Molecular Genetics in Clinical Practice

Trinucleotide Repetition and Fragile X Syndrome

STEPHEN T. WARREN
Emory University

By 2005, or sooner, the three billion code letters of a representative human genome will be known, along with the locations of all of its genes. Even today, however, the work is greatly accelerating identification of disease-related genes. One outcome will be tests for genetic components of risk in the majority of common illnesses. In the longer run, genetic discoveries will surely lead to new treatments.

As much as 3% of the human population is mentally retarded, and at least 300 genetic disorders include mental retardation. In seeking the genes involved, it has long been clear that males are at special risk--as much as three times greater than females. Thus, there has long been reason to suspect X-linked mutations. To the extent that the mutations are recessive, females would be protected by having two X chromosomes. Initially, it was assumed that no single mutation would overshadow the others. Then, in the 1970s, an individual locus was identified cytogenetically as a constriction called a fragile site, near the tip of the long arm of the X chromosome (Figure 1). Eventually, fragile X syndrome would be recognized as among the most frequent causes of mental retardation, affecting one in 1,500 males and one in 2,500 females. By itself, it accounts for nearly 10% of all inherited retardation and almost half of the male preponderance. For a sex-linked mendelian trait, its inheritance pattern is odd. Genotypically, it is an X-linked dominant disorder, as confirmed by its capacity to affect heterozygous females. Yet some males with a fragile X chromosome appear to be unaffected. Overall, the penetrance is only about 80% in males and 35% in females. Hence, despite its genetic dominance, the trait has many carriers.
Even more oddly, the risk of mental impairment varies with position in a pedigree. Men unaffected by having a fragile X chromosome are called transmitting males because they pass the defective chromosome to all of their daughters, who have almost no risk of mental impairment. In marked contrast, the sons of carrier women have a 76% penetrance, and the daughters a 32% penetrance. In the children of impaired women, the figures are 100% for boys and 56% for girls. In brief, disease expression depends on inheritance from the mother, and as the disease trait descends through a kindred, the proportion of affected boys becomes greater in each generation (Figure 2). No standard type of mutation could have accounted for any such changes.
Indeed, the peculiarities resisted all explanation until the mutation was understood. By 1991, fragile X syndrome had been mapped to a small interval marked by induced translocations at the fragile X site in somatic cell hybrids. Intensive investigation of the region disclosed two forms of abnormality in fragile X chromosomes: a span of DNA hypermethylation and also a length variation, soon traced to a great number of repetitions of the trinucleotide CGG. The area affected was a gene now designated FMR1 (fragile X mental retardation 1). It codes for FMRP, a protein predicted to consist of 657 amino acids. The risk of being a patient or only a carrier has since proved to relate not to chromosomal constriction or fragility per se but to the length of the trinucleotide repetition and to the degree of methylation; the latter appears to be the final event that silences gene expression and causes fragile X syndrome. In turn, the discoveries raise new questions about how the mutation evolves and gets transmitted, how the disease phenotype develops, and how a therapy might be devised. Meanwhile, abnormally lengthened trinucleotide repeats have been identified in 11 other genetic diseases.

**Phenotype and Genotype**

Clinically, fragile X syndrome is hard to recognize, especially in newborns. Even in fully affected males, its outward appearance is unremarkable. Generally, adult male patients have a long, narrow face with a prominent jaw and forehead, large ears, and a somewhat increased head circumference. In some additional features—hyperextensible joints, a high, arched palate, flat feet, mitral-valve prolapse—the syndrome may resemble a connective-tissue disorder. Testicular enlargement is common. The more significant features are delayed development and mental retardation, so that boys may present with avoidance behavior, hyperactivity, or attention deficit. Similarly, mild communication difficulties are common. The average IQ is in the moderate-retardation range. Females are usually much less markedly involved, both somatically and in measures of retardation.

Genotypically, the abnormality is far more clear-cut. Well over 99% of all patients share the same type of mutation, a greatly expanded trinucleotide repetition in FMR1 (Figure 3). Among normal persons the repeat number is
polymorphic, ranging from seven to 52. Most commonly, the number is 30. In fragile X syndrome, the number is dramatically expanded, to hundreds or even thousands, and the expansion is often variable, with different repeat lengths in the same or different tissues. Patients with especially marked variability in this respect have been termed mosaics. When the repeat number exceeds about 230, the DNA becomes abnormally methylated, and the gene becomes nonfunctional. Thus, repeat expansion plus methylation are said to constitute a full mutation. Most carriers have only a so-called premutation: an unmethylated, transcriptionally active allele with 60 to 200 repeats. Unfortunately, a premutation can be extremely unstable. Indeed, maternally transmitted premutations with more than 100 repeats almost always expand into full mutations.

Figure 3. Trinucleotide expansion responsible for fragile X syndrome lies in an unexpressed part of the X-linked gene FMR1. The gene itself (top) divides its code into 17 exons spread over 26 kilobases. Its first and last exons include regions transcribed into messenger RNA that are not represented in the final translated protein. In turn, the 5' untranslated region (in exon 1) includes a sequence of CGG repeats (bottom). Normally, the tract is polymorphic, ranging from 7 to 52 repetitions. The example shown is the most common, with 30. In a premutation, the number is 60 to 200; the example has 96. In a full mutation, the number is almost always several hundred; the example has 720. When the number exceeds 230, the entire region is hypermethylated (inset), receiving a methyl group at the C in each CG dinucleotide along both strands of the DNA double helix. The gene's promoter is deactivated, and the gene becomes silent. Arrows in and near FMR1 mark the locations of three polymorphisms.
How does the expansion occur? Perhaps the best way to approach this question—a central enigma in fragile X syndrome—is to begin by asking another: Are some alleles predisposed? Identification of polymorphic markers in and near FMR1 aids in probing the latter question. Such idiosyncrasies occur in a number of places, as they would around any gene. If they are irrelevant to a disease, a population of normal persons and a population of patients will show no differences in their frequencies (assuming the two groups have the same mix of ethnic backgrounds). In some instances, however, the frequencies are dramatically skewed. Some haplotypes—that is, sets of markers—are far more common in patients than in a normal population; others are far more common in normals. The implication is that something special about a particular set of chromosomes causes predisposition, or that something confers protection, leaving the rest at heightened risk.

For such an analysis, my colleagues and I chose three markers, designated AC1, FMRa, and FMRb. One lay upstream, and two downstream, from the trinucleotide repeat (see Figure 3). Among 85 normal X chromosomes from Caucasian subjects, 24 had short repeat lengths (less than 27 triplets). All were of haplotype 2 (consisting of the C, A, and A alleles of AC1, FMRa, and FMRb, respectively). By contrast, the 61 chromosomes with longer repeat lengths (34 to 52) were chiefly of haplotype 3 (ABB). However, in contrast to both groups, the fragile X chromosomes of 97 Caucasian patients were most commonly haplotype 1 (DAA). Indeed, something special about particular chromosomes appeared to have influenced the trinucleotide repeat, both in its normal variation and in the transition to fragile X. Yet in themselves, longer repeats were not necessarily predisposing.

Accordingly, we examined in detail the sequence of the trinucleotide repeat. Fundamentally, the repeating unit is CGG, but with occasional cryptic AGG triplets interspersed. They are called cryptic because they go undetected in any assay that simply measures the length of the tract. In 82 normal Caucasian X chromosomes, the sequence showed marked variation, including 19 differences in the number and location of AGG triplets (or their complete absence). If cryptic variation and total length were both considered, 47 patterns emerged.

The patterns nevertheless were nonrandom (Figure 4). In nearly all haplotype 1 and 3 chromosomes, the first interruption in CGG triplets was at the tenth triplet. In haplotype 2, the more common first interruption was at triplet 11. Across the three haplotypes, a second AGG was frequent at triplet 20 or 21. Evidently, in several alleles a replication error caused gain of a CGG triplet, shifting the AGG triplets by one position. Otherwise, there was no sign of change in length. In a few alleles, a third cryptic interruption occurred at position 30. Toward the 3' end of the DNA, the variability increased, especially in regard to tract length. The overall impression was that the rate of a mechanism that lengthened the 3' end of the repeats had exceeded the rate at which the 5' end changed. Perhaps the presence of AGG triplets had somehow increased the stability of the first part of the region.
Figure 4. Genetic variability found in 82 normal Caucasian X chromosomes offers clues to the history of the fragile X mutation. The analysis depends on the identification of three polymorphisms near the trinucleotide repeat (FRAXA); these are AC1, a dinucleotide repeat 7 kilobases upstream, and FMRa and b, two single-nucleotide polymorphisms 6 and 17 kb downstream. At these sites, three combinations, or haplotypes, are found. Haplotype 2 accounts for the shortest FRAXA tracts (less than 27 triplets) and haplotype 3 for many long ones (34 and over). Yet haplotype 1 is known to be the most common among fragile X patients. Attention turns, therefore, to a form of variability consisting of AGG triplets amid the CGGs. Across all haplotypes, the first such feature is usually at triplet 10 or 11, with a second at 20 or 21, suggesting that local stability is undermined by no more than occasional gain of a CGG, shifting the AGGs by one position toward the 3' end of the DNA; the variability increases, especially in regards to length. The overall impression is that AGGs somehow bolster the first part of the region.
We then turned our attention to the FMR1 alleles in a family with no incidence of fragile X syndrome but with an unstable repeat length of 54 to 60, at the boundary between normal and premutation. In alleles with 40, 56, and 58 repeats, from two brothers, there were no AGG triplets. The chromosomes were all of haplotype 1—for which, among 23 chromosomes in our normal sample, only one lacked AGG interruptions.

### A Hypothetical History

Overall, fragile X chromosomes may derive from an ancestral pool of normal X chromosomes by a mechanism dependent not solely on the trinucleotide repeat length in the normal chromosome but also on the number and position of cryptic interruptions. The available data suggest a multistep history in which a normal allele becomes a predisposed allele, which eventually becomes a premutation, which then expands into full mutation.

At the outset, safety resides in the shortness of the CGG repetition or in its interruption by one or more cryptic AGGs. Thus, a normal allele becomes a predisposed allele if it lengthens or loses an interruption, creating in either case a longer perfect tract than before. Conceivably, the most common such event is the deletion of an AGG triplet, or perhaps a point mutation of A to C. FMR1 genes containing a perfect run of more than, say, 25 repeats may be predisposed alleles. Such a surmise at least accords with what is now known of other triplet-expansion-related disorders. In Huntington disease, the culprit CAG tract does not show cryptic interruptions even in normal alleles, but normal alleles have less than 25 repeats and predisposed alleles appear to have a few more than 25.

For its part, a predisposed allele has a propensity for eventual lengthening into a premutation. The precise explanation remains elusive. One hypothesis is that risk accrues during DNA replication, because perfect repeats may prevent the replicative enzymes from recognizing a slippage along the template DNA (Figure 5). During replication, the DNA double helix is unwound, producing the forks at which a polymerase complex can bind to begin making new DNA. One of the unwound strands—the leading strand—gets a new companion strand continuously. The other—the lagging strand—gets its new counterpart discontinuously, in short spans called Okazaki fragments, synthesized along the length of the lagging strand.
Figure 5. Slippage mechanism proposed as an explanation of the expansion of a normal trinucleotide repeat into premutation size relies on an idiosyncrasy in DNA replication, namely that one parent strand—the lagging strand—gets its companion strand discontinuously, as a sequence of Okazaki fragments (top). If the repeat tract were perfect but short (A), the matching fragment (or fragments) would extend into unique flanking sequence (multiple colors), which offers a strong basis for registration. Likewise, even a single AGG within a run of CGGs might be a signpost (B). If, on the other hand, the tract were long and perfect, an Okazaki fragment might start and end within its length and have only the repetitions as a means for alignment (C). In such circumstances, a loop in an Okazaki fragment might go unrecognized, and repair enzymes might add nucleotides to match the loop’s length. The illustration omits some details of DNA replication; for example, each Okazaki fragment begins with a short RNA “primer.” Also, DNA synthesis proceeds base by base, not triplet by triplet, as schematized, and each base is complemented, not duplicated.

If such a fragment must register itself against a sequence unique at many points, such as TCCATCGCGCT... (to cite at random some code from exon 1 of FMR1), a mistake arising from polymerase slippage along the template will be rapidly caught when the polymerase tries to impose an incorrect base pairing. (To form the DNA double helix, A links with T and G links with C.) If, on the other hand, the fragment must register itself against a rhythm of CGCGCGCGCGG..., a slippage may go undetected, because the base pairings may continue to seem correct. True, if the rhythm were short enough, the fragment would extend beyond it and into unique flanking sequence. This might offer a basis for ensuring correct registration. Likewise, even a single A within a run of CGGs might be a signpost. If, however, the tract were long and perfect, an Okazaki fragment might start and end within its length and have only the repetitions as a basis for alignment. In such circumstances, the polymerase might introduce a loop into the fragment by unknowingly working twice along the same template sequence. The outcome might be a DNA elongation.

Initially, the risk of such a mishap may not be great. But over successive
generations the length of a predisposed allele may increase a bit, and then a bit more. At a length of about 40 repeats, or 120 nucleotides, the tract would attain the typical length of an Okazaki fragment. At such a stage, further changes might predictably be observed from parents to their children. Ultimately, the tract would become a premutation.

In turn, a premutation alters abruptly into a full mutation. The expansion can be from, say, 80 repeats in a premutation mother to 900 or even 1,200 in her child. It is hard to imagine that polymerase slippage alone can cause so dramatic a change. Although 70 repeats is far longer than a typical Okazaki fragment, it is far too short to permit an undetected slippage hundreds of nucleotides in length. There must be some other process. One idea is that a polymerase might stall somewhere along CGGCGG..., continuing to synthesize DNA while it fails to advance along the template.

Precisely when might this happen? In the most popular current idea, the oocyte of a premutation woman still has only a premutation. The full mutation first appears soon afterward, mitotically, as cells divide in the early embryo. There is, however, an alternative possibility in which the oocyte of the premutation woman already has the full mutation, created meiotically. In this view, for which new evidence has recently been found by my laboratory, premutations in mosaic patients would represent length reductions. Such reductions are known to occur. Indeed, they may account for the curious fact that persons with fragile X syndrome inherit it only from their mother. It now appears that males with a full mutation (or even only a premutation) cannot tolerate full-mutation sperm. In their testes, occasional length reductions are sufficient to reanimate FMR1. In turn, germline cells expressing FMRP outcompete those that cannot express it.

Overall, the stability or instability of a given FMR1 allele is probably more complex than the preceding account may suggest. Differences in flanking sequence may have an effect. For instance, a polymorphism placing a replication fork near the trinucleotide repeat may be more dangerous than a polymorphism placing one farther away. In addition, individuals may have different capabilities for DNA repair, so that some persons tolerate polymerase slippage better than others. Still, the entire sequence from the loss of an AGG to the final dramatic expansion of a premutation into full fragile X syndrome might span perhaps 80 generations. Thus, the disease clinicians see today may have had its genetic origin two millennia ago.

By the same token, there seem today to be chromosomes that do not cause fragile X syndrome but are destined to progress along the sequence. In particular, haplotype 3 includes X chromosomes with perfect runs of more than 24 CGGs about as often as haplotype 1 does (1% vs. 0.8%). It accounts, however, for only 25% of fragile X chromosomes (haplotype 1 accounts for 46%), perhaps reflecting a relatively recent evolutionary origin—a suspicion strengthened by findings that haplotype 3 is not represented in all populations. In view of such findings, it has been predicted that in the future the frequency of fragile X syndrome may increase slightly.

FMR1

Genetically, the final event in the transmission of fragile X syndrome is the abnormal methylation of FMR1, probably early in embryonic development. It occurs at virtually every site where a methylation is possible—namely, each CG dinucleotide. (The C receives the methyl group). Thus, the repeat itself is massively methylated at virtually every C. In addition, the flanking DNA undergoes methylation for many kilobases. At the downstream side, the methylation may extend past the gene's final exon. At the upstream side, the gene's promoter is well within reach. It lies not more than 250 bases away.
As an intracellular process affecting DNA, methylation is well recognized. There is evidence implicating it in imprinting, the geneticist's term for the differences through which DNA "remembers" whether it derived from a male or a female. In Prader-Willi syndrome, mental retardation is accompanied by obesity, short stature, a rounded face, and small hands and feet. The syndrome results from disruption of a gene or genes on the long arm of chromosome 15. In Angelman syndrome, retardation is accompanied by a quite different set of features, including a large mandible and an open-mouthed expression. This syndrome derives from the same chromosomal region. The difference appears to lie in whether a disruption affects the paternally (Prader-Willi) or the maternally (Angelman) derived chromosome. In at least one gene in the region, distinctive methylation patterns have been identified in alleles of paternal versus maternal origin.

Ample evidence also implicates methylation in gene inactivation. There is, for example, the Barr body--the transcriptionally silent, heavily methylated X chromosome found in each female somatic cell. More broadly, throughout the genome, genes inactivated during development usually show a methylated promoter (as contrasted with the unmethylated promoter of an activated gene).

Why it happens in fragile X syndrome is not understood. One hypothesis is that methylation is always occurring, but to only a small extent. In fragile X syndrome, a greatly expanded trinucleotide repeat causes the DNA to take on an unusual local conformation, providing a substrate for exaggerated methylase activity.

Another idea is that the abnormal methylation represents the activity of what might amount to a primitive, intracellular immune system. To protect themselves against invasion by foreign DNA, bacteria such as Escherichia coli deploy restriction endonucleases and methylases that recognize and attack features foreign to the host's own genetic material. Conceivably, mammalian cells retain a similar defense involving methylation. If so, the hundreds of CGG repeats characteristic of a full fragile X mutation might make the overall sequence recognizably different from any span of similar length in a normal human genome. Perhaps enzymes see it as being so different that it must be foreign, say an invading virus, and therefore demands methylation.

In any case, the methylation spans the gene's promoter. The gene is rendered inactive, and in the absence of protein expression, the clinical syndrome develops.

The Fragile X Protein

But how does the absence of a protein cause the disease? Normally, FMR1 is expressed not only in brain but also in testes, lung, kidney, and heart. In the brain, the gene is active chiefly in neurons. Its expression is abundant, particularly in the hippocampus and in the cerebellum's granular layer. In the mouse, the counterpart gene shows impressive structural similarity: 95% for nucleotides in the coding region, and 97% for amino acids. (The trinucleotide repeat is also conserved, though it consists in the mouse of only eight CGGs interrupted by one CGA.) Still, when the gene was first cloned and sequenced and the predicted protein analyzed, initial protein-database searches found no instructive homologies.

Subsequent efforts have been more revealing (Figure 6). At the center of its amino acid sequence, FMRP shows two copies of a 30-amino-acid motif called the KH domain. Conserved across evolution, from bacteria to humans, the motif is known to occur in proteins that interact with RNAs. Toward the carboxy end of FMRP is an RGG box, another motif implicated in RNA binding. Indeed, FMRP binds selectively to messenger RNAs expressed in the human brain. Overall, about 4% of messengers isolated from fetal human brain tissue display such interaction. The
protein also interacts with ribosomes, the cell's machinery for translating mRNA into protein.

Figure 6. FMRP is the protein whose absence causes fragile X syndrome. Although its overall amino-acid sequence (top) betrays no resemblance to any other known protein, it incorporates three domains involved in binding to RNA: two copies of a so-called KH domain and one RGG box. More recently, the protein has been discovered to incorporate two domains involved in intracellular trafficking: both a nuclear localization signal (NLS) and a nuclear export signal (NES), which enable a polypeptide to move into and out of the cell nucleus. An NES has been identified so far in only a few proteins, of which one is mammalian (protein kinase inhibitor α) but the other is retroviral: the Rev protein of human immunodeficiency virus. The matches are most impressive if the mammalian proteins are "stretched" by one amino acid.

In addition, new work demonstrates that FMRP has both a nuclear localization signal (NLS), near its amino terminal, and a nuclear export signal (NES), encoded by exon 14. Given an NLS, essentially any polypeptide can be imported from the cytoplasm into the nucleoplasm via passage through a nuclear pore. Likewise, given an NES, a polypeptide can be made subject to nuclear export. (In our own experiments the NLS from FMRP bestowed that property on bovine serum albumin.) Both the NLS and the NES interact with other proteins involved in transport across the nuclear membrane. Apparently they can also be masked, keeping the parent polypeptide sequestered in a particular intracellular compartment. The details of their function are only now being explored—not surprising, since FMRP is among only a handful of proteins known to have an NES, the first of which were reported in August 1995.

One of these was a mammalian intracellular macromolecule, protein kinase inhibitor alpha. The other was a retroviral invader: the Rev protein of human immunodeficiency virus. In fact, Rev, like FMRP, has both an NLS and an NES. Consequently, the two proteins appear to have some functional similarities. Both can enter and exit the nucleus. Moreover, both can bind RNAs. By its binding to RNA, however, Rev appears to upregulate the transcription of viral genes. In particular, it is thought that Rev binds to nascent transcripts beginning to come free from the proviral DNA, and that in this position it recruits the host's transcription factors. FMRP shows no evidence of an effect on gene transcription. Conversely, Rev shows no evidence of associating with ribosomes.

From all of these observations, a hypothetical life cycle for FMRP can be assembled (Figure 7). From the cytoplasm FMRP enters the nucleus. Thus, the NLS is the first signal to function. In the nucleoplasm, it binds to RNAs and perhaps to other proteins, forming a complex called a ribonucleoprotein particle. Through the action of its NES, it gets exported from the nucleus. Back in the cytoplasm, it associates with ribosomes—in particular, ribosomes in the neuronal dendrites.
On the basis of these intracellular shuttling, we hypothesize that a neuron employs FMRP as transportation for what amounts to rapid deployment forces. In this view, the crucial function of a neuron is its readiness to signal and be signaled, activities requiring timely protein supplies, as shown by experiments in which the ability of one hippocampal neuron to receive signals from another is abolished by inhibitors of protein biosynthesis. In this respect, the neuron faces a fundamental problem: Its dendrites and axon can be distant from the cell nucleus. To replenish supplies, a remote site can certainly signal the nucleus, which in return can transcribe a gene. Still, the nucleus will have no way of “knowing” which site needed the protein. It seems instead that the neuron stations certain mRNAs in its dendrites, prepositioned on ribosomes (where FMRP, having brought the messengers, may serve an additional, perhaps regulatory function). In response to local conditions, the periphery can simply make what it needs.

If so, the absence of FMRP would not be lethal to a neuron. Any proteins the periphery requires can still be supplied from blueprints back in central storage. Nevertheless, the cell would presumably suffer a subtle deficit in its communicability with other neurons, perhaps being not quite as "available" as it should be. Conceivably this explains why fragile X syndrome is a relatively subtle disorder. Were it not for the mental retardation, its nonneurologic features--flat feet, an unusual ability to bend back the fingers--would hardly be called a disease. The worst of it is merely mitral valve prolapse. Even neurologically, the brain tissue shows no gross abnormality. Moreover, among its neurologic manifestations, fragile X syndrome has no motor or sensory signs.

The neurologic implication is that FMRP may transport a group of mRNAs selectively involved in higher brain function. Thus, besides yielding knowledge--and perhaps a specific treatment--of fragile X syndrome, continued study of FMRP has the potential to identify, among all brain proteins, a class related to cognition.

**Clinical Implications**

Diagnostic issues in fragile X syndrome arise from two basic facts: the mutation is common but the phenotype is subtle, especially early in life. First and foremost, clinicians should be aware of the syndrome's existence, so that the possibility of its presence can be considered in any child with mental retardation, developmental delay, or solely a learning disability or other behavioral problem, including even autism. Affected children do not look strikingly abnormal. Hence, in the absence of a diagnosis, parents who know that their child is not developing properly may expend much money and heartache fruitlessly seeking an explanation.
Fortunately, the diagnosis is now straightforward. Before 1991, when the fragile X mutation was characterized, the only genetic test for fragile X syndrome had been for the fragile site, called FRAXA, where, in special culture, the metaphase chromosome can sometimes be induced to show a characteristic gap. The test requires skill, and results can sometimes be inconclusive, especially prenatally. For one thing, a number of fragilities resembling FRAXA occur with varying population frequencies throughout the human genome. Two such sites lie near FRAXA and may be confused with it. One of these, FRAXE, has been associated with mild mental retardation.

Today, no such problems remain. Indeed, since virtually all patients share a mutation at precisely the same site, as opposed to a range of mutations scattered along the length of a gene, the genetic diagnosis (or exclusion) of fragile X syndrome is remarkably reliable. Even in the absence of a specific therapy, there is benefit in its identification, which ends the search for the cause of a child's disability. Moreover, when the syndrome has been identified, family members can be evaluated for risk of bearing affected children. Those in whom a repeat expansion is found may opt for prenatal testing of future pregnancies. Conversely, many persons who fear that they are at risk of bearing a child with mental retardation (because, for example, they themselves have a retarded brother) may learn they are not at risk.

Currently, most centers testing for fragile X syndrome investigate both the length of FRAXA and the methylation of FMR1. The chief study is of length. In this respect, there has been discussion of whether a finding might be ambiguous. For example, should 55 repeats be judged normal, or is it a premutation? A careful answer would require investigation of the tract's cryptic interruptions, and even then a precise answer might not be possible. Clinically, however, it has not been much of an issue. A fetus with an allele containing 55 repeats will not have retardation owing to such an inheritance. The worst one would predict is that the fetus's grandchildren might be at risk (if a treatment had still not emerged at that distant time). Much the same is true of overlap between premutation and full-mutation length. Although 200, or 230, or 250 repeats might be cited as a cutoff, only rare persons actually have such alleles. Indeed, one seldom sees patients with less than 400 repeats.

When genetic testing first became feasible, there was also discussion of whether the expansion of a trinucleotide repeat might continue even after a prenatal test. Commercial labs would call researchers to say that they had found an allele with 100 repeats. What did that mean? It can now be said with assurance that in this pregnancy the allele will not expand further. In practice, the most complicated prognostic issues arise for mosaic males--perhaps 10% of all male cases of fragile X--with their complex mix of allele lengths. Such persons may express FMRP, but often in less than normal quantity. Clinically, some are well functioning, whereas others are plainly retarded. In such cases, study of the protein is more informative than study of the gene alone.

**Conclusion**

Before 1991, trinucleotide expansions were unknown. Today, 12 genetic diseases have been traced to this type of mutation. Two additional fragile sites do not lie near a gene and thus are not linked to disease. Among the total of 14 sites, the repeat differs from one to another. The effect of expansion differs as well. In fragile X syndrome, the expansion affects CGG in a gene's 5' untranslated region. The result is transcriptional suppression. In myotonic dystrophy, the expansion affects CTG in a gene's 3' untranslated region. The result is mRNA instability. In Kennedy disease, Huntington disease, and spinocerebellar ataxia types 1, 2, and now 6, the expansion affects CAG within a gene's coding region. The result is a lengthened polyglutamine tract in a nascent translated protein, perhaps constituting a gain-of-function mutation.

In all 12 diseases, the expansion constitutes the biomolecular basis of genetic anticipation. In general, the term refers to a gene's intensified capacity to cause disease as it descends through a kindred. In fragile X syndrome, the phenomenon takes the unusual form of an increased disease frequency in males, and a concomitant decrease in carriers. The earliest chroniclers of such patterns were clinicians intrigued by families that showed severe or early disease in their
youngest generations. Geneticists may have discounted the patterns, knowing of no mechanism that could produce them, and perhaps also supposing that severe disease in a child might prompt diagnoses of the same disease, only milder, in family members with equivocal findings. In sum, genetic anticipation got to be little studied. Only when trinucleotide expansions were identified could clinicians be confirmed right. Today's ongoing discoveries make it hard to remember that genetic anticipation was judged impossible only five or six years ago.

Selected Reading

Eberhart DE et al: The fragile X mental retardation protein is a ribonucleoprotein containing both nuclear localization and nuclear export signals. Hum Mol Genet 5:1083, 1996


Kunst CB, Warren ST: Cryptic and polar variation of the fragile X repeat could result in predisposing normal alleles. Cell 77:853, 1994


