



BASIC SCIENCE FOR THE CHEST PHYSICIAN

Non-coding RNA as lung disease biomarkers

Sebastian F Vencken, Catherine M Greene, Paul J McKiernan

Respiratory Research Division,
Department of Medicine, Royal
College of Surgeons in Ireland,
Education and Research
Centre, Beaumont Hospital,
Dublin, Ireland

Correspondence to

Professor Catherine M Greene,
Respiratory Research Division,
Department of Medicine, Royal
College of Surgeons in Ireland,
Education and Research
Centre, Beaumont Hospital,
Dublin 9, Ireland;
cmgreene@rcsi.ie

Received 19 August 2014
Revised 19 November 2014
Accepted 1 December 2014

ABSTRACT

Biomarkers are quantifiable indicators of disease. These surrogates should be specific, sensitive, predictive, robust and easily accessible. A major class of RNA described as non-coding RNA fulfils many of these criteria, and recent studies have demonstrated that the two major subclasses of non-coding RNA, long non-coding RNA and, in particular, microRNA are promising potential biomarkers. The ability to detect non-coding RNAs in biofluids has highlighted their usefulness as non-invasive markers of lung disease. Because expression of specific non-coding RNAs is altered in many lung diseases and their levels in the circulation often reflect the changes in expression of their lung-specific counterparts, exploiting these biomolecules as diagnostic tools seems an obvious goal. New technology is driving developments in this area and there has been significant recent progress with respect to lung cancer diagnostics. The non-coding RNA biomarker field represents a clear example of modern-day bench-to-bedside research.

One particularly surprising finding that has arisen from sequencing the human genome is that the number of protein-coding genes is considerably lower than expected. Although not precisely known, <2% of the genome actually encodes proteins. In light of this fact, the expression of non-coding RNA (ncRNA) is being examined with increased importance. ncRNA is a super-class of endogenous, non-protein coding RNA transcripts (figure 1A), many of which have essential and very diverse biological functions in almost all cell processes. Based on their size, sequence and function, ncRNA can be classified into various discrete subclasses, the two most notable of which are microRNAs (miRNA) and long non-coding RNAs (lncRNA). Both species of ncRNA have vital roles in regulation of the genome, from transcription into messenger RNA (mRNA), to mRNA processing, to translation into protein.

MicroRNAs are endogenous 18–25 nucleotide RNAs whose dominant function is to inhibit protein production (figure 1B). These bind to target mRNAs in a sequence-specific manner; this leads to cleavage and degradation of the target mRNA and/or inhibition of translation of the target mRNA into protein. A single miRNA can regulate multiple genes, and most miRNAs have the ability to act as master regulators because each tends to target sets of genes that participate in molecular networks. Occasionally, miRNAs can actually activate translation or directly regulate transcription. The more recently classified lncRNA group is loosely categorised by its length, which is any ncRNA over 200 nucleotides in length that is

non-coding. lncRNAs have broad and wide-ranging regulatory functions in chromosomal configuration, RNA transcription and translation (figure 1B)—this is possible due to their ability to bind to DNA, RNA or protein. Collectively, miRNA and lncRNA are believed to play roles in regulation at almost every level of gene expression and are crucial regulators of cell and tissue development and homeostasis.

In the respiratory system, many ncRNAs responsible for normal development and maintenance can become disease contributors if expressed aberrantly. This has important implications for major lung diseases, such as lung cancer, COPD, pulmonary fibrosis, asthma and cystic fibrosis. A good example is the malexpression of miR-21, the most commonly expressed miRNA in the human lung, which at elevated levels can contribute to the symptomatic development of asthma, idiopathic pulmonary fibrosis and non-small-cell lung cancer (NSCLC) through its effects on inflammatory and cancer-suppressing genes. Among lncRNAs, the overexpression of MALAT-1 has a malignant impact on lung adenocarcinoma by regulating genes involved in metastasis. The disease-specific expression profiles of miR-21, MALAT-1 and many other ncRNAs can be detected in tissue and biofluids and can be used as disease biomarkers.

MIRNA BIOMARKERS AND LUNG CANCER

miRNA expression profiles in lung tumours are different compared with non-malignant tissue. Most research has focused on the pathological reasons and consequences for these aberrant expression profiles, however, much effort has been made to exploit differential miRNA expression in lung cancer for diagnostic or prognostic purposes. This has resulted in the discovery of various miRNA expression ‘signatures’ that can identify specific types of lung cancer in lung biopsies.

An ideal biomarker should, among other properties, be measurable with minimal invasiveness, risk and discomfort to the patient. Thus, screening biofluids, such as blood, sputum or urine is favoured over more invasive techniques required to access lung biopsies, bronchial brushings or bronchoalveolar lavage fluid. Interestingly, miRNAs have been detected in all human biofluids, despite the hostile conditions of these environments which generally promote RNA degradation. Their relative stability in biofluids has given miRNAs a high potential as non-invasive biomarkers, particularly as an alternative to tumour biopsies.

So far, more than 40 studies have been conducted to identify circulating miRNA profiles as

To cite: Vencken SF,
Greene CM, McKiernan PJ.
Thorax Published Online
First: [please include Day
Month Year] doi:10.1136/
thoraxjnl-2014-206193

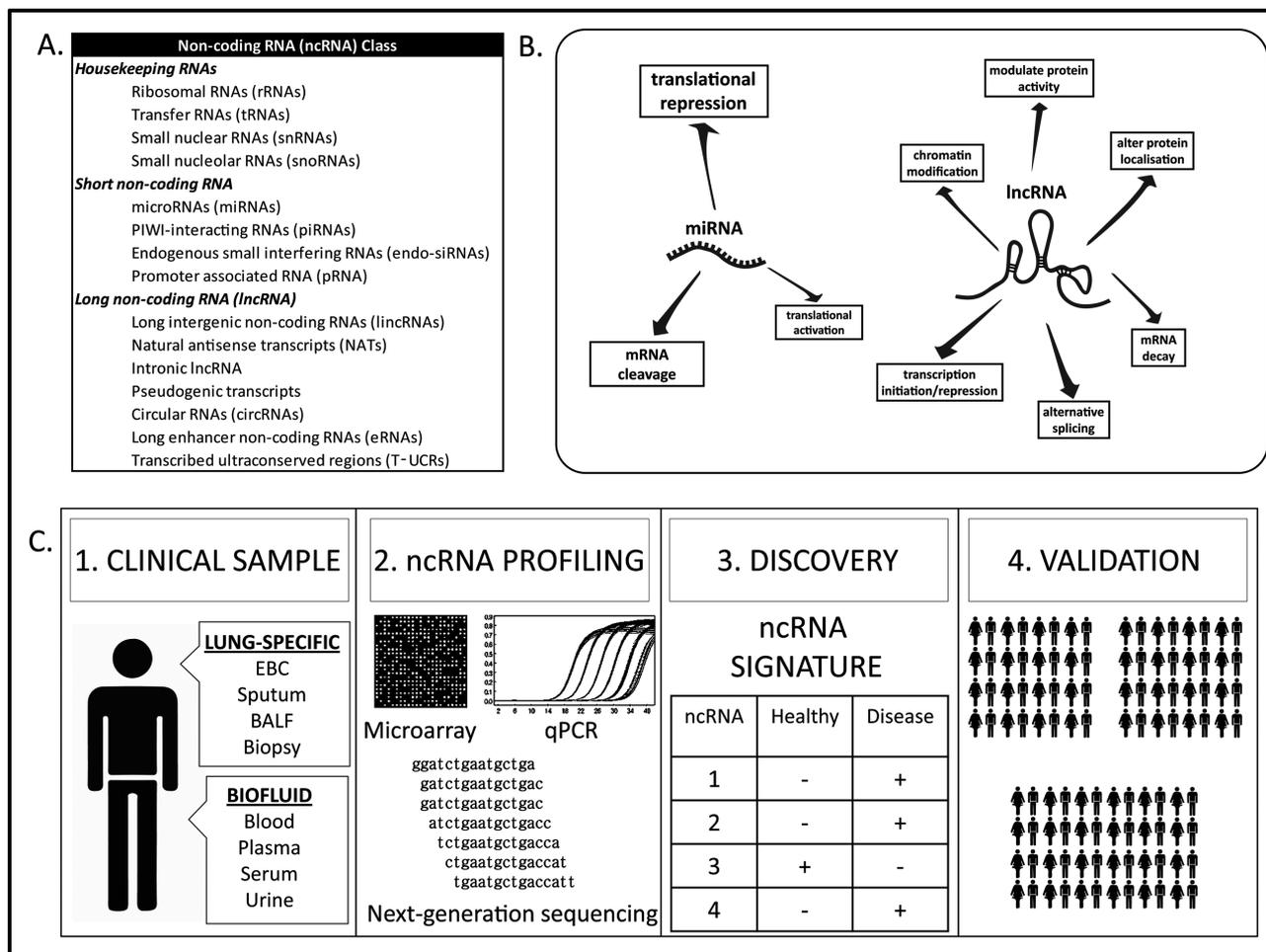


Figure 1 Non-coding RNA (ncRNA) classes, functions and biomarker potential. (A) Classes of ncRNA, (B) function of miRNA and lncRNA and (C) workflow for ncRNA biomarker discovery and validation. (1) Clinical samples are (2) profiled for ncRNA expression by various methods. (3) A clear ncRNA signature is identified in disease versus healthy samples and (4) validated in large independent patient and healthy control cohorts. BALF, bronchoalveolar lavage fluid; EBC, exhaled breath condensate; qPCR, quantitative PCR.

diagnostic markers for lung cancer, with some showing specificities and sensitivities of >90% for differential profiles involving 10–40 miRNAs.¹ Some of the most prominently featured miRNAs in these profiles are miR-21, miR-210 and miR-155, all of which also have biological functions in lung cancer.

ncRNA biomarker discovery is generally performed with broad-range profiling methods, such as microarrays, RT-PCR and next-generation sequencing that can measure the differential expression of up to thousands of ncRNAs (figure 1C). With the advent of diagnostic ncRNA profiling, disease-specific assay panels are now being designed that measure only the differential expression of miRNAs that are clinically relevant to the disease in question. This reduces both the cost and complexity of the assay.

A number of companies have started to develop these clinical diagnostic test panels to classify various lung cancer types. Recently, a lung cancer test has been developed which measures the expression of eight miRNAs in lung biopsies. This test has the ability to differentiate between small-cell lung cancer and NSCLC types with very high sensitivity and specificity. Elsewhere, plans are currently afoot to introduce a diagnostic circulating miRNA assay that can detect different lung cancers with 90% sensitivity and 80% specificity. Besides having the advantage of being a non-invasive method of lung cancer diagnosis, nucleic acid assays, such as miRNA assays, enjoy

favourable Food and Drug Administration (FDA) regulatory guidelines for multiplexing compared with protein assays. This could make miRNA profiling more accurate and far more cost effective than protein assays.

LNCRNA BIOMARKERS

Although the expression of many lncRNAs has been evaluated in cancer and other respiratory diseases, their validation as diagnostic or prognostic biomarkers is at an earlier stage than that of their miRNA counterparts. Progress in non-lung cancers has been quicker in terms of the lncRNA biomarker discovery pipeline; the most advanced development being an FDA-approved PCA3 assay which detects levels of the lncRNA PCA3 (*prostate cancer antigen 3*) in urine as a biomarker for prostate cancer.

It is promising to note that lncRNAs have been easily isolated from bronchial brushings, biopsies and sputum. One caveat with respect to lncRNAs is that only a few species have been detected in biofluids. Thus, more invasive sampling methods may be required, which is not to say that a proportion that exist in biofluids may turn out not to be very useful circulating biomarkers due to their stability and ease of quantification. Good progress has been made in the discovery of potential lung cancer lncRNA biomarkers. At the time of writing, MALAT-1 (metastasis-associated lung adenocarcinoma transcript 1) appears to be one of the most promising candidates, especially as a biomarker of NSCLC² in

tissue specimens. *MALAT-1* is highly expressed in NSCLC and is a predictive marker for metastasis development and poor prognosis in this cancer, especially that arising from squamous cell carcinoma. Other promising lncRNAs whose expression is highly altered in lung cancer include *CCAT2* and *HOTAIR*.

OTHER LUNG DISEASES

The development of ncRNA biomarker signatures for other respiratory diseases has not progressed at the same pace compared with that of lung cancer. However, several studies are setting the groundwork for this. Specific circulatory miRNA signatures have been identified and validated in small sample sizes for chronic thromboembolic pulmonary hypertension, COPD, asthma, lung fibrosis and several infectious respiratory diseases.^{3–5} For some of these, differential miRNA profiles are evident in bronchoalveolar lavage fluid and exhaled breath condensates (EBC). Whether lncRNAs are detectable in EBCs has not yet been reported.

PERSPECTIVE

Clearly, more studies are required for the robust identification and validation of diagnostic, prognostic and therapeutic ncRNA biomarkers. Current pharmaceutical developments are focused on lowering the cost and increasing the accuracy of ncRNA bioassays with new technologies such as single-molecule detection assays. Increased standardisation within the field of ncRNA

biomarkers, such as finding reliable data normalisation factors, will also further enhance accuracy and reproducibility of results. The next 5 years are likely to produce a flurry of studies in this area.

Acknowledgements The authors thank Dr Kevin Molloy for constructive feedback regarding the manuscript.

Contributors SV, CMG and PM planned the content and drafted the article.

Funding Funding for research in CMG's group from the Health Research Board in Ireland (PHD/2007/11), the National Children's Research Centre (C/13/1) and the European Respiratory Society/GlaxoSmithKline Award for Rare Pulmonary Disease 2013 is gratefully acknowledged.

Competing interests None.

Provenance and peer review Not commissioned; internally peer reviewed.

REFERENCES

- 1 Ullivi P, Zoli W. miRNAs as non-invasive biomarkers for lung cancer diagnosis. *Molecules* 2014;19:8220–37.
- 2 Ji P, Diederichs S, Wang W, *et al.* MALAT-1, a novel noncodingRNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* 2003;22:8031–41.
- 3 Banerjee A, Luettich K. MicroRNAs as potential biomarkers of smoking-related diseases. *Biomark Med* 2012;6:671–84.
- 4 Kishore A, Borucka J, Petrakova J, *et al.* Novel insights into miRNA in lung and heart inflammatory diseases. *Mediators Inflamm* 2014;1:1–27.
- 5 Guo L, Yang Y, Liu J, *et al.* Differentially expressed plasma microRNAs and the potential regulatory function of Let-7b in chronic thromboembolic pulmonary hypertension. *PLoS ONE* 2014;9:e101055.