COMMENTARY

Genomic Technologies and the New Era of Genomic Medicine

Arunkanth Ankala and Madhuri Hegde

From the Department of Human Genetics, Emory University School of Medicine, Atlanta, Georgia

Genetic technologies have been used extensively in the discovery and diagnosis of constitutional disorders and cancer. These technologies range from cytogenetic analysis and standard PCR-based methods to the more informative quantitative PCR, Southern blot analysis, microarrays, and Sanger sequencing. In the last decade, we have witnessed the emergence of next-generation sequencing (NGS) technology with rapidly evolving sequencing chemistries and instruments, which has led to a boom in clinically available NGS-based assays. The research community at large has since adopted rapidly NGS technology for the discovery of new genes as well as genomewide association studies. More importantly, NGS technology has made its way into the arena of clinical diagnostics for genetically heterogeneous constitutional disorders (such as neuromuscular disorders, cardiomyopathy, and intellectual disability), with several disease-specific diagnostic assays that are now available. However, application of NGS to clinical oncology testing, especially those involving somatic (acquired) mutations, is still in its infancy. In this issue of The Journal of Molecular Diagnostics, Cottrell et al discuss the design, clinical validation, and utility of a targeted NGS panel involving 25 cancer-associated genes for somatic mutation detection.

Molecular Diagnosis of Cancer

Until recently, the molecular diagnosis of cancer relied mainly on interrogating the most frequent mutations, or hotspot mutations or regions, and involved low-throughput molecular and cytogenetic methods, all of which are limited by the number of mutations or target regions that can be tested in one assay. Nonetheless, over the years, a growing number of mutations and genes have been identified in a variety of cancers such as lung, breast, melanoma, and brain cancer.

Often, the presence or absence of variants in a certain gene, like the KRAS mutations in colorectal cancers, modulates the efficacy and therapeutic outcome of an otherwise highly effective drug. Therefore, regardless of the cancer type, establishing a molecular diagnosis and knowing the exact causative gene and mutation are crucial for directing therapeutic options and stratifying the overall risk as well as in predicting the prognosis of cancer patients. This has driven a paradigm shift from single-gene, or hotspot, testing, to parallel multi-gene testing using NGS technology. However, given the nature of cancer mutations (which can be inherited or somatically acquired) and the complexity of the disease, the performance of NGS assays needs thorough evaluation before clinical application.

As is apparent from recent reports, the clinical application of NGS assays and, more specifically, targeted panels has improved the diagnostic yields of molecular testing rapidly for constitutional diseases tremendously, but the clinical utility of the panel approach for diagnosis of cancer, especially those involving somatic mutations, is still being explored. Alternatively, the clinical utility and diagnostic yields of the more comprehensive whole-exome sequencing (WES) and whole-genome sequencing (WGS) tests have yet to be established for either constitutional diseases or cancers. An overview of the clinical utilities of different test types for both constitutional diseases and cancer is shown in Figure 1. Although WES and WGS have proved to be revolutionary for new gene discovery in a research scenario, their clinical application is limited by several factors including such challenges as involving the interpretation of newly identified variants and genes, missing coverage for at least 10% of the targeted region, the magnitude of data generated per sample, and the efforts and time required for interpretation of this data for clinical
relevance. Compared to WES, however, WGS has the advantage of being able to detect novel rearrangements in cancer testing.9 Given all these factors, disease-specific targeted panels are constantly being developed and validated for use in routine clinical diagnostics.

In contrast to some recently reported targeted cancer panels that focus more on hotspot mutations in multiple cancer genes,10 the Washington University Cancer Mutation Profiling (WUCaMP) assay developed by Cottrell et al4 interrogates the entire coding region of 25 cancer-associated genes for mutation detection. Such a comprehensive assay is likely to have a higher diagnostic yield compared to those that only target hotspots within these genes, as mutations outside hotspots are equally important.

Clinical Design of the WUCaMP Assay

The WUCaMP assay involves deep sequencing of the entire 300-kb targeted 25-gene region with an average coverage of $>1000 \times$ across the entire targeted region. The assay has 100% sensitivity and specificity for single nucleotide variant detection as well as 100% sensitivity for detecting variants at an observed 10% allele fraction, confirming strong reliability of the assay for cancer diagnostics. Such thorough validation of performance and sensitivity is critical since these are high-complexity assays as defined by Clinical Laboratory Improvement Amendments and require an extensive investment in technology and informatics.

Several organizations, such as the College of American Pathologists and the CDC, have released detailed guidelines and recommendations for validation and clinical implementation of NGS assays.11 Nevertheless, given the differences in sequencing chemistries, platforms, and the assortment of steps involved in the entire assay, from sample preparation to data analysis and interpretation, clinical labs developing these assays face significant challenges and, hence, seek guidance. To address these concerns, Cottrell et al4 report the validation of the WUCaMP assay in detail.

Assay Validation

As is critical in cancer diagnostics, which frequently involve the use of highly heterogeneous tissue specimens, Cottrell et al4 thoroughly validated the WUCaMP assay for its performance in detecting clinically actionable somatic cancer mutations at low variant allele fraction (VAF). VAF refers to the fraction by which a certain variant at an allele is represented in the cell population within a tissue specimen. The VAF for somatic tumor variants in the tissue specimen being tested depends on tumor cellularity as a result of its invasion into stroma and infiltration by immune cells as well as on the precision with which the tumor tissue was isolated from adjacent nontumor tissue.

As part of their assessment of the assay’s analytical performance for single nucleotide variant detection in clinical cancer specimens, Cottrell et al4 analyzed 81 individual samples (78 tumor specimens and 3 HapMap DNA cell lines) using WUCaMP. Unlike the case of constitutional variants, where the VAF is most often as high as 50% (heterozygous mutations; 100% for hemizygous or homozygous variants), the VAF for somatic pathogenic tumor variants or mutations can vary widely, frequently 20% or less. Therefore, although the
sensitivity for VAF is not a major issue for non-cancer constitutional diseases with the exception of rare mosaic cases, it can be a significant limitation in the detection of acquired or somatic cancer-causing variants. In developing the WUCaMP assay, the authors addressed this issue by using proportionately pooled DNA samples and strategic in silico mixtures of two different HapMap samples and assessing the analytic performance in detecting variants with a resultant VAF as low as 10% or less. These dilution experiments showed 100% sensitivity for variants with VAF greater than 10% in the targeted regions, which is compelling. They then validated the reproducibility and versatility of the assay by comparing results from different specimen types (formalin-fixed, paraffin-embedded tissue, fresh-frozen tissue, blood, bone marrow, and cell lines), as well as interrun and interpersonnel performances.

**Advantages**

In cancer genomes, the number of somatic mutations is typically on the order of thousands, with most adult cancers being in the 1000 to 10,000 range.\(^2\) Of these mutations, only a small number are driver mutations (ie, those that confer a growth advantage to the cell by altering a fundamental cellular process). Most are so-called passenger mutations, which may not contribute to clonal selection and tumorigenesis but accumulate over the lifetime of the tumor.\(^3\) The ability to distinguish driver mutations from passenger mutations is therefore critical to establishing diagnosis. An evidence-based gene panel approach, which specifically detects variants in clinically relevant genes, skirts this issue to a certain extent, unlike WES and WGS that often detect a significant number of variants and genes of unknown significance. One caveat, however, is that certain driver mutations harbored by untargeted genes are missed by the panel approach. The false-negative rate is equally important in terms of the clinical utility of an assay. Data from 51 clinical patients tested after successful launch of the WUCaMP test indicated that there were no false negatives in the targeted coding region, and all mutations detected by Sanger sequencing were detectable. Also, there were no false-positive variants detected in the region targeted by WUCaMP. High false-positive and false-negative rates have plagued most of the current NGS-based assays, despite rapid advances in target capture and NGS technology. Although these rates can be reduced significantly by increasing the depth of sequencing, it raises the cost of the assay. The relatively small target size of panels (compared to WES and WGS) allows for high coverage depth without significant increases in the sequencing cost. In silico sensitivity analysis performed for the WUCaMP assay indicates that as high as 1000x unique coverage levels pick 100% of all variants with an allele fraction of at least 10%.

Another significant finding of the Cottrell et al\(^4\) study is the detection of a high-quality, low-allelic fraction single nucleotide variant that was not detected by Sanger sequencing. Detection of this true positive variant by a targeted NGS approach emphasizes the potential of the technology over the relatively less sensitive Sanger sequencing. With optimal coverage depth and thorough validation of assays, the expectation is that NGS panels eventually can be offered as stand-alone tests without complementary Sanger confirmation. Though this is currently not the practice in clinical labs, studies such as this and others\(^1,6\) will soon make it a reality.

**Limitations**

Knowing the limitations of a test is critical to making an informed decision about ordering that test for any disease, and Cottrell et al\(^4\) do indeed elaborate on the limitations of the assay. One such limitation is the lack of sufficient coverage (at least 50x) for 14 of 393 exons (3.6%), most of which are GC-rich first exons of several of the targeted genes. This is true for most of clinically available NGS panel tests and is in fact a limitation of the available sequencing chemistries.\(^1,6\) However, the authors indicate that these missed exons may not be clinically significant since causative mutations have rarely been reported in these exons.

A second limitation discussed by the authors is the sensitivity of somatic variant detection. Although the WUCaMP assay can detect variants with an allele fraction as low as 10% with a sensitivity of 100%, the sensitivity for variants with allele fraction 5% and 1% is zero. Knowing this is important, especially when dealing with impure and heterogeneous tumor specimens or with challenging specimen types, like formalin-fixed, paraffin-embedded samples, as formalin fixation is known to gradually degrade nucleic acids.

One added advantage of NGS assays, is their ability to process pooled multiple samples and perform parallel sequencing, allowing for an appreciable cost reduction. The maximum number of samples that can be pooled together nevertheless is dependent on the size of the targeted region and the read and coverage depth required to make a confident call on the detected variant. However, this must be evaluated and validated for each individual NGS panel assay. Once these parameters are determined, multiple individual samples, each referred for the test, are processed simultaneously and pooled together after tagging each DNA sample with a short unique sequence index, specific for each patient. Though this is a routine practice in NGS labs, it is compelling that Cottrell et al\(^4\) saw as low as 0.18% crossover during library preparation and 0.03% to 0.06% estimated crossover during multiplex sequencing of pooled samples.

**Perspectives on the WUCaMP Assay**

The validation procedures reported in this study can be helpful for organizations formulating regulations and recommendations for clinical NGS practice, as well as for laboratories developing similar assays. NGS technology has truly shifted the paradigm of medical genetics from the concept of one gene/one disorder to one gene/many disorders and many disorders/many genes.
Along with the increasing throughput and the ability to massively sequence large sections of the genome come the challenges of interpreting novel variants (including both disease associated and unassociated) and identifying false positives (which are technical artifacts). Large-scale WES data from various groups, such as the National Heart Lung and Blood Institute’s Exome Sequencing Project, have helped to some extent in interpreting these variants. Moreover, recent NIH awards to Clinical Genome Resource to annotate variants for their clinical relevance and disease association and to disseminate this information through the National Center for Biotechnology Information’s ClinVar are expected to further improve our ability to interpret variants as we move into the NGS era of genomic medicine. Meanwhile, internally curated variant databases and complementary software packages such as those developed to support the WUCaMP assay are crucial to analyzing the data generated from NGS assays. Clinical laboratories venturing into NGS testing should thus be prepared accordingly, so as to reduce analysis burden and improve test efficiency.

Conclusions

Overall, the WUCaMP assay serves as a good proof-of-principle study for somatic cancer mutation detection on clinical cancer specimens and outlines the critical validation strategies that can be adopted in developing an NGS-based target panel test for cancer diagnostics. Ongoing efforts to extend the assay’s potential to detect additional sequence changes, such as copy number variations, structural rearrangements, and small insertions and deletions, will significantly improve the overall clinical utility of targeted NGS assays, both for constitutional disease and cancer diagnosis. This is a major concern for cancers, more than for constitutional diseases, as structural rearrangements and copy number variations leading to loss of heterozygosity are more common.

We can state that new technology is evolving rapidly, and great strides are being made to bring it to bear on diagnostics for the betterment of patient care. However, due to a lack of supporting evidence and thorough validation, these emerging technologies are still being complemented with additional Sanger sequencing for the confirmation of observed variants. Thorough optimization to yield high positive-predictive values, such as those observed by Cottrell et al.⁴ (98% to 100%), and to detect true positives missed by less sensitive conventional methods will make these NGS assays quite dependable. We fully expect comprehensive NGS-based testing to emerge quickly as a standalone test in routine cancer clinical diagnosis and patient care.

References