Molecular biomarkers of lung carcinoma

Bozica Vrabec Branica¹, Srecko Gajovic²

¹Department of Pathology and Cytology, University Hospital of Zagreb, Jordanovac 104, 10000 Zagreb, Croatia, ²University of Zagreb School of Medicine, Croatian Institute for Brain Research, Salata 12, 10000 Zagreb, Croatia

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Types of lung carcinoma molecular biomarkers
   3.1. DNA-based lung carcinoma molecular biomarkers
      3.1.1. Genetic biomarkers
      3.1.2. Epigenetic biomarkers
      3.1.3. Chromosomal changes
   3.2. RNA-based lung carcinoma biomarkers
   3.3. Protein-based lung carcinoma biomarkers
   3.4. Metabolomics
4. Applications of lung carcinoma molecular biomarkers
   4.1. Lung carcinoma molecular biomarkers in detection of precancerous lesions and early carcinomas
   4.2. Lung carcinoma molecular biomarkers as potential therapeutic targets
5. Lung carcinoma molecular biomarkers in various specimens
6. Various technologies in lung carcinoma molecular biomarkers
7. Our experience in PCR applications for detection of lung carcinoma biomarkers in cytology specimens
8. Perspective
9. Acknowledgements
10. References

1. ABSTRACT

Lung carcinoma is still the leading cause of cancer deaths in men and women. There is a constantly increasing need to find molecular biomarkers for lung cancer which can be used for risk stratification, early detection, treatment selection, prognosis and monitoring for recurrence. Recent advances in imaging and improved bronchoscopic techniques have intensified interest in lung carcinoma screening techniques, especially in new molecular markers which can help cytopathologists to make a definitive diagnosis on very small specimens in non-invasive, non-expensive, simple and efficient manner. Several decades of intensive research have originated numerous potential lung carcinoma molecular biomarkers but only few turned out to be useful in clinic. The review describes types of biomarkers, sources and techniques for their identification in cancer diagnosis and therapy. A deep understanding of each biomarker will be a key to efficiently diagnose lung carcinomas and direct patients toward beneficial drugs based on individual patient profile.

2. INTRODUCTION

The completion of human genome project expanded our knowledge of the pathogenesis of lung cancer at the molecular level while the recent advances in imaging and bronchoscopic techniques stirred intense interest for lung carcinoma molecular biomarkers (LCMB) in order to detect and recognize very small lung lesions. The cytopathologists find themselves being pressured to make a definitive diagnosis on very small specimens and to investigate different molecular markers which would serve as predictors of response to therapy in lung carcinomas.

3. TYPES OF LUNG CARCINOMA MOLECULAR BIOMARKERS

3.1. DNA-based lung carcinoma molecular biomarkers

DNA-based molecular biomarkers include detection of genetic, epigenetic and chromosomal changes as the main targets for biomarker discovery in tumor cells.
### 3.1. Genetic biomarkers

Detection of genetic alterations in the multistep process of carcinogenesis includes detection of mutations which lead to activation of oncogenes or inactivation of tumor-suppressor genes and subsequent uncontrolled cell growth development of cancerous cells (1). The genetic mutations most commonly studied as possible lung cancer biomarkers are located in K-ras, p53, Rb and myc gene. K-ras gene mediates signals from growth factor receptors and other signaling inputs. The mutations convert normal K-ras proto-oncogene to an oncogene, causing the oncoprotein to become overactive in transmitting growth factor-initiated signals (2). Mutations occur most often in exon 12. They are detected in 30% of lung adenocarcinoma in smokers and they are correlated with a poor prognosis (3-5). Both Rb and p53 have critical roles in the control of cell cycle, and their abnormalities are the most frequently associated with cell cycle dysregulation in malignancy (6). The well-known „guardian of the genome”, p53 induces apoptosis and reduces phosphorylation of retinoblastoma (Rb) protein, which leads to cell cycle arrest at the G1-S check point. When this tumor-suppressor gene is deleted or inactivated, cells cannot control DNA repair and undergo aberrant growth. Mutations of p53 are present in 50% of non-small cell lung carcinoma (NSCLC) and 75% of small cell lung carcinoma (SCLC) (7-11). Rb is inactivated in more than 80% of high-grade neuroendocrine lung tumors, especially SCLC (1). Myc is located on chromosome 8 and has transcriptional regulation activity. Gene amplification analysis showed it is overexpressed in 10-40% of SCLC.

The loss of one allele or loss of heterozygosity (LOH) is another predisposing factor to lung cancer (12,13) causing the inactivation or silencing of genes critical for growth regulation and homeostasis. More than 90% of small cell carcinomas and more than 70% of NSCLCs demonstrate loss of heterozygosity (14-16). In NSCLC loss of heterozygosity generally involves genetic loci on chromosome arms 1p, 3p, 5p, 9p, 13q, 17p, 19p, Xp and Xq. In small cell carcinomas, loss of heterozygosity generally involves chromosome arms 3p, 4q, 5q, 10q, 13q, 15q and 17p (14-19).

Telomerase activation represents a potential candidate for lung cancer molecular biomarker (20). Telomere shortening is the earliest and most prevalent genetic change identified in the precursor lesion (21). Telomere attrition is thought to predispose to chromosome fusion and the resulting translocations, followed by their missegregation during mitosis. Later during tumorigenesis, telomerase activity is resumed, compensating the telomere erosive processes while still permitting continued chromosomal instability. The ability to image and quantitatively assess telomeres makes these alterations a potential diagnostic tool, particularly to characterize early or preinvasive lesions (22). Telomerase activity has been detected in samples of abnormal bronchial epithelium: hyperplasia (in 71%), metaplasia (in 80%), in situ carcinoma (in 100%). High levels of telomerase activity are seen in nearly all small cell lung carcinoma and in 80% to 85% of all non-small cell lung carcinoma (20).

### 3.1.2. Epigenetic biomarkers

Epigenetic changes like hypermethylation of CpG (cytosine phosphate guanosine rich regions) islands is a common mechanism by which tumor suppressor genes are inactivated in most tumors including lung carcinomas (23-30). Hypermethylation of gene promoter regions is associated with silencing of transcription and it affects genes involved in various aspects of normal cell function. Therefore it is a critical trigger for malignant transformation and progression. Aberrant methylation of p16/NK4a, APC, TMS1, RASSF1, and DAPK has been reported in lung carcinoma (31-36).

### 3.1.3 Chromosomal changes

Inactivation of tumor suppressor genes can be induced by chromosomal changes like chromosomal deletions or rearrangements accidently happening during cell division. The most frequently abnormality is deletion of the short arm of chromosome 3 (3p) where several tumor suppressor genes are present (37-40).

### 3.2. RNA-based lung carcinoma molecular biomarkers

Compared to DNA biomarkers, which are mostly evaluated individually, most RNA-based biomarkers undergoing clinical evaluation are screened as multi-gene molecular patterns or „fingerprints“. This could be more accurate but selection of genes to include in the pattern needs new developments in biostatistics, bioinformatics and data visualization (41).

Non-coding RNAs (ncRNAs) are functional transcripts that do not code for proteins but play a major role in regulating almost every level of gene expression (42). They play an important role in tumorigenesis and could represent biomarkers for carcinoma (43). Except abundant and functionally important transfer of ribosomal RNAs, ncRNAs include other RNAs such as small nucleolar RNAs ( snoRNAs), micro RNAs (miRNAs), short interfering RNAs (siRNAs), piwi-associated RNAs, small Cajal body-specific RNAs ( scaRNAs), small nuclear RNAs (snRNAs) and long ncRNAs that are still partially understood. From the small ncRNAs, miRNAs and siRNAs have been extensively been studied in carcinogenesis (42,44-46). The expression of specific populations of miRNAs in a tissue- and time-dependent manner correlates with clinical characteristics for several cancer types including lung cancer (47-49). miRNAs expression profiles can be used to classify human cancers and to suggest a correlation between disease prognosis and therapeutic outcome. Investigation of metastasis-associated miRNA markers is expanding and these markers were referred to as metastamirs (50). miRNA can act as a tumor suppressor as well as an oncogene (51). For example, let-7 is a suppressor for RAS in lung cancer and mir17 and mir21 clusters modulate PTEN, TGFβ-RII and are oncogenic in lung cancers (52).

### 3.3. Protein-based lung carcinoma molecular biomarkers

Protein-based biomarkers are very important because proteins are the main executioner bio-molecules in cells (53-54). The biomarkers currently of interest for
Lung carcinoma molecular biomarkers

Changes in clinical use are single protein molecules and most are serum-derived. They are carcinoembryonic antigen (CEA), CYFRA-21-1 and tissue polypeptideantigen (TSA) for non-small cell lung carcinoma and progastrin-relasing peptide (ProGRP) and neuron-specific enolase (NSE) for neuroendocrine lung carcinoma. There are also many potential lung carcinoma molecular biomarkers not yet available in clinical usage: serum amyloid A, haptoglobin-α-2, a fragment of apolipoprotein A-1 (APOE) and plasma kalikrein B1 (KLKB1) (55). Some proteins are important for prognosis and selection of therapy, for example epidermal growth factor receptor (EGFR). The altered activation of EGFR activates the Ras-Raf-MEK-MAPK pathway. Together with diagnostic purpose it is important for therapy with several drugs which target this pathway and are specific for EGF receptor (EGFR) (56-57). Another example is bel-2 protein, which acts as apoptosis inhibitor and indicates aberrant growth of cells. Its high expression is detected in 10-25% of NSCLC.

A different approach is provided by global protein assessment or proteomic analyses. Proteomic technologies analyze body fluids for protein constituents and identify biomarkers for early detection of cancers. Traditional approach for quantitative protein profiling is based on two-dimensional electrophoresis for protein separation, followed by mass spectrometric identification of selected proteins or all proteins detected (58). Another approach for disease biomarker discovery is using proteomic pattern analysis. This approach primarily compares the pattern of signals observed within a mass spectrum to identify differentially abundant peaks within normal and disease samples for distinguishing of two groups (target and control).

3.4. Metabolomics

Metabolomics detect low molecular weight molecules or metabolites such as amino acids, peptides, lipids and carbohydrates (1). During the progression of some cancers, the expression of certain N-linked and O-linked glycans changes. These altered glycoforms can serve as biomarkers for cancer detection (59-61).

4. APPLICATIONS OF LUNG CARCINOMA MOLECULAR BIOMARKERS

4.1. Lung carcinoma molecular biomarkers in detection of precancerous lesions and early carcinomas

Morphology is today a gold standard in diagnosing premalignant squamous lesions which are goblet cell hyperplasia, basal cell hyperplasia, immature squamous metaplasia, mature squamous metaplasia, mild squamous dysplasia, moderate squamous dysplasia, severe squamous dysplasia, squamous carcinoma in situ. Several studies detected molecular characterization of the premalignant changes. The earliest changes are loss of heterozygosity at multiple 3p chromosome sites and loss of heterozygosity at site 9p21 (62-64). Later changes are alterations at 8p21-23, 13q14 and 17p13 (62). The pattern of p53 expression in precursor lesions suggests gene alteration, but the detection of gene mutation seems to be more important (65-66). Loss of heterozygosity and p53 changes can be seen as well in normal bronchial mucosa of smokers, so it probably indicates the smoking related genetic damage, which with multiple stepwise accumulations of the mutations can still develop into lung cancer. A similar observation was made for methylation of p16NK4a which is demonstrated in 75% of in situ carcinoma adjacent to invasive squamous cell carcinoma of the lung. The frequency of this event increases from basal cell hyperplasia (17%) to squamous metaplasia (24%) to carcinoma in situ and invasive carcinoma (50-75%) (67-68).

In glandular lesions, recently, atypical adenomatous hyperplasia (AAH) is considered to represent the adenoma in a putative “adenoma-carcinoma” sequence in the lung periphery leading to development of nonmucinous bronchoalveolar carcinoma and invasive peripheral adenocarcinoma. K-ras mutations are detected in 24% to 50% of lung adenocarcinomas and represent an early event that precedes malignant growth (69). Recent studies suggest AAH could develop by either K-ras or EGFR gene mutation pathways, but some AAHs harboring K-ras gene mutation do not progress further to an invasive cancer (70). Loss of heterozygosity on chromosome 3p (FHT gene) and 9p (p16 gene) have been identified in 18% and 13% of AAH p53 protein expression has been demonstrated with increasing frequency in the progression from AAH through bronchoalveolar carcinoma (BAC) to invasive carcinoma (71-72).

In neuroendocrine carcinoma of the lung currently there are no genetic markers that could separate neoplastic from nonneoplastic proliferations.

4.2. Lung carcinoma molecular biomarkers as potential therapeutic targets

The most known molecular biomarkers used as predictors for therapy with tyrosine kinase inhibitors (gefitinib and erlotinib) are epidermal growth factor receptor (EGFR) and K-ras. EGFR is frequently overexpressed in non-small cell carcinoma (NSCLC) and development of small molecule inhibitor of EGFR resulted in clinical trials that have shown a good response in lung adenocarcinomas with somatic mutations in exons 18 to 21 of the tyrosine kinase domain of EGFR together with clinical characteristics like female sex, Asian ethnicity and absent or low smoking history (73-76). The K-ras mutations have been shown as negative prognostic factor and as predictors of failure of EGFR tyrosine kinase inhibitors therapy (73, 74, 76). Testing for EGFR and K-ras status in lung carcinoma involves a variety of different methodologies including DNA mutational analysis, fluorescence in situ hybridization (FISH), and immunohistochemistry but still there is no standardization of the methods and the criteria for interpretation. Also there is a discussion about the most appropriate sample (surgical versus cytological) that should be submitted for analysis and whether to test primary tumor and/or metastasis as they could differ from each other (77). The EML4-ALK fusion gene is one of the most recently identified molecular targets for treatment of non small cell lung carcinoma. It has been identified in approximately 4% of patients with non small
Lung carcinoma molecular biomarkers

cell lung carcinoma, commonly in younger males, never smokers and patients with adenocarcinoma. Its presence is strongly associated with resistance to EGFR-tyrosine kinase inhibitors and sensitivity to ALK inhibitors (78). One of the newest anticancer therapies is antiangiogenic therapy and there are now a large number of antiangiogenic drugs. Among the most advanced of these drugs are those that block the proangiogenic function of vascular endothelial growth factor (VEGF). Some studies suggest that plasma VEGF should be considered as a possible surrogate pharmacodynamic marker for determining the optimal biological dose of antibody drugs that block VEGFR-2 activity in a clinical setting (79).

5. LUNG CARCINOMA MOLECULAR BIOMARKERS IN VARIOUS SPECIMENS

Lung carcinoma molecular biomarkers can be detected in blood (serum), tissue specimens and cytology specimens. Blood is the most commonly used specimen obtained in clinics owing to the ease of its collection (1). Although blood contains a large number of proteins and represents a health status of a patient, it is very difficult biological sample to analyze. Six of the most abundant protein impede the detection of other protein signals in the serum or plasma and a very small amount of total proteins is cancer related (80, 81). There is no strict standard protocol for collection of blood and serum, therefore variations among many researchers may occur. Blood lung carcinoma molecular biomarkers are not specific only for lung carcinoma but are also associated with colorectal carcinoma and ovarian carcinoma (82). Currently, the most well-known markers are CEA, CA-125, cytokeratin 19, CYFRA 21-1, TPA, ProGRP, NSE and Tumor M2-pyruvate kinase (55).

Lung carcinoma tissue specimens the best represent the cancer constituents in human biological samples but the biopsy or resection is aggressive, sometimes impossible and expensive approach. Thorax puncture biopsy is also riskier than other cancer biopsies. There is no standardized method for collecting and handling biopsy samples, so variations among samples can occur (1). In biopsied lung tissue not only carcinoma cells but also nontumor cells and other molecules like immune cells, cytokines and derivates from immune or inflammatory response are present, which limits the value of sample as a source for biomarkers identification.

Diagnostics from cytology specimens has an advantage of non-invasive method, can be repeated several times and can be easily combined with other molecular techniques. In cytology, sputum was used for the longest time, since 1930s in diagnostic of lung carcinoma. Sputum cytology has high specificity (98%) but lower sensitivity (less than 65%). Similarly bronchoalveolar lavage cytology shows low sensitivity, about 27.5%. Several researches demonstrated that combination of several different cytohistological techniques increases sensitivity (83). Sputum contains biomaterial from the original site of the cancer which can be analyzed and the most representative biomarkers from sputum are K-ras, EGFR, p53 mutations, hnRNP and A2/B1 which can be also identified in pathological tissue samples (81). Many recent studies have investigated the use of cytology samples for EGFR (84-89) and K-ras (77) testing. For example patients may only have metastatic disease documented on exfoliative or aspiration cytology material like pleural fluids, fine needle aspirate of the liver or of the pleura or fine needle aspirate of lymph nodes. Still even this cytology samples were suitable for EGFR FISH or mutational analysis (85,87,88,90-92) and for K-ras mutation analysis (77).

6. VARIOUS TECHNOLOGIES IN LUNG CARCINOMA MOLECULAR BIOMARKERS

For nucleic acid-based lung carcinoma molecular biomarkers identifications three of the most important techniques are FISH, PCR and recently also DNA-microarray. Fluorescence In situ Hybridization (FISH) is commonly used with cytolgy, biopsy and surgical specimens to detect tumor cells and certain microorganisms. It has prognostic value and it can be used to determine treatment response (93-99). In situ hybridization uses DNA or RNA probes to evaluate intact cells for genetic changes. Probes are visualized using fluorescent labels. In CISH (chromogenic in situ hybridization) probes are visualized with a chromogen that produces a colored chemical at the reaction site (5).

Polymerase Chain Reaction (PCR) is efficient and sensitive method for studying the molecular pathology of primary and metastatic neoplasms, inflammatory mechanisms and infectious diseases (100-107). This method is used to detect the modifications of cancer-related genes or chromosomal abnormalities in diagnosis or prediction of lung carcinomas (108). Due to its high availability it is expected to be useful in predicting individuals who are likely to develop lung carcinoma (109).

DNA-microarray is high-throughput technique used to study molecular alterations at the level of gene expression, DNA-protein interactions and whole genome analysis, as well as gene expression levels on thousands of genes simultaneously. Methods used to detect lung cancer biomarkers at the RNA expression level include Quantitative Reverse Transcription Polymerase Chain Reaction (RT-qPCR), Serial Analysis of Gene Expression (SAGE), differential display, bead-based methods and microfluid card and micro-array analysis (110).

Identification of the biomarkers depends on accessibility to sufficient amount of material from different specimens. Another limitation is the level of RNA present in RT-PCR or microarray, which does not always indicate the level of present protein because mRNA and protein levels do not necessarily correspond. Therefore the nucleic acid- based biomarkers may not reflect the underlying characteristics of the cancer cell (55).

Protein-based “fingerprints” may outperform individual protein markers. Technologies as differential in-gel electrophoresis (DIGE), two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and
Lung carcinoma molecular biomarkers

multidimensional protein-identification technology (MudPIT) can be used for higher-throughput profiling with microgram quantities of protein (111). Other high-throughput technologies as Mass Spectroscopy (MS), Matrix Associated Laser Absorption Desorption Ionization Time of Flight (MALDI-TOF), Surface Enhanced Laser Absorption Desorption Ionization Time of Flight (SELDI-TOF), and reverse phase microarray are more sensitive (47,54,112-114). Some other promising techniques are still are experimental: nanotechnology immuno-PCR, Field Effect Transistor (FET)-based Protein Detection and Quantum Dots (115-118).

Mass spectrometry can be used to detect disease-associated carbohydrate markers related to glycomics present in lung cancers (119,120).

7. OUR EXPERIENCE IN PCR APPLICATIONS FOR DETECTION OF LUNG CARCINOMA BIOMARKERS IN CYTOLOGY SPECIMENS

As it was mentioned before, genomic based biomarkers identification depends on sufficient amount of good quality DNA from specimen which is important for DNA amplification. Our primary interest was the analysis of lung cytological samples obtained during bronchoscopy. Obtaining genomic DNA for successful PCR amplification is critical for molecular analyses. Routinely used methods for DNA isolation include commercially available kits, or extensive procedures that usually involve cell lysis by proteinase K, followed by isopropanol and/or phenol/chloroform extraction as an additional step of DNA purification (121,122). The complexity and expenses of the procedure represent a burden in systematic large-scale DNA isolation; so every simplification in DNA isolation procedure is welcomed and saves money and time (123). We tested a simple method which combined alkaline and temperature lyses for isolation of DNA solution, and use it to extract DNA from cytology samples from bronchial aspirates. Alkaline lysis is a method for DNA isolation already used in the analysis of samples from human and mouse buccal swabs (124-126). It is based on solubilization of membrane and coagulation of proteins which enables the unaffected DNA to leave the cells and be available for further analysis by PCR in less than an hour. This method was a starting point to be adapted for use in cytology and it was tested for amplification of sex chromatin gene fragments in bronchial aspirates. The method was showed as highly reliable for DNA isolation, as it preserves high-quality DNA suitable for PCR-based assays. Its main advantages are noninvasive approach, simplicity of the procedure as only unsophisticated chemicals are used, and a high success rate of PCR (123). It was successfully used for testing by PCR molecular markers related to lung cancer, for example, the presence of “high-risk” human papilloma virus types which are associated with cervical cancers and subset of head and neck cancers and can be, according some authors, cocarcinogen for lung carcinoma development (52,123,127).

8. PERSPECTIVE

Early diagnosis is crucial for decreasing lung cancer mortality. Gene and protein expression profiling has advanced our understanding of lung carcinoma. The increasing knowledge of the pathway activation profile in lung carcinoma indicates new targets but also new markers to select patients and predict therapy efficiency. Many studies report numerous lung carcinoma molecular biomarkers but they are often made on small number of patients difficult to reproduce and not sensitive or specific enough for clinical use. It is clear that all lung carcinoma molecular biomarkers must be validated in a large set of clinical samples. Lung carcinomas are very heterogeneous and morphologic heterogeneity is reflected at the molecular level, so several lung carcinoma molecular biomarkers must be combined with different approaches according to the different lung carcinoma types. A deep biological understanding of each molecular biomarker is crucial and can be a key to correct diagnose and also a predictor for therapy efficacy. Larger new studies with clinical data and follow-up are required to achieve these goals.

9. ACKNOWLEDGEMENTS

This work was supported by grants of Ministry of Science and Technology, Republic of Croatia (108-1081870-1902), and Unity for Knowledge Fund, Republic of Croatia (UKF 35/08).

10. REFERENCES


Lung carcinoma molecular biomarkers


Lung carcinoma molecular biomarkers


46. B.E. Jady and T. Kiss: A small nucleolar guide RNA functions both in 2'-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. *EMBO J* 20, 541-551 (2001)


Lung carcinoma molecular biomarkers


Lung carcinoma molecular biomarkers

KRAS status in primary lung carcinoma and matched metastases. *Hum Pathol* 41, 94-102 (2010)


100. I. Raad, H. Hanna, A. Huaringa, D. Sumoza, R. Hachem and M. Albair: Diagnosis of invasive pulmonary aspergillosis using polymerase chain reaction-based
Lung carcinoma molecular biomarkers

detection of aspergillus in BAL. *Chest* 121, 1171-1176 (2002)


124. L. Rudbeck and J. Dissing: Rapid, simple alkaline extraction of human genomic DNA from whole blood, buccal epithelial cells, semen and forensic stains for PCR. *Biotechniques* 25, 588-592 (1998)
Lung carcinoma molecular biomarkers


Key Words: Lung, Carcinoma, Molecular Biomarkers, Review

Send correspondence to: Srecko Gajovic, University of Zagreb School of Medicine, Croatian Institute for Brain Research, Salata 12, 10000 Zagreb, Croatia, Tel: 385 1 4566948, Fax: 385 1 4566795, E-mail: srecko.gajovic@hiim.hr

http://www.bioscience.org/current/vol4E.htm