Publishing Negative Data: β-Tubulin Mutations in Lung Cancer

Frederic J. Kaye

It is widely known that the threshold for publishing negative data in the scientific literature is high, and those manuscripts that do survive peer-review are generally restricted to observations that refute a hypothesis that is currently in vogue. This is the case for a study reported by Kelley et al. (1) in this issue of the Journal, which analyzed a series of lung cancer samples for the presence of β-tubulin mutations. The publication of these data, however, highlights a much broader and more important topic. Excusing an obvious truism, this issue relates to the timely need to rigorously validate the ever-growing number of genes and gene pathways that have been linked to the etiology and/or progression of lung cancer.

Kelley et al. (1) sought to confirm earlier reports (2,3) that had identified mutations within the most common isoform of the β-tubulin gene family (called the TUBB gene) in tumor biopsy specimens from 16 (33%) of 49 patients with non-small-cell lung cancer (NSCLC) (2) and in serum DNA isolated from 55 (42%) of 131 patients with NSCLC (3). Because normal tissues did not have these nucleotide alterations, these earlier reports (2,3) suggested that many lung cancers, irrespective of cytotoxic treatments, were associated with acquired tumor-specific mutations within the TUBB gene. These observations were compelling for several reasons. First, they partly resembled in vitro data accumulated during the past two decades that had detected β-tubulin gene alterations in cancer cell lines selected for resistance to a variety of microtubule-disrupting agents, such as colchicine or paclitaxel (4–7). No primary tumors with β-tubulin mutations, however, had been previously identified, raising the clinical significance of these findings in patients with lung cancer. Second, most of the mutations identified in the tumors were localized to a conserved GTP-binding domain, which is predicted to alter microtubule assembly in vivo. Finally, the presence of these TUBB mutations was strongly associated with a reduced clinical response to the antitubulin drug paclitaxel, suggesting that the mutations may be a valuable tool for planning anticancer strategies (2).

The brief communication by Kelley et al. (1), however, detected only two silent polymorphisms (no change in the encoded amino acid) in 20 primary lung cancer samples and in 25 independent lung tumor cell lines. The authors, therefore, concluded that TUBB mutations in the GTP-binding domain are rare in lung cancer tumors or in derived cell lines. They extended their observations by subjecting the same genomic DNA samples to polymerase chain reaction amplification using either their own intron-anchored primers or the paired exonic primers reported in the earlier positive study (2). Although no TUBB alterations were observed with the intron-based primers, several heterozygous TUBB alterations were detected when DNA harvested from the same tumor was amplified with the exonic primers. The most straightforward explanation for these findings was that the exon-based primers were not capable of discriminating between the highly related members of the TUBB gene family, including a diverse group of clustered and nonclustered pseudogenes. The fact that some, but not all, of the “mutations” could be tagged to a known β-tubulin pseudogene perhaps suggests that not all of the β-tubulin variant sequences have been tabulated in the human genome database and/or that spotty nucleotide errors have not yet been edited properly. Regardless, the lack of TUBB gene alterations in 45 different lung cancer samples with the use of the primers outlined by Kelley et al. (1) emphasizes that alterations in this β-tubulin isoform are rare, if not absent, in lung cancer.

The strength of the study by Kelley et al. (1), however, lies less with the negative data of gene mutations and more with the ability of the authors to provide a plausible (and testable) hypothesis to explain the discrepancy with the earlier reports. In their study (1), a convincing explanation was clear to the authors because the confusion of pseudogene sequences with putative cancer gene alterations is not an uncommon event, particularly because some pseudogenes (psi) may even retain messenger RNA expression, as noted for the psiPTEN locus (8–10). In

Affiliations of author: Genetics Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, and National Naval Medical Center, Bethesda, MD.

Correspondence to: Frederic J. Kaye, M.D., National Naval Medical Center, Bldg. 8, Rm. 5101, Bethesda, MD 20889 (e-mail: fkaye@helix.nih.gov).
contrast, the difficulty to confirm other compelling, but unproven, gene associations as validated etiologic/prognostic factors in lung cancer is an ongoing challenge for both basic scientists and for clinical investigators seeking laboratory correlations and intermediate endpoints. For example, the recognition that DNA hypermethylation may be associated with gene inactivation and that haploinsufficiency may variably inactivate other genes with an otherwise wild-type sequence has added to the complexity and uncertainty of defining a minimal genetic etiology of lung cancer (11). Moreover, no one functional assay for tumorigenicity can be applied to all candidate cancer genes and even within a single parameter, such as in vitro colony suppression or in vivo nude mouse tumorigenicity assays, quantitative differences can be obtained in the results. These and other variables led to studies that have yet to define the primary role for several candidate genes in both small-cell lung cancer and NSCLC, including the β-retinoic acid receptor, several protein phosphatase genes, the FHI1T gene, the RASSFI1A gene, and the retinoblastoma-related gene p130, components of the PTEN/AKT pathway, alterations in expression of ErbB1, ErbB2, and c-Kit, elevated levels of the survival factor BCL2, and expression of simian virus 40 large T antigen in mesothelioma and in a subset of other lung tumors (12). This uncertainty is further emphasized in analyses of other genes that may be linked to susceptibility risks and/or prognostic factors in lung cancer as: 1) there are more issues in data collection and uncontrolled patient variables, 2) the assays employed, such as immunohistochemistry, are imprecise for scoring meaningful cutoff ranges, 3) the influence of publication bias is greater because negative studies are generally published only when there is a compelling hypothesis for the authors to argue against, and 4) the focus is often not based on a rigorous biologic genotype:phenotype experiment. For example, many studies with conflicting results have analyzed the risk of developing a wide range of human cancers, including lung cancer, using polymorphisms linked to either the L-MYC gene (13), the p53 gene (14), or a range of candidate metabolism gene products (15). In addition, other studies (12) have assessed the prognostic value of genetic alterations in the retinoblastoma or p53 tumor suppressor genes and, in other genes, in lung cancer with no definitive conclusions. As a result, investigators (16) have begun to use specialized meta-analysis tools to reanalyze these datasets.

Four years ago, a minimal definition for a cancer gene was proposed that required a candidate gene to demonstrate either a tumor-specific loss-of-function or gain-of-function mutation (11). Although it is beyond the scope of this editorial to debate the need to extend this definition by recommending additional “levels of evidence” guidelines for defining both cancer genes and cancer-related prognostic factors, it is clear that such an approach may help to organize and to prioritize the growing list of markers that are currently under consideration for correlative clinical studies in lung cancer.

In summary, much of the low-lying fruit has been picked, and the important work of defining additional key pathways in lung cancer will be difficult and tedious. A good start, as recommended by Kelley et al. (1), is to postpone clinical correlations with TUBB “mutations” in lung cancer patients until the genetic basis for the preferential detection of these sequences has been demonstrated convincingly.

REFERENCES