

Molecular Pathogenesis of Non-Small Cell Lung Carcinomas

Non-small cell lung carcinoma (NSCLC) from histological and biological perspectives is a complex neoplasm. The sequential preneoplastic changes have been defined for centrally arising squamous cell carcinomas (SCCs) of the lung, and they are less documented for peripherally arising adenocarcinomas. The main morphologic forms of preneoplastic lesions recognized in the lung are squamous dysplasias for SCC, and atypical adenomatous hyperplasia for adenocarcinoma. Several studies have provided information regarding the molecular characterization of lung preneoplastic changes, especially for SCC. These molecular changes have been detected in the histologically normal and abnormal respiratory epithelium of smokers and patients with lung cancer, phenomenon known as field of cancerization. Our improved understanding of the changes and origins of the field of cancerization can be applied clinically to improve early detection of lung cancer. In the last decade, significant progress has been made in the characterization of molecular abnormalities in NSCLC tumors that are being used as molecular targets and predictive biomarkers for patients' selection for targeted therapy. As our understanding of the biology of the molecular pathogenesis of lung cancer evolves, there is an opportunity to use this knowledge for the development of novel chemoprevention strategies using those molecularly targeted agents used to treat advanced lung cancer, a concept coined as reverse migration. The rapid development of technologies for large-scale molecular analysis, including microarrays and next-generation sequencing will facilitate high-throughput molecular analysis of lung cancer preneoplastic lesions and the field of cancerization. (**J Lung Cancer 2012; 11(1):12–20**)

Key Words: Lung neoplasms, Non-small-cell lung carcinoma, Adenocarcinoma, Squamous cell carcinoma, Epidermal growth factor receptor, Precancerous conditions

Ignacio I. Wistuba, M.D.

Departments of Pathology and Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Received: June 12, 2012
Revised: June 14, 2012
Accepted: June 14, 2012

Address for correspondence

Ignacio I. Wistuba, M.D.
Departments of Thoracic/Head and Neck Medical Oncology and Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
Tel: 1-713-563-9184
Fax: 1-713-730-0309
E-mail: iwistuba@mdanderson.org

INTRODUCTION

Lung cancer is the leading cause of cancer deaths worldwide (1). The vast majority of the lung cancers are diagnosed at advanced stages when the options for treatment are not curative. The majority of patients when diagnosed have either locally advanced unresectable lung cancer or metastatic lung cancer (2). From histopathological and biological perspectives, lung cancer is a highly complex neoplasm. Accurate histological classification of lung cancer is essential for patients to receive appropriate therapy (3). Although classification of the

vast majority of lung cancers is straightforward, areas of controversy and diagnostic challenges remain (3). Lung cancer consists of several histological types, being the most frequent the non-small cell lung carcinomas (NSCLC) types squamous cell carcinoma (SCC, 30%) and adenocarcinoma (45%) (3).

The advances in molecular methodologies are providing new insights into the biology involved in the pathogenesis of NSCLC (4). It has been shown that clinically evident lung cancers are the results of the accumulation of a number of genetic and epigenetic changes, including abnormalities for the inactivation of tumor suppressor genes and the activation of oncogenes (2,4). All those molecular abnormalities involve the

“hallmarks of cancer,” including abnormalities in self-sufficiency of growth signals, insensitivity to anti-growth signals, sustained angiogenesis, evading apoptosis, limitless replicative potential, and tissue invasion and metastasis (5). The recent molecular advances in our understanding of the pathogenesis of NSCLC will provide a unique opportunity to improve the prevention and treatment of lung cancer (6).

In this review, the molecular abnormalities associated to the early pathogenesis of the two most common histological types of lung cancer, SCC and adenocarcinoma, will be described. Two important concepts of the pathogenesis of lung cancer (*field of cancerization*) and the development of novel chemoprevention strategies (*reverse migration*) will be discussed. Further research in the early pathogenesis of lung cancer may provide new methods for assessing the likelihood of developing invasive lung cancer in smokers and allow for early detection and monitoring of their response to chemoprevention regimens using molecularly targeted agents.

MOLECULAR ABNORMALITIES OF SQUAMOUS CELL CARCINOMAS AND ADENOCARCINOMAS OF THE LUNG

1) Squamous cell carcinoma (SCC)

This tumor type frequently harbors multiple genetic abnormalities, including activation of several oncogenes (e.g., *SOX2*, *FGFR1*, *PIK3CA*, *EGFR*, and *DDR2*), and inactivation of tumor suppressor genes (e.g., *TP53*, and *p16^{INK4a}*) (Table 1, Fig. 1) (7,8). The sex determining region Y-box 2 (*SOX2*) gene has

been suggested to be a lineage survival oncogene in lung SCC (9). *SOX2* is a transcription factor coded by a gene located at 3q26.33. This protein is implicated in the maintenance of embryonic stem cells and the induction of pluripotent stem cells from various different cell types, and plays important roles in multiple developmental processes, including lung branching (9). *SOX2* protein is expressed in normal bronchial epithelium and preneoplastic squamous lesions (10), and is more frequently detected in SCC (90%) compared with adenocarcinoma (20%) (9,10). *SOX2* gene is amplified in 20% SCC of the lung, and correlates with the protein overexpression (9). *SOX2* amplification has not been detected in lung tumors with adenocarcinoma histology.

Recently, two clinically relevant molecular abnormalities have been identified to occur in SCC: amplification of *FGFR1* gene (chromosome 8p11-12), and mutation of the *DDR2* gene. The *FGFR1* gene encodes for a transmembrane tyrosine kinase (TK) and is amplified in ~20% of SCCs and only in 1~3% of adenocarcinomas (11). Mutations of another TK, *DDR2* gene, occur in ~4% of SCCs, and it has not been reported in adenocarcinomas of the lung (12). Importantly, lung tumors with *FGFR1* amplification respond to FGFR TK inhibitors (TKIs) (13), and lung cancers with *DDR2* mutation are sensitive to dasatinib and nilotinib treatments (12).

2) Adenocarcinoma

Lung adenocarcinomas are histologically and biologically heterogeneous tumors (14,15). Most adenocarcinomas (~90%) have mixed histological growth patterns, including acinar,

Table 1. Summary of the Frequency of Molecular Abnormalities on Oncogenes Associated with the Lung Adenocarcinoma and Squamous Cell Carcinoma Histologies

Gene	Molecular change	Squamous cell carcinoma %	Adenocarcinoma %
<i>EGFR</i>	Mutation	Very rare	10~40
	Amplification/Copy gain	40	15
<i>HER2</i>	Mutation	Very rare	2~4
	Amplification	Very rare	5
<i>EML4-ALK</i>	Translocation	Very rare	5~15
<i>KIF5B-RET</i>	Translocation	Not reported	2
<i>ROS1-FIG</i>	Translocation	Not reported	2
<i>KRAS</i>	Mutation	Very rare	10~30
<i>BRAF</i>	Mutation	Very rare	1~3
<i>FGFR1</i>	Amplification	20	Not reported
<i>DDR2</i>	Mutation	4	Not reported
<i>PIK3CA</i>	Amplification/Copy gain	30	2~6
	Mutation	2~6	2~6

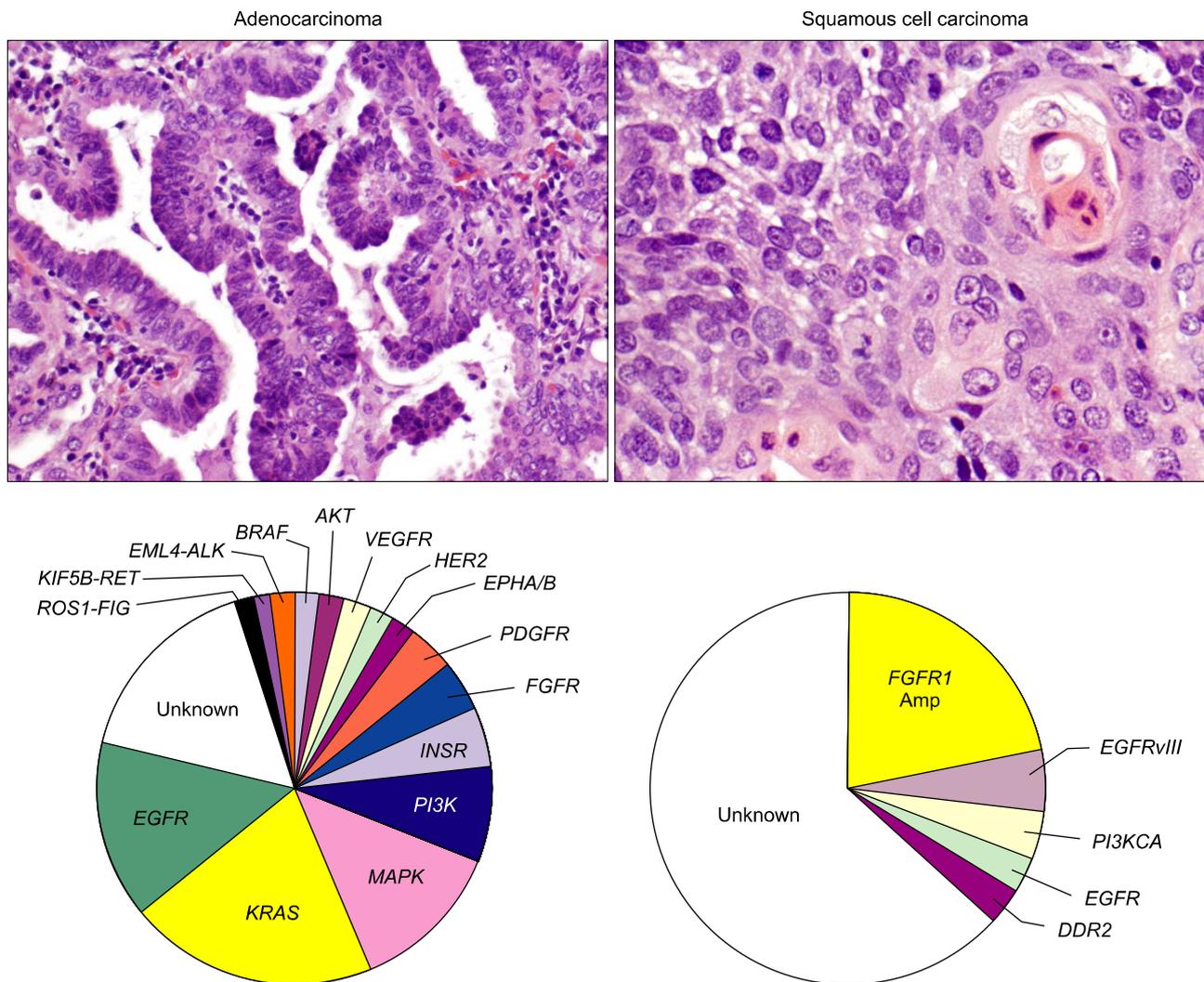


Fig. 1. Paradigm changes in the classification of non-small cell lung carcinoma. Targetable abnormalities of multiple oncogenes, particularly mutations, translocations and gene amplifications have been recently discovered in both adenocarcinoma and squamous cell carcinoma histologies. Most abnormalities are referred to gene mutations, except *EML4-ALK*, *KIF5B-RET* and *ROS1-FIG* rearrangement and *FGFR1* amplification (Amp).

papillary, lepidic (former bronchioloalveolar carcinoma, BAC), micropapillary, and solid patterns. Attempts to establish correlations between histological features and molecular abnormalities have been unsuccessful (3).

Over the past decade, significant progress has been made in the characterization of molecular abnormalities of lung adenocarcinoma (4). Similarly to SCC, adenocarcinomas of the lung show frequent inactivation of tumor suppressor genes, including, among others, *TP53*, and *p16^{INK4a}* (7,8). DNA sequencing of 188 lung adenocarcinomas including 623 genes with known or potential relationships to cancer revealed more than 1,000 somatic mutations across the samples, and 26 genes were

mutated at significantly high frequencies (16). The frequently mutated oncogenes included TKs *HER2*, multiple ephrin receptor genes (notably *EPHA3*), vascular endothelial growth factor receptor *KDR*, and neurotropic TK (*NTRK*) genes (16).

In the last decade, lung adenocarcinomas arising in never or light smokers have shown to be characterized by significantly higher frequencies of a series of targetable oncogenes abnormalities (4), including *EGFR* and *HER2* TK domain activating mutations, and *EML4-ALK* translocation (Table 1, Fig. 1) (4). Recently, two additional potentially targetable gene translocations, *KIF5B-RET* (10p;11q)(p11.22; q11-21) (17-19) and *ROS1-FIG* (chromosome 6q22) (20), have been identified in

lung adenocarcinoma from never and ever smokers. Other potentially targetable genetic abnormalities have been detected in both major NSCLC histologies, including among others, *PIK3CA* mutation and amplification, *MET* amplification (7q21-q31), *AKT1* and *MAP2K1* mutations (4,16).

EGFR mutations in lung cancer are limited to the first four exons of the TK domain (exons 18~21), being the most frequent mutations in-frame deletions in exon 19 (44% of all mutations) and missense mutations in exon 21 (41% of all mutations) (21,22). In addition, in-frame duplications/insertions occurring in exon 20 have been described in about 5% of the mutant cases, and rare missense mutations occur in multiple sites (21,22). *EGFR* mutations occur predominantly in adenocarcinoma (~15% to 48% vs. other NSCLC histologies ~2%), and are more frequent in never-smokers (54% vs. ever-smokers 16%) and female patients (49% vs. male patients 19%) (21,22). Activating *EGFR* mutations strongly correlate with clinical response to TKIs, and currently is the most important criterion to select patients for *EGFR* TKI therapy in lung cancer (23). *HER2* mutations have been described predominantly in adenocarcinoma histology (~2 to 4%) and patients with East-Asian ethnic background and never-smoker history (21,22).

EML4-ALK translocation (2;5)(p23q35), with subsequent activation of *ALK*, has been detected in ~6% of lung adenocarcinomas (24,25). This translocation occurs particularly in patients with never or light smoking history, and associated with young onset of tumor. Tumors with *EML4-ALK* translocation usually lack *EGFR* and *KRAS* mutations (25-27). Histologically, *EML4-ALK*-rearranged adenocarcinomas have been described to have a frequently a solid pattern with signet ring cells (27), but this translocation can be identified in all histological subtypes of lung adenocarcinomas.

KRAS mutations are most common in lung adenocarcinoma than other NSCLC histology types, and are more frequently found in tumors from patients with smoking history (~30%) (28). In lung cancer, *KRAS* mutations are found in codons 12, 13 and 61, and they (42% of all mutations) are mainly GGT to TGT transversions producing a Glycine to Cysteine amino acid changes (29,30). *KRAS* mutations are usually not detected in *EGFR*-mutant tumors, and have been associated to low response rates to *EGFR*-TKI therapies (31). *BRAF* oncogene can be activated in NSCLC, particularly adenocarcinoma histology (1~3%), by gene point mutations affecting exons 11

and 15 (4).

Most (70~80%) lung adenocarcinomas express the thyroid transcription factor-1 (TTF-1) protein encoded by the gene *TTF-1*, also termed *NKX2-1* (32). TTF-1 is a homeodomain-containing transcription factor expressed at the onset of lung morphogenesis in epithelial cells (33). In the postnatal lung, TTF-1 expression is restricted primarily to alveolar type II cells and subsets of non-ciliated bronchiolar epithelial cells (33). The *TTF-1* gene has been suggested to be a lineage-specific oncogene, and gene amplification has been detected in ~11% of lung adenocarcinomas, which correlated with poor survival rates in surgically resected tumors (32,34). However, conversely to the putative oncogene properties, high level of TTF-1 protein expression in adenocarcinomas has been correlated with smaller tumor size, lower pathologic tumor-nodes-metastasis (TNM) stage, and better survival rates (15,34).

PRECURSOR LESIONS OF NSCLC

Lung cancers are believed to arise after a series of progressive pathological changes (preneoplastic or precursor lesions) in the respiratory mucosa (Fig. 2). For NSCLC, the 2004 World Health Organization (WHO) histological classification of preinvasive lesions of the lung lists two main morphologic forms of preneoplastic lesions (14): a) squamous dysplasia and carcinoma *in situ* (CIS); and, b) atypical adenomatous hyperplasia (AAH) (Table 2). While the sequential preneoplastic changes have been defined for SCC, they have been poorly documented for adenocarcinomas and other lung cancer histologies (e.g., large carcinomas, small cell lung carcinoma, etc.) (35). Mucosal changes in the large airways that may precede invasive SCC include squamous dysplasia and CIS (35). Adenocarcinomas may be preceded by morphological changes including AAH in peripheral airway (36).

1) Squamous preneoplastic lesions

The mucosal changes in the large airways that precede invasive SCC include: hyperplasia, squamous metaplasia, squamous dysplasia and CIS (Fig. 2). Dysplastic squamous lesions may be of different intensities (mild, moderate, and severe); however, these lesions represent a continuum of cytologic and histologic atypical changes that may show some overlapping between categories (35). Little is known about the

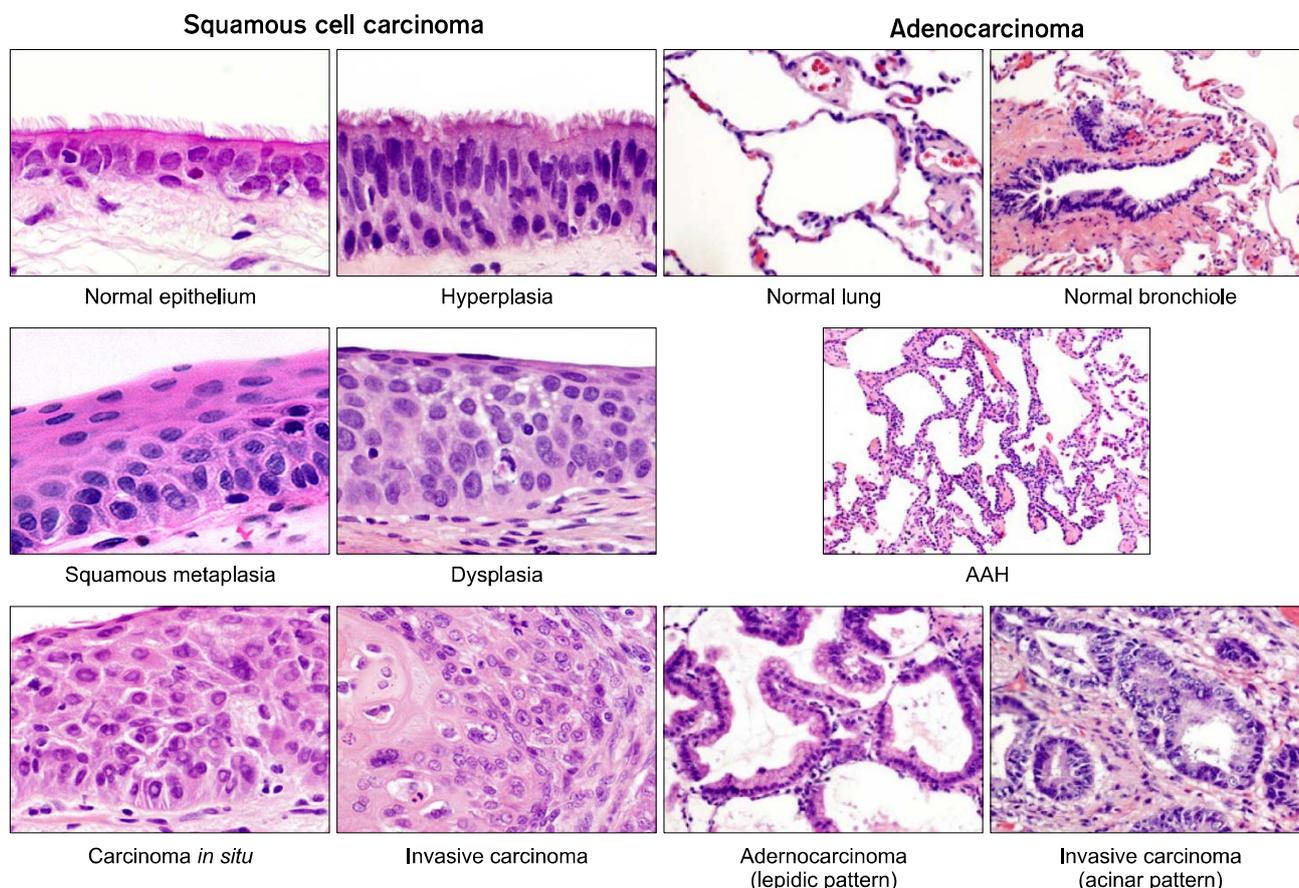


Fig. 2. Sequential histological abnormalities involved in the pathogenesis of squamous cell carcinoma and adenocarcinoma of the lung. The mucosal changes in the large airways that precede invasive squamous cell carcinoma include: hyperplasia, squamous metaplasia, squamous dysplasia and carcinoma *in situ*. Dysplastic squamous lesions may be of different intensities: mild, moderate and severe. Atypical adenomatous hyperplasia (AAH) is considered a putative precursor of adenocarcinoma, particularly those with extensive lepidic pattern (former bronchioloalveolar [BAC] pattern). AAH is a discrete parenchymal lesion arising in the alveoli close to terminal and respiratory bronchioles.

Table 2. Histopathological and Molecular Abnormalities of NSCLC Precursor Lesions

Abnormality	Adenocarcinoma	Squamous cell carcinoma
Histopathology		
Lesion	AAH	Squamous dysplasia and carcinoma <i>in situ</i>
Location	Lung parenchyma (peripheral)	Bronchial mucosa (central)
Molecular		
Gene abnormalities	<i>KRAS</i> mutation <i>EGFR</i> mutation <i>p16</i> methylation	<i>TP53</i> LOH and mutation <i>SOX2</i> amplification <i>p16</i> methylation
Genetic instability	Low	Intermediate
LOH	Low	Intermediate
Chromosomal regions with LOH	9p21, 17p/ <i>TP53</i>	8p21-23, 9p21, 17p/ <i>TP53</i>

NSCLC: non-small cell lung carcinoma, AAH: atypical adenomatous hyperplasia, LOH: loss of heterozygosity.

rate and risks of progression of squamous dysplasia to CIS and ultimately to invasive SCC.

The current working model of the sequential molecular

abnormalities in the pathogenesis of SCC indicates that the genetic abnormalities (gene deletions, genetic instability, and other changes) commence in histologically normal epithelium

and increase with increasing severity of histologic change, and that the genetic changes follow a sequence (37). Several studies have shown that multiple clonal and subclonal patches of airway cells with molecular abnormalities can be detected in the normal and slightly abnormal bronchial epithelium of smokers and patients with lung cancer (37,38). The molecular abnormalities associated with the early pathogenesis of SCC include allelic losses at multiple regions of chromosome 3p and 9p21 (*p16^{INK4a}*), subsequent losses of chromosome regions 8p21-23, 13q14 (*RB*), and 17p13 (*TP53*) (39) (Table 2). Additionally, it has been shown that *p16^{INK4a}* methylation increases during the progression of SCC preneoplastic lesions (39).

Recent studies showed progressive amplification of 3q amplification in the evolution of squamous dysplasias and implicate *SOX2* as a key target of this process (40). Interestingly, in a small cohort of squamous dysplasias lesions the degree of 3q amplification associates with clinical progression of the precursor lesions (41). However, despite all these advances, neither histological features nor molecular changes can be reliably used for accurate risk of progression of squamous preneoplastic lesions and be used to assess the risk of developing SCC (42).

2) Adenocarcinoma precursor lesions

It has been suggested that adenocarcinomas may be preceded by AAH in peripheral airway cells (Fig. 2) (36,43); however, the respiratory structures and the specific epithelia cell types involved in the origin of most lung adenocarcinomas are unknown.

AAH is a discrete parenchymal lesion arising in the alveoli close to terminal and respiratory bronchioles. Because of their size, AAH are frequently incidental histological findings, but they may be detected grossly, especially if they are 0.5 cm or larger. The increasing use of high resolution computed tomography scans for screening purposes has led to an increasing awareness of this entity, as it remains one of the most important differential diagnoses of air filled peripheral lesions (called “ground glass opacities”). The origin of AAH is still unknown, but the differentiation phenotype derived from immunohistochemical and ultrastructural features suggests an origin from the progenitor cells of the peripheral airways, such Clara cells and type II pneumocytes (44,45).

Several molecular changes frequently present in lung adeno-

carcinomas are also present in AAH lesions, and they are further evidence that AAH may represent true preneoplastic lesions (44). The most important finding is the presence of *KRAS* (codon 12) mutations in up to 39% of AAHs, which are also a relatively frequent alteration in lung adenocarcinomas (36,46). Other molecular alterations detected in AAH are overexpression of cyclin D1 (~70%), p53 (range, 10~58%), survivin (48%), and HER2/neu (7%) proteins over-expression (36,47,48). Interestingly, *EGFR* mutations have been detected in some cases of AAH accompanying resected peripheral adenocarcinomas, providing further evidence that they represent precursor lesions of peripheral adenocarcinomas (49).

However, the respiratory structures and the specific epithelial cell types involved in the origin of most lung adenocarcinomas are unknown, and candidates include Clara cells, alveolar type II cells, and an epithelial cell located at the junction between the terminal bronchiole and alveolus, termed bronchioloalveolar stem cell (BASC). In addition, lung adenocarcinoma may arise from bronchial or bronchiolar epithelia, since *EGFR* activating mutation has been found in normal bronchial epithelium adjacent to *EGFR*-mutant adenocarcinomas in 43% of patients (50,51).

LUNG CANCER PROGENITOR OR STEM CELLS

Classically, stem cells are located in specialized niche and are preserved from overuse in maintaining tissue homeostasis for their conservation throughout the organism’s life. As such, stem cells undergo infrequent cycling and stem cell compartments are identified based upon this characteristic. Currently, it is believed that at least four stem cell niches are present in adult lung airway epithelium: gland duct basal cells, inter-cartelagenous zone surface basal cells, bronchiolar Clara cells associated with pulmonary neuroendocrine cell bodies, and bronchioloalveolar duct junction variant Clara cells (52).

From mouse model studies, it has been shown that BASCs are putative stem cell initiators of *KRAS* induced NSCLC adenocarcinoma (53). In addition, to the identification of BASCs, several stem cell markers for lung tumors have been identified. The most studied lung cancer stem cell marker has been CD133, which has been shown to express in NSCLC and other lung tumors (e.g., small cell carcinoma) (54). Recently,

the stem cell markers aldehyde dehydrogenase 1A1 (ALDH1A1) has been also shown to be frequently expressed in NSCLC associated to patients' outcome (55). However, despite multiple studies, the progenitor or stem cells for human lung cancers have not been identified yet.

LUNG AIRWAY FIELD CANCERIZATION

It has been shown that human tumors and adjacent histologically normal tissue share certain molecular abnormalities, a phenomenon coined as *field cancerization* (56). The concept of field of cancerization was first introduced in 1953 in the setting of oral cancer (57). Later, field of cancerization involving histological and molecular changes of the lung airway of patients with lung cancer and smokers began to be characterized (56). A detailed analysis of histologically normal, premalignant and malignant epithelia from smokers and patients with SCC, have indicated that multiple, sequentially occurring, allele-specific chromosomal deletions and genomic instability begin in dispersed clonally independent foci early in the pathogenesis of this malignancy, including histologically normal bronchial epithelium (39). In addition, it has been demonstrated that *KRAS* (58) and *EGFR* (50,51) mutations are found in peripherally located histologically normal lung tissue adjacent to lung adenocarcinomas. The finding of higher frequency of *EGFR* mutation in normal cells obtained from bronchioles close to the tumor compared to more distant sites, suggested the presence of a localized field effect in the pathogenesis of lung adenocarcinomas (50). These and other observations suggest a compartmentalization of the field of cancerization, and the dissimilarity between SCCs and adenocarcinomas may reflect the differential mechanisms of pathogenesis of these NSCLC subtypes.

Our improved understanding of the changes and origins of the field of cancerization can be applied clinically to improve our ability of risk assessment, screening, diagnosis and treatment of patients with or at risk for lung cancer. Recently, global gene expression (mRNA) studies using microarrays and pathway signatures analysis have been derived from cells obtained from normal bronchial epithelium of smokers, and showed to exhibit properties to predict the development of lung cancer (59,60). Global alterations in gene expression differences between normal bronchial epithelium of healthy cancer-free

smokers and that of never smokers have been described (59). Interestingly, irreversible changes in gene expression in airways of former smokers after years of smoking cessation were described to underlay the still increased risk of developing of lung cancer observed in this individuals (59). Interestingly, an 80-gene expression signature developed from bronchial epithelial cells obtained from the large airway can distinguish smokers without overt cancer from smokers without lung cancer (60).

TARGETED PREVENTION AND REVERSE MIGRATION IN LUNG CANCER

Positive chemoprevention trials usually involve known targets that can be inhibited effectively by drugs. The significant progress made in the characterization of targetable molecular abnormalities in NSCLC tumors offers a unique opportunity to use this knowledge for the development of novel chemoprevention strategies using molecularly targeted agents. However, our knowledge of the molecular pathogenesis of lung cancer must continue to evolve.

Recently, it has been proposed to reverse migration of findings on molecular markers, pharmacogenomics and targeted therapy from the advanced NSCLC setting to the chemoprevention setting (61). For reverse migration to be an effective strategy, several assumptions must be true: 1) pre-malignancy shares molecular properties with malignancy; 2) effective targeted therapy of malignancy is applicable to precursor lesions; and 3) predictive biomarkers can be applied to patients at risk for malignancy. The current knowledge on the early pathogenesis of NSCLC suggests that these assumptions are probably true, therefore, there is an opportunity for the utilization of targeted agents for primary, secondary and tertiary chemoprevention in lung cancer, particularly NSCLC.

FUTURE DIRECTIONS

Applying advanced high-throughput molecular methodologies currently used to study advanced lung tumors to the molecular analysis of the field of cancerization is expected to expand our understanding of the biology of the lung cancer. The rapid development of technologies for large-scale molecular analysis, including the newest next-generation sequencing

(NGS) of RNA (transcriptome) and DNA (whole genome or exome) is facilitating high-throughput molecular studies holding various advantages over traditionally molecular analysis of nucleic acids abnormalities. NGS can simultaneously detect deletions, insertions, copy number alterations, translocations, and exome-wide base substitutions (including known hot-spot mutations) in all known cancer-related genes (62). An important step to apply NGS in lung cancer pathogenesis was a recent study in which RNA of bronchial airway epithelial cells brushings from never-smoker and ever-smoker individuals with and without lung cancer was analyzed by RNA-sequencing providing additional insights besides those provided when using RNA microarrays (63). A comprehensive analysis of the early molecular events in NSCLC pathogenesis utilizing novel methodologies will undoubtedly unravel novel molecular targets and biomarkers that can aid prevention deliver its longstanding promise in lung cancer.

REFERENCES

- Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. *CA Cancer J Clin* 2010;60:277-300.
- Herbst RS, Heymach JV, Lippman SM. Lung cancer. *N Engl J Med* 2008;359:1367-1380.
- Travis WD, Rekhtman N. Pathological diagnosis and classification of lung cancer in small biopsies and cytology: strategic management of tissue for molecular testing. *Semin Respir Crit Care Med* 2011;32:22-31.
- Pao W, Girard N. New driver mutations in non-small-cell lung cancer. *Lancet Oncol* 2011;12:175-180.
- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011;144:646-674.
- Kadara H, Behrens C, Yuan P, et al. A five-gene and corresponding protein signature for stage-I lung adenocarcinoma prognosis. *Clin Cancer Res* 2011;17:1490-1501.
- Gold KA, Wistuba II, Kim ES. New strategies in squamous cell carcinoma of the lung: identification of tumor drivers to personalize therapy. *Clin Cancer Res* 2012;18:3002-3007.
- Perez-Moreno P, Brambilla E, Thomas R, Soria JC. Squamous cell carcinoma of the lung: molecular subtypes and therapeutic opportunities. *Clin Cancer Res* 2012;18:2443-2451.
- Bass AJ, Watanabe H, Mermel CH, et al. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nat Genet* 2009;41:1238-1242.
- Yuan P, Kadara H, Behrens C, et al. Sex determining region Y-Box 2 (SOX2) is a potential cell-lineage gene highly expressed in the pathogenesis of squamous cell carcinomas of the lung. *PLoS One* 2010;5:e9112.
- Weiss J, Sos ML, Seidel D, et al. Frequent and focal FGFR1 amplification associates with therapeutically tractable FGFR1 dependency in squamous cell lung cancer. *Sci Transl Med* 2010;2:62ra93.
- Hammerman PS, Sos ML, Ramos AH, et al. Mutations in the DDR2 kinase gene identify a novel therapeutic target in squamous cell lung cancer. *Cancer Discov* 2011;1:78-89.
- Dutt A, Ramos AH, Hammerman PS, et al. Inhibitor-sensitive FGFR1 amplification in human non-small cell lung cancer. *PLoS One* 2011;6:e20351.
- Travis WD, Brambilla E, Müller-Hermelink HK, Harris CC. World Health Organization classification of tumours: pathology and genetics of tumours of the lung, pleura, thymus and heart. Lyon: IARC Press; 2004. p.9-124.
- Solis LM, Behrens C, Raso MG, et al. Histologic patterns and molecular characteristics of lung adenocarcinoma associated with clinical outcome. *Cancer* 2012;118:2889-2899.
- Ding L, Getz G, Wheeler DA, et al. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 2008;455:1069-1075.
- Kohno T, Ichikawa H, Totoki Y, et al. KIF5B-RET fusions in lung adenocarcinoma. *Nat Med* 2012;18:375-377.
- Takeuchi K, Soda M, Togashi Y, et al. RET, ROS1 and ALK fusions in lung cancer. *Nat Med* 2012;18:378-381.
- Lipson D, Capelletti M, Yelensky R, et al. Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. *Nat Med* 2012;18:382-384.
- Bergethon K, Shaw AT, Ou SH, et al. ROS1 rearrangements define a unique molecular class of lung cancers. *J Clin Oncol* 2012;30:863-870.
- Shigematsu H, Gazdar AF. Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. *Int J Cancer* 2006;118:257-262.
- Shigematsu H, Lin L, Takahashi T, et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J Natl Cancer Inst* 2005;97:339-346.
- Mok TS, Wu YL, Thongprasert S, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 2009;361:947-957.
- Soda M, Choi YL, Enomoto M, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature* 2007;448:561-566.
- Inamura K, Takeuchi K, Togashi Y, et al. EML4-ALK lung cancers are characterized by rare other mutations, a TTF-1 cell lineage, an acinar histology, and young onset. *Mod Pathol* 2009;22:508-515.
- Shaw AT, Yeap BY, Solomon BJ, et al. Effect of crizotinib on overall survival in patients with advanced non-small-cell lung cancer harbouring ALK gene rearrangement: a retrospective analysis. *Lancet Oncol* 2011;12:1004-1012.
- Inamura K, Takeuchi K, Togashi Y, et al. EML4-ALK fusion is linked to histological characteristics in a subset of lung cancers. *J Thorac Oncol* 2008;3:13-17.
- Dacic S, Shuai Y, Yousem S, Ohori P, Nikiforova M. Clinicopathological predictors of EGFR/KRAS mutational status in primary lung adenocarcinomas. *Mod Pathol* 2010;23:159-168.

29. Siegfried JM, Gillespie AT, Mera R, et al. Prognostic value of specific KRAS mutations in lung adenocarcinomas. *Cancer Epidemiol Biomarkers Prev* 1997;6:841-847.
30. Riely GJ, Kris MG, Rosenbaum D, et al. Frequency and distinctive spectrum of KRAS mutations in never smokers with lung adenocarcinoma. *Clin Cancer Res* 2008;14:5731-5734.
31. Massarelli E, Varella-Garcia M, Tang X, et al. KRAS mutation is an important predictor of resistance to therapy with epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancer. *Clin Cancer Res* 2007;13:2890-2896.
32. Tang X, Kadara H, Behrens C, et al. Abnormalities of the TTF-1 lineage-specific oncogene in NSCLC: implications in lung cancer pathogenesis and prognosis. *Clin Cancer Res* 2011;17:2434-2443.
33. Whitsett J. A lungful of transcription factors. *Nat Genet* 1998;20:7-8.
34. Perner S, Wagner PL, Soltermann A, et al. TTF1 expression in non-small cell lung carcinoma: association with TTF1 gene amplification and improved survival. *J Pathol* 2009;217:65-72.
35. Kerr KM. Pulmonary preinvasive neoplasia. *J Clin Pathol* 2001;54:257-271.
36. Westra WH. Early glandular neoplasia of the lung. *Respir Res* 2000;1:163-169.
37. Wistuba II. Genetics of preneoplasia: lessons from lung cancer. *Curr Mol Med* 2007;7:3-14.
38. Park IW, Wistuba II, Maitra A, et al. Multiple clonal abnormalities in the bronchial epithelium of patients with lung cancer. *J Natl Cancer Inst* 1999;91:1863-1868.
39. Wistuba II, Mao L, Gazdar AF. Smoking molecular damage in bronchial epithelium. *Oncogene* 2002;21:7298-7306.
40. McCaughan F, Pole JC, Bankier AT, et al. Progressive 3q amplification consistently targets SOX2 in preinvasive squamous lung cancer. *Am J Respir Crit Care Med* 2010;182:83-91.
41. McCaughan F, Pipinikas CP, Janes SM, George PJ, Rabbitts PH, Dear PH. Genomic evidence of pre-invasive clonal expansion, dispersal and progression in bronchial dysplasia. *J Pathol* 2011;224:153-159.
42. Wistuba II. Histologic evaluation of bronchial squamous lesions: any role in lung cancer risk assessment? *Clin Cancer Res* 2005;11:1358-1360.
43. Colby TV, Wistuba II, Gazdar A. Precursors to pulmonary neoplasia. *Adv Anat Pathol* 1998;5:205-215.
44. Kitamura H, Kameda Y, Ito T, Hayashi H. Atypical adenomatous hyperplasia of the lung. Implications for the pathogenesis of peripheral lung adenocarcinoma. *Am J Clin Pathol* 1999;111:610-622.
45. Osanai M, Igarashi T, Yoshida Y. Unique cellular features in atypical adenomatous hyperplasia of the lung: ultrastructural evidence of its cytodifferentiation. *Ultrastruct Pathol* 2001;25:367-373.
46. Westra WH, Baas IO, Hruban RH, et al. K-ras oncogene activation in atypical alveolar hyperplasias of the human lung. *Cancer Res* 1996;56:2224-2228.
47. Tominaga M, Sueoka N, Irie K, et al. Detection and discrimination of preneoplastic and early stages of lung adenocarcinoma using hnRNP B1 combined with the cell cycle-related markers p16, cyclin D1, and Ki-67. *Lung Cancer* 2003;40:45-53.
48. Nakanishi K, Kawai T, Kumaki F, Hiroi S, Mukai M, Ikeda E. Survivin expression in atypical adenomatous hyperplasia of the lung. *Am J Clin Pathol* 2003;120:712-719.
49. Yatabe Y, Kosaka T, Takahashi T, Mitsudomi T. EGFR mutation is specific for terminal respiratory unit type adenocarcinoma. *Am J Surg Pathol* 2005;29:633-639.
50. Tang X, Shigematsu H, Bekele BN, et al. EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients. *Cancer Res* 2005;65:7568-7572.
51. Tang X, Varella-Garcia M, Xavier AC, et al. Epidermal growth factor receptor abnormalities in the pathogenesis and progression of lung adenocarcinomas. *Cancer Prev Res (Phila)* 2008;1:192-200.
52. Rawlins EL, Okubo T, Que J, et al. Epithelial stem/progenitor cells in lung postnatal growth, maintenance, and repair. *Cold Spring Harb Symp Quant Biol* 2008;73:291-295.
53. Kim CF, Jackson EL, Woolfenden AE, et al. Identification of bronchioalveolar stem cells in normal lung and lung cancer. *Cell* 2005;121:823-835.
54. Eramo A, Lotti F, Sette G, et al. Identification and expansion of the tumorigenic lung cancer stem cell population. *Cell Death Differ* 2008;15:504-514.
55. Sullivan JP, Spinola M, Dodge M, et al. Aldehyde dehydrogenase activity selects for lung adenocarcinoma stem cells dependent on notch signaling. *Cancer Res* 2010;70:9937-9948.
56. Kadara H, Wistuba II. Field cancerization in non-small cell lung cancer: implications in disease pathogenesis. *Proc Am Thorac Soc* 2012;9:38-42.
57. Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer* 1953;6:963-968.
58. Nelson MA, Wymer J, Clements N Jr. Detection of K-ras gene mutations in non-neoplastic lung tissue and lung cancers. *Cancer Lett* 1996;103:115-121.
59. Spira A, Beane J, Shah V, et al. Effects of cigarette smoke on the human airway epithelial cell transcriptome. *Proc Natl Acad Sci U S A* 2004;101:10143-10148.
60. Spira A, Beane JE, Shah V, et al. Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nat Med* 2007;13:361-366.
61. Gold KA, Kim ES, Lee JJ, Wistuba II, Farhangfar CJ, Hong WK. The BATTLE to personalize lung cancer prevention through reverse migration. *Cancer Prev Res (Phila)* 2011;4:962-972.
62. Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. *Nat Rev Genet* 2010;11:685-696.
63. Beane J, Vick J, Schembri F, et al. Characterizing the impact of smoking and lung cancer on the airway transcriptome using RNA-Seq. *Cancer Prev Res (Phila)* 2011;4:803-817.