# Rearranging Detection of Gene Rearrangements

*Kurtis D. Davies, PhD,\* Robert C. Doebele, MD, PhD† and Dara L. Aisner, MD, PhD\**

Recent years have witnessed a dramatic expansion in our knowledge of the genetic diversity of non–small-cell lung cancer (NSCLC), and in particular adenocarcinoma. This appreciation of the genetic landscape of the disease has resulted in the development of targeted therapies that are directed toward activated oncogenes. Several of these drugs have been approved for use in patients and many more are being investigated in clinical trials. However, there have been only been limited studies that have focused on histologic subtypes within NSCLC and more complete analyses of these specific disease entities are needed.

In this issue of the Journal of Thoracic Oncology, Shim et al. explore the genomic alterations and clinical characteristics associated with invasive mucinous adenocarcinoma, a relatively rare histologic subtype of NSCLC. In confirmation of previous studies that have examined the genetic basis of this disease, the authors observed a high rate of *KRAS* mutation (63%).[1] Interestingly, while the G12C variant is the most prevalent *KRAS* alteration in lung adenocarcinoma as a whole, G12D and G12V accounted for the majority (73%) of KRAS mutations in this cohort of mucinous adenocarcinomas, with G12C making up only 11%. The authors then examined cases lacking *KRAS* mutation in more detail using two targeted deep sequencing approaches. In one, the RNA component of extracted total nucleic acid was evaluated for gene rearrangements. They observed that nine of the *KRAS* wild-type cases were positive for rearrangements involving genes that have previously been found to be involved in fusions that create oncogenic chimeric proteins. As has been seen in other studies examining mucinous adenocarcinoma, rearrangements involving the *ERBB* family ligand *NRG1* were particularly prevalent.[2–4] By examining genomic DNA by deep sequencing using a separate targeted approach, well-characterized activating mutations in *ERBB2*, *BRAF*, a *PIK3CA* were detected in four of the *KRAS* wild-type cases. Notably, these oncogenic alterations all occurred independently of each other. However, since the deep sequencing approaches were only applied to the initial *KRAS* wild-type cohort, the exclusivity between these activated oncogenes requires further analysis. Nevertheless, when taking the initial *KRAS* mutant cohort into account, the majority (81%) of examined cases in the entire cohort were found to be positive for an activated oncogene.

Remarkably, mutations in the tumor suppressor gene *TP53*, which occur in approximately 50% of lung adenocarcinomas and commonly overlap with activating oncogenic alterations, were rare in this cohort of mucinous adenocarcinomas, occurring in only two of 50 evaluated cases. The authors expanded on this finding by querying publicly available TCGA lung adenocarcinoma data and found only one *TP53* mutated case out of 12 mucinous tumors. In the TCGA dataset, the low rate of *TP53* mutation correlated with a low mutational burden in mucinous versus nonmucinous tumors. While smoking status alone correlated with mutational burden in adenocarcinoma as a whole, multivariate analysis suggested that *TP53* status likely contributed to low mutational burden in mucinous tumors but that smoking status did

not. This has potentially interesting implications for *TP53* status acting as a surrogate biomarker for mutational burden, which itself has been shown to be a potential predictive signature for response to immune checkpoint therapy in NSCLC.[5]

While this study has confirmed the presence and distribution of previously reported alterations in this disease, several items are particularly noteworthy. First, it adds to the body of literature which increasingly demonstrates that mucinous adenocarcinoma of the lung has distinct molecular features compared with nonmucinous tumors. Second, it underscores the expanding perception that gene rearrangements, while existing at relatively low frequencies individually in lung cancer, cumulatively constitute a significant fraction of this disease. In mucinous lung adenocarcinoma, lung cancer as a whole, and many other solid and hematological malignancies, it is becoming increasingly apparent that chimeric proteins created by gene rearrangements are the primary oncogenic driver in a substantial proportion of tumors. Many of these fusion proteins can be directly targeted or targeted downstream by approved or investigational drugs, making accurate detection of rearrangements by diagnostic tests of vital clinical importance. Therefore, perhaps the most notable impact of this study lies in the methodology used to detect gene fusion events.

Evaluation for rearrangements has traditionally been accomplished by cytogenetic techniques, predominantly "break-apart" fluorescence in-situ hybridization (FISH). However, while FISH assays can be multiplexed using multiple fluorophore colors, such assays can be difficult to develop and have limitations to the extent to which multiplexing can be accomplished. This can make FISH an inefficient approach to examine a large number of fusion genes (for example *NRG1, BRAF, NTRK1, ALK, ERBB4,* and *RET* fusions identified so far by this and other mucinous adenocarcinoma studies).[2-4] Reverse transcriptase-polymerase chain reaction (RT-PCR) is another traditionally employed rearrangement detection method, but this approach also lacks the ability to efficiently multiplex for analysis of many targets. In addition, this technique typically relies upon knowing the identity of the fusion partner, which is problematic as novel fusion partners for the above oncogenes are being continually identified. Next-generation sequencing (aka deep sequencing) has emerged as a strategy to simultaneously assay multiple genetic loci. Indeed, two of the previous studies that discovered gene fusions in mucinous adenocarcinoma of the lung did so through whole-transcriptome sequencing.[2,3] While transcriptome sequencing is ideal for discovery efforts, it is not yet an efficient clinical diagnostic approach in most laboratories. This is due to a need for high depth of coverage and a need for high throughput (to accommodate many patient samples at once) in the clinical realm. Therefore, most next-generation sequencing assays used for clinical diagnostic purposes rely on target enrichment approaches and focus on subsets of genes that are potentially clinically relevant. While this approach is relatively straight-forward for identifying single nucleotide variants and small insertions and deletions, the detection of larger structural rearrangements (including gene fusions) is more challenging. For example, if using an amplicon-based target enrichment technique to detect a fusion event, primer sets are needed that span the breakpoint, meaning that a priori knowledge of the fusion partner is required (similar to RT-PCR). The study by Shim et al. circumvents this problem by employing a recently reported variation of amplicon-based target enrichment: anchored multiplex PCR (AMP).[6]

In AMP, cDNA made from extracted RNA or total nucleic acid is processed and ligated to a "half-functional adapter" (similar to an adapter used for standard next-generation sequencing although incomplete). Then, a PCR technique using one primer specific to the oncogene component of the fusion and another primer specific to the half-functional adapter is performed. If a particular sequence contains a fusion break point, then the sequence surrounding the breakpoint is amplified by virtue of being between the gene-specific and adapter-specific primers. This step is followed by an additional nested PCR round that completes the creation of functional adapters on both ends of the sequence, making clonal amplification for next-generation sequencing possible. Importantly, by employing a universal primer to the half-functional adapter, the need for a primer complimentary to the fusion partner is bypassed. Thus, most fusions, including fusions to previously unknown partners, are detected for each fusion oncogene that is examined. One caveat of this assay is that some foreknowledge of breakpoints is still needed, because some rearrangements may occur within unexpected introns or fuse the partner gene at the opposite end (i.e., 3′ instead of 5′); although this can easily be overcome by broader and bi-directional design. Importantly, this assay can be performed in multiplex fashion (using primers to multiple oncogenes involved in fusions), and, due to the amplification steps, can be performed on low nucleic acid inputs, making it highly adaptable for the clinical setting.

As the list of known activated oncogenes increases and the arsenal of targeted therapies for cancer grows, there will be an ever-increasing need to simultaneously detect multiple types of oncogenic alterations. The study by Shim et al. exemplifies how multiple next-generation sequencing library preparation techniques can be combined to obtain a more complete assessment of the activated oncogenes in a tumor sample. Using this approach, they reveal a distinct portrait of the genetic landscape of mucinous adenocarcinoma of the lung and highlight differences between this histologic subtype and NSCLC as a whole. Ultimately, approaches such as these will serve a critical role in the identification of patients for oncogene-targeted therapies.

## REFERENCES

1. Marchetti A, Buttitta F, Pellegrini S, et al. Bronchioloalveolar lung carcinomas: K-ras mutations are constant events in the mucinous subtype. *J Pathol* 1996;179:254–259.
2. Fernandez-Cuesta L, Plenker D, Osada H, et al. CD74-NRG1 fusions in lung adenocarcinoma. *Cancer Discov* 2014;4:415–422.
3. Nakaoku T, Tsuta K, Ichikawa H, et al. Druggable oncogene fusions in invasive mucinous lung adenocarcinoma. *Clin Cancer Res* 2014;20:3087–3093.
4. Gow CH, Wu SG, Chang YL, Shih JY. Multidriver mutation analysis in pulmonary mucinous adenocarcinoma in Taiwan: Identification of a rare CD74-NRG1 translocation case. *Med Oncol* 2014;31:34.
5. Rizvi NA, Hellmann MD, Snyder A, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 2015;348:124–128.
6. Zheng Z, Liebers M, Zhelyazkova B, et al. Anchored multiplex PCR for targeted next-generation sequencing. *Nat Med* 2014;20:1479–1484.