarcinomas and squamous cell lung cancers [49, 61]. The overexpression of MALAT1 in NSCLC was reported to be associated with poor prognosis, shorter overall survival, and metastasis development. An RNAi-mediated suppression of MALAT1 performed in A549 cells led to the suppression of cell migration and clonogenic growth. Reversely, enforced upregulation of MALAT1 resulted in an increased NSCLC cell growth and colony formation in vitro [61]. In a later study, a highly efficient knockdown of MALAT1 in a NSCLC cell line, through zinc finger nuclease-based technique, confirmed the MALAT1 positive influence on the cell metastasis in vitro and in vivo without affecting cell proliferation. Furthermore, blocking MALAT1 by antisense oligonucleotides (ASO) prevented the tumour metastasis formation in a tumour implanted mouse xenograft module, which provides a potential therapeutic approach to prevent lung cancer metastasis [62].
Numbers of studies have been performed on elucidating the mechanism of MALAT1 regulation. Alternative splicing is one of the major topics in the MALAT1 functions. One example was that MALAT1 could interact with some alternative splicing factors such as Serine/Arginine (SR) proteins, thus affected the gene expression [52]. Another example in human diploid fibroblasts cell lines demonstrated the depletion of MALAT1 might led to the aberrant alternative splicing of the pre-mRNA of oncogenic transcription factor B-MYB, thus reduced the expression of it [63]. Another major topic is the gene expression regulation. MALAT1 displayed a strong association with genes involved in cellular growth, movement, proliferation, signalling, and immune regulation [61]. In lung cancer, MALAT1 could activate the expression of some metastasis-associated genes without affecting the alternative splicing [62]. A study involved in the epigenetic field reported that MALAT1 could interact with demethylated Polycomb 2 protein (Pc2), and controlled the re-localization of growth control genes between polycomb bodies and interchromatin granules [64].

In summary, the poor prognostic role and multiple functions of MALAT1 indicate its potential as predictable biomarker and therapy target. However, detailed studies are still needed as the mechanisms of MALAT1 regulation differ in various situations.

2.2.3. H19

H19 is a paternally imprinted and maternally expressed gene localized on human chromosome 11p15 [65]. Beside the H19 transcripts, the H19 locus also harbours miR-675, an antisense transcript, and an antisense protein-encoding transcript [66]. H19 was first reported as a tumour suppressor gene in mice [65]. However, later studies point out the oncogenic potential of H19.

Loss of imprinting (LOI) in paternal allele and the resulting overexpression of H19 was found in various cancers including lung cancer [67], oesophagus cancer [68], osteosarcoma [69], and bladder cancer [70]. H19 upregulation was found related to a range of risk factors such as smoking, carcinogens exposing, and hypoxia. Cigarette smoking could induce a LOI-independent upregulation of H19 by activating the H19 maternal allele [71]. This observation was also confirmed in a later in vitro study, which reported the increased expression of H19 in human respiratory epithelial cells treated with cigarette smoking condensate [72]. Hypoxia could also induce upregulation of H19 in cell lines with mutated p53, through a critical factor HIF1-alpha [73]. H19 was also directly induced by MYC oncogene in different cell types including fibroblast cells. A study demonstrated that c-Myc selectively increased H19 transcription from the maternally derived allele, and downregulated the reciprocally imprinted gene insulin-like growth factor 2 (IGF2) at the H19/IGF2 locus. This study indicated that c-Myc and H19 expression had strong association in lung carcinomas [74]. The mineral dust-induced gene (MDIG) could conduct the demethylation of H3K9me3 in the promoter region of H19 and activate the H19 expression. High expression of MDIG and H19 were found correlated with poorer survival of the lung cancer patients [75].

In NSCLC, the expression levels of lncRNA H19 in tissues and cells were significantly higher than adjacent tissues and normal cells, overexpression and knockdown of c-Myc could change
the H19 expression level significantly. Moreover the higher expression of H19 was positively correlated with advanced tumour-node-metastasis (TNM) stage and tumour size [76].

In summary, the expression level of H19 is related to many risk factors including smoking, which is an important lung cancer factor. Overexpression of H19 may contribute to the cell proliferative in many cancers and is associated with poor prognosis. Since the deregulation of H19 expression may occur from different mechanism, future studies should focus on the different functions of H19 in physiological and pathological processes and evaluate the potential of H19 as biomarkers and therapy targets under different situations.

2.2.4. Cancer-associated region long non-coding RNA 5 (CARLo-5)

LncRNAs cancer-associated region long non-coding RNA 5 (CARLo-5) is transcribed from the (−) strand of the 8q24.21 genomic region, where two other transcripts, sharing significant overlap in their sequences, colon cancer-associated transcript 1 (CCAT1) and CCAT1 long isoform (CCAT1-L) are also transcribed [66].

CARLo-5 was originally reported to be overexpressed in colorectal cancer patient tissues [77]. Later study revealed the overexpression of CARLo-5 in NSCLC and in some other cancers such as gastric cancer [78, 79]. In NSCLC patients, high CARLo-5 expression level was associated with advanced pathological stage and lymph node metastasis and was a significant predictor of shorter overall survival. An in vitro knockdown experiment showed a significant inhibition of proliferation in tumour cells, mainly due to the induction of the G0/G1 arrest [77, 78]. Furthermore, silencing CARLo-5 could result in the inhibitory effects in the cell invasion and migration, possibly by modulating the EMT process [78].

The regulatory mechanism of CARLo-5 may be related to the adjacent region of the cancer-associated variant rs6983267, as the region including rs6983267 has enhancer activity and can interact with the proto-oncogene MYC [80]. Evidences were provided in a colon cancer study that demonstrated a strong connection between the cancer-associated variant rs6983267 and the expression of CARLo-5. The chromosome conformation capture method revealed the MYC enhancer region could physically interact with the active regulatory region of the CARLo-5 promoter and enhanced the expression of CARLo-5, this finding suggested there was a long-range interaction of MYC enhancer with the CARLo-5 promoter [77]. Since CARLo-5 is proved to have an oncogenic function, further studies focus on elucidating the mechanism of CARLo-5 regulation may provide potential therapy target for cancer treatment.

2.2.5. Other functional LncRNAs in NSCLC

LncRNA colon cancer-associated transcript 2 (CCAT2) is expressed from a highly conserved MYC enhancer region within chromosome 8q24.21. It was initially reported to be involved in metastatic progress and chromosome instability in colorectal cancer [81]. Further studies reported its association with poor prognosis in various cancers including NSCLC [82]. CCAT2 was significantly overexpressed in NSCLC tissues, in particular, the overexpression of CCAT2 was associated with adenocarcinomas specifically but not with squamous cell carcinoma.
Silencing CCAT2 by siRNA led to the inhibition of proliferation and invasion in NSCLC cell lines in vitro, supporting the role of CCAT2 in the metastatic progression. Further analysis found that CCAT2 combined with CEA could predict lymph node metastasis. These findings implied the potential of CCAT2 as a specific ADC biomarker for lymph node metastasis.

Growth arrest-specific transcript 5 (GAS5) is expressed from human chromosome 1q25 [83]. GAS5 was found significantly downregulated in NSCLC tissues, which was correlated with advanced TNM stage and increased tumour size [84]. GAS5 could compete with the glucocorticoid response elements (GRE) on DNA by directly interacting with the DNA-binding domain of glucocorticoid receptor (GR), thus prevented the activation of glucocorticoid-responsive genes. This competition resulted in the reduction of cell growth and metabolism, while sensitizing cells to apoptosis [85]. Recent study demonstrated that downregulation of GAS5 was associated with cisplatin resistance in NSCLC. GAS5 could inhibit autophagy and therefore enhance cisplatin sensitivity in NSCLC cells [86]. Another study found that GAS5 overexpression was inversely correlated with EGFR pathway and the expression of IGF-1R proteins in human ADC cell line, indicating its role in reversing EGFR-TKIs resistance [87]. These findings indicate the tumour suppressor lncRNA GAS5 may represent a potential biomarker for diagnosis and therapy target for NSCLC intervention.

SRY-box containing gene 2 overlapping transcript (SOX2OT) locates in the chromosome region 3q26.33, and is transcribed form the same orientation of gene SOX2 [88]. SOX2OT was reported upregulated in NSCLC, along with the upregulation of SOX2, meanwhile, the expression level was significantly higher in lung SCCs than ADCs [89]. Further study found high SOX2OT expression predicted poor survival in lung cancer patients. In lung cancer cell lines, knocking down SOX2OT inhibited the cell proliferation. These finding suggest the oncogenic SOX2OT may be prognostic indicator for NSCLC [89].

BRAF-activated non-coding RNA (BANCR) is a 693 bps lncRNA located on (−) strand of chromosome 9q21, which is initially found as a tumour suppressor factor involved in melanoma cell migration [90]. BANCR expression level was reported to be significantly decreased in NSCLC tumour tissues samples, the reduction of BANCR was related to the larger tumour size, advanced TNM stage, metastasis development, and shorter overall survival. BANCR was also an independent poor prognostic predictor of poor survival for NSCLC. An investigation on the mechanisms of tissue-specific expression revealed that histone deacetylase might be involved in the repression of BANCR. Furthermore, upregulation of BANCR inhibited NSCLC cell viability, migration, and invasion, while promoting the apoptosis process. Reversely, knockdown of BANCR promoted migration and invasion of NSCLC cells in vitro. These inhibitory effects were reported to be associated with EMT [91]. Interestingly, although BANCR was downregulated in NSCLC, some studies reported the significant upregulation of BANCR in other cancers, which suggested a tissue-dependent regulation mechanism of BANCR [66].

Maternally expressed gene 3 (MEG3) is expressed in chromosome 14q32.3 with a full length of 1.6 kb nucleotides [92]. Alternative splicing process was found associated with the gene MEG3, which consisted of 10 exons and could generate multiple transcripts [93]. It was reported that the expression level of MEG3 in NSCLC tissues was significantly lowered than
normal tissues, which might due to the higher methylation rate of MEG3-DMR in NSCLC cells. Downregulation of MEG3 in NSCLC patients was associated with poor prognosis. In addition, overexpression of MEG3 by transfecting exogenous pCDNA-MEG3 into NSCLC cells inhibited cell proliferation and induced cell apoptosis in vitro, partially through activating the p53 signalling pathway [94].

3. Long noncoding RNA as novel NSCLC biomarkers

The high mortal rate of NSCLC may be mainly attributed to the late diagnosis and tumour metastasis. In addition, the heterogeneity of disease also increases the difficulty in the diagnosis and treatment, the molecular characters are different from each subtypes. Early detection, precise diagnosis, and treatment may increases the survival rate of NSCLC. To meet these ends, it is of great importance to identify novel NSCLC biomarkers.

As a new class of functional RNA molecules, IncRNAs are involved in a wide range of cellular and biological processes. Dysregulation of IncRNAs is associated with many cancer-related processes. In addition, the expression of IncRNA can be very tissue specific. These advantageous features imply a potential role of IncRNAs in cancer detection and treatment. Reduced BANCR expression was found to be an independent prognostic factor for NSCLC [91]. Huang et al. found small amount of IncRNAs (3.36%) in circulating vesicles [95]. Later research detected IncRNA HOTAIR, MALAT1, and H19 in the plasma of patients with gastric cancer and identified the expression level of plasma H19 was significantly higher than normal samples, furthermore, plasma H19 level was reduced in postoperative samples, which suggested H19 might be a biomarkers for gastric cancer [96]. Ren et al. identified fragments of IncRNA MALAT1 in plasma of prostate cancer (Pca) and named one of them as MALAT1-derived miniRNA (MD-miniRNA). Researchers then evaluated the diagnostic performance of MD-miniRNA in plasma samples of 192 patients. The results showed a sensitivity of 58.6% and specificity of 84.8% for discriminating PCa from non-PCa [97]. Although the functional IncRNAs mentioned above have been well-characterized, only few of them have been evaluated as biomarkers for diagnosis and prognosis in NSCLC, further validations is still need.

With the development of high-throughput technology, an increasing number of previously unidentified IncRNAs have been found. More and more researchers started to explore novel biomarkers from these unidentified IncRNAs. MiTranscriptome is a database, which derived from computational analysis of high-throughput RNA sequencing (RNA-Seq) data comprising 6500 samples spanning diverse cancer and tissue types. In database, 1128 ADC-related IncRNAs and 1309 lung SCC-related IncRNAs are identified, among these, 4 IncRNAs in ADC and 11 IncRNAs in lung SCC are predicted to be tissue specific, indicating that IncRNAs can discriminate not only between tumour and normal samples, but also between different subtypes [98]. Although most of these IncRNAs remain to be annotated and validated, the large number of cancer-related IncRNAs provides great hope for further screening of biomarkers and therapy targets. Some groups have investigated the potential of IncRNAs as biomarkers in early detection of NSCLCs. Wang et al. examined the expression of IncRNAs in three pairs of early stage ADC samples by high-throughput microarray technology and identified
<table>
<thead>
<tr>
<th>LncRNA</th>
<th>Regulation</th>
<th>Region</th>
<th>p-Value</th>
<th>AUC</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Type of sample</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BC034684</td>
<td>Up</td>
<td>Chr1:203,148,063–203,148,611</td>
<td>1.486E−06</td>
<td>0.719</td>
<td>79.4</td>
<td>60.3</td>
<td>Tissue</td>
<td>Wang et al. [99]</td>
</tr>
<tr>
<td>RP11-1008C21.2</td>
<td>Down</td>
<td>Chr15:38,363,827–38,364,884</td>
<td>1.193E−07</td>
<td>0.843</td>
<td>81</td>
<td>79.4</td>
<td>Tissue</td>
<td></td>
</tr>
<tr>
<td>AK094413</td>
<td>Down</td>
<td>Chr9:104,235,441–104,237,132</td>
<td>6.634E−08</td>
<td>0.821</td>
<td>85.2</td>
<td>62.4</td>
<td>Tissue</td>
<td></td>
</tr>
<tr>
<td>RP11-598F7.5</td>
<td>Down</td>
<td>Chr12:273,829–275,487</td>
<td>4.108E−11</td>
<td>0.882</td>
<td>79.4</td>
<td>84.1</td>
<td>Tissue</td>
<td></td>
</tr>
<tr>
<td>VNN2</td>
<td>Down</td>
<td>Chr6:133,008,008–133,079,022</td>
<td>1.063E−05</td>
<td>0.835</td>
<td>77.8</td>
<td>79.4</td>
<td>Tissue</td>
<td></td>
</tr>
<tr>
<td>Combination</td>
<td></td>
<td></td>
<td>0.987</td>
<td>0.92</td>
<td>98</td>
<td></td>
<td>Tissue</td>
<td></td>
</tr>
<tr>
<td>SPRY4-IT1</td>
<td>Up</td>
<td>Chr5:142,317,620–142,318,322</td>
<td>&lt;0.01</td>
<td>0.603</td>
<td>/</td>
<td>/</td>
<td>Plasma</td>
<td>Hu et al. [102]</td>
</tr>
<tr>
<td>ANRIL</td>
<td>Up</td>
<td>Chr9:21,994,791–22,120,646</td>
<td>&lt;0.001</td>
<td>0.798</td>
<td>/</td>
<td>/</td>
<td>Plasma</td>
<td></td>
</tr>
<tr>
<td>NEAT1</td>
<td>Up</td>
<td>Chr11:65,422,800–65,423,368</td>
<td>&lt;0.001</td>
<td>0.693</td>
<td>/</td>
<td>/</td>
<td>Plasma</td>
<td></td>
</tr>
<tr>
<td>Combination</td>
<td></td>
<td></td>
<td>0.876</td>
<td>0.88</td>
<td>81</td>
<td></td>
<td>Plasma</td>
<td></td>
</tr>
<tr>
<td>MALAT1</td>
<td>Up</td>
<td>Chr11:65,497,762–65,506,469</td>
<td>&lt;0.0001</td>
<td>0.79</td>
<td>56</td>
<td>96</td>
<td>Peripheral blood</td>
<td>Weber et al. [101]</td>
</tr>
</tbody>
</table>

Table 1. List of NSCLC-associated lncRNA biomarkers identified in different researches.
1170 differentially expressed lncRNAs (DE-lncRNAs) between early stage ADC tissues and their adjacent normal tissues. Further analysis identified 20 candidates of lncRNAs from 1170 DE-lncRNAs through a screening pipeline, the pipeline could be summarized briefly as follows: if an lncRNA’s average inter-group difference between tumour group and normal group was 10 times bigger than the inner group difference, it would be selected as a candidate. These 20 candidates were then validated by real-time quantitative PCR (qPCR) on a total of 102 pairs of early stage ADC samples. A panel of five lncRNAs (Table 1) was finally identified which can distinguish early stage adenocarcinoma samples from normal samples with high sensitivity (97%) and specificity (92%) [99]. Another study, which integrated two NSCLC microarray datasets comprising 165 and 90 patients, reported a list of 64 significantly deregulated lncRNAs in NSCLC tumours compared with normal lung tissues and a panel of 181 lncRNAs that were specific to histological subtypes of NSCLC [100].

An ideal biomarker should be of high sensitivity and specificity, and it should be easy to detect, better with non-invasive methods from body liquids. Weber et al. detected the expression level of MALAT1 in the cellular fraction of blood of a small NSCLC patients group (Table 1), they found that MALAT1 was detectable in peripheral human blood and the expression level between cancer patients and cancer-free controls was different, the sensitivity and specificity for discrimination was 56% and 96%, respectively [101]. Another study reported circulating SPRY4-IT1, ANRIL, and NEAT1 were significantly increased in plasma samples of NSCLC patients (Table 1). Combination with the three factors indicated a high power of discrimination (AUC, 0.876; sensitivity, 82.8%; and specificity, 92.3%) [102].

4. Conclusion

Since the researchers have identified that most of the genome is actively transcribed, while only small part of the human genome has the coding potential as protein-coding genes, the roles of non-coding RNAs have been transferred from transcriptional noises to the important functional molecules. This finding has led the classical view of the central dogma, which considers that the RNA functions mainly as an intermediate bridge between DNA sequences and protein synthesis, into a deeper understanding.

The roles of lncRNAs in the upstream of whole cellular signal system indicate that lncRNAs are closely associated with cellular differentiation, mitosis, and apoptosis. In the view of epigenetics, the functions of lncRNAs are mainly involved in three levels, including chromatin remodelling, transcriptional control, and post-transcriptional processes. LncRNA can act as transcriptional regulators and modulate the expression of protein-coding genes in cis or trans manner through directly or indirectly binding to DNA or protein molecules. Thus, the occurrences of diseases including cancers are always along with the dysregulation of lncRNAs. In this chapter, we have discussed several lncRNAs that have been reported to be associated with NSCLCs. Some of them are well-characterized and also identified in other cancers, while some still remain to be studied. Most of these lncRNAs have the potential as biomarkers in diagnosis and prognosis of NSCLCs. However, these lncRNAs lack NSCLC tissue specificity, thus great efforts are still needed to identify NSCLC tissue-specific lncRNAs.
Despite the high performance of these lncRNA panels in diagnosis, most of them are identified from statistical analysis, which means the biological meanings of lncRNAs have not been taken into consideration. In addition, the candidate screening methods are mainly based on p-value, fold change, absolute expression level, and PAM method, outcomes of the candidates may differ for different methods [103]. Thus, a better candidate screening method combining the biological meaning of lncRNAs and robust statistical pipeline is needed for future studies. Also, up to now, most of the samples in these studies are collected from patient tissues through invasive methods, more works are still needed to explore circulating lncRNA expressions in blood plasma, urine, or sputum, which can meet the non-invasive demands.

Considering the sophisticated functions and large number of lncRNAs, we have now identified just the tip of the lncRNA iceberg. Lots of questions are waiting to be clarified, for example, what does the classification of lncRNAs looks like, and what the mechanistic basis of their functions is. Huge gaps are still in front of us in understanding the big picture of the lncRNA world. Fortunately, new technology such as the third generation sequencing, which allows the longer read length, are now providing more reliable and accurate information of lncRNAs. In future, we believe that understanding the lncRNA world will bring us new answers to old questions in evolution, development, and the understanding of NSCLCs. There may be a long way before the clinical application of lncRNAs in NSCLC, however, fast progressing in the lncRNA filed opens up numerous opportunities for diagnosis and therapeutic intervention against NSCLC.

Author details

Zule Cheng1,2 and Hongju Mao*  
*Address all correspondence to: hjmao@mail.sim.ac.cn  
1 State Key Laboratory of Transducer Technology, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Science, Shanghai, China  
2 University of Chinese Academy of Sciences, Beijing, China

References


