20 Standards and Standardization of Molecular Diagnostics

JOHN P. JAKUPIECAK AND CATHERINE D. O’CONNELL

1. INTRODUCTION

DNA has become a major target for clinical laboratory testing over the past 5 yr, and RNA testing is emerging for infectious disease and gene expression (1). To normalize laboratory results across different technology platforms as well as between laboratories, standardized reagents will become increasingly important. Reliable standards promote the speed at which a diagnostic test can be offered, as well as third-party reimbursement. Standardized control reagents ensure the ability of diagnostic laboratories to pass proficiency testing and quality assurance/quality control (QA/QC) measurements (2).

Consensus guidelines endorsed by professional societies and governmental agencies provide a framework for determining standardization needs. In its role supporting US science and industry, the National Institute of Standards and Technology (NIST), a nonregulatory agency of the US Department of Commerce, provides physical and chemical standards in support of national commerce, manufacturing, and science (3). These materials are available internationally as Standard Reference Materials (SRMs) for use by industry developing assays and/or technology platforms for diagnostic use, by regulatory agencies ensuring the quality and efficacy of these assays, and by clinical laboratories providing diagnostic tests for patients. Traditionally, NIST responds to standard needs as defined by these communities. Consensus is developed through NIST workshops attended by representatives of these communities as well as direct request by other governmental agencies. Specific examples of ongoing programs within the Biotechnology Division at NIST are described in this chapter.

2. TRINUCLEOTIDE REPEAT MEASUREMENTS

Basic research and technological advances in human genetics, biochemistry, and model systems have brought much progress toward the understanding of human infectious, hereditary, and somatically acquired diseases. In fact, whereas in 1991, triplet repeat expansion diseases could be described in a single article, now the remarkable developments within each disease has created volumes of work (4,5). One factor contributing to triplet repeat diseases is the adoption of unusual non-B DNA structures by the repetitive elements. These structures are associated with disease progression and severity. Despite all of the advances, an effective therapy is yet unrealized. However, the defining element of these diseases, the unstable DNA mutation, can accurately be measured. Some diseases already have diagnostics methods available [i.e., Huntington’s disease and Fragile X syndrome (6,7)]. Yet, in the case of Fragile X syndrome, accurate size determination of the triplet repeat region is not easily determined because of the nature of the repeat elements associated with this disease; consisting entirely of the C and G nucleotides. This sequence content increases the error rate of polymerase chain reaction (PCR) amplification and is commonly methylated in the disease state.

2.1. MEASUREMENT TECHNOLOGIES FOR ACCURATE SIZING: FRAGILE X SYNDROME

Current measurement technologies for repetitive elements and their adjacent flanking sequences are Southern blot hybridization, PCR amplification and electrophoretic separation, and DNA sequencing (7). The need for reference materials for Fragile X syndrome was discussed at a workshop held at NIST in 1998 entitled “Standards for Nucleic Acid Diagnostic Applications” (8). Reference materials for this syndrome are needed to provide accurate repeat size measurements across technology platforms and interlaboratory diagnostic and prognostic agreement. The specific size range for such a standard reference material was based on the likelihood of full mutation transmission (9).

The Biotechnology Division established a measurement program in this area, focusing on accurate size measurements after PCR amplification and sequencing. The accuracy of an optimized PCR amplification protocol to correctly measure the number of (CGG) repeats from normal, gray zone, and premutation length alleles was determined (4). The DNAs used in this study were reported to contain CGG repeat elements ranging from 29 to 110 repeats. Both slab-PAGE (polyacrylamide gel electrophoresis) and capillary measurements were conducted, and the factors impacting sensitivity, accuracy, and reproducibility of
results were examined. Long repeats (full mutation length alleles), which typically have clear clinical phenotypes, were not included in this study. DNA obtained from both previously tested clinical specimens as well as cell lines obtained from the Coriell Cell Repository were analyzed after PCR amplification. These samples were used as in-house controls by two clinical laboratories. Our initial measurements focused on measurement variability: (1) between slab-PAGE and capillary (CE) separation systems, (2) interlane variability, (3) intergel variability (slab-PAGE), and (4) variability during PCR amplification. We performed statistical analyses on system reproducibility and interlane and intergel variability. Samples were run in triplicate for all measurements and the analysis was performed using GeneScan analysis software. DNA sequencing was performed to verify repeat lengths.

2.1.1. Slab-PAGE Analysis As expected, the shorter alleles were more easily amplified and sequenced than longer alleles. The standard deviations for interlane measurements in slab-gels ranged from 0.05 to 0.35. The variation in size measurements performed on different gels and PCR amplifications ranged from 0.06 to 0.30 (7). This suggests that these measurements varied by up to a single nucleotide (0.33 of a three-nucleotide repeat).

2.1.2. CE Analysis The CGG repeat measurements performed by capillary electrophoresis were slightly more precise, with standard deviations ranging from 0.02 to 0.29. However, allele sizes observed after CE separations were significantly smaller than those obtained after slab-gel electrophoresis (Table 1, samples 6910 and 6968). DNA sequence analysis confirmed that the size measurements were correct for the slab-gel data and inaccurate for the CE results. It was hypothesized that the proprietary gel matrix used for capillary electrophoresis (POP-4™) leads to anomalous rapid electrophoretic mobility of CG-rich sequences (10).

2.1.3. Allele Size Analysis The peak ratio of each allele was compared within each female sample (Table 1). The detection method with peak ratios closest to 1.00 is more accurate but not necessarily more precise. The results indicate a gradual decrease in peak ratio with increase of allele size. As shown, premutation alleles contain the highest peak ratio discrepancy (comparison of samples 6910 and 6968 with 7541 and 13664). This bias would be even more pronounced in the detection of larger, premutation and full mutation length alleles (7).

Table 1 also shows that CE analysis has greater error detecting the presence of long alleles than slab-PAGE, as shown by the difference between peak ratios (i.e., shaded areas samples 6968; 0.59 for slab-PAGE vs 0.25 for CE). This suggests that the electrokinetic injection used in CE results in a bias toward capillary loading of smaller alleles obtained from female specimens. Therefore, both amplification and amplicon loading appear to contribute to peak ratio discrepancy.

2.2. SUMMARY OF ACCURATE SIZING METHODS FOR FRAGILE X SYNDROME The accuracy of our sizing data for Fragile X measurements within the normal, gray zone, and premutation allele sizes was within one repeat length for slab-PAGE measurements. The precision was equally high in lane-to-lane comparisons, comparisons of PCR results between gels, and in multiple PCR amplifications. There is no statistically significant evidence for heterogeneity of size determination after separation by either slab-PAGE or CE measurements. Hence, the gel matrix and running conditions for slab-PAGE were suitable for accurate size determinations (as confirmed by DNA sequencing) despite known migration anomalies (10). A GC-rich sizing standard would improve the accuracy of sizing by capillary electrophoresis, as this separation method results in high-precision measurements but incorrect size determinations.

Our data reveal several important considerations in the performance of Fragile X testing by PCR. First, size measurements were not directly comparable between the two separation systems for the larger, premutation, and presumably full mutation length alleles. Second, the POP4 polymer used in this study resulted in premutation size measurements that varied from actual size (as measured by DNA sequencing) by four to eight repeat elements. The greatest error (three repeats) detected in slab-PAGE measurements was found in the sample with the longest allele size—112 repeats as determined by DNA sequencing. All other normal, gray zone, and premutation measurements agreed with DNA sequencing. In addition, electrokinetic injection by CE resulted in allele-biased loading of premutation alleles. Because PCR methods are developed that robustly amplify full mutation alleles, the bias for loading smaller alleles could impact the ability to detect these alleles. High-precision measurements for samples containing long premutation and full mutation length alleles are currently under validation. Sizing standards that effectively allow cross-platform and interlaboratory comparisons are under development as a NIST SRM.

3. MEASUREMENT TECHNOLOGIES FOR MUTATION DETECTION; TP53 STANDARDS

According to the latest information provided by the publicly funded GeneTests website (http://www.genetests.org/), 928 diseases are currently tested for in medical genetics laboratories. Most of the current methods for the detection of mutations use PCR amplification technologies with subsequent analyte detection after electrophoretic separation or chromatographic separation of the products (11–15). Microarray technologies have also been used to detect DNA mutations (14). The US Food and Drug Administration (FDA) evaluates many of the innovations for commercial development of diagnostic tests. Although some have received approval, molecular diagnostic kits are rare. FDA approval of these technologies has been hampered, in part, by the lack of universal standards. New diagnostic assays emerge each year and their validation as accurate measurement technologies is critical for establishing the assay.