Review Article
Mechanistic Roles of Noncoding RNAs in Lung Cancer Biology and Their Clinical Implications

Katey S. S. Enfield,1, 2 Larissa A. Pikor,1, 2 Victor D. Martinez,1, 3 and Wan L. Lam1, 2, 3

1 British Columbia Cancer Research Center, Vancouver, BC, Canada V5Z 1L3
2 Interdisciplinary Oncology Program, University of British Columbia, Vancouver, BC, Canada V5Z1L3
3 Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada V6T2B5

Correspondence should be addressed to Katey S. S. Enfield, kenfield@bccrc.ca

Received 16 September 2011; Accepted 8 March 2012

Academic Editor: Elfride De Baere

Copyright © 2012 Katey S. S. Enfield et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Lung cancer biology has traditionally focused on genomic and epigenomic deregulation of protein-coding genes to identify oncogenes and tumor suppressors diagnostic and therapeutic targets. Another important layer of cancer biology has emerged in the form of noncoding RNAs (ncRNAs), which are major regulators of key cellular processes such as proliferation, RNA splicing, gene regulation, and apoptosis. In the past decade, microRNAs (miRNAs) have moved to the forefront of ncRNA cancer research, while the role of long noncoding RNAs (lncRNAs) is emerging. Here we review the mechanisms by which miRNAs and lncRNAs are deregulated in lung cancer, the technologies that can be applied to detect such alterations, and the clinical potential of these RNA species. An improved comprehension of lung cancer biology will come through the understanding of the interplay between deregulation of non-coding RNAs, the protein-coding genes they regulate, and how these interactions influence cellular networks and signalling pathways.

1. Introduction

The human genome is comprised of less than 2% protein coding genes; however, more than 90% of the genome is transcribed, suggesting that the majority of the transcriptome is comprised of noncoding RNAs—transcripts that lack an open reading frame and as such do not encode a protein [1–4]. However this by no mean implies that ncRNAs lack function, but rather highlights the importance of looking beyond protein-coding genes in order to improve our knowledge of normal and disease biology. ncRNAs are loosely classified into two main categories: small non-coding RNAs (18–200 nucleotides), which includes transcripts such as miRNAs, transfer RNAs (tRNAs), small interfering RNAs (siRNAs), piwi-interacting RNAs (piRNAs) and some ribosomal RNAs, and long non-coding RNAs (lncRNAs) (200+ nucleotides), a family comprised of pseudogenes, antisense RNA and transcribed ultraconserved regions to name a few (Table 1) [4]. ncRNAs comprise a class of transcripts that until the last few decades was largely overlooked. While some are known to play important roles in the regulation of gene expression, splicing, epigenetic control, chromatin structure and nuclear transport, the function of most ncRNAs remains unknown [5, 6]. Of the species of ncRNAs identified to date, miRNAs, siRNAs, and piRNAs are the most thoroughly investigated. With roles in a number of cellular functions, it is not surprising that the deregulation of ncRNAs has been linked to human disease, including a number of cancers, such as breast, prostate, lung, colon, and liver. Increasing evidence that ncRNAs, beyond miRNAs, may be primary genetic regulators has led to the hypothesis that they may be ideal diagnostic markers and therapeutic targets [4].

Lung cancer is the leading cause of cancer deaths worldwide. The consistent poor 5-year survival rate of 15%, owing largely to the late stage of diagnosis and a lack of effective therapeutics, underscores the need for novel therapeutic modalities as well as early detection and prognostic markers [7–9]. While protein coding genes remain the primary focus of current genomic and proteomic studies, deregulation of ncRNAs has a demonstrated role in the regulation of gene
### Table 1: Classes of human non-coding RNAs.

<table>
<thead>
<tr>
<th>Type</th>
<th>Class</th>
<th>Characteristics and function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small ncRNA (&lt;200 nt)</td>
<td>Small Interfering RNAs (siRNAs)</td>
<td>21–22 nt double-stranded RNAs produced by Dicer and involved in gene silencing and viral defence</td>
<td>[163]</td>
</tr>
<tr>
<td></td>
<td>microRNAs (miRNAs)</td>
<td>18–25 nt RNAs that modulate gene expression posttranscriptionally</td>
<td>[163, 164]</td>
</tr>
<tr>
<td></td>
<td>Transfer RNAs (tRNA)</td>
<td>An adaptor molecule with an inverted L structure involved in translation of mRNA into protein</td>
<td>[165]</td>
</tr>
<tr>
<td></td>
<td>PIWI-interacting RNAs (piRNAs)</td>
<td>Dicer independent 26–31 nt RNAs located in the germline and adjacent somatic cells, involved in germline development and stability through the regulation of transposons</td>
<td>[163]</td>
</tr>
<tr>
<td></td>
<td>Small nucleolar RNAs (snoRNAs)</td>
<td>Guide molecules for modification and processing of rRNA, specifically site-specific methylation and pseudouridylation</td>
<td>[164]</td>
</tr>
<tr>
<td></td>
<td>microRNA-offset RNAs (moRNAs)</td>
<td>RNAs derived from the ends of pre-miRNAs, predominantly from the 5' end, independent of the mature miRNA. The function of moRNAs are currently unknown</td>
<td>[16]</td>
</tr>
<tr>
<td></td>
<td>Ribosomal 5.8S</td>
<td>Transcribed by pol I as a part of the 45S precursor, 5.8S is a component of the large ribosomal subunit in eukaryotes, and thus involved in protein translation</td>
<td>[166]</td>
</tr>
<tr>
<td></td>
<td>Promoter-associated short RNAs (PASRs)</td>
<td>Transcripts within a few hundred bases of protein coding or noncoding transcription start site that may regulate gene expression</td>
<td>[167]</td>
</tr>
</tbody>
</table>

| Long ncRNA (>200 nt) | Long ncRNA | A broad class of RNAs > 200 nt with functions in epigenetic regulation, splicing, and cellular localization | [40] |
| | Transcribed ultraconserved regions (T-UCR) | Non-coding sequences 100% conserved among humans, mice, and rats, with roles in the regulation of alternative splicing and gene expression, and altered in a number of human cancers | [168] |
| | Pseudogenes | Nonfunctional sequences of genomic DNA originally derived from functional genes but with mutations or premature stop codons that prevent their expression. Known to regulate gene expression and recombination | [161, 169] |
| | Promoter associated long RNAs (PARs) | Transcripts 250–500 nt long within a few hundred bases of protein coding or non coding transcription start sites that may regulate gene expression | [167] |
| | Antisense RNAs | Single stranded RNA complementary to a transcribed mRNA, capable of binding and blocking translation of its complementary mRNA, and promoting target decay | [170] |

Expression and warrants continued investigation. Clinically, ncRNAs are emerging as potential tools and targets in lung cancer. miRNA expression profiles have been associated with lung cancer prognosis, disease progression, survival, and outcome prediction as well as discrimination of subtypes [10–13]. In fact, a recent study deemed miRNA expression signatures superior to global mRNA expression profiles in the accurate classification of NSCLC subtypes [14]. While IncRNAs in lung cancer is still an emerging field, several have been shown to be involved in tumorigenesis, including *HOTAIR, H19, ANRIL*, and *MALAT1* [4, 15]. Other ncRNAs implicated in lung cancer include microRNA offset RNAs (moRNA) and *piRNA-651* [16, 17]. While new ncRNAs important to lung cancer continue to be discovered, miRNAs and IncRNAs constitute the majority of known cancer related non-coding transcripts. This paper therefore focuses on miRNAs and IncRNAs; their mechanisms of disruption, current technologies for detection and analysis, their role in lung cancer, and their impact on lung cancer diagnosis and treatment.

### 2. miRNAs and Their Role in Lung Cancer

miRNAs are small non-coding RNAs approximately 18–25 nucleotides in length that negatively regulate gene expression posttranscriptionally [18]. miRNAs have been shown to regulate a number of critical biological processes, including but not limited to, proliferation, apoptosis, metabolism, epithelial to mesenchymal transition, differentiation, and cellular development, acting as both oncogenes and tumor suppressors [14, 18]. miRNAs are transcribed by RNA polymerase II (pol II) into long, double-stranded stem-loop containing
primary (pri)-miRNAs, typically hundreds-to-thousands of nucleotides in length. The pri-miRNA is processed into a shorter double-stranded RNA of 70 nucleotides (pre-miRNA) by the endonuclease Drosha, exported to the cytoplasm via XPO5, and further processed to a length of 22 nucleotides (mature duplex) by the endonucleases and Dicer. Dissociation of the miRNA double strand duplex and incorporation of the mature strand into the RNA-induced silencing complex (RISC) guides RISC to the target mRNA, where the miRNA targets the 3’ UTR, or less frequently the 5’ UTR [19], of the mRNA based on sequence similarity. Translation of the mRNA is ultimately prevented either by transcript degradation, inhibition of translation, or mRNA decay, and depends on sequence complementarity between the miRNA and its mRNA target, the particular Ago protein in the RISC, and possibly the position and number of complementary nucleotides [20, 21]. Perfect complementarity leads to Ago2-mediated miRNA cleavage [22], while imperfect complementarity can lead either to transcript decay or translational inhibition via either Ago1, Ago3, or Ago4 [14, 23–25].

To date, over 1400 human miRNAs have been identified [26]. A single miRNA is capable of affecting multiple protein coding genes, while similarly a gene can be targeted by more than one miRNA. It is believed that over one-third of the genome is regulated by at least a single miRNA [24]. Frequently located at chromosomal breakpoints, fragile sites, regions of LOH or amplification, miRNAs are highly susceptible to genomic alterations and subsequently deregulated expression [27, 28]. Changes in miRNA expression have been detected in a variety of malignancies as well as preinvasive cancer and have been associated with clinical features such as prognosis and survival. As such, many miRNAs are currently under investigation as diagnostic and prognostic biomarkers, therapeutic targets, and as markers of disease subtypes [29].

The pathogenesis of lung cancer has been associated with the deregulation of several miRNAs (Table 2), altering cellular processes including angiogenesis, cell differentiation, invasion, and metastasis. Let-7, the first miRNA identified to be aberrantly expressed in lung cancer, targets KRAS and HMG2, resulting in suppression of proliferation, with reduced let-7 expression correlating with poor clinical outcome [12, 30]. Garofalo et al. showed that overexpression of miR-221 and -222 enhances cellular migration through activation of AKT, impairs TRAIL-dependent apoptosis by targeting PTEN and correlates with aggressive nonsmall cell lung cancer (NSCLC) [31]. Similarly, the downregulation of the miR-34 family leads to increased proliferation and inhibition of apoptosis through the p53 pathway and clinically correlates with a higher risk of relapse [32]. Studies in lung cancer cell lines have also revealed a number of important miRNAs, including mir-125a, -126, and -206, the overexpression of which have all been shown to be associated with invasive and metastatic potential [33, 34]. While miRNAs are now appreciated as key regulators of gene expression in lung cancer, capable of classifying histological subtypes, and predicting recurrence, progression, and prognosis, they are not the only class of ncRNA implicated in lung tumorigenesis [13, 35].

### 3. LncRNAs: Emerging Players in Lung Cancer

LncRNAs are largely polyadenylated RNAs greater than 200 nucleotides in length that regulate gene expression through epigenetic regulation, splicing, imprinting, transcriptional regulation and subcellular transport [5, 6, 36, 37]. Although originally regarded as transcriptional noise, IncRNAs function in both cis, such as antisense non-coding RNA in the INK4 locus (ANRIL) which complexes with Polycomb Repressive Complex 2 (PRC2) to act on the same chromosome [38], and trans, such as HOX antisense intergenic RNA (HOTAIR) whose association with PRC2 affects different chromosomes [39]. LncRNAs demonstrate developmental stage and tissue specificity, indicating their expression is tightly regulated [5, 40–44]. They are loosely categorized by their position relative to coding genes as intergenic, intragenic/intronic and antisense [40]. Current estimates of lncRNA content range from 7000–23,000 unique lncRNAs, with a growing cohort being validated as having a role in human disease processes [42, 45]. Included in this list are a number of human cancers, suggesting that aberrant expression of lncRNAs contributes to tumorigenesis, and highlighting the need to better understand the mechanisms through which these transcripts exert their function.

The first lncRNAs identified were the imprinted H19 gene and X-inactive-specific transcript (XIST) critical to X chromosome inactivation, although at the time of discovery they were not coined “lncRNAs.” Since then, lncRNAs have been associated with Alzheimers, Fragile X Syndrome, blepharophimosis syndrome (BPS), and cancer [46–48]. HOTAIR, located in the HOXC locus on 12q13.13, was one of the first lncRNAs to be described as having a fundamental role in cancer [49]. HOTAIR associates with PRC2, silencing a portion of the HOXD locus and inducing H3 lysine 27 trimethylation, subsequently remodelling the gene expression pattern of breast epithelial cells to more closely resemble embryonic fibroblasts [39, 49]. HOTAIR is overexpressed in roughly one quarter of human breast cancers and was found to be an independent prognostic marker of poor survival and metastasis as well as a driver of metastasis in murine models [39]. At this time, a handful of lncRNAs have been associated with lung cancer (Table 3), most significantly metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), a lncRNA that is associated with high metastatic potential and poor patient prognosis in NSCLC [15]. MALAT1 is upregulated in a variety of other human cancers including breast, colon, prostate, and liver cancer. Its exact mechanism of action is under investigation, although it is thought to regulate the level of phosphorylated splicing factors, thereby affecting alternative splicing in the cell [50].

### 4. Mechanisms of miRNA and LncRNA Disruption in Lung Cancer

The identification of genetic and epigenetic mechanisms of ncRNA deregulation provides insight into which ncRNAs are aberrantly expressed and therefore relevant to lung cancer biology. Similar to protein coding genes, ncRNAs can
<table>
<thead>
<tr>
<th>miRNA affected</th>
<th>Significance</th>
<th>Technology</th>
<th>Source tissue</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overexpression of miR-155, miR-21, and miR-106a. Decreased expression of let-7a</td>
<td>Prognostic biomarker of adenocarcinoma patient survival</td>
<td>Oligonucleotide microchip</td>
<td>Primary LC cases and corresponding noncancerous tissues</td>
<td>[12]</td>
</tr>
<tr>
<td>Overexpression of miR-21</td>
<td>Candidate for molecular targets in treatment for LC in never-smokers</td>
<td>miRNA microarray assay on a CodeLink platform (miRNA oligo probe)</td>
<td>Matched pairs of LC and noncancerous lung tissues from never-smokers</td>
<td>[134]</td>
</tr>
<tr>
<td>Decreased expression of let-7</td>
<td>Shortened postoperative survival in NSCLC</td>
<td>RT-PCR</td>
<td>Tumor specimens</td>
<td>[30]</td>
</tr>
<tr>
<td>Decreased expression of miRNA-451</td>
<td>Expression negatively associated with lymph node metastasis, the stage of TNM classification and poor prognosis of NSCLC patients</td>
<td>qPCR, confirmed by northern blot analysis</td>
<td>Fresh tissue of NSCLC samples and the adjacent histologically normal tissue.</td>
<td>[171]</td>
</tr>
<tr>
<td>Overexpression of miR-92-1</td>
<td>This might cause a decrease in the lung surfactant secretion, and loss of the protection of lung cells against external carcinogens</td>
<td>2D electrophoresis profiling and mass spectrometric analysis</td>
<td>SBC-3 cell line</td>
<td>[172]</td>
</tr>
<tr>
<td>Decreased expression of miR-30a</td>
<td>May function as a tumor suppressor, by targeting Snai1 and inhibiting migration, invasion, and metastasis</td>
<td>qPCR</td>
<td>A549 cell line and fresh snap-frozen surgical specimens of tumor tissues and of the corresponding normal specimens</td>
<td>[173]</td>
</tr>
<tr>
<td>Overexpression of miR-21</td>
<td>Overexpressed in tumor tissues relative to adjacent nontumor tissues. Negative regulation of PTEN and promotion of cellular growth and invasion in NSCLC cells</td>
<td>qRT-PCR</td>
<td>Paired NSCLC and adjacent non-tumor lung tissues</td>
<td>[174]</td>
</tr>
<tr>
<td>Overexpression of miR-126</td>
<td>Inhibition of tumor growth in vivo by targeting EGFL7</td>
<td>Flow cytometry assay, qRT-PCR, and Western blot</td>
<td>A549 cell line</td>
<td>[175]</td>
</tr>
<tr>
<td>Decreased expression of miR-133B</td>
<td>Increased apoptosis in response to gemcitabine and reduced MCL-1 and BCL2L2 expression</td>
<td>Quantitative-reverse transcriptase (q-RT) PCR profiling</td>
<td>Frozen lung tumors (adenocarcinoma) and noninvolved adjacent lung and LC cell lines</td>
<td>[78]</td>
</tr>
<tr>
<td>Overexpression and gain of miR-17-92</td>
<td>Enhanced cell proliferation</td>
<td>Northern blot confirmed with RT-PCR, followed by southern blot</td>
<td>SCLC tissue and cell lines</td>
<td>[55]</td>
</tr>
<tr>
<td>Loss of miR-1</td>
<td>Inhibition of cell proliferation and invasion in vitro, and tumor growth in vivo, by targeting MET and FoxP1</td>
<td>qRT-PCR</td>
<td>Lung cancer tissue and cell lines</td>
<td>[59]</td>
</tr>
<tr>
<td>Hypermethylation of miR-34a</td>
<td>Avoidance of senescence</td>
<td>Methylation-specific PCR followed by qRT-PCR</td>
<td>Lung cancer cell lines</td>
<td>[57]</td>
</tr>
<tr>
<td>Overexpression of miR-25 and miR-223</td>
<td>Biomarker of NSCLC found in sera of NSCLC patients, but not in that of healthy donors</td>
<td>Solexa sequencing</td>
<td>Sera from NSCLC compared to healthy donor controls</td>
<td>[127]</td>
</tr>
<tr>
<td>Overexpression of miR-21 and miR-210, decreased expression of miR-486-5p</td>
<td>Biomarkers of malignant nodules identified by CT</td>
<td>qRT-PCR</td>
<td>Plasma from patients with malignant nodules compared to those with benign nodules and healthy controls</td>
<td>[133]</td>
</tr>
</tbody>
</table>
be deregulated by multiple genetic and epigenetic mechanisms. Deregulation occurs both directly at the miRNA or lncRNA loci, and indirectly through disruption of processing components or alterations to target transcripts. As miRNA deregulation has been studied more comprehensively than lncRNAs, it will be discussed in greater detail.

### 4.1. Copy Number Alterations

miRNA loci are often located at regions of genomic instability and as such are highly susceptible to genonomic alterations [28]. Perhaps the most well-described example of direct alteration to a miRNA is the loss of let-7 [30, 51–54]. Acting as tumor suppressive miRNAs, the let-7 family members are located in chromosomal regions frequently deleted in lung and other cancers, including 3p, 9q, and 21p. A well-characterized example of oncogenic miRNAs is the miR-17-92 cluster, amplified in small cell lung cancer (SCLC) tumors and cell lines [55]. Expression of this cluster has been shown to be regulated by MYC, an oncogene frequently overexpressed in lung cancer [51]. Conversely, gene dosage alterations affecting lncRNA expression have yet to be reported.

### 4.2. Epigenetic Mechanisms

Epigenetic influences, including the effects of global hypomethylation and site-specific hypermethylation found in cancer genomes, have been investigated with reference to miRNAs. Diederichs and Haber showed that treatment of the A549 cell line with demethylating agent 5-azacytidine did not result in significantly altered miRNA expression by microarray analysis [56]. However, increasing evidence suggests that miRNAs are subject to epigenetic regulation. For example, the promoter region of miR-34a is known to be hypermethylated in lung and other cancer cell lines [57], while let-7a-3 has been found to be hypomethylated in lung adenocarcinoma compared to normal lung tissue [58]. Downregulation of miR-1, can be reversed by histone deacetylase activity, suggesting its and possibly other miRNAs deregulation occurs via histone modification [59]. Indirect alterations to DNA methylation patterns can also occur through miRNA targeting of DNA methyltransferases. For example, miR-29 a, b, and c were shown to directly target both DNMT3A and DNMT3B, two genes that are often overexpressed in lung cancer [60], resulting in aberrant DNA methylation.

### 4.3. Single-Nucleotide Polymorphisms

Unlike protein-coding genes, SNPs within the functional seed sequences of miRNAs are rare, occurring in <1% of miRNAs [61]. Exact sequence matches observed in specific regions of let-7 among different species demonstrate that miRNAs have evolutionarily conserved functions [62] and demonstrate the presence of negative selection against sequence variations. However, in lung cancer, SNPs have been identified within pri- and pre-miRNA sequences [63–67], within or near miRNA binding sites [68], and within genes encoding miRNA processing machinery [69–72]. A SNP located in pre-miRNA region of miR-196a2 (rs11614913 homozygous variant) was associated with significantly increased risk and poor survival among Chinese lung cancer patients [65]. Additionally, a G-to-T variant (rs3134615) in the 3′ UTR of L-MYC can inhibit the interaction between miR-1827 and the L-MYC target region, resulting in a constitutively higher expression level of L-MYC and an increased risk of SCLC in the Chinese population studied [73]. Furthermore, a SNP in the 3′ UTR of KRAS was able to alter its let-7-mediated regulation and was linked to increased risk of NSCLC among moderate smokers [74]. Interestingly, some of these SNPs seem to be lung cancer-specific. For example, while individuals carrying the CC genotype of the rs11614913 hsa-mir-196a2 variant have increased susceptibility to lung cancer, this relationship was not observed in hepatocellular carcinoma, gastric, or esophageal cancer patients [65, 74–76]. Finally, some SNPs relating to miRNA function can potentially influence the processing or target selection of miRNAs [77], adding another level of complexity to deregulation induced by the occurrence of SNPs in miRNAs sequences.

### 4.4. Deregulated Noncoding RNAs Identified at the Expression Level

The search for miRNAs and lncRNAs deregulated in lung cancer has frequently involved expression comparisons between cancer specimens and corresponding normal tissues [78], or between various clinico-pathological groupings, such as subtype and therapy response [79–82]. A recent NSCLC study revealed a 41-miRNA signature that could distinguish lung cancer tissues from noncancerous lung tissues, and a 6-miRNA signature that could differentiate the two major NSCLC subtypes: adenocarcinoma (AC) and squamous cell carcinoma (SqCC) [12]. Moreover, miRNA expression

<table>
<thead>
<tr>
<th>IncRNA</th>
<th>Significance</th>
<th>Technology</th>
<th>Source tissue</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overexpression of MALAT1</td>
<td>Predict metastasis and survival in early-stage NSCLC</td>
<td>Subtractive hybridization method, sequencing and quantitative RT-PCR</td>
<td>Shock frozen primary nonsmall cell lung tumors</td>
<td>[15]</td>
</tr>
<tr>
<td>Deregulated expression of BC200</td>
<td>Detectable at significant levels in tumors. Normal tissue from the same patient was found to be BC200-negative</td>
<td>In situ hybridization</td>
<td>Tumour and normal tissue frozen in liquid nitrogen</td>
<td>[87]</td>
</tr>
<tr>
<td>Overexpression of H19</td>
<td>Loss of imprinting in lung adenocarcinoma</td>
<td>RT-PCR</td>
<td>LC tissue and normal lung</td>
<td>[176]</td>
</tr>
</tbody>
</table>

### Table 3: Involvement of IncRNAs in lung cancer and technologies used for identification.

- MALAT1: Metastasis-associated lung adenocarcinoma Transcript 1
- BC200: None
- H19: Hypermethylated in brain

### 4.5. Summary

In conclusion, the dysregulation of lncRNAs and miRNAs plays a crucial role in the development and progression of lung cancer. The identification of these molecules and their targets provides new opportunities for diagnosis, prognosis, and therapeutic intervention.
studies have identified miRNAs associated with increased metastatic potential [83] and survival [12, 13, 30]. Aberrant ncRNA expression may also be driven by the deregulation of transcription factors. Like miR-17-92, expression of the non-coding RNA H19 has been shown to be induced by MYC in lung carcinomas [84]. Additional IncRNAs whose expression has been shown to be deregulated in lung carcinoma include H19, which undergoes loss of imprinting and overexpression [85], MALAT-1 [15], cancer up-regulated drug resistant (CUDR) [86], and BC200 [87].

4.5. Deregulation of miRNAs by Alterations to Processing Machinery. The general decrease in abundance of mature miRNAs is a common event in cancer and can be at least partially attributed to alterations in the miRNA processing machinery [88]. Described here are such examples that have been documented in lung cancer. Nuclear export of pre-miRNA requires XPO5, which occasionally suffers inactivating mutations resulting in a lower abundance of cytoplasmic miRNA [89]. Inactivating mutations have also been documented in TARBP2, encoding TRBP, a cofactor that functions in conjunction with Dicer to cleave pre-miRNA in the cytoplasm [90, 91]. Mutations in these loci appear to be mutually exclusive [91]. In fact, DICER, TARBP2, and XPO5 all represent haploinsufficient tumor suppressors in lung cancer [30, 92–94]. Loss of DICER has even been linked to shorter time to recurrence and poorly differentiated tumors [30, 95].

5. Technologies for the Identification and Quantification of Noncoding RNA

5.1. Expression Profiling. Several high-throughput genome wide approaches have been used to discover, establish targets, and predict functions of ncRNAs, including microarrays, serial analysis of gene expression (SAGE), next-generation sequencing, immunoprecipitation (IP)-based, and computational analyses. Tiling path arrays offer some advantages for discovery of ncRNAs over oligonucleotide-based arrays, as they are not dependent on current gene annotations; however, their resolution can be a limiting factor [96]. SAGE libraries, originally used to assess levels of protein-coding transcripts, can be queried to measure SAGE tag counts that map to IncRNA sequences [97, 98]. Gibb et al. queried SAGE libraries from both cancer and normal tissues, namely, breast, brain, and lung and identified 90 IncRNAs to be specifically deregulated in lung cancers compared to normal lung [98]. Next-generation sequencing technology has facilitated the identification of previously undescribed miRNAs and other small RNAs often missed by the limited depth of traditional sequencing methods [99, 100]. For example, the study by Meiri et al. identified seven novel lung tumor-specific miRNAs, one of which was the star strand of miR-663 [99]. In addition to identification of novel miRNAs, deep sequencing has also been proven to be a powerful tool for quantifying miRNA expression and defining variation within miRNA sequences. Small RNA sequencing offers a number of significant advantages over microarray platforms. Sequencing approaches are not limited by an a priori knowledge of the miRNAs being queried and are not constrained by lower (or upper) limits of detection. The ability to determine absolute expression values that can be compared within and across experiments represents another distinct advantage of sequencing data over the relative quantifications given by microarray measurements [96, 101]. Furthermore, sequencing approaches do not suffer from cross-hybridization artefacts observed in microarray experiments. In order to validate suspected target transcripts and identify enrichment of mRNA subsets potentially regulated by miRNAs, microarray expression analyses and/or sequencing methods are often complemented with IP-based approaches [96].

5.2. Computational Prediction. Several computational approaches, based on the conformational characteristics of ncRNAs, have been designed in an attempt to predict non-coding transcripts and their targets [96, 102, 103]. Secondary structure is one of the hallmarks for ncRNA prediction and/or identification software, and most algorithms are based on hairpin structures of precursor forms of miRNAs [104–106]. Different approaches have been developed in order to face the main challenges related with ncRNA research, including secondary structure prediction, comparison, and identification [107]. Computational methods were first focused on folding and prediction of RNA secondary structures, while assuming an RNA molecule is folded using minimum free energy. Some algorithms developed under this context are shown in Table 4. Secondary structure-based approaches can overlap some classes of ncRNAs. For example, IncRNAs can contain structural regions; however, they in general are not densely structured [108]. In this context, accuracy has been improved by comparative analysis based on structure preserving changes of base pairs. Among other methods, the search for ncRNAs can be carried out mainly through sequence and structure similarity, sequence-based alignments, and local searches [108]. Examples of tools using these approaches are also shown in Table 4.

Other computational methods for the identification of ncRNAs and interacting RNA molecules are based on the identification of short conserved motifs in the 3’ UTRS of protein-coding genes, as potential target sequences. Following this, queries are conducted searching for conserved sequences complementary to these motifs [109–111]. Programs based on these characteristics, such as “Pictar,” “TargetScan,” and “MiRanda,” have been used to predict miRNA functions and miRNA targets, resulting in the creation of a number of ncRNA databases [4, 70, 112–114].

5.3. Measuring Expression in Archival Tumor Materials. Formalin-fixed, paraffin-embedded (FFPE) samples are the most commonly available clinical specimens for histological and pathological analysis and represent a vast resource of samples for the identification of novel molecular markers as well as therapeutic targets [115, 116]. FFPE tissues have long been considered a challenge for nucleic acid analyses, specifically gene expression studies, as they contain cross-linked and fragmented nucleic acids, and RNA species are
Table 4: Examples of computational approaches used for ncRNA characterization.

<table>
<thead>
<tr>
<th>Method</th>
<th>Brief description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Secondary structure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MFOLD</td>
<td>Folding prediction using a thermodynamic model, returning a structure of minimal free energy (MFE)</td>
<td>[177]</td>
</tr>
<tr>
<td>RNAfold</td>
<td>Algorithm which finds optimal pseudoknotted RNA structures</td>
<td>[103]</td>
</tr>
<tr>
<td>PKNOTS</td>
<td>Finds the best RNA structure including the pseudoknot (based on MFE-model)</td>
<td>[178]</td>
</tr>
<tr>
<td>pknotsRG</td>
<td></td>
<td>[179, 180]</td>
</tr>
<tr>
<td><strong>Sequence similarity search</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INFERNAL</td>
<td>Generates consensus RNA secondary structure, then searches for homologous RNAs, or creates new sequence- and structure-based multiple sequence alignments.</td>
<td>[181]</td>
</tr>
<tr>
<td><strong>Sequence-based alignments</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNAz</td>
<td>Performs de novo searches for RNA structure</td>
<td>[182]</td>
</tr>
<tr>
<td>qRNA</td>
<td>Predicts structured RNAs from sequence alignments (only works on pair-wise alignments)</td>
<td>[183]</td>
</tr>
<tr>
<td>EvoFold</td>
<td>Functional RNA-structure identification in multiple sequence alignments</td>
<td>[184]</td>
</tr>
<tr>
<td>Dynalign</td>
<td>A free energy minimization algorithm for joint alignment and secondary structure prediction</td>
<td>[185]</td>
</tr>
<tr>
<td><strong>Local searches</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOLDALIGN</td>
<td>Alignment of RNA sequences and selection of subsets containing the most significant alignments.</td>
<td>[186]</td>
</tr>
<tr>
<td>CMfinder</td>
<td>Finds conserved RNA motifs in a set of unaligned sequences</td>
<td>[187]</td>
</tr>
</tbody>
</table>

often degraded [117]. However, the small size of miRNAs significantly reduces degradation by fixation, and numerous studies have shown miRNAs to be unaffected and well preserved in FFPE samples [118, 119]. Studies have indicated no significant difference in miRNA expression between matched FFPE and fresh frozen (FF) samples, suggesting that FFPE specimens are suitable for miRNA expression analyses [120–123]. Current methods for investigation of miRNA expression levels derived from FFPE specimens include quantitative real-time PCR [70, 116, 124], as well as microarray platforms by Agilent, Affymetrix, and Exiqon [116, 120, 121], all of which have shown comparable results between FFPE and FF tissue.

For IncRNAs, the use of FFPE tissue is not well documented. Due to their longer length, degradation and fragmentation associated with fixation may be an issue, much the same as for mRNA [117]. In an attempt to address these issues, Beck et al. developed a novel method of gene expression profiling termed 3′-end sequencing for expression quantification (3SEQ), which is applied to next-generation sequencing technologies [125]. Comparative analysis of a panel of FFPE and FF desmoid type fibromatosis and solitary fibrous tumor samples demonstrated that 3SEQ of FFPE samples outperformed microarray technologies and was comparable to 3SEQ results from corresponding FF samples. Although further work is still required, 3SEQ may be a useful method for the detection of IncRNAs in FFPE specimens.

6. Clinical Application of Non-Coding RNAs

Advancements in our understanding of the mechanisms driving aberrant ncRNA expression in lung cancer and other cancer types may yield significant clinical utility. Commonly overexpressed miRNAs and SNPs within miRNA sequences or target sequences could be exploited diagnostically, as biomarkers of disease. Therapeutically, ncRNAs silenced by methylation could potentially be restored with existing demethylating agents. MiRNAs are well preserved in FFPE tissues and circulate in bodily fluids with substantial stability. They can be used to accurately identify primary and metastatic cancer tissue origins, distinguish lung cancer subtypes, and predict outcome [126]. These qualities highlight their potential as both diagnostic and prognostic biomarkers in lung cancer. Studies examining miRNA levels in blood serum or plasma of patients with cancer appear promising, and there have already been several reports of miRNAs specific to the sera of lung cancer patients that are not detectable in disease-free individuals [127]. Typically, longer RNA species are not stable in blood; however, the short length of miRNAs makes them resistant to degradation and thus robust candidates for blood-based biomarkers [128].

6.1. Noncoding RNAs as Diagnostic Biomarkers. Several studies searching for miRNAs capable of separating individuals with lung cancer from those that are lung cancer-free have
been conducted. Yu et al. identified a four-miRNA signature (miR-21, miR-486, miR-375, and miR-200b) in sputum capable of distinguishing patients with lung adenocarcinoma from normal subjects with reportedly 80.6 sensitivity and 91.7% specificity [129]. Similarly, a five-miRNA signature (miR-210, miR-182, miR-486-5p, miR-30a, and miR-140-3p) was developed to distinguish squamous cell carcinomas from matched normals [130]. Efforts to improve early detection technologies are ongoing and frequently involve blood-based analyses of miRNA levels. Using Solexa sequencing, Chen et al. detected two overexpressed serum miRNAs (miR-25 and miR-223) that could be used as biomarkers for early detection of NSCLC [127], while another more recent study identified a 10-miRNA signature for the same purpose [131]. Foss et al. recently reported that miR-1254 and miR-574-5p were detected in the sera of patients with early-stage NSCLC compared to controls with a sensitivity and specificity of 82% and 77%, respectively [132]. Early detection has improved with the advent of CT technologies; however, the false positive rate is quite high. To address this issue, a recent study by Shen et al. identified that plasma miRNAs capable of distinguishing between lung cancer and benign nodules in CT-detected solitary pulmonary nodules [133].

Similarly, lung cancer subtypes can also be accurately defined by their characteristic miRNA expression profiles. For example, miR-205 is a highly specific marker for squamous cell lung carcinoma, capable of distinguishing squamous from nonsquamous NSCLC with high sensitivity and specificity [10]. Subtypes can be further subcategorized by miRNA expression patterns unique to specific genotypes. In lung adenocarcinoma, miR-155 is upregulated exclusively in tumors without KRAS or EGFR mutations. miR-25 and miR-21 are upregulated in EGFR mutation positive tumors, which are typically found in lung cancer never-smoker cases, while KRAS mutation positive tumors are associated with miR-495 up-regulation [29, 134]. To date, no lncRNAs have demonstrated diagnostic potential in lung cancer; however, diagnostic lncRNAs have been identified in other cancer types. The prostate-specific lncRNA DD3 is a highly specific marker of prostate cancer detectable in urine [135, 136], whereas the hepatocellular carcinoma-specific lncRNA highly up-regulated in liver cancer (HULC), is detectable in the blood of patients with liver cancer [137]. Collectively, these findings support the clinical potential of lncRNAs as diagnostic tools.

6.2. Non-Coding RNAs as Prognostic Biomarkers. To date, there are far more prognostic ncRNAs than those for diagnostic purposes, and in this paper, we will focus on only a few of the most well-established biomarkers. miR-21 is overexpressed in a number of human cancers, including lung cancer, where it has been shown to be an independent negative prognostic factor for overall survival as it stimulates growth and invasion through the inhibition of PTEN [11]. Recently, Saito et al. showed that increased miR-21 expression is associated with disease progression and survival in stage 1 lung cancer [138]. The Myc-activated miRNA cluster, miR-17-92, first identified as potential oncogenes in B-cell lymphoma, plays an important role in lung development, with high expression in embryonic development that declines throughout development into adulthood. Overexpression of this cluster is associated with the inactivation of RB and improved tumor vasculature through the inhibition of antiangiogenic thrombospondin-1 [139, 140]. Hu et al. derived a four miRNA signature that was significantly associated with overall survival of NSCLC patients; this signature derived from serum samples is a demonstration of miRNA stability in blood as well as their potential use as non-invasive biomarkers [141]. Additionally, miRNA signatures of recurrence free survival in stage I NSCLC patients were established for both NSCLC and AC patients, encompassing 34 and 27 miRNAs, respectively [142]. The scope of lncRNAs as prognostic markers is limited, but growing. MALAT1 is currently the only lncRNA with prognostic significance in lung cancer and is associated with high metastatic potential and poor patient prognosis in NSCLC [15].

6.3. Therapeutic Potential of ncRNAs. miRNAs are implicated in almost every process of lung tumorigenesis, from tumor development to metastasis and drug resistance, underscoring their therapeutic potential. Expression of let-7 inhibits growth of lung cancer cell lines and xenografts, reducing tumor burden, while lipid-based delivery systems of miR-34 have been shown to block tumor growth in murine models and downregulate the inhibitor of apoptosis protein, survivin, expression in lung metastases [143, 144]. More recently, miR-145 was found to inhibit cell proliferation through down-regulation of c-Myc in EGFR positive tumors [83], and miR-200c abrogated the capacity of metastatic lung adenocarcinoma to undergo epithelial to mesenchymal transition, invade and metastasize. This suggests that ectopic expression of miR-145 and -200 may be useful as novel therapeutic agents in lung cancer [145]. Expression of miR-29 displays an antivascular and anti-proliferative effect on lung cancer cells in vitro through the restoration of normal patterns of DNA methylation, supporting the notion that miRNAs may have an application as novel demethylating agents [60, 146].

miRNA expression is also known to play a significant role in drug sensitivity and resistance. Sensitivity to cisplatin has been linked with up-regulation of miR-181a, whereas resistance is conferred through the up-regulation of miR-630 [147]. Improved response and survival following gefitinib treatment has been correlated with loss of miR-122b [82], while overexpression of miR-137, -134, and let-7a has been shown to increase drug sensitivity for a number of anticancer drugs [148]. The clinical utility of lncRNAs as therapeutics has yet to be fully realized; however, overexpression of CUDR is associated with resistance to doxorubicin and apoptosis in SqCC cell lines A431 and A10A [86], indicating these transcripts may be useful clinically.

A number of current strategies to manipulate miRNA expression are currently being investigated and tested. These include but are not limited to antagonomers, miRNA sponges and small molecules to reduce miRNA expression, locked nucleic acids (LNAs), lipid-formulated mimics, and adenovirus vectors to over/reexpress down-regulated miRNAs [149, 150]. A number of hurdles remain before miRNAs can
Figure 1: Schematic depiction of ncRNA deregulation and its impact on regulatory proteins. (a) (i) normal levels of EZH2 are maintained by a balance of EZH2 transcription and miR-101 regulation. (ii) EZH2 is overexpressed as a result of copy number gain of the EZH2 locus. (iii) EZH2 is overexpressed as a result of miR-101 loss. (b) hypothetical scenario of crizotinib resistance. miR-628 is overexpressed resulting in increased degradation of putative target transcript, CASP3, required for crizotinib-induced cell death. (c) (i) normal levels of PTEN are maintained through the ability of its pseudogene, PTENP1, to act as a miRNA sponge. (ii) mutation in the 3′-UTR of PTENP1 results in loss of miRNA binding and redirection of the miRNA to degrade PTEN.

be widely established as therapeutic targets. For instance, the targeted delivery of RNA therapeutics to the site of interest, such as primary tumors or their metastases, is a major challenge. Many of the current strategies rely on the increased half-life and stability that chemical modifications offer molecules such as antagonirs and LNAs for systemic circulation and eventual uptake by the target tissue [151–153]. However, this strategy is associated with the possibility of negatively impacting healthy tissues. The ability to specifically deliver RNA therapeutics to the sites of interest would avoid this problem and has been demonstrated in a mouse model of lung metastatic melanoma [44]. Chi et al. developed a liposome-polycation-hyaluronic acid nanoparticle modified with a tumor targeting antibody to deliver the contained siRNA and miRNAs to the lung metastases, resulting in reduced tumor burden. In addition to the challenges of targeted delivery, the possibility of off-target effects is an equally relevant issue and stems from the ability of a single miRNA to target multiple mRNAs. This issue is not simplified by the use of target prediction algorithms, which are far from perfect. While the use of multiple target prediction algorithms improves sensitivity, false positives remain a significant problem, requiring filtering and experimental validation. New methods such as high-throughput sequencing of RNAs isolated by crosslinking immunoprecipitation (HITS-CLIP) are emerging to identify bonafide RISC-associated miRNA-mRNA interactions [154]. Despite these many complicating factors, miRNAs have begun to enter the clinical setting. Preclinical trials are currently underway examining the effectiveness of let-7 reintroduction into murine NSCLC models, while miR-122 antagonistic technologies are in Phase II clinical trials for the treatment of Hepatitis C [155].

7. Impact of ncRNA Deregulation on Biological Networks

Significant advances in our understanding of lung cancer biology will result from an improved understanding of the interplay between the deregulation of protein-coding genes and ncRNAs. Such an example is well illustrated by the relationship between miR-101 and EZH2 (Figure 1(a)). miR-101 is frequently deleted, with a higher preponderance of loss in NSCLC as compared to SCLC [156], while EZH2,
miRNAs (Figure 1(c)). Mutations to the miRNAs that target miRNAs, thus UTR sequences results in common targeting by the same transcriptional repression of target genes such as CDKN2A. EZH2 is a target of mir-101 [157], with loss of mir-101 resulting in derepression of EZH2 expression, and providing yet another mechanism of EZH2 activation. Importantly, this relationship informs us of mechanisms of oncogene activation beyond copy number gain, mutation, and hypomethylation. This provides evidence that identification of novel cancer-related genes may result from examination of ncRNA deregulation, which may have been missed by studying conventional means of disruption.

As miRNAs are major regulators of gene signalling pathways, it is not surprising that they have been shown to be deregulated in association with specific drug response phenotypes. A multidimensional analysis of a panel of lung cancer cell lines examining miRNA copy number, expression and mRNA expression found gain and overexpression of mir-628 to be associated with crizotinib drug resistance [80] (Figure 1(b)). Crizotinib-induced cell death occurs through activation of the caspase-3 pathway, and interestingly, one of the predicted targets of mir-628 was CASP3 [158]. These data suggest that mir-628 may play a critical role in crizotinib resistance through the repression of a key effector molecule required for drug function. The addition of miRNAs, and eventually lncRNAs, to gene networks will further our understanding of the biological mechanisms governing drug response, and potentially influence treatment choices or even identify novel therapeutic targets.

These two examples represent the classical view of miRNA repression of protein-coding transcripts. However, a growing area of research pertains to cellular interaction networks, namely, the interaction between miRNAs and lncRNAs. LncRNAs can act as natural “miRNA sponges” reducing levels of free miRNAs and relieving inhibition of other target transcripts [159]. One such example is HULC, an lncRNA that contains a cAMP response element binding protein (CREB) binding site in its promoter. HULC creates an autoregulatory loop by acting as a molecular decoy that sequesters miR-372 and prevents repression of PRKACB, a kinase that targets CREB. Through this mechanism HULC facilitates CREB induction, which subsequently leads to increased HULC transcription [160]. Another example of lncRNA-miRNA regulation was identified in prostate cancer. The 3′ UTR of the tumor suppressor PTEN, frequently lost in prostate and other cancer types, shares sequence similarity with the lncRNA PTENP1. The homology in 3′ UTR sequences results in common targeting by the same miRNAs, thus PTENP1 can act as a molecular sponge for miRNAs that target PTEN, limiting PTEN repression by miRNAs (Figure 1(c)). Mutations to the PTENP1 3′ UTR disrupt miRNA binding eliminating the protective ability of the transcript and leading to repression of PTEN and promotion of tumor growth [161, 162]. This is a novel finding, with no similar examples described in lung cancer. Further study into miRNA-lncRNA relationships could lead to the identification of novel miRNA, lncRNA, and protein-coding gene signalling triads.

8. Conclusions

ncRNAs play a role in nearly every biological process and therefore have the potential to serve as diagnostic and prognostic biomarkers as well as therapeutic targets. Findings from recent studies strongly support this notion, with ncRNAs being implicated in survival, prognosis, and drug response while also being capable of discerning cancerous from benign lesions and discriminating between subtypes of lung cancer. It is evident that a comprehensive understanding of tumor biology must therefore include both coding and non-coding transcripts. Only through the inclusion of these transcripts in molecular studies it will be possible to better understand tumor biology and human disease.

Authors’ Contribution

K. S. Enfield and L. A. Pikor contributed equally to this paper.

Acknowledgments

This work was supported by funds from the Canadian Institutes for Health Research (MOP 86731, MOP 94867, and MOP-110949), Canadian Cancer Society (CCS20485), U.S. Department of Defense (CDMRP W81XWH-10-1-0634), NCI Early Detection Research Network and the Canary Foundation. L. A. Pikor was supported by Vanier Canada Graduate Scholarship. The authors would like to thank Ewan Gibb and Emily Vucic for insightful comments regarding the paper.

References


P. Han, T. Sato, N. Sugito et al., “Counterbalance between RB inactivation and miR-17-92 overexpression in reactive


