INTRODUCTION

Lung cancer is the number one cause of cancer-related death in the Western world. Lung cancer remains the leading cause of cancer mortality for both men and women in the United States. Its incidence is highly correlated with cigarette smoking, and lung cancer is eventually diagnosed in about 10% of long-term smokers. Among the 10% to 15% of patients without a smoking history in whom lung cancer develops, environmental or inherited causes of lung cancer contribute to the risk. As large as the clinical challenges patients and physicians face in managing this disease, equally enormous are the challenges in detailing the complex molecular pathogenesis of lung cancer. Major progress has been made, however, and we are beginning to see this knowledge translated into the clinic. In this chapter, we summarize our understanding of how lung epithelial cell injury leads to signaling pathway activation that triggers the uncontrolled proliferation, resistance to apoptosis, metastasis, and escape from immune surveillance seen in lung cancer. Whereas all subtypes of lung cancer have historically been divided between non–small cell lung cancer (NSCLC) and small cell lung cancer (SCLC), the field is stepping away from purely histologic classification and moving into molecular classification of tumors based on better understanding of driver genes and the potential benefit of therapeutic strategies using targeted therapies that interfere with mechanisms of tumor growth and progression.

The pathogenesis of lung cancer involves the accumulation of multiple molecular abnormalities over a long period of time. Alterations in gene expression can result from abnormal microRNA expression or methylation, DNA sequence changes, DNA segment amplification, deletion, or whole chromosome gains or losses. These lesions allow cells to escape the normal regulation of cell division, apoptosis, and invasion and/or alter its interaction with the host. The historical focus of much of this research has been to identify and study the role of specific genetic abnormalities in tumor cells related to chromosomal abnormalities, inactivation of specific tumor suppressor genes, the activation of specific oncogenes, and the expression of hormone receptors and growth factor production associated with the development of cancer. More recently, the contribution of stromal interactions, the induction of angiogenesis, the control of apoptosis, and epigenetic phenomena such as posttranslational modification of critical genes have been the subject of intense research. The completion of the first draft of the human genome sequence and the availability of high-throughput technologies (e.g., microarrays) have also prompted investigators to propose studies to discover...
molecular alterations of individual tumors with incredible sensitivity and unprecedented scope. Common molecular abnormalities in both precancerous and invasive lung cancers are being integrated and hold the promise for potential utility in prevention, early detection, and therapy.

PREDISPOSITION TO LUNG CANCER

SMOKING

Tobacco use continues to be a major cause of cancer in the world. Worldwide, tobacco use causes more than 5 million deaths per year, and current trends show that tobacco use will cause more than 8 million deaths annually by 2030. Tobacco control efforts are leading to a decrease in cancer mortality in some areas of the world, but globally there are still more than 1 billion smokers. More than 85% of all lung cancers are attributable to cigarette smoking; however, lung cancer develops in only a fraction of long-term cigarette smokers, suggesting a role for variable inherited genetic susceptibility in lung tumorigenesis.

Lung cancer develops through a multistage process in a background of increasing genomic instability and inflammation. For example, cytogenetic alterations arise throughout the airways with monoclonal and trisomic clonal patches observed. 

Table 51-1 Characteristics of Lung Cancer Found in Never-Smokers

<table>
<thead>
<tr>
<th>Condition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peripheral disease in distal airways</td>
<td>Adenocarcinoma</td>
</tr>
<tr>
<td>Female</td>
<td>Younger age</td>
</tr>
<tr>
<td>Environment tobacco smoke, human papillomavirus 16 and 18 infection</td>
<td>Familial and genetic risk, Frequent EGFR and TP53 mutations, Rare KRAS mutations</td>
</tr>
</tbody>
</table>

Genetic susceptibility may be seen with rare autosomal dominant genes that explain cases of early-onset lung cancer but more common genetic variants or polymorphisms are more likely to affect lung cancer risk. Inherited differences in DNA repair capacity in response to tobacco-mediated damage have been implicated as underlying some of this susceptibility. When lymphocytes from lung cancer patients and age-matched controls are exposed to bleomycin, the lymphocytes from lung cancer cases develop more chromatid breaks than the control lymphocytes.

A similar assay has been developed using benzo(a)pyrene diol-epoxide, a reactive substrate that is derived by in vitro processes from benzo(a)pyrene, a major carcinogen in tobacco smoke. These results suggest that the DNA repair capacity influences risk for lung cancer. Telomere length has also been inversely associated with lung cancer risk. Nicotinic acetylcholine receptor subunit genes may contribute to the risk of smoking as well as the risk of lung cancer. In three lung cancer genome-wide association studies, polymorphic differences in chromosome 15q25 have been associated with the risk of developing lung cancer. Single nucleotide polymorphisms mapped to a region of strong linkage disequilibrium within 15q25.1 contain nicotinic acetylcholine receptor subunit genes. The genetic variants were associated both with nicotine dependence as well as with certain lung cancer phenotypes, including lung cancer developing in younger patients and in patients with lower amounts of smoking: the variants were not associated with lung cancer in nonsmokers or with other smoking-related cancers like renal or bladder.

Since then, subsequent analyses have identified additional variants that influence lung cancer risk, which will require further investigation in a larger sample size. The relative impact of the nicotinic acetylcholine receptor gene variants on the propensity to smoke or on the response to the direct carcinogenic effects of smoke is undergoing validation in large numbers of smokers and nonsmokers.

Familial risks of lung cancer provide another route for uncovering susceptibility genes. In one genome-wide linkage analysis, a genomic locus on chromosome 6q23-25 has now been associated with a familial risk of lung cancer.
To date, no tumor suppressor genes inactivated by mutation have been identified in this locus, although the frequent inactivation of multiple candidate tumor suppressor genes within chromosome 6q likely contributes to development of sporadic lung cancer. Finally, a genome-wide association analysis was conducted to investigate associations between single nucleotide polymorphisms and the risk of lung cancer in 194 patients with familial lung cancer and 219 cancer-free control subjects. A strong association was found between common sequence variants at 15q24-25.1 and lung cancer. The risk of lung cancer was more than fivefold higher among those subjects who had both a family history of lung cancer and two copies of high-risk alleles in the 15q24-25.1 locus as compared with control subjects. Thus, studies into sporadic and familial cancer risk may identify some of the same genetic predispositions, such as 15q24-25.1 or germline EGFR T790M mutations, and have the potential to identify novel genetic predispositions.

Mathematical models integrating clinical, physiologic, imaging, and biologic variables are now proposed for evaluating risk for lung cancer in clinical practice. This field has made tremendous progress, and prediction models are being proposed for implementation in selecting populations for screening and/or chemoprevention trials.

### EARLY EVENTS IN LUNG TUMORIGENESIS

#### FIELD CANCERIZATION EFFECT

Whereas lung cancer may originate from only one or a few airway epithelial cells, it is clear that exposure of the whole airway mucosa to tobacco smoke can cause the entire bronchial tree to be at increased risk for development of lung cancer, leading to the concept of field cancerization (Fig. 51-1). Field cancerization was first proposed in the 1950s, and its molecular correlates were later confirmed in the airways of human smokers. Field cancerization is also demonstrated by the elevated Ki-67 labeling index in the airways of smokers at more than one site. In addition, evidence has been found in a single smoker with diffuse dysplasia that a single identical point mutation of TP53 could be found diffusely in both lungs, but not in the blood or other solid organs, suggesting that a single progenitor bronchial epithelial clone populated the bronchial mucosa. Although the risk of developing lung cancer increases with the presence of such preinvasive lesions, clearly the entire at-risk epithelium does not undergo malignant transformation, and no one has identified the molecular determinants of preinvasive lesions that may predict irreversible progression to lung cancer. Thus, if one considers the total number of bronchial epithelial cells and the proliferation rate of patches of clonal abnormalities, cancer remains a rare event.

The concept of field cancerization has implications for approaching diagnosis, prevention, and treatment of lung cancer. For example, diagnostic information may be contained even in the normal-appearing airway epithelium; in patients with and without lung cancer, changes in the airway field have been found to correlate with the presence of lung cancer elsewhere in the airways or with an elevated risk for transformation. Similarly, prevention efforts should be directed to the entire epithelium, which can then also provide biomarkers for effective prevention. Finally, the field cancerization concept has major implications for therapy, because one should consider treating the entire field as opposed to limiting treatment efforts to preinvasive lesions that may never develop into lung cancer.

#### GENOMIC INSTABILITY

Lung cancer develops along a multistage process in a setting of increasing genomic instability. Genomic instability is universally found during accumulation of these genetic abnormalities. Genome instability is a fundamental characteristic of cancer initiation and progression. However, our understanding of the onset of instability, the rate of instability, and the mechanisms leading to instability is far from complete. We do know that—from initiation, to promotion, to the development of an invasive phenotype—a series of molecular changes contribute to a molecular instability that puts the airway epithelium of some smokers at greater risk for lung cancer.
Two main theories of the genomic changes underlying lung tumorigenesis have emerged: the stochastic (random) and the gene-centric (nonrandom) theories. The stochastic theory argues that lung cancer arises by genomic aberrations that are stochastic or random in nature. Indeed, in many tumor types, there are relatively few cancer-specific mutations, whereas there can be a high degree of chromosomal aberrations. Some of these low-level aberrations have been called noise but may reflect the underlying instability. Nonclonal chromosomal alterations such as defective mitotic figures, chromosomal fragmentation, or noncurrent genomic alterations indicate a dynamic process leading to instability. Nonclonal chromosomal alterations may provide a survival advantage, whereas clonal chromosomal alterations at a later stage of progression may confer a growth advantage. In contrast, the gene-centric theory argues that sequential accumulation of epigenetic and genetic aberrations is important for lung cancer development. This gene-centric theory is particularly appealing to explain clonal alterations in genes implicated in tumor cell growth and survival that represent possible targets for the management of cancer. These alterations include single nucleotide point mutations, changes in chromosome copy number (aneuploidy), and specific genomic amplifications or deletions. The genetic changes are implicated in the pathogenesis of tumor development in part through the activation of oncogenes and inactivation of tumor suppressor genes and are considered one of the key hallmarks of cancer. Some genomic signatures seem to persist after tumor development and throughout their progression and their histologic differentiation. The fact that smoking history leads to specific genomic alterations across different tumors also suggests a specific pathogenesis with genomic alteration resulting from a series of dysregulations in DNA repair mechanisms and chromosomal segregation.

The two theories are not necessarily mutually exclusive, and both random and nonrandom changes may be important. The progressive accumulation of mutations, loss of apoptotic control and regulation of cell proliferation, and the appearance of an abnormal number of chromosomes (aneuploidy) are associated with worsening dysplasia and may reflect underlying dysregulation of mechanisms controlling genomic fidelity. In some lung tumors, mismatch repair deficiency leads to microsatellite instability at the nucleotide sequence level, whereas in other tumors, aneuploidy is the dominant feature. Specific defects in DNA repair may also be present. Polymorphisms in DNA repair genes XPD (codon 312 Asp/Asp vs. Asp/Asn) are associated with impaired efficiency of DNA repair and apoptotic function in lung cancer. Measures of genomic instability parallel rates of loss of heterozygosity and accumulation of other genomic abnormalities. In the airways, progressively more severe and more frequent abnormalities are seen in preinvasive lesions. The progressive accumulation of genomic abnormalities associated with clonal growth among populations of tumor cells is well described. Recent reports use array comparative genomic hybridization (aCGH) to assess genomic instability; in these reports, in addition to numerical chromosome changes, smaller aberrations are found at specific chromosome loci. Clearly, genomic instability itself drives malignant progression, and this instability may itself be a target for prevention or therapy of lung cancer.

**MUCOSAL RESPONSE TO INJURY, THE EMERGENCE OF CRITICAL MUTATIONS**

Since the late 1970s, somatic mutations have been identified and associated with the development of cancer. These mutations, involving those of tumor suppressor genes or oncogenes, may or may not be critical events in the formation of the cancer. Epidemiologic data support the hypothesis that clones of cells accumulate several key mutations during oncogenesis. Proposed by Loeb, the model of the mutator phenotype, a phenotype in which mutations develop faster than in normal cells, suggests that this propensity for rapid mutation may be required for multistage carcinogenesis and may be acquired by cells early on. This phenotype may also be hereditary.

DNA mutations may arise from a failure to repair DNA damage. In addition to damage and mutations due to environmental carcinogens, spontaneous errors of replication attributed to DNA polymerase take place at a rate of 1/10,000 to 1/100,000 base pairs. These intrinsic mutations may also be an important component underlying genomic instability and eventually tumor development. A recent analysis of 623 genes in 188 lung adenocarcinomas revealed more than 1000 somatic mutations, and those in 26 different genes were found to be mutated at significant frequencies. Many of these mutated genes fall into a handful of common signaling pathways. We briefly comment on a few examples of these commonly acquired mutations in KRAS, TP53, PIK3CA, epidermal growth factor receptor (EGFR), and anaplastic lymphoma kinase (ALK).

KRAS mutations are among the most common mutations found in lung cancer. The mutations are found in 30% of adenocarcinomas of the lung, but much less frequently in other subtypes. Mutations in KRAS lead to activation of several pathways. Ras proteins activate the RAF/MEK/ERK pathway, which mediates cell growth and cell cycle entry by phosphorylation of transcription factors (e.g., such as c-FOS, ERK1, and MYC); phosphorylation of the RSK (ribosomal protein S6 kinase) and MNK (mitogen-activated protein kinase [MAPK]-interacting serine/threonine kinase) family of kinases; and the phosphatidylinositol-3-kinase (PI3K)/protein kinase B (AKT) pathway, which controls cell survival, cell growth, and metabolism. Mutant KRAS (most frequently, codon 12 G-T transversions) can transform airway epithelial cells by activating the extracellular signal-regulated kinase (ERK)-MAPK and PI3K/AKT pathways. Because the KRAS mutation is found early in alveolar atypical hyperplasia, a presumed precursor lesion to adenocarcinomas, KRAS mutation may be an important step in the genesis of this subtype of lung cancer. The fact that mutant ras transgenic mice develop adenocarcinomas of the lung supports this hypothesis.

TP53 is a prototype tumor suppressor gene that is the most common genetic lesion in human cancers and is thus well suited for analysis of the mutational spectrum in human cancers. The gene product p53 is a potent regulator of cell growth and DNA damage responses, and directly regulates the cell cycle via induction of p21. Mutations of
**PART 3  •  Clinical Respiratory Medicine**

**TP53** are most commonly seen in squamous carcinoma and small cell carcinoma of the lung. Mutations predominantly representing G to T transversions consistent with causation by bulky DNA adducts such as the polycyclic hydrocarbons are frequently found in the lungs of smokers.\(^{23}\) When the gene is mutated, the p53 protein can function as an oncogene. Mutated p53 exhibits a prolonged half-life, and the protein can be found to be overexpressed in about 50% of lung cancers by immunohistochemistry.\(^{24}\) Although not consistently associated with prognostic significance, there is little doubt that TP53 mutations play a key role in tumor development by dysregulation of cell cycle control and apoptosis.

p16, a cyclin-dependent kinase inhibitor, functions as a tumor suppressor and critical member of the retinoblastoma (Rb) cell cycle control pathway. The gene is inactivated in more than 40% of NSCLCs. Previous studies have demonstrated that point mutations, deletion or loss of heterozygosity (when one parental copy of a region is lost) on 9p21, or hypermethylation of the gene provides alternate mechanisms of inactivation in 30% to 50% of NSCLCs. In smokers, the loss of P16 appears to be from different mechanisms than in nonsmokers. In tumors arising in smokers, P16 was found to be lost due to point mutations or homozygous deletions, whereas in tumors from nonsmokers, P16 was lost via promoter hypermethylation.\(^{25}\) The relationship between the loss of P16 and tobacco points to P16 as another genetic target of cigarette smoke in the pathogenesis of lung cancer.

BRAF is a serine/threonine kinase that belongs to the RAS/RAF/MEK/ERK/MAPK pathway, and is important for transducing mitogenic signals from the cell surface. BRAF mutations were found in thyroid, colorectal, and lung cancer as well as in a majority of malignant melanomas.\(^{31}\) The percentage of BRAF-mutant lung cancers appears to be less than 5% and to be limited to lung adenocarcinomas.\(^{26}\) This observation has led to drug discovery programs dedicated to specific targeting of BRAF-dependent tumors. Trials investigating the efficacy of these agents are ongoing in several tumor types.\(^{31}\) Whether BRAF-mutant lung cancer can be successfully treated with inhibitors targeting downstream pathway members of BRAF remains to be established.

**PIK3CA**, the gene encoding the catalytic subunit of the class-1a phosphoinositide-(3,4,5)-kinase, is an important component of the AKT antiapoptotic pathway. Mutations in this gene have been found in a large fraction of epithelial cancers.\(^{34}\) Although the most frequently mutated oncogene in breast cancer, PIK3CA was found to be mutated in less than 5% of lung cancers. In several functional analyses, these mutations were found to be oncogenic by their ability to activate the AKT survival pathway.\(^{35}\) Most recently, preclinical studies involving new generation PI3K-kinase inhibitors suggest that tumors carrying these mutations might be exquisitely sensitive to pathway inhibition.\(^{36}\)

The gene for the receptor tyrosine kinase EGFR is mutated in approximately 10% of white and 40% of East Asian patients with NSCLC,\(^{67,69}\) leading to constitutive activation of this growth factor receptor. These mutations are more common in patients who have never smoked, whose tumors are of adenocarcinoma histology, and who are of East Asian ethnicity. Exon resequencing in responders to the tyrosine kinase inhibitors, gefitinib or erlotinib, revealed that most of them had deletion mutations in exon 19 or point mutations in exon 21 of the EGFR kinase domain, but patients not responding to these drugs only rarely had these mutations.\(^{69}\) Inhibition of mutant EGFR is an example of successful therapeutic targeting of oncogenically activated tyrosine kinases in cancer. The presence and type of EGFR mutations indicate which patients will respond to therapy with EGFR inhibitors. Trials selecting patients on the basis of EGFR mutations have yielded response rates to EGFR inhibition therapy exceeding 70% and led to median overall survival exceeding 20 months.\(^{91-93}\)

Even though dramatic responses are frequently seen in patients with the common EGFR mutations, most patients initially responding to EGFR inhibitors eventually relapse. These relapses may be explained by additional mutations in EGFR or activation of other growth pathways. In certain cases, resistance has been found to be due to an activating mutation of EGFR at a second site. This mutation, T790M, is analogous to the T315I mutation of the murine leukemia viral oncogene homolog (ABL) that causes acquired resistance in chronic myelogenous leukemia patients who had initially responded to imatinib.\(^{94,95}\) An additional mechanism of acquired resistance is the de novo amplification of the MET receptor tyrosine kinase gene.\(^{96,97}\) Understanding the sources of acquired resistance may reveal strategies for prevention of resistance.

The ALK is a receptor tyrosine kinase that is aberrant in a variety of malignancies. ALK was originally discovered in anaplastic large cell lymphoma as part of a chromosomal translocation t(2;5), which fuses the C-terminal kinase domain of ALK encoded on chromosome 2p23 to the N-terminus of nucleophosmin on chromosome 5q35.\(^{98}\) Subsequently, a variety of ALK fusion proteins have been found in multiple malignancies, including inflammatory myofibroblastic tumor and NSCLC.\(^{99,100}\) ALK fusion proteins are transforming and are highly susceptible to ALK inhibitors.\(^{100}\) Indeed, the ALK tyrosine kinase inhibitor (TKI) crizotinib was recently approved for the treatment of ALK fusion-positive lung cancer.\(^{100}\) Because crizotinib has “off-target” activity against the ROS1 kinase, somatic activating ROS1 fusions may also be susceptible to crizotinib.\(^{101}\) Rearrangement involving another kinase, RET, may be targeted with RET TKIs.\(^{102}\) Because of their ability to guide therapeutic decision making, EGFR, KRAS, ALK, ROS1, RET, BRAF and PIK3CA mutation testing have become part of standard pathologic analysis of lung adenocarcinomas in many centers.

The efforts to characterize the genomes of major human cancer types\(^{104}\) have revealed that, beyond known cancer-related genes, many additional genes are mutated in individual cancers. However, most of these mutations are found in only a single tumor, suggesting that each tumor contains an individual set of mutations contributing to tumorigenesis. Ding and coworkers\(^{13}\) found between zero and 40 mutations in individual lung adenocarcinomas, with an average of 24.3 for those with a mutant PRKDC gene (involved in DNA repair) and 4.7 for those without this mutation. The average number of mutations for different tumors range significantly from 0.39 mutations per one million bases (mut/MB) to the highest in small cell lung cancer (7.34 mut/MB) followed closely by melanoma.\(^{105-108}\)
Mutagens

Cigarette smoking may account for 85% of lung cancers. Approximately 1 in 10 life-smokers will develop lung cancer, suggesting individual differences in susceptibility. Environmental carcinogenesis resulting from tobacco smoke exposure is a complex process that can involve activation of procarcinogens that lead to DNA adduct formation and subsequent failure of DNA repair, which should normally remove these adducts. Studies comparing DNA repair capacity among newly diagnosed lung cancer patients and age-matched controls indicate significant differences between the two groups.

The majority of lung cancers are now diagnosed among ex-smokers. This suggests that the accumulation of molecular damage during cigarette exposure initiates a cascade of events that lead to cancer even decades after smoking cessation. Risk factors for lung cancer include smoking, including total consumption, age of initiation, and years of smoking; occupational and environmental exposure (asbestos, uranium, radiation), diet (vitamin A, vitamin E, cholesterol), and host (familial aggregation) and genetic factors (see Chapter 52). Cigarette smoke is a complex mixture and includes substances that are responsible for DNA adduct formation such as polycyclic aromatic hydrocarbons, aromatic amines, and tobacco-specific nitrates. These form DNA adducts that may escape normal adduct repair mechanisms and result in heritable alterations in DNA sequence. For example, the conversion of G-C base pairs to T-A is involved in the activation of the KRAS oncogene and inactivation of the TP53 tumor suppressor gene. The activated form of benzo(a)pyrene is the diol epoxide benzo(a)pyrene that can cause DNA adducts, leading to point mutations and single-strand chromatid breaks. A major concern is that people who start smoking at young ages seem to have greater amounts of permanent DNA alterations than smokers who start smoking later.

ROLE OF INFLAMMATION IN LUNG TUMORIGENESIS

The tumor cell signaling pathways that trigger the uncontrolled proliferation, resistance to apoptosis, metastasis, and escape from immune surveillance are partially understood. In contrast, how inflammation and its control may participate in lung tumorigenesis remains more poorly understood.

The reason why some smokers develop chronic obstructive pulmonary disease (COPD), some develop lung cancer, and some develop both diseases remains unclear. Cigarette smoke contains high concentrations of oxidants and free radicals, together with thousands of particles. Local antioxidant and metabolizing enzymes inactivate many toxic species. Nuclear factor-KB activation and subsequent trans-activation of inflammation-related genes appear to play a central role in both COPD and cancer. How some smokers remain free of disease while others develop COPD or cancer may be determined by the activation of genes in response to cigarette smoke. Following exposure to the same toxic smoke, individuals appear to progress toward disease along different paths. In those who develop lung cancer, genomic instability develops that causes further chromosomal abnormalities resulting in clonal expansion of cells that have a growth advantage, whereas in those who develop COPD, an intense immune response and further inflammation predominate. The process by which these events diverge is unclear. More likely, heritable genetic polymorphisms that influence susceptibility to DNA or connective tissue damage, efficiency of DNA or connective tissue repair, the intensity of immune responses to constituents of tobacco smoke, intrinsic or acquired genomic instability, or the induction of factors that suppress immune surveillance likely determine the disease pathway taken.

Toll-like receptors (TLRs), which recognize a variety of pathogen-associated molecular patterns, are centrally involved in the initiation of the innate and adaptive immune responses. Recent evidence shows that functional TLRs are also expressed in a wide variety of tumors, suggesting that TLRs may play important roles in tumor biology, and are being evaluated as therapeutic targets. Activation of tumor cell TLRs not only promotes tumor cell proliferation and resistance to apoptosis but also enhances tumor cell invasion and metastasis by regulating metalloproteins and integrins. Moreover, the activation of TLR signaling in tumor cells induces the synthesis of proinflammatory factors and immunosuppressive molecules, which enhances the resistance of tumor cells to attack by cytotoxic lymphocytes and leads to immune evasion. Thus, the TLR signaling pathways may be usurped by the neoplastic process to advance cancer progression, which suggests that targeting tumor TLR signaling pathways may open novel therapeutic avenues.

Inflammation and tumor-infiltrating inflammatory cells have been shown to induce and help sustain tumor angiogenesis and sustain cellular proliferation. COPD and its underlying chronic airway inflammation provide support of lung cancer growth. Chemokines, a component of cancer-related inflammation, predispose to cancer by affecting multiple pathways of tumor progression, including leukocyte recruitment and function; cellular senescence; tumor cell proliferation and survival; invasion and metastasis. The inflammatory system is expected to provide valuable targets for the development of innovative therapeutic strategies.

ROLE OF VIRUSES IN LUNG TUMORIGENESIS

Viruses can cause lung cancer in animal models; such viruses, including SV40 large T antigen and polyomavirus large and middle T antigens, cause lung cancer in transgenic models. No common respiratory viruses have been conclusively incriminated in the development of human lung cancer, but several have been implicated. Human papillomavirus, for example, has been associated with lung cancer and, in particular, lung cancer arising in women. These results remain controversial because the Cancer Genome Atlas, a coordinated effort supervised by the National Cancer Institute to apply high-throughput genome analysis to the study of cancers, has not reported viral sequences in tumors. Similarly, controversy persists on the role of other viruses. Simian virus 40 has been suspected in the development of mesothelioma. Epstein-Barr virus has been suspected to be involved in the development of papillomas, mesotheliomas, and lymphomas of the lung.
the Jaagsiekte sheep retrovirus or a variant has been thought to cause adenocarcinoma in situ in humans as it does in sheep.124 However, many polymerase chain reaction–based assays have attempted to correlate bronchogenic carcinomas with respiratory viruses without success. Recent advances in proteomics may be useful in studying the role of viral infection in airway epithelial cell transformation. The proteomic analysis of tumors may allow the identification of peptide sequences specific to pathogens otherwise unknown in tumorigenesis.

In recent reports, infectious agents associated with diseased human tissues have been discovered by high-throughput sequencing technologies.125 However, identification of a nucleic acid sequence in a sample does not necessarily implicate an organism as causal in a disease. Viral nucleic acid may be found because of contamination of tissues, because of a subclinical infection unrelated to the disease, or because it is latent in human tissue but not actively causing illness (e.g., human herpesvirus and human papillomavirus).

NEUROENDOCRINE TUMOR DEVELOPMENT, GENOMICS OF SCLC

SCLC develops from neuroendocrine cells in the lung. The evidence for this is best established from mouse models with deletion of Rb and TP53.126,127 In these studies, mouse SCLCs often arise in the lung epithelium, where neuroendocrine cells are located, and the majority of early lesions were composed of proliferating neuroendocrine cells. When Rb and TP53 were deleted in nonneuroendocrine lung epithelial cells, mice did not develop SCLC.128,129

Recent publications offer new molecular characterization of SCLC. Using comprehensive genomic analysis of SCLCs130,130a with high-throughput sequencing techniques, Rudin and colleagues have demonstrated that SCLCs are composed of multiple known and new molecular abnormalities. They confirmed known SCLC molecular alterations to genes such as TP53, RB1, PIK3CA, CDKN2A, and PTEN. In addition, they identified other genes coding for proteins with a wide array of cellular functions such as the RAS family regulator, chromatin-modifying enzymes, transcription regulators, kinases, protein phosphatases of G-protein coupled receptors. Further genetic mutational clustering identified specific pathways at the epicenter of SCLC pathogenesis, which included PI3K, Hedgehog, NOTCH, mediator complex, glutamate receptor, DNA repair, and the SOX family. This study showed a high mutational frequency of several SOX genes and overexpression of SOX2 in close to 25% of the SCLCs analyzed. Lastly, they identified new mutations to genes coding for tyrosine kinases such as FLT1, FLT4, KDR, and KIT as well as several fusion proteins such as RLF-MYC/L1 which could be potentially targeted therapeutically.

In another integrative genomic analysis of 99 SCLC tumors and lines, Feiler and coworkers confirmed recurrent losses of the tumor suppressor genes TP53 and RB1 and novel amplifications of FGFR1, MYC/L1, and MYCN in SCLC.131 These findings were found to be similar to findings in TP53/Rb1 conditional knockout mice that, in the absence of p53 and RB1 protein expression, developed several SCLCs expressing amplifications of MYC/L1, MYCN, and NF1B. Exome sequencing, transcriptome sequencing, and genome sequencing confirmed that SCLCs have a high mutation rate (7.4 protein-altering mutations per MB) that corresponds with tobacco-associated carcinogen exposure. From these studies, they identified TP53, RB1, CREBBP, EP300, SLIT2, MLL, COBL, and EPHA7 as potential oncogenic drivers, which were classified into five major groups: (1) receptor tyrosine kinases (RTK), (2) P13-kinase and p53 pathways, (3) cell cycle control, (4) histone modifiers, and (5) SLIT-ROBO signaling. Using reverse phase protein arrays, Byers and associates identified PARP1, a DNA repair protein and E2F1 coactivator, to be highly expressed at the mRNA and protein levels in SCLCs. PARP inhibition down-regulated key components of the DNA repair machinery and enhanced the efficacy of chemotherapy.132 Justifying future clinical studies evaluating PARP and EZH2 inhibition, together with chemotherapy or other agents. These gene-centric approaches will have to be complemented by genome control studies, studies aimed at understanding gene regulatory mechanisms by investigating the whole genome structure, including chromatin structure and epigenetic alterations, and function. Understanding both the genes and their control will be necessary to resolve the molecular networks explaining the heterogeneity found in tumors and their functional implications.

Small cell lung cancer has worse clinical outcomes and no currently proven personalized therapy. SCLC and NSCLC account for 15% to 20% and 80% to 85% of annual lung cancer cases, respectively. SCLC is clinically more aggressive with an untreated median survival of 2 to 4 months.133 Despite initial complete responses exceeding 50%, the best combinations of chemotherapy and radiation result in less than 5% of patients surviving 5 years from initial diagnosis.134 Efforts to identify driver mutations, gene amplifications, or signatures with clinical utility in SCLC have thus far not translated to new therapies. In addition, prognostic or diagnostic markers for SCLC are scarce.135,136 Over the past decade, several molecular abnormalities have been described as indicators of chemotherapy resistance and poor prognosis in NSCLCs.137 As described earlier, recent seminal studies have identified genomic aberrations in SCLC that will hopefully lead to successful therapeutic strategies.

Why is no targeted therapy available in SCLC? The discovery of new targets for early detection and for therapeutics has been hindered by the fact that SCLC is often detected late and tissue is rarely obtained for study because surgical resection is not standard therapy. Another major shortcoming associated with the slow progress has been the lack of sensitive and robust technology to detect the signatures of cancer cells from minute quantities of available tissues or serum. Recently, advances in genomic technology provide high-resolution, high-throughput tools to measure copy number, gene expression, DNA methylation, and nucleotide sequence in cancer have started to move the field forward.138,139,140 These reports predict rapid evolution in the understanding and management of SCLC.

OTHER MOLECULAR ALTERATIONS DRIVING THE TUMOR PHENOTYPE

CHROMOSOMAL CHANGES

Cancer cells are characterized not only by mutations but also by a series of other chromosomal aberrations including
deletions and amplifications.\textsuperscript{118-140} Although chromosomal aberrations have been linked to most tumors and are a characteristic of cancer,\textsuperscript{64} chromosomal aberrations are recognized as increasingly complicated.\textsuperscript{11,141} Many of the observed alterations have been considered consequences instead of causes of lung cancer development. However, there are chromosomal regions with frequent losses found in regions coding for essential tumor suppressor genes and DNA repair genes that may be involved in the pathogenesis of several tumor types.\textsuperscript{142}

Chromosomal alterations in NSCLC have been measured by aCGH.\textsuperscript{86,87,143} Specific areas of amplification and deletion can distinguish squamous from adenocarcinoma of the lung and other subtypes. Among many areas of genomic abnormality, amplification of chromosomal region 3q26-3q29 was the most prevalent abnormality in squamous carcinoma of the lung. These studies confirm and refine previously documented copy number alterations in lung cancer using CGH analysis.\textsuperscript{84,144} and show particularly common amplified regions on chromosomal arms 1q, 3q, 5p, 8q, 11q, 12p, 17q, and 20q.

Chromosomal abnormalities may have a role in classifying tumors. Attempts have been made to look comprehensively at the genome to identify tumors with common groups of genetic features that might provide biologic or clinical guidance beyond traditional classification by light microscopy. Single nucleotide polymorphism arrays have been developed that are able to analyze loss or gain of genetic material at very high resolution\textsuperscript{57} and cancer genome sequencing efforts have uncovered common mutations.\textsuperscript{106,107,145,146} From genomic analyses, relatively small differences have been observed between squamous and adenocarcinoma of the lung. The most prevalent differences were found on chromosome 3q and include the p63 gene.\textsuperscript{144,147,148} Accumulation of specific chromosomal abnormalities has been correlated with clinical and pathologic data in NSCLC. Chromosomal abnormalities have been correlated with clinical outcome for a variety of cancers,\textsuperscript{60} but often, the genes or networks responsible for the observed biology are very difficult to define and are only partly understood.\textsuperscript{106,149,151}

Chromosomal abnormalities may underlie the progression of cancer. Loss of chromosomal regions on chromosomes 3p and 9p have been recognized as early events and identified in preinvasive lesions and in the normal-appearing epithelium of smokers.\textsuperscript{152,153} In contrast, \textit{TP53} and \textit{KRAS} mutations have been seen primarily in later stages of preneoplasia or frank invasive lesions.\textsuperscript{154} Amplification of large regions on the q arm of chromosome 3 has been characterized in invasive carcinomas\textsuperscript{4,64,144,155} and only recently in preinvasive lesions.\textsuperscript{147} Patterns of chromosomal change may be specific to cigarette smoke. In fact, lessons have been learned from the molecular analysis of lung cancers from patients with and without a history of smoking. For example, \textit{loss of heterozygosity} (LOH), the loss of a chromosome containing a wild-type gene when the gene on the other allele was previously mutated, at chromosome 3p14 was evaluated in smokers and ex-smokers. LOH at this site was found to be more frequent in current smokers (22/25 cases) than in former smokers (5/11 cases), and the high frequency in smokers correlated with a high metaphasia index.\textsuperscript{152,153} This implies that not only are these chromosomal changes frequent in normal-appearing bronchial epithelia but also that cells with these changes may regress after smoking cessation and be replaced by cells without this damage. Patterns of lung cancer allelic loss have been investigated in detail and new regions of allelic loss have been identified using high-throughput technologies.\textsuperscript{117} Interestingly, chromosome losses such as LOH and chromosomal gains are more prevalent at all sites in cancers from smokers than from nonsmokers, indicating a greater chromosomal instability in the cancers from smokers.\textsuperscript{15} aCGH has been used to discover specific patterns of genomic alterations found in NSCLC that might be related to smoking history. Smoking-related genomic signatures were identified in NSCLCs and could be predicted with an overall 88% accuracy.\textsuperscript{66,67,143} Lung tumors arising from current smokers had the greatest number of copy number alterations. The genomic regions most significantly associated with smoking were located within 32 regions and functionally associated with genes controlling the M phase of the cell cycle, the segregation of chromosomes, and the methylation of DNA. Understanding chromosomal abnormalities specific to cigarette smoke exposure may point to the etiology of smoking-induced lung cancer.

Specific translocation is another chromosomal abnormality found in lung cancer, but much less commonly than in hematologic or mesodermal tumors.\textsuperscript{156,157} Chromosomal translocations modify gene function through the deregulated expression of cellular proto-oncogenes without altering the structure of the protein product or by generating and expressing a chimeric protein with growth-promoting activities. Unbalanced translocations are also responsible for a large number of LOHs, which is a genetic event known to cause inactivation of tumor suppressor genes.\textsuperscript{159} Dang and colleagues\textsuperscript{159} identified a chromosome 19-15 balanced translocation associated with overexpression of NOTCH3, a protein that establishes an intercellular signaling pathway that plays a key role in development. The authors developed a transgenic mouse model overexpressing NOTCH3 causing neonatal mortality with a phenotype suggestive of alveolar cell hyperplasia. These data suggest that NOTCH3 overexpression prevents epithelial differentiation and this may play a significant role in promoting oncogenesis in a subset of lung cancers.\textsuperscript{160}

Similarly, recent findings are pointing towards new fusion genes such as \textit{ALK}, \textit{ROS1}, and \textit{RET} (discussed earlier) with rapid translation toward novel therapeutics in lung cancer.\textsuperscript{100,102,161-163}

**EPIGENETIC ALTERATIONS OF GENE EXPRESSION IN LUNG CANCER**

**DNA Adducts**

DNA adducts are covalent modifications of the DNA that result from exposure to specific carcinogens and thus, the level of DNA adducts in normal cells can serve as a biomarker for a significant exposure to carcinogens. In addition to being markers of carcinogen exposure, DNA adducts may directly alter regulation of transcription of tumor suppressors or oncogenes.\textsuperscript{156} Because DNA adduct levels in tumor tissue and in blood lymphocytes have been associated with lung cancer,\textsuperscript{165} and because these levels correlate with daily or lifetime cigarette consumption and do not
DNA adducts have been proposed as potential biomarkers of risk for lung cancer. In an attempt to identify risk factors associated with the level of DNA adduct accumulation, Wiencke and associates studied DNA adducts in current and former smokers and found that, for current smokers, the most important variable in determining the level of DNA adducts was the number of cigarettes smoked per day. In contrast, they found that for ex-smokers, the most important variable was age at initiation. Mechanisms responsible for the relationship between DNA adduct levels and age of initiation are unknown: the relative contribution of increased adduct formation at younger ages or perhaps decreased adduct removal by impaired DNA repair is yet to be determined. Prospective studies are needed to follow current and ex-smokers over time to determine the value of adduct levels in risk assessment.

DNA adducts have also been associated with risk of lung cancer. In a matched case-control study nested within the prospective Physicians' Health Study, there was an increased level of DNA adducts in active smokers who later developed lung cancer than in smokers who did not develop lung cancer. The level of adducts has also been found to be more prevalent among female than male smokers. Women smokers may be at higher risk of developing lung cancer for a given tobacco exposure and women also seem to accumulate aromatic/hydrophobic DNA adducts at a faster rate than men. (See Chapter 52.)

**DNA Methylation**

DNA methylation alters gene expression by increasing the density of methylation at promoter regions. In contrast to genetic mutations that require two hits to inhibit both alleles, aberrant methylation is a dynamic process over multiple division cycles and may cause increasing degrees of gene function loss over time. “CpG islands,” the major targets of DNA methyltransferase, are associated with the transcription start sites in almost half of human genes. Dense methylation of cytosines within CpG islands can cause heritable gene silencing. Genomic DNA hypomethylation, leading to genomic instability, as well as promoter hypermethylation, leading to inactivation of tumor suppressor genes, have been shown to be common events in human cancer. Gain of methylation in normally unmethylated CpG islands surrounding gene transcription start sites is an increasingly recognized and important means by which gene expression is altered in tumors. The genes affected include more than half of the tumor suppressor genes that are known to cause familial cancers. Aberrant methylation can begin early in tumor progression and cause loss of cell cycle control (e.g., p16), loss of mismatch repair function (e.g., MLH1), and loss of cell-cell interaction (e.g., E-cadherin). The exact mechanism by which hypermethylation may cause tumor progression is still unknown. In fact, there is still debate about whether methylation is a result of or a cause of gene function loss.

Promoter region hypermethylation has also been proposed as an excellent tumor marker. In lung cancer, methylated loci were found in both tumor and sputum DNA and have been detected in the sputum for up to 3 years before the diagnosis of cancer. Recently, methylation of the promoter region of four genes (TP16, CDH13, RASSFIA, and APC) in patients with stage I NSCLC was shown to be associated with early recurrence. The recent effort in The Cancer Genome Atlas also reported the landscape of alterations in methylation across a large number of lung cancers.

**Histone Deacetylation**

Histone deacetylation is another mechanism of epigenetic control. Histones are nuclear proteins that package DNA and allow ribosomal access to the DNA. Acetylation of histone tails on the nucleosome is associated with chromatin unfolding and increased regional transcriptional activity. Histone deacetylases modulate chromatin structure by regulating acetylation of core histone proteins. Deacetylation of histones is thus associated with compacting the DNA and repressing transcription. In lung cancer cell lines, for example, deacetylation of histone 3 correlated with refractoriness to retinoic acid, a phenomenon related to RARbeta promoter methylation in a subset of cell lines. Inhibitors of histone deacetylases have already been shown to decrease the level of a series of oncoproteins, suggesting a potential role as antitumor therapeutic agents.

**Regulation by miRNAs**

Micro-RNA (miRNA) are a large family of single-stranded noncoding RNAs that direct the posttranscriptional repression of protein-coding. In a study analyzing 104 pairs of NSCLC and corresponding normal lung tissues, an expression profile of 43 miRNAs discriminated lung cancers from noncancerous lung tissues. Six miRNAs (hsa-mir-205, hsa-mir-99b, hsa-mir-203, hsa-mir-202, hsa-mir-102, and hsa-mir-204-prec) were differentially expressed in adenocarcinomas when compared with squamous cell carcinomas. Furthermore, high hsa-mir-155 and low hsa-let-7a-2 expression correlated with poor survival in lung adenocarcinomas. In another study of 143 surgically resected NSCLCs, low let-7 expression was also significantly associated with shorter survival, whereas overexpression of let-7 in the A549 lung adenocarcinoma cell line inhibited lung cancer cell growth in vitro.

Differentially expressed miRNA genes in NSCLC are frequently located in fragile sites of chromosomes highly prone to mutation and/or chromosomal regions with frequent copy number alterations, suggesting that differences in miRNA expression may be induced by genomic alterations. Because greater than 50% of miRNAs are present at cancer-related chromosomal regions, miRNAs are also suspected to play a role as oncogenes or tumor suppressor genes. miRNA expression profiles represent potential markers for lung cancer diagnosis, classification, and prognosis. Methodologic issues related to the reproducibility of the analytic platforms should be resolved soon.

**PROTEOMIC ALTERATIONS**

Changes in DNA and in RNA may not be reflected in changes in protein expression. In fact, recent advances in protein profiling have suggested a poor correlation between gene expression and protein expression. It is now well established that protein activity is often highly regulated by reverse after smoking cessation. DNA adducts have been proposed as potential biomarkers of risk for lung cancer.
posttranslational modifications such as proteolysis and phosphorylation. Neither protein expression levels nor posttranslational modification can be assessed by genomic or cDNA microarray technologies, prompting interest in evaluation of protein expression, commonly referred to as “proteomics.”

Lung cancer has been studied by several proteomic methods including two-dimensional gel electrophoresis, mass spectrometry, and immunohistochemistry to identify biomarkers in tumors, or in biologic fluids such as bronchoalveolar lavage of patients with or without cancer. Matrix-assisted laser desorption ionization profiling is rapid, high throughput, but detects only the most abundant proteins of relatively low molecular weight and does not enable direct identification when applied to complex proteomes. Two-dimensional gel-based analysis suffers problems of interlaboratory reproducibility and throughput. Mass spectrometry has an even lower throughput, but yields confident identification of large numbers of proteins from every cellular compartment.

Advances in proteomic analysis of human samples are driving critical aspects of biomarker discovery and the identification of molecular pathways involved in disease etiology. Our group has used a standardized shotgun proteomic analysis method, which consists of identifying proteins in complex mixtures using a combination of high performance liquid chromatography and mass spectrometry, for profiling the two major subtypes of NSCLC and normal lung tissues (Fig. 51-2). With this approach, 3621 proteins were identified from the analysis of pooled human samples of squamous cell carcinoma, adenocarcinoma, and control specimens. In addition to proteins previously shown to be implicated in lung cancer, multiple new proteins were found of potential interest as therapeutic targets or diagnostic biomarkers, including some that were not identified by transcriptome profiling.

Up-regulation of these proteins was confirmed by multiple reaction monitoring mass spectrometry. This proteomic technology platform allows deep mining of lung tumor proteomes, enabling the identification of novel, previously undetected biomarker candidates and potential targets for therapy, such as SLC1A5, a neutral amino acid transporter that is responsible for more than 50% of glutamine transport into lung cancer cells. SLC1A5 is also located at the cytoplasmic membrane and its inhibition decreased cell growth and viability in lung cancer cells. Such approaches may uncover novel and hopefully useful biomarkers of lung cancer.

Figure 51-2 The shotgun proteomic method reproducibly identifies large numbers of proteins from clinical samples. Left, Heat map view of all identified proteins. Peptides were generated from three different histology pools including two control pools (Cont-1 and Cont-2), squamous cell carcinoma (SCC), and adenocarcinoma (ADC) pools. The number of protein groups (2863) found across all pools (top), are shown at the upper right side of the panel. Protein groups present only in the SCC and ADC tumor pools (598) (bottom), are shown at the lower right. Control: Noninvasive lung tissue. Right, We identified proteins in the SCC pool from four repeated experiments as indicated by the numbers 1 through 4 shown above each row. Out of the 5310 protein groups identified, 3303 (62.2%) protein groups were observed in every experiment (A), and 487 (9.2%) protein groups were observed in three experiments (B). The number of proteins observed in two experiments (C) and one experiment (D) was 597 (11.2%) and 923 (17.3%), respectively. The blue lines in each row represent identified protein groups. The intensity of each line is indicative of the spectra counts for each protein group with the scale shown at the bottom of the figure. (Adapted from Kikuchi T, Hassanein M, Amann JM, et al: In-depth proteomic analysis of nonsmall cell lung cancer to discover molecular targets and candidate biomarkers. Mol Cell Proteomics 11:916–932, 2012.)
Reverse phase protein arrays are a means to validate protein biomarkers in biologic specimens. The major advantage is that it allows rapid evaluation of known signaling pathways in lung cancer and it requires a minute amount of tissue. The limitation is that this method is a targeted one and dependent on the validation of specific antibodies.

**STRATEGIES TO DEEPEN OUR UNDERSTANDING OF LUNG CANCER**

**HIGH-THROUGHPUT PROFILING TECHNIQUES**

**Expression Arrays to Next Generation Sequencing**

Because DNA ultimately affects cellular behavior via proteins translated from the RNA, RNA expression patterns may be more relevant than either DNA copy number or epigenetic DNA changes. The microarray technology developed in the mid-1990s offers the hope that a fingerprint of the genetic expression of these tumors can be developed and then associated with clinical features. Beyond allowing for better classification of lung cancers, this technical advance in profiling gene expression opens a window into the world of tumor behavior (disease progression, recurrence, response to therapy) as well as to the mechanisms of tumor development. Tumor gene expression profiles are also influenced by the surrounding nonmalignant cells and so profiling both the tumor and non-tumor will allow for the study of the regulatory role of both entities.

Expression array data have provided new ways to distinguish and classify tumors. Selected genes allow the discrimination between primary lung cancer and metastasis of extrapulmonary sites. Studies of expression profiles of adenocarcinomas of the lung using commercially available gene chips or custom arrays have identified different classes of tumors, albeit with some overlap. For example, four classes of adenocarcinomas were found to have specific prognoses and expression signatures. These were characterized respectively by their (1) expression of cell cycle or proliferation genes, (2) expression of neuroendocrine markers, (3) expression of markers of alveolar origin, and (4) expression of ornithine decarboxylase or glutathione S-transferase. The neuroendocrine subclass was found to have a clinical outcome significantly worse than the others. These subclass differences may suggest possible new therapies targeted to these subsets and explain why tumors that superficially seem similar may respond quite differently to therapy. For example, when cDNA microarrays were used to study neuroendocrine tumors, a poor correlation was found between genes expressed in carcinoid and SCLC, suggesting the potential role for PARP1 and EZH2 in SCLC. In general, however, because of a lack of reproducibility between platforms and institutions, proteomic profiling, unlike genomic signature profiling, has not yet made a major impact in clinical practice.

Proteomic profiling in biological fluids or tissue samples has also been studied, but is not without challenges. The complexity of the sample composition and the predominance of few abundant proteins in the sample that can mask lower abundant proteins currently limit the utility of this approach. A newer serum proteomic platform has been tested in retrospective studies and shown to have the ability to predict whether or not NSCLC patients will benefit from the EGFR TKI erlotinib. Future efforts to analyze specific proteins or protein signatures in serum or other fluids has potential for diagnosis, prediction of therapeutic response and monitoring.

**Advanced Proteomic Strategies**

Mass spectrometry-based technologies are capable of very high throughput, allowing a sample to be analyzed in seconds and with a higher tolerance for salts, buffers, and other biologic contaminants. Because of these attributes, matrix-assisted laser desorption ionization mass spectrometry has been used to study proteins/peptides in serum, urine, tissue extracts, whole cells, and laser-captured microdissected cells. These profiling experiments have been applied to a series of biologic specimens. In one matrix-assisted laser desorption ionization mass spectrometry study from our group, hierarchical clustering of data from lung cancers and normal tissues allowed the identification of patterns distinguishing between tumor and normal as well as between histologic subgroups. Shotgun proteomic approaches allow the survey of increasing numbers of proteins and uncovers new molecular characteristics of lung cancers, as recently demonstrated in NSCLC. Using an integrative proteomic and transcriptomic analysis, molecular differences between SCLCs and NSCLCs have been identified, supporting the potential role for PARP1 and EZH2 in SCLC.

In general, because of a lack of reproducibility between platforms and institutions, proteomic profiling, unlike genomic signature profiling, has not yet made a major impact in clinical practice. Proteomic profiling in biological fluids or tissue samples has also been studied, but is not without challenges. The complexity of the sample composition and the predominance of few abundant proteins in the sample that can mask lower abundant proteins currently limit the utility of this approach. A newer serum proteomic platform has been tested in retrospective studies and shown to have the ability to predict whether or not NSCLC patients will benefit from the EGFR TKI erlotinib. Future efforts to analyze specific proteins or protein signatures in serum or other fluids has potential for diagnosis, prediction of therapeutic response and monitoring.

**MOLECULAR NETWORKS—SYSTEM BIOLOGY**

To determine the most informative molecular targets from high-throughput assays, increasingly sophisticated analytic tools are being developed. These analytic tools help define biologic processes, cellular components, or molecular pathways as well as provide gene ontology information about genes of interest; such analytic tools include WebGestalt, Pathway Studio, and Ingenuity. The Kyoto Encyclopedia of Genes and Genomes is a knowledge base for systematic analysis of gene functions in terms of the networks of genes and molecules. Although most pathway and network analysis tools are gene-centric, tools for combining and integrating multidimensional information at various levels are rapidly developing. Network analysis helps identify central genes or proteins, called hubs, that link many others in the network. One can assume that mutations in regulatory (hub) genes or proteins, for example, are more likely to cause disease than mutations in those less connected and peripheral in the network. These hubs may also be target sites for therapy. Therefore, quantitative systems biology applied to cancer offers a unique approach for defining the pathogenesis, and developing
Biomarkers may also provide prognostic and predictive information. Expression array signatures and serum proteomic assays have been discussed previously. Polymorphisms or expression of DNA repair genes (e.g., ERCC1 and RMM1) have been associated with predicting response to chemotherapy, but also have been shown to have significant practical obstacles in terms of the performance reliability of antibody-based immunohistochemical tests, raising questions about their clinical utility. Cyclooxygenase-2 expression and urinary prostaglandins have been associated with benefit from COX inhibitors and are being further studied.

PERSONALIZED MEDICINE AND MOLECULAR THERAPEUTICS

The biggest impact of biomarkers in NSCLC clinical care has been in testing patients with advanced or metastatic disease for the presence of “driver” genotypes, which can then be matched to corresponding targeted therapies (Fig. 51-4). The term driver genotype refers to a crucial, necessary biological change, usually a gain-of-function imparted by a gene mutation, amplification, or translocation. This driver is oncogenic and plays a critical role in the signaling that drives the malignant phenotype of the cancer. A wide range
of such driver genotypes have been found consistently in NSCLC, predominantly in adenocarcinoma tumors but to some extent in squamous cell cancers as well.\textsuperscript{232,233} Genotype screening of large patient populations has demonstrated that the driver genotypes tend to be mutually exclusive, suggesting that indeed these changes play a role in the early development of cancer and confirming the notion that genotypic biomarkers can be used to distinguish one subtype of lung cancer from another. Clinical standards of diagnosis for NSCLC are evolving to take genotyping into account at the time of diagnosis; multiplexed platforms are being developed to accomplish testing in a time, money, and tissue-efficient manner.\textsuperscript{234} The genotype information is then used to match patients with targeted therapies, typically small molecular inhibitors that specifically block signaling of the driver pathway. A friendly tool available online at "www.mycancergenome.org" provides such support in an unprecedented way.\textsuperscript{235}

There are several EGFR mutations (most commonly deletions in exon 19 and the L858R point mutation in exon 21) that have been found in adenocarcinomas among patients with a low smoking history.\textsuperscript{87-89} EGFR mutations are much more frequent in patients originally from East Asian nations, for reasons that are unknown; about 10% to 15% of North American and European lung cancer patients have EGFR mutations, whereas in Japan the proportion is 50% or greater.\textsuperscript{232} These mutations lead to a constitutive activation of the EGF receptor and its downstream growth pathways and patients with EGFR mutations can respond in a dramatic way to the approved EGFR TKIs gefitinib, erlotinib, and afatinib (Fig. 51-5 and Table 51-3).\textsuperscript{92,211,236} For context,

![Figure 51-4 Pie chart showing the percentage distribution of genetic abnormalities in lung adenocarcinoma. Treatment decisions about lung cancer have shifted from those based only on histology to those incorporating genetic alterations. As seen in this pie chart, the genome of lung cancer is being chipped away, with discovery of driver mutations that lead to constitutively active signaling proteins and provide a target for therapy. The most recently identified KIF5B-RET fusion subset, which accounts for approximately 1% of this distribution, is boxed. NRAS, neuroblastoma RAS viral (v-ras) oncogene homolog; MAP2K1, mitogen-activated protein kinase kinase 1; AKT1, v-akt murine thymoma viral oncogene homolog 1; PIK3CA, phosphoinositide-3-kinase, catalytic, α polypeptide; BRAF, v-raf murine sarcoma viral oncogene homolog B1; HER2, human epidermal growth factor receptor 2; EGFR, epidermal growth factor receptor; KRAS, v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog. (From Pao W, Hutchinson KE: Chipping away at the lung cancer genome. Nature Med 18:349–351, 2012.)

![Figure 51-5 A response to anti-epidermal growth factor receptor (EGFR)-targeted therapy. Axial chest CT scan of a woman with chemotherapy-refractory adenocarcinoma shows a dramatic response to gefitinib (EGFR tyrosine kinase inhibitor) after 2 months of treatment (A, before treatment; B, after treatment). This woman was a never-smoker of Asian ethnicity with an activating EGFR mutation.]
the response to first-line chemotherapy for lung cancer patients is approximately 20% to 35% and the median progression-free survival (PFS; how long patients remain on treatment without progressing) is about 5 to 6 months. An EGFR mutation-positive patient typically has a 75% response rate to an EGFR TKI with a corresponding PFS of 10 to 13 months. In a series of prospective randomized trials, it has been shown convincingly that initial treatment of advanced lung cancer in patients with EGFR mutations with an EGFR TKI instead of standard chemotherapy significantly improves PFS and quality of life; this has now become the standard of care.

Translocations that result in the apposition of the ALK receptor tyrosine kinase gene near the EML4 or other similar promoter regions have been identified in about 5% of NSCLC cases, again most commonly found in patients with a low smoking history. ALK rearrangements confer sensitivity to the ALK TKI crizotinib with an approximate 70% response rate and a PFS of 10 months (Table 51-4). Patients with ALK translocations previously treated with chemotherapy have been randomized to crizotinib versus standard second-line chemotherapy; crizotinib treated patients have an improved overall survival compared to patients that were unable to receive crizotinib or were diagnosed and treated prior to the availability of crizotinib. This is an important issue because targeted therapies proven to yield dramatic responses in a genotype-defined population pose an ethical dilemma for future randomized trials, because it is no longer feasible to randomize patients away from the cancer histology from adenocarcinoma to small cell lung cancer has been noted in some patients with EML4-ALK rearrangements in about 1% of NSCLC patients. Given the homology of ROS1 to ALK, the TKI crizotinib also happens to be a specific and effective inhibitor for this subset of lung cancer patients. Interestingly, crizotinib also inhibits the MET tyrosine kinase, and about 1% or less of NSCLC patients have a driver genotype characterized by high-level amplification of MET. Case reports have described patients with this genotype who respond to crizotinib therapy.

Even though more common in melanoma, BRAF mutations are also found in about 1% of NSCLC patients. Both HER2 mutations and HER2 amplification are seen in a small percentage (1% to 2%) of lung cancer patients and are associated with response to the HER2 monoclonal antibody trastuzumab and the TKI afatinib. Other rare driver genotypes are emerging in rapid succession.

KRAS mutations are the most common genotype but have so far been refractory to therapeutic targeting. In colon cancer, the presence of KRAS mutation has been a robust negative predictor for EGFR-targeted therapies (mainly monoclonal antibodies), but in NSCLC it has been less clear that KRAS mutations should disqualify patients for EGFR therapy. Development of downstream signaling inhibitors looks like a potential promising strategy for KRAS mutants and a randomized phase 2 study of docetaxel chemotherapy with or without the MEK inhibitor selumetinib showed an interesting increase in response rate and in PFS. Harnessing the immune system to attack cancer has long been a dream of physicians, but progress has historically been slow. However, recently, new treatment options in the form of “checkpoint” modulators have emerged as highly promising therapeutic options. Programmed death 1 (PD-1) protein, a T-cell coinhibitory receptor, and one of its ligands, PD-L1, play a pivotal role in the ability of tumor cells to evade the host’s immune system. Blockade of interactions between PD-1 and PD-L1 enhances immune function in vitro and mediates antitumor activity in preclinical models. In two pivotal phase I trials, antibodies against PD-1 and PD-L1 enhance immune function and mediate antitumor activity in preclinical models. In two pivotal phase I trials, antibodies against PD-1 and PD-L1 have led to remarkably high response rates with durable clinical benefit. CTLA-4 is another protein receptor that down-regulates the immune system. The anti-CTLA-4 antibody, ipilimumab, was approved for use in melanoma and is being tested in lung cancer. Combinations of anti-PD-1 and anti-CTLA-4 antibodies may be even more active, as has been shown in melanoma. Anti-PD-1 and -PD-L1 antibodies may be most effective in tumors expressing the proteins, but the biomarkers have not yet been established as definitive predictors of response.

MECHANISMS OF RESISTANCE

One of the most active areas of research currently in the field of targeted therapy for lung cancer is how to overcome acquired resistance. The genotypes with the most proven success in treatment with targeted agents (EGFR mutations and rearrangements in ALK and ROS1) have also demonstrated that resistance is acquired after 1 to 2 years for most patients. Resistance is often in the form of a point mutation at the gatekeeper location within the RTK, but may also be in the form of a bypass track, or detour using another cancer signaling pathway. Evolution of the cancer histology from adenocarcinoma to small cell lung cancer has been noted in some patients with EGFR mutations with acquired resistance to EGFR TKIs but the mechanism is not well understood. Several therapies for acquired resistance to therapy are being studied, primarily in NSCLC with EGFR mutations and ALK rearrangements.
Key Points

- Lung cancer in smokers is preceded by an accumulation of genetic, epigenetic, and posttranslational abnormalities in the entire bronchial epithelium, showing the validity of the field cancerization concept.
- Lung cancer in never-smokers is less well understood. Nonetheless, the risk for either type of cancer (among smokers or nonsmokers) appears to have a hereditary predisposition, as seen, for example, with genetic alterations in the nicotinic acetylcholine receptors.
- Sequencing of hundreds of candidate genes in adenocarcinomas has shown that there are between zero and 40 lesions in individual lung tumors, which often fall into a handful of previously defined pathways.
- The most common genes altered in lung cancer are TP53 and KRAS, for which effective targeted therapies do not yet exist.
- For some genetic lesions, particularly driver mutations or translocations in EGFR, ALK and ROS1, targeted therapies exist and striking clinical effectiveness has been observed.

Complete reference list available at ExpertConsult.

Key Readings


