Profiles in Variation: Lung Carcinogenesis

Perspective on Kadara et al., p. 702

David P. Carbone

Abstract This perspective on Kadara et al. (beginning on p. 702 in this issue of the journal) examines the critical development of genomic and proteomic signatures of lung cancer risk, prognosis, and sensitivity to chemoprevention or chemotherapy. The novel work of Kadara et al. represents the first demonstration that a molecular signature developed in a premalignancy model (in this case, cultured normal human bronchial epithelial cells and increasingly transformed derivative cells) is clinically relevant to invasive lung cancer.

Lung cancer is a major health problem and cause of mortality worldwide. According to estimates for 2009, in the United States, it will cause almost twice as many deaths in women compared with breast cancer (72,000 versus 40,000) and thrice as many in men compared with prostate cancer (89,000 versus 27,000; ref. 1). Lung cancer is also unique among the common malignancies in not having a proven screening or early detection strategy. Even patients with resected clinical stage I lung cancer have a 5-year survival rate of <50% (2). Lung cancer has also historically lacked the national will to make it a research priority, at least in part due to the stigma of smoking. This attitude contrasts with the high research priority placed on cardiovascular disease, much of which is also self-inflicted, by poor diet and lack of exercise in this case. Furthermore, it is not generally appreciated that ~25,000 patients who never smoked develop and die from lung cancer each year in the Unites States, very similar to the number of deaths from prostate cancer. These grim statistics support increasing the research priority on lung cancer control, not only with better, targeted therapies but with early detection and prevention as well.

Important strides are now being made in improving the quality and quantity of life for patients with lung cancer. Twenty years ago, it was common practice not even to attempt to treat patients with unresectable lung cancer, and the comparatively primitive surgical procedures of that time had very poor outcomes. Now, there is clear evidence that chemotherapy not only improves the quality and quantity of life for patients with advanced lung cancer but that it also improves outcomes in the adjuvant setting of relatively early-stage disease as well. Ten years ago, most patients were offered some sort of therapy, but it was the same for all types of non–small cell lung cancer. Today, new therapies targeting mutant epidermal growth factor or HER2 receptors and ALK kinase seem to produce dramatic tumor responses in small subsets of patients with non–small cell lung cancer. The day is coming when routine pathology labs will provide analyses of molecular alterations in lung cancer that will help guide routine medical decision-making. Detecting the cancer early, when it is surgically curable, or preventing it altogether, would be even better.

Now and for the foreseeable future, however, and especially for the 85% of lung cancer patients in the West who do not have targetable tumor mutations, chemotherapy will remain the mainstay of early/adjuvant and advanced-stage lung cancer therapy. Targeted therapies as well as traditional chemotherapies are toxic, expensive, and effective only in a subset of patients. As poor as surgical outcomes are, about half of early-stage patients are cured without additional therapy. Therefore, identifying patients requiring adjuvant therapy and the optimal therapy for them is an area of intense research interest.

In this issue of the journal, the Lotan group (Kadara et al.) extend their earlier proteomic analyses of protein expression profiles in cultured normal human bronchial epithelial cells and increasingly transformed derivative cells (3) with new analyses of gene expression profiles (transcriptomes) derived in the same model system (4). They derived a 584-gene signature of neoplastic progression and applied it to the Shedden data set of resected adenocarcinomas (5), finding that patients with a "normal"-like tumor signature had a better survival than did patients with an "invasive"-like tumor signature. Six signature genes they selected from the overall pathway analysis identified significant differences in survival among subsets of patients within each of three different published cohorts of patients with resected adenocarcinoma. These investigators then showed that the expression of a single protein, UBE2C, by RNA and immunohistochemistry correlated with progression to cancer and survival.

This work represents the first time that a molecular signature developed in a premalignancy model was assessed for

Author's Affiliation: Vanderbilt-Ingram Cancer Center, Departments of Medicine, Cancer Biology, and Cell Biology, Vanderbilt University, Nashville, Tennessee

Received 6/12/09; accepted 7/1/09; published OnlineFirst 7/28/09.

Requests for reprints: David P. Carbone, Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt University, 685 Preston Research Building, 2220 Pierce Avenue, Nashville, TN 37232. Phone: 615-936-1279; Fax: 615-936-3322; E-mail: david.carbone@vanderbilt.edu.

doi:10.1158/1940-6207.CAPR-09-0142

clinical and biological relevance in a clinical data set of cancer in any setting. The earlier proteomic signature of this group (3) correlated well with human premalignancy specimens but was not studied for correlations with clinical cancer outcome.

The multifaceted importance of this work includes its development of the concept of deriving key signatures for clinical cancer by looking at a defined series of cell lines. This approach allows the isolation of tumor cell–specific signaling and eliminates the unavoidable noise due to necrosis or normal cell infiltration encountered in analyzing clinical tumor samples. The statistically significant survival prediction of this purely molecular signature (no clinical variables) in the Shedden data set is a success that eluded multiple other signatures tested in the same data set and implies the power of this approach.

This approach also may be applicable to noncancer/premalignant bronchial epithelium for defining features that are associated with progression to malignancy and independent of histologic features. This application is not only one of the most important but also one of the most difficult to pursue, requiring testing in sets of bronchial samples collected years (and potentially decades) before the development of cancer, with complete clinical outcome data. Such samples are being collected now but are still extremely scarce. Another potential opportunity for assessing signatures in normal and/or premalignant tissues would be in randomized, placebo-controlled lung cancer prevention trials in high-risk smokers or the adjuvant/second-primary-tumor prevention setting. Accomplishing this goal would potentially allow the identification of patients for more intensive surveillance, earlier surgical intervention, and chemoprevention, in which relevant signatures could perhaps act as surrogate end point biomarkers in addition to acting as potential risk markers and agent targets. It would be potentially informative to combine these markers with epithelial proteomic markers such as those reported earlier by the Lotan group (3) and by Massion and colleagues (6).

Much of the Lotan group data (4) focuses on patient subsets with better or worse survival identified from multiple published and in-house cohorts of patients with resected lung cancer. The success of this identification in multiple testing sets speaks to the robustness of the signature, although it is not entirely clear whether recursive testing was used to select the smaller panel of markers in these "test" cohorts. There also are many caveats in pursuing and interpreting this type of research. For example, the signature was developed in bronchial epithelial cell lines and was tested primarily in adenocarcinomas, which often develop in the peripheral parenchyma rather than in the large airways. Whether signatures of progression in large-airway epithelial cells will be useful in predicting the risk of developing peripheral adenocarcinoma remains to be shown, although the work of Spira et al. in this regard is encouraging (7). This large-airway marker set may more powerfully predict squamous cell lung cancer, which more typically arises in large-airway epithelium and for which cohorts are in current development. Notwithstanding the earlier mentioned strength (noise avoidance), basing a signature on epithelial cell lines also has the disadvantage of ignoring host-tumor interactions, including growth-modulating cytokines, angiogenic factors, immune factors, and invasive/metastatic interactions. The interplay of these factors is clearly important in both the development and progression of lung cancers.

The only two groups to have reported gene expression profiles in nonmalignant lung tissue are those of Spira and Mao (8, 9). The present work of the Lotan group differs from that of Spira and Mao not only in deriving signatures *in vitro* but also in applying them to clinical cancer specimens and data in order to assess biological and clinical relevance by correlating with disease outcome. This dual approach in the setting of early-stage disease reflects a convergence of profiling for the risk of both new cancer development and recurrence of an established cancer. Signatures of prognosis in early-stage cancer may resemble or be the same as signatures of risk in premalignant conditions.

In more general terms, the real utility of developing cancer signatures (such as the one reported in this issue of the journal) for patients who already have cancer is the personalization of therapy to improve the quality and quantity of life. This use often is framed as enabling the identification of patients at a higher or lower risk of relapse and thus, by implication, the subset that would benefit from adjuvant therapy, a line of thought that the authors wisely did not pursue. A powerful argument can be made that none of these signatures have robustly identified a subset of patients with lung cancer who do so well that attempts to improve their outcomes are unnecessary. None of the classifications shown in this article tease out such a subset. More important, a poor prognosis does not equate to the likelihood of a benefit from chemotherapy. Indeed, as the authors point out, their star marker, UBE2C, is overexpressed in the poor-prognosis patients, and this overexpression may predict resistance to taxane-based therapy (10). Studies in other cancer settings have shown that a marker (e.g., HER2 in breast cancer) may indicate a poor prognosis and resistance to one therapy but sensitivity to another. Therefore, the poor-prognosis subgroup would be predicted to do badly on adjuvant taxane therapy (or without it), but the "good" prognosis subgroup would do better. These predictions are opposite to the usual (and likely erroneous) assumptions concerning prognosis subgroups.

The development of clinical prognostic signatures in resected lung cancer, especially in publicly available data sets, often gives little attention to the clinical characteristics of the patient population, whereas the pathologic criteria are often very strict and rigorously reviewed. For example, such studies frequently are not clear on whether all of the resected patients underwent mediastinal node sampling. It is astounding how many lung cancer resections occur in the United States without a single mediastinal nodal station being sampled. What the researcher is calling a stage I lung cancer might actually have been a stage III cancer with a worse outcome because it was treated by the clinician as stage I. More subtle is the lack of data set information on adjuvant therapy (and lack of appreciation by the investigators, or data set compilers, of its importance). One of the most important reasons for defining prognostic subgroups of patients with resected lung cancer is to identify poor-prognosis patients who then could be assessed for predictive markers of sensitivity to, or benefit from, a particular adjuvant therapy. Deriving signatures from a mixture of patients who did and did not receive adjuvant therapy will fatally confound the results. For example, if adjuvant chemotherapy converted poor-prognosis patients into

good-prognosis patients and this converted population is mixed together with poor-prognosis patients in the signature derivation, no clinically meaningful signature will emerge. The Shedden data set exemplifies this problem, although the evidence is provided in the supplementary information, not in the article (5). About one-third of the data set patients had adjuvant chemotherapy, one-third had no information, and a large number had adjuvant radiation therapy. Yet Shedden et al. mixed all of these patients together for the analyses they reported (5). The flaws of variable staging and a lack of accurate and complete clinical follow-up are common to most of the published data sets.

For adjuvant therapy, the much more relevant search would be for signatures of treatment effect rather than of simple prognosis. Indeed, patients with resected lung cancer usually relapse by developing metastases, and it could be argued that the most useful signatures for adjuvant therapy would be derived from predicting response to an agent in cohorts of patients with metastatic lung cancer. Another potential direction of this research would be to apply these metastasis-derived signatures to primary/second-primary-tumor chemoprevention. Annotated sets of adequate tissue samples from patients with metastatic disease are very scarce, and the small number of efforts now under way to produce more metastatic disease data sets include data collections within the M. D. Anderson-based BATTLE (Biomarker-based Approaches of Targeted Therapy for Lung Cancer Elimination), Vanderbilt-based Lung SPECS (Strategic Partnering to Evaluate Cancer Signatures), Addario Medical Research Institute's CASTLE, and other trials or programs.

In summary, it is very clear that the study of the Lotan group advances the lung cancer biomarker field and gives us hypotheses for future testing in the clinic, particularly in relation to premalignancy and biomarkers of progression. Substantial efforts should be directed at collecting appropriate sample sets with long clinical follow-up for testing the ability of this study's and other biomarker signatures to identify patients at risk of progressing to lung cancer. No one would argue that correlating these signatures with overall early-stage disease outcome is biologically relevant, but this correlation does not provide any data on predicted outcomes of specific targeted and nontargeted therapies. With respect to classifying patients already diagnosed with lung cancer, it is clear that technology has outstripped the science, and biomarker signature development for diagnosed patients would benefit from moving forward with the development of signatures predicting response to and benefit from specific interventions. Advances in this area will depend on framing thoughtful clinical questions tested in sets of appropriate and clinically annotated samples.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Society AC. Cancer facts & figures 2009. Atlanta: American Cancer Society; 2009.
2. Goldstraw P, Crowley J, Chansky K, et al. The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. J Thorac Oncol 2007;2:706–14.
3. Shen J, Behrens C, Wistuba II, et al. Identification and validation of differences in protein levels in normal, premalignant, and malignant lung cells and tissues using high-throughput Western array and immunohistochemistry. Cancer Res 2006;66:11194–206.
4. Kadara H, Lacroix L, Behrens C, et al. Identification of gene signatures and molecular markers for human lung cancer prognosis using an *in vitro* lung carcinogenesis system. Cancer Prev Res 2009;2:702–11.
5. Shedden K, Taylor JM, Enkemann SA, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. Nat Med 2008;14:822–7.
6. Jamshedur Rahman SM, Shyr Y, Yildiz PB, et al. Proteomic patterns of preinvasive bronchial lesions. Am J Respir Crit Care Med 2005;172:1556–62.
7. 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;13:361–6.
8. Zhang L, Lee JJ, Tang H, et al. Impact of smoking cessation on global gene expression in the bronchial epithelium of chronic smokers. Cancer Prev Res 2008;1:112–8.
9. 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;101:10143–8.
10. Reddy SK, Rape M, Margansky WA, Kirschner MW. Ubiquitination by the anaphase-promoting complex drives spindle checkpoint inactivation. Nature 2007;446:921–5.