Dynamic mutations pose unique challenges for the molecular diagnostics laboratory

RONALD C. McGLENNEN

Routine clinical molecular testing of diseases associated with unstable or dynamic trinucleotide repeat syndromes poses unique technical, medical, and ethical challenges to the laboratory. Although the pathophysiology of these disorders is to date still largely undefined, the uniformity of their genetics has led to the development of highly informative diagnostic tests. In general, amplification techniques, such as the polymerase chain reaction (PCR), are used to determine the size of alleles within the genes linked to these disorders. Technically, these assays require empirical optimization so that the PCR reactions are both robust and reproducible, and occasionally other methods must be used to confirm diagnoses. Beyond execution of the test, however, the molecular diagnostics laboratory needs also to be fundamentally involved in the process of interpreting these tests in the correct clinical context and in setting policy as to how these data are presented to patients.

INDEXING TERMS: heritable disorders • trinucleotide repeat syndromes • fragile X syndrome • myotonic dystrophy • Huntington disease

In the wake of the rapid new gene discoveries in the past several years, a new paradigm in the pathogenesis of certain genetic diseases has emerged, namely, dynamic mutations. Distinct from the classical Mendelian models of dominant, recessive, and sex-linked inheritance, dynamic mutations are caused by ongoing alterations in repetitive DNA sequences within or adjacent to genes associated with specific disorders. Molecular characterization of some of these diseases has demonstrated a form of genetic instability that preferentially affects microsatellites such as those involving the repetitive di-, tri-, and tetranucleotide structures. The dynamic nature of these mutations is evidenced by the fact that they continue to change in subsequent generations—sometimes only in specific somatic tissues through cell mitosis, but more often in meiotic cell division, with preference for either maternal or paternal gametes. Clinically, some of these disorders are described as showing anticipation, classically defined as the progressive worsening of a phenotype in serial generations within an affected kinship, including increased severity of symptoms and deficits as well as an earlier age of onset. Among these disorders, which include several familial forms of cancer, is a unique group of diseases associated with mental retardation, movement disorders, and dementia that have as a fundamental genetic defect an expansion in the number of trinucleotide repeat units.

The trinucleotide repeat syndromes are the best characterized of the dynamic mutations. To date, seven distinct diseases, most associated with some form of neurodegeneration, have been shown to contain alterations in one of several characteristic trinucleotide motifs (Table 1): fragile X syndrome (CGG), spinal cerebellar ataxia types 1 and 3 (CAG), spinal bulbar muscular atrophy (CAG), dentatorubral-pallidoluysian atrophy (CAG), myotonic dystrophy (MD; CTG), Huntington disease (HD; CAG), and, most recently, Friedreich ataxia (GAA) [1, 2].

Regarding the reading frame and strand within the respective genes, all of the known trinucleotide repeat syndromes (except Friedreich ataxia) are associated with expansions of either CGG or CAG repeats. In the families of normal individuals, these trinucleotide repeat alleles are inherited stably. However, when expansions of these motifs result in disease, the disorder transmits as a dominant trait, although the precise allele sizes may vary considerably between different affected family members and even between different sources of tissue from one individual. Correspondingly, at the molecular level, genetic anticipation in the trinucleotide repeat syndromes is hypothesized to result from the continued increase in size of the disease-conferring allele.

Mechanistically, the causes of these dynamic mutations are largely unknown, and probably the fundamental molecular defect leading to the expansion of these microsatellites differs for each disorder. For example, expansion of the CGG repeat in the 5'-untranslated region of the FMR-1 gene causing fragile X

1 Nonstandard abbreviations: HD, Huntington disease; MD, myotonic dystrophy; PCR, polymerase chain reaction; HDSA, Huntington Disease Society of America; and DMSO, dimethyl sulfoxide.
syndrome is believed to arise from a progenitor premutation allele that acquires additional instability in passing through a maternal meiosis, whereupon it expands markedly from an allele with several tens of repeats to one of hundreds and occasionally thousands of repeats. The genesis of the premutation alleles, and why they are incrementally larger than the size distribution of alleles in the population as a whole, is completely unknown but may relate to the relative extent of methylation of DNA at this locus. Similarly, DM is the result of the expansion of the trinucleotide CTG, located in the 3' end of the myotonic kinase gene; this expansion can occur incrementally or very dramatically. DM serves as a classic example of genetic anticipation, wherein mild expansions, beginning with alleles of >50 repeats, cause aforme fruste of the disease; eventually, typically within two or three generations, very large expansions of up to 2000 repeats arise that are associated with congenital onset.

By contrast, HD is caused by a more moderate expansion of the trinucleotide repeat CAG located within exon 1 of the gene \( IT-15 \), which encodes for a tract of glutamine residues in the huntingtin protein [3]. In this case, transmission of the paternal allele imparts a greater degree of genetic instability, resulting typically in further expansion of the CAG motif in the range of 6–20 repeats, and leading to disease that is earlier in onset of symptoms. This apparent predilection for the paternal allele to expand is particularly true when the father himself is or will become clinically affected and has an allele in the affected size range (>37 or 38 repeats). Recent research has also focused on what, if any, risk of developing HD a person who inherits an allele at the high end of the normal range or in the “intermediate allele” range (30–37 repeats) [4]. Unlike fragile X syndrome and MD, the origin of HD has, until recently, been believed to derive from an ancestral founder, with new mutations in the \( IT-15 \) gene being exceptionally rare. However, reports of sporadic cases of HD arising from intermediate alleles in parents who have never shown signs of disease provide proof of the dynamism of this CAG repeat. The size of the progenitor allele alone may not explain the risk of HD, as suggested by the finding that intermediate alleles in families of sporadic cases of HD are more unstable than intermediate alleles in the general population [5]. Moreover, whether the inherent risk of genetic instability in HD alleles is due to a sequence alteration, e.g., a difference between perfect and imperfect repeats, or because of some other factor that lies cis or trans to the \( IT-15 \) gene, such observations have been made possible only by the increased availability of testing to a broader variety of human populations.

Although the scientific basis for the cause of dynamic mutations may be elusive, their very nature provides an enormous advantage in designing diagnostic tests. Unlike testing for other genetic diseases, analyses of the mutations in the trinucleotide repeat syndromes are highly uniform. Thus, >95% of cases of fragile X syndrome are due to the expanded CGG repeat in \( FMR-1 \), although a very rare example of coding-region point mutations has been described. In fragile X syndrome, expansion of CGG leads to hypermethylation at this locus and consequently a lack of \( FMR-1 \) transcripts. In the example of HD, the CAG expansion does not affect gene transcription, even in homozygotes, but results in an altered protein structure, which gains unique function. Thus, investigators believe that, in 100%

---

**Table 1. Characteristics of trinucleotide repeat syndromes.**

<table>
<thead>
<tr>
<th>Disease</th>
<th>Dynamic mutation (gene)</th>
<th>Size range of alleles: no. of trinucleotide repeats</th>
<th>Pathology</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRPLA</td>
<td>CAG (atrophin)</td>
<td>Normal 7–23 Premutation na Expanded 49–75</td>
<td>Degeneration of cerebellum, brainstem, basal ganglia, spinal cord, and cerebral cortex</td>
</tr>
<tr>
<td>Fragile X syndrome</td>
<td>CGG (FMR-1)</td>
<td>6–48 52–200&gt;200</td>
<td>Methylation of regulatory region of ( FMR-1 ), resulting in absence of expression in all body tissues</td>
</tr>
<tr>
<td>Friedreich ataxia</td>
<td>GAA (X25)</td>
<td>7–22</td>
<td>Degeneration of dorsal root ganglia, large sensory neurons, posterior columns, spinocerebellar tracts; hypertrophic cardiomyopathy</td>
</tr>
<tr>
<td>Huntington disease</td>
<td>CAG (( IT-15 ))</td>
<td>11–27 28–36</td>
<td>Expansion of a polyglutamine tract within exon 1 of ( IT-15 ); degeneration of neurons in basal ganglia and cerebral cortex</td>
</tr>
<tr>
<td>Myotonic dystrophy</td>
<td>CTG (myototonin kinase)</td>
<td>5–30</td>
<td>Muscle loss, cardiac arrhythmia, cataracts, frontal balding</td>
</tr>
<tr>
<td>Spinal bulbar muscular atrophy</td>
<td>CAG (androgen receptor)</td>
<td>11–33</td>
<td>Expansion of a polyglutamine tract within the human androgen receptor, leading to degeneration of neurons in the spinal cord, brainstem, and sensory neurons</td>
</tr>
<tr>
<td>Spinal cerebellar ataxia (SCA) type I</td>
<td>CAG (ataxin 1)</td>
<td>6–44</td>
<td>Increased function of ataxin 1 associated with loss of Purkinje cells in cerebellum and inferior olivary nuclei</td>
</tr>
<tr>
<td>SCA type III</td>
<td>CAG (MJD 1)</td>
<td>13–40</td>
<td>Degeneration of multiple motor control regions of brain and spinal cord</td>
</tr>
</tbody>
</table>

DRPLA, dentatorubral-pallidoluysian atrophy; na, not applicable.
of cases of HD, a single test detecting the large CAG repeat allele can provide for diagnosis of HD with near perfect sensitivity and specificity. The issues remaining for the clinical laboratory then are to ensure a more accurate diagnosis of HD and of each of the other dynamic mutations by implementation of a robust and standardized protocol.

In the clinical laboratory, much work is still to be done before these assays are robust enough to boast the degree of diagnostic precision that is expected in testing for HD. In general, the routine performance of molecular genetic assays for trinucleotide repeat syndromes is time consuming, technically difficult, and in practice involves occasional batch failures that necessitate repeat analyses, often using more than a single methodology to resolve issues of homozygosity or allele sizing not amenable to techniques such as polymerase chain reaction (PCR) (Figs. 1-3). In the cases of the fragile X syndrome and MD, complete molecular evaluation involves first a PCR-based assay to precisely measure the sizes of the normal and the premutation alleles. In the routine performance of these assays, positive controls that challenge the limits of routine PCR, typically in the range of 120–180 repeats, are an important gauge of a laboratory's ability to diagnose a high premutation or mild form of disease (Fig. 1, 2A, and 2B). To diagnose a patient suspected of being clinically affected, one must be able to assess the alleles, which are invariably markedly expanded (usually to >1.5 kb or >500 repeats) and therefore are not detectable by conventional PCR amplification. Consequently, confirmation of the diagnosis is established by performing a Southern transfer of genomic DNA that has been cut with the appropriate restriction enzymes and hybridized with probes overlying the locus of triplet repeat DNA (Fig. 2C).

In testing for HD and other diseases associated with moderately sized CAG repeat expansions, one or several sets of PCR-based reactions can detect all but the very largest alleles. However, despite the simplicity of these algorithms, the experience of performing a large number of PCR-based assays for HD and other trinucleotide repeat disorders has shown that each test has its specific nuances and limitations. In particular, two problems—PCR optimization and choice of a sensitive PCR detection system—are the central technical issues of this type of testing. At present, both problems can be solved by various methodological approaches, but which approach to use is a decision predicated on the volume of work, the menu of tests, and budgetary constraints.

In this issue of Clinical Chemistry, Muglia et al. report one more in the series of valuable technical innovations in applying PCR to diagnostic testing for HD [6]. Their work principally describes the utility of a sensitive silver-staining technique as the means of detecting PCR-derived DNA products. Although detection of DNA by silver nitrate staining is known to be less sensitive than radioisotopic or automated fluorescence detection, the authors have emphasized the importance of modifying the PCR protocol to achieve more-efficient amplification, which in turn allows for the use of silver staining. Correspondingly, their work illustrates that, regardless of the merits of any one published approach, each laboratory will need to optimize both the amplification chemistry and the choice of detection method for each test according to its own experience.

**Current Laboratory Approaches to Testing for HD**

Clinical laboratory testing for the trinucleotide repeat syndromes are among the more complex types of testing practiced in molecular diagnostics laboratories. The complexity of these tests lies not only in the technical execution of the assay, but also in dealing with all of the consent forms and acquisition of the vital clinical data needed to interpret the test. Finally, in reporting the results of the tests, the laboratory has the responsibility to ensure that patients are dealt with in an empathetic and confidential manner. The number of laboratories that currently offer this test routinely is limited. In the most recent testing survey for molecular genetics offered by the College of American Pathologists, 30 laboratories from North America participated in the HD testing section. The survey tested the ability of each laboratory to correctly size the HD alleles in cell
formamide.

Analysis

Fig. 2. Two-part analysis of the CTG repeat alleles at the 3' region of the myotonic kinase gene associated with myotonic dystrophy.

Analysis of the normal and slightly expanded alleles was first achieved by PCR of this region in two different protocols: amplification with 32p-labeled CTP (A) with no denaturing additives and (B) with 50 mL/L formamide. The PCR products were electrophoresed in 6% denaturing polyacrylamide gels, dried under reduced pressure, and exposed to x-ray film overnight with double-intensifying screens. The results demonstrate the improved signal of the expanded alleles in reactions containing formamide. Other denaturants such as dimethyl sulfide (DMSO) or glycerol may also improve results. (C) Southern transfer analysis of the samples illustrated in A and B shows the characteristic normal alleles sizes (9 and 10 kb) and the smeared band(s) >10 kb typical of a markedly expanded (affected) allele (patient B).

lines derived from clinical samples and to render a diagnosis. In the majority of laboratories, the routine approach toward HD testing involved the exclusive use of PCR to amplify the sequences flanking the polymorphic CAG repeat in exon 1 of IT-15. The most commonly used oligonucleotide primers were those that immediately flank the polymorphic CAG tract (so-called HD 1/3), which create DNA fragments of 80–143 bp (11–32 CAG repeats) for the normal-sized alleles and 158 bp and larger for the affected alleles [7]. A minority of the participating laboratories used an alternative selection of primers, principally the “HD 1/2” set, which includes the additional polymorphic CCG motif immediately adjacent to the disease-associated CAG tract. This region of the HD gene has previously been shown to contain at least five alleles, ranging from 7 to 12 CCG repeats, that can impact the interpretation of the CAG repeats linked to HD [8]. This proved to be true in the CAP HD survey, wherein the laboratories that relied exclusively on the HD 1/2 primer set tended to report allele sizes between 1 and 3 repeats larger than those obtained with the HD 1/3 primers. One particularly useful feature of the polymorphic CCG tract is that at least one population studied shows a marked degree of linkage disequilibrium toward the 7-repeat allele and the expanded chromosome [9]. Thus, HD 1/2 primers are of great value when the results of the HD 1/3 analysis demonstrate homozygosity or when there is concern about missing an occult or ambiguous high allele in the affected size range (Fig. 3).

Fig. 3. Southern analysis of PCR-amplified HD locus to detect very high-repeat-number alleles.

Genomic DNA was amplified with primer sets HD 1/3 or HD 1/2 but with no radioactive nucleotide used in the PCR incorporation. The products were electrophoresed in 2% agarose and then transferred to a nylon membrane by capillary transfer. The membrane was then hybridized and probed with a 32p end-labeled oligonucleotide, (GTC)n, washed under stringent conditions, and exposed to x-ray film at room temperature for 15 h. The very high-repeat HD alleles are not typically visualizable by conventional protocols. Illustrated are two examples of juvenile HD with genotypes 19/17-240 (patient A) and 17/17-160 (patient B), which were not amenable to conventional HD protocols.
Another advantage of using multiple primer sets is to establish a short haplotype when nonpaternity might be an issue. In the case of de novo HD mutations, this concern is relevant because, by definition, the allele sizes on the affected chromosome are expanded to a greater extent than expected—either from dynamic change between the generations or from an assay artifact. To this end, use of a variety of linked dinucleotide repeat markers, both centromeric and distal to \( IT-15 \), helps to set the phase of haplotypeing and thus imparts added confidence in the diagnosis [10].

Therefore, a rational approach to HD testing involves the routine use of several oligonucleotide primer sets for PCR-based allele sizing, as well as the use of anonymous linked markers to aid in the evaluation of families. In some cases, a genomic Southern transfer (Fig. 3) can aid in the detection of large expanded alleles, but this is not a widely used technique [11].

Surprisingly, only a handful of laboratories in the CAP survey responded that they use a multifaceted approach for HD testing. Unlike many of the routine analyses available for some constitutional disorders, testing for HD, whether in a symptomatic patient or in presymptomatic evaluation, includes an additional layer of concern for the emotional impact and ethical ramifications of making such testing results available. The apparent reliance on a single PCR-based assay approach to diagnose all cases underscores how important it is that the assay work.

TECHNICAL ASPECTS OF PCR FOR HD TESTING

In the first publications describing a PCR-based assay for detection of HD alleles, the investigators acknowledged that, because of the peculiar sequence of the DNA at this locus, conventional amplification schema were usually inefficient and thus not reliable for routine clinical testing. A series of short publications followed that focused on how to improve the sensitivity of the detection system. Relatively little has been written on improving the efficiency of the PCR reaction itself to provide adequate amplification of the expanded allele without the compromise of creating a large number of nonspecific products that can obscure interpretation. Muglia et al. [6] place major emphasis on the silver nitrate detection system, but the results also illustrate the importance of optimizing the PCR reaction. The two central problems with these particular PCR reactions are that (a) the target sequence is characteristically more difficult to amplify because of the tendency of the template to form secondary structures and thus make for low primer-annealing efficiencies, and (b) the highly repetitive sequence at this locus negatively affects the efficiency of Taq polymerase in base incorporation and thus results in a low processivity.

The first problem, which encompasses finding the balance point between primer multimerization and efficient primer-template interaction, can be addressed by conducting experiments that investigate optimal primer melting and annealing temperatures \( T_m \). Determination of the theoretical \( T_m \) whether by various commercial software algorithms or by hand calculations, yields results generally below the permissible temperature for multimerization, such that primers are also able to combine with each other because of the terminal transferase activity of Taq polymerase. Optimal \( T_m \) values can be determined by testing different temperatures—in the absence of target—in 2–3 °C increments around the theoretical \( T_m \) and choosing the lowest temperature at which no dimer products are visible. The primers used in the HD assay require high \( T_m \) because of their high proportions of the bases guanosine and cytosine (GC). Similarly, the HD locus, which is \(~86%\) GC rich, is exceptionally prone to forming hairpin and other secondary structures in the DNA [12]. The issue of decreasing secondary structure in the template is addressed by including various chemical additives that effectively "iron out" the DNA and make it more amenable to primer annealing. Such chemicals include DMSO, formamide, and glycerol, which act principally as template denaturants. Certain protocols call for adding only one of these agents, e.g., DMSO or glycerol, in amounts \( \leq 10% \) of the reaction volume—although DMSO in particular has been shown to decrease the activity of Taq by as much as \( 50% \) [13].

In describing their PCR protocol, Muglia et al. determined that formamide at a final concentration of 35 mL/L in the reaction volume was "necessary" to achieve adequate amplification of the upper repeat-number alleles. Moreover, they found DMSO not helpful in the assay. However, in the spirit that characterizes PCR assays for the trinucleotide repeat syndromes, our laboratory and others have found just the opposite to be true for the fragile X assay (personal communications). Adding these agents in combination may or may not have synergistic effects on the reaction efficiency, but, again, the results of these modifications are easily observed by simply monitoring the end products.

Regarding the second problem of enzyme efficiency in amplifying the highly repetitive sequence of the trinucleotide motifs, another group of innovations improves the performance of the Taq polymerase in the PCR reactions. The first is in the use of 7-deaz deoxyguanosine triphosphate (7-deazoG\(_3\)TP). In theory, add \( n \) of 7-deazapurine deoxynucleotides alters the primary structure of DNA within the major groove and imparts resistance of the fragment to digestion by select endonucleases and even the intrinsic exonuclease activity of Taq polymerase [14]. However, only 7-deazoG\(_4\)TP, but not 7-deazoA\(_n\)TP or 7-deazoL\(_n\)TP, is accepted into the elongating DNA chain by Taq polymerase. In practice, 7-deazoG\(_3\)TP is incorporated into PCR-amplified DNA less efficiently than G\(_3\)TP but is particularly useful in highly GC-rich DNA, where it may reduce the unfavorable aggregation properties known for G\(_3\)TP and does reduce the secondary structures of the DNA chain. Many protocols, including the one described by Muglia et al., have found it invaluable in the HD assay, typically in a 1:3 or 1:4 ratio with G\(_3\)TP. Maintaining the fidelity of the PCR reaction, however, requires maintaining an equimolar concentration of each nucleotide. These modifications, in concert with such additives as bovine serum albumin and 2-β-mercaptoethanol, which help to stabilize the enzyme and the enzyme/template complex, improve not only the processivity of the DNA synthesis but also the fidelity of the nucleotide incorporation reaction.
LABORATORY COMMITMENT BEYOND TECHNICAL INNOVATIONS

Providing comprehensive service in molecular diagnostics goes beyond the excellent performance of the test. The challenge to provide comprehensive testing for HD and other trinucleotide repeat disorders mandates that the laboratory work beyond the walls of the laboratory and be involved in the network of clinical geneticists and genetic counselors who are caring for these patients. The laboratory can set policies to ensure that the best interests and the rights of patients being tested are protected from those who neither understand the disease (or the test results) nor care about the consequences—both medical and ethical—of making these diagnoses. This protection can be achieved in part by striving to maintain a high level of communication between those offering the results of the test and the clinicians who will in turn interpret the results to the patient. In practice, these additional challenges imposed on the laboratory offering HD testing actually begin at the initial physician/patient encounter. Ethical practice guidelines, as established by the Huntington Disease Society of America (HDSA), require an appropriate amount of preanalytical counseling with the patient suspected of carrying the HD gene. This amounts to the need to obtain a complete family history, complete with a pedigree, followed by a thorough physical examination by a physician appropriately trained in the clinical signs of HD and knowledgeable in the interpretation of neuropsychiatric testing directed at diagnosing HD. In some areas of the US, clinical centers specializing in diagnosing HD provide these services. According to HDSA guidelines, patients not local to an HD center should be referred to one, and this is to take place before a blood sample is obtained for molecular testing. The HDSA also recommends that, after the clinical evaluation, time be specifically dedicated for the physician and patient to discuss the option of direct gene testing, with an emphasis on what the diagnosis is likely to be and what the results will mean to that individual’s immediate and future plans.

The laboratory must play a key role in assuring that these preanalytical steps are taken. One recommendation is that the laboratory supply testing kits that include an information brochure about the test, its cost, and to whom the bill will be submitted. In addition, a consent form (also supplied in the kit), which is to be signed by the patient and the referring physician, would be required before the sample is processed. Some molecular diagnostic laboratories rely on genetic counselors, who work closely with the laboratory, the clinical geneticists, and the patients, to follow through on these items [13]. Additionally, genetic counselors may provide their experience in counseling per se to physicians unfamiliar with dealing with genetic diseases and genetic test data. Ideally, each of the physicians requesting testing should be known to the genetic counselors and to the laboratory so as to facilitate the continuity of concern inherent to these disorders. In practice, this can be difficult, particularly because only a few specialized centers offer these analyses and, therefore, are likely to be located geographically distant from the clinical center.

Finally, to return to the theme that the dynamic mutations represent a unique area of clinical genetics, our knowledge base is clearly far from complete in the technical issues of these tests, as well as the regulative and ethical aspects. It remains for the laboratory that is truly dedicated to providing comprehensive services in these areas to be fundamentally involved in the attempts to standardize testing, monitor quality control, and advise empowered institutions about how this evolving discipline of medicine is to be regulated. Unlike other areas of the clinical laboratory, where automation and standardization are more highly developed than in molecular diagnostics, at present a laboratory has relatively few opportunities to compare itself with others with regard to the growing number of tests being offered clinically. One of the first steps is for each laboratory to develop a correlative database so that the collective experience of many laboratories can be compared. Of equal importance is the need to provide physicians and, occasionally, families with data regarding experience with high- and low-repeat alleles and with intermediate alleles. To do this, the laboratory must take deliberate steps to retrieve all of the relevant clinical and family data and to periodically examine these data.

This exercise is also a convenient excuse to find out how the laboratory-generated test data were dealt with once they were reported. Specifically, the issue of ensuring patient confidentiality is one that each laboratory must strive unceasingly to maintain. One of the disheartening, but increasingly common, consequences of providing genetic testing is that patients’ data are being used not only for purposes other than for medical judgment but occasionally to discriminate against people in insurance applications or in the workplace [16]. Perhaps nowhere is this more true than in testing for diseases such as HD. One of the obligations, then, of the invested laboratorian is to help direct the growth of molecular diagnostics in ways that are positive not only in technique, but also in its great potential to aid humankind.

References
8. Rubinsztejn DC, Barton DE, Davison BC, Ferguson-Smith MA.


