Annu. Rev. Neurosci. 1996. 19:79–107

TRINUCLEOTIDE REPEATS IN NEUROGENETIC DISORDERS

H. L. Paulson and K. H. Fischbeck
Department of Neurology, University of Pennsylvania Medical School, 3400 Spruce Street, Philadelphia, Pennsylvania 19104

KEY WORDS: neurodegeneration, genetic instability, polyglutamine repeats, Huntington's disease, expanded repeats

ABSTRACT

Trinucleotide repeat expansion is increasingly recognized as a cause of neurogenetic diseases. To date, seven diseases have been identified as expanded repeat disorders: the fragile X syndrome of mental retardation (both *FRAXA* and *FRAXE* loci), myotonic dystrophy, X-linked spinal and bulbar muscular atrophy, Huntington's disease, spinocerebellar ataxia type 1, dentatorubral-pallidoluysian atrophy, and Machado-Joseph disease. All are neurologic disorders, affecting one or more regions of the neuraxis. Moreover, five of the seven (the last five above) are progressive neurodegenerative disorders whose strikingly similar mutations suggest a common mechanism of neuronal degeneration. In this article we discuss specific characteristics of each trinucleotide repeat disease, review their shared clinical and genetic features, and address possible molecular mechanisms underlying the neuropathology in each disease. Particular attention is paid to the neurodegenerative diseases, all of which are caused by CAG repeats encoding polyglutamine tracts in the disease gene protein.

INTRODUCTION

Simple sequence repeats occur commonly in the human genome. Also called microsatellites or tandem repeats, these normally polymorphic sequences have proved useful as markers in linkage studies. In the past five years, it has become clear that they also underlie an entirely new class of human mutations. The expansion of one group of simple repeats, trinucleotide or triplet repeats, is now known to cause seven inherited disorders, a number that is likely to continue growing (Table 1). To date, the list includes the fragile X syndrome of mental retardation (both the *FRAXA* and *FRAXE* loci), myotonic dystrophy (DM), X-linked spinal and bulbar muscular atrophy (SBMA), Huntington's disease (HD), spinocerebellar ataxia type 1 (SCA1), dentatorubral-pallidoluysian atrophy (DRPLA), and Machado-Joseph disease (MJD).

0147-006X/96/0301-0079$08.00

Table 1 Trinucleotide repeat diseases

Disease	Repeat	Repeat length	Gene product
Spinal and bulbar muscular atrophy	CAG	Normal: 11 to 34 Disease: 40 to 62	Androgen receptor (transcription factor)
Fragile X syndrome (FRAXA locus)	CGG	Normal: 6 to 52 Premutation: ~60 to 200 Full mutation: ~200 to >2000	FMR-1 (RNA binding protein)
FRAXE mental retardation	GCC	Normal: 6 to 25 Disease: >200	?
Myotonic dystrophy	CTG	Normal: 5 to 37 Protomutation: ~50 to 180 Full mutation: ~200 to >2000	myotonin protein kinase (serine-threonine kinase)
Huntington's disease	CAG	Normal: 11 to 34 Disease: 36 to 121	huntingtin
Spinocerebellar ataxia type 1	CAG	Normal: 6 to 39 Disease: 41 to 81	ataxin-1
Dentatorubral-pallidoluysian atrophy	CAG	Normal: 7 to 25 Disease: 49 to 88	atrophin
Machado-Joseph disease	CAG	Normal: 13 to 36 Disease: 68 to 79	MJD1 gene product

Expanded repeat mutations do not adhere strictly to rules of Mendelian inheritance. They are unstable mutations that change size in successive generations and thus are especially intriguing to geneticists. But expanded repeat diseases should interest neuroscientists and neurologists as well. All of these diseases are neurologic disorders, affecting one or more regions of the neuraxis. Moreover, five of the seven (the last five above) are progressive neurodegenerative disorders whose strikingly similar mutations suggest a common mechanism of neuronal degeneration. In this article we discuss specific characteristics of each trinucleotide repeat disease and then review their shared clinical and genetic features. Finally, we address the molecular mechanisms that may underlie the neuropathology in each disease. Particular attention is paid to the neurodegenerative diseases, all of which are caused by CAG repeats encoding polyglutamine tracts in the disease gene protein. [For other recent reviews, see Bates & Lehrach (1994), Kuhl et al (1993), LaSpada et al (1994), Mandel (1994), Ross et al (1993), and Richards & Sutherland (1994).]

GENERAL FEATURES AND CLASSIFICATION

Several generalizations can be made regarding the shared features of trinucleotide repeat diseases. 1. Inheritance is autosomal dominant or X-linked. 2. Disease severity varies widely and correlates directly with increasing repeat length. 3. Expansions arise from CG-rich triplet repeats that are polymorphic

Table 2 Classification of the trinucleotide repeat diseases

	Type 1	Type 2
Representative disorders	SBMA, HD, SCA1 DRPLA, MJD	FRAXA, FRAXE and DM
Repeat	CAG	CGG or CTG
Degree of instability	Moderate; small expansions	Highly unstable; prone to very large expansions
Repeat translated into protein?	Yes	No
Pattern of pathology	Specific neurons affected	Multisystem disorders
Proposed mechanism	Toxic gain of function	Altered expression of gene product or RNA

in the normal population. 4. Expanded trinucleotide repeats are unstable, changing in size when transmitted to successive generations. In contrast, normal-sized repeats are usually transmitted stably despite their polymorphic nature. 5. Although either further expansions or contractions may occur with transmission of expanded repeats, expansions are more common. 6. The diseases frequently show anticipation—the tendency for disease severity to increase in successive generations of a family. Anticipation is explained by the tendency for expanded repeats to further enlarge with transmission, coupled with the fact that longer repeats correlate with more severe disease. 7. Parent-of-origin effects are common. With several of the neurodegenerative diseases, for example, the most severely affected individuals usually inherit the disease from the father. In contrast, the most severe forms of myotonic dystrophy and Fragile X syndrome are, with rare exceptions, maternally transmitted.

Despite their shared features, trinucleotide repeat diseases clearly fall into two groups (Table 2). The first, which we have designated Type 1 (LaSpada et al 1994), is comprised of the five progressive neurodegenerative disorders—SBMA, HD, SCA1, DRPLA, and MJD. These disorders are essentially confined to the nervous system (SBMA's hormonal effects excluded) and are caused by small, constrained CAG repeat expansions that encode polyglutamine tracts in the disease gene protein. These mutations likely alter one or more properties of the disease protein, leading (as is discussed later) to a gain of function that is particularly deleterious to neurons.

The second group, which we have called Type 2, includes DM, FRAXA, and FRAXE. These are multisystem disorders characterized by much larger expansions of trinucleotide repeats that fall outside the protein coding region. To date, these have been non-CAG repeats: CTG, CGG, and GCC in DM, FRAXA, and FRAXE, respectively. (CTG designates the coding strand in DM; the noncoding strand is CAG.) The underlying mechanism of disease does not seem to be altered protein function but instead altered expression of the gene

in at least one (FRAXA) and possibly all three diseases. These are not progressive neurodegenerative disorders; instead, CNS manifestations of Type 2 diseases may in part reflect abnormalities in brain and neuromuscular development. Another feature distinguishing this group from CAG repeat diseases is marked somatic mosaicism. Whereas repeat sizes vary widely in somatic tissues from DM and FRAXA patients (Anvret et al 1993), relatively little somatic mosaicism occurs in the CAG repeat diseases (MacDonald et al 1993, Telenius et al 1994).

THE DISEASES

Here we review the clinical, neuropathologic, and genetic features of each disease. First, the five Type 1 diseases are presented in order of discovery, followed by the Type 2 diseases.

X-Linked Spinal and Bulbar Muscular Atrophy

SBMA, or Kennedy's disease, is a rare progressive neuromuscular disorder characterized by proximal muscle weakness, atrophy, and fasciculations (Kennedy et al 1968). The X-linked pattern of inheritance, adult onset, and bulbar involvement distinguish it from autosomal forms of spinal muscular atrophy (Harding et al 1982). Because SBMA patients often have signs of androgen insensitivity (gynecomastia, reduced fertility, and testicular atrophy), the androgen receptor (AR) gene became a candidate gene for SBMA once it was mapped to the same region of the X chromosome (Fischbeck et al 1991). La Spada et al (1991) found that a trinucleotide repeat, $(CAG)_n$, was expanded in the first exon of the AR gene in all SBMA patients but never in controls. This CAG repeat normally encodes a polyglutamine tract of 11–34 amino acids in the AR protein. In SBMA patients, the polyglutamine tract is lengthened to 40–62 amino acids.

SBMA is characterized pathologically by degeneration of anterior horn cells, bulbar neurons, and dorsal root ganglion cells (Sobue et al 1989). The AR is expressed in these neurons as well as in several other regions of the brain. What role androgens play in the normal function and viability of motor neurons is presently unclear, although Yu (1989) showed that androgens can attenuate neuronal degeneration following axotomy.

Huntington's Disease

Probably the most common Type 1 disease, HD is an autosomal dominant neurodegenerative disorder characterized by involuntary movements, cognitive impairment, and emotional disturbance (Harper et al 1991, Wexler et al 1991). Symptoms of HD usually first appear in the third to fifth decades, but

may begin in childhood or after age 60. Anticipation is not uncommon, and juvenile-onset cases are usually paternally transmitted.

As with all CAG repeat disorders, adult-onset HD begins insidiously and progresses slowly. The first motor sign of HD is usually chorea—uncontrolled flickering movements that become widespread and disabling as the disease progresses. In advanced disease, rigidity, bradykinesia, and dystonia may surface. Juvenile-onset disease differs from adult-onset disease in that juvenile patients often present with a combination of parkinsonian signs, dystonia, chorea, and seizures.

Degenerative changes in HD are most marked in the striatum. Cell death follows a gradient in three axes, occurring first in the tail of the caudate nucleus, then proceeding mediolaterally through the caudate and dorsolaterally through the putamen (Vonsattel et al 1985). Early in disease, there is selective degeneration of the most prevalent striatal neurons, the GABAergic medium spiny neurons that give rise to the striatal output to the globus pallidus and substantia nigra. In advanced stages neuronal loss becomes widespread, involving the cortex and cerebellum.

In 1983, the HD gene was mapped to chromosome 4 (Gusella et al 1983). Ten years later, the defective gene was identified after an intensive collaborative effort (Huntington's Disease Collaborative Group 1993). The defect is an expanded trinucleotide CAG repeat within the coding region of a large novel gene whose gene product is called huntingtin. The CAG repeat encodes a polyglutamine tract 11 to 34 amino acids long in normal individuals and 36 to 121 in HD patients (Andrew et al 1993, Duyao et al 1993, Huntington's Disease Collaborative Group 1993, Novelletto et al 1994, Snell et al 1993; reviewed in Goldberg et al 1994). A smaller polymorphic CGG repeat lies immediately 3′ to the CAG repeat and encodes a polyproline stretch in the protein. The HD gene is critical for embryonic development, as targeted disruption of the gene results in embryonic lethality (Nasir et al 1995).

The primary structure of huntingtin does not provide obvious clues to its function. However, much information already exists on possible biochemical defects in HD [reviewed in Beal (1992) and Wexler et al (1991)]. Positron emission tomography studies, for example, have demonstrated decreased glucose metabolism in the basal ganglia of HD patients and asymptomatic at-risk individuals, and proton NMR spectroscopy has shown increased lactate in the brains of symptomatic patients. These findings suggest that neuronal energy metabolism is impaired in HD, and this may play a role in the neurodegenerative process. Beal (1992) has proposed that impaired energy metabolism underlies the cell death in many neurodegenerative disorders. Altered bioenergetics may result in slow excitotoxic injury as neurons, depleted of energy and partially depolarized, become vulnerable to the excitatory amino acid glutamate. Support for this model in HD has come

from animal studies in which injections of either *N*-methyl-D-aspartate receptor agonists (e.g. quinolinic acid) or the mitochondrial toxin 3-nitropropionic acid result in pathologic changes resembling those seen in HD (Beal et al 1993, Ferrante et al 1993).

Spinocerebellar Ataxia

SCA1 is an autosomal dominant neurodegenerative disease characterized by ataxia, ophthalmoparesis, and weakness (Harding 1993, Schut 1954). It is one of several genetically distinct disorders comprising the autosomal dominant cerebellar ataxias (Harding 1993, Rosenberg 1995), a group that includes at least one other trinucleotide repeat disorder, MJD (or SCA3). Many of the dominant hereditary ataxias may prove to be trinucleotide repeat disorders.

Like other CAG repeat disorders, SCA1 usually begins in the third to fifth decades and progresses over 10 to 20 years, although it can appear in childhood or much later in life. Anticipation has been reported (Schut 1954); juvenile-onset cases most often occur with paternal transmission. The disease typically begins with cerebellar signs such as clumsiness and ataxia but progresses to include dysarthria, bulbar dysfunction, oculomotor disturbance, and pyramidal tract signs. In its late stages, the disease may include dystonia and choreoathetosis, but dementia and parkinsonian signs are uncommon (Ranum et al 1994). Neuropathologic findings include neuronal loss in the cerebellum (principally Purkinje cells) and brain stem and degeneration of spinocerebellar tracts. Little is known about biochemical abnormalities in this disease.

The discovery of the genetic defect in SCA1, mapped earlier to chromosome 6, was facilitated by specifically screening genomic DNA for trinucleotide repeats. This screening identified a trinucleotide CAG repeat that is expanded in all SCA1 patients (Orr et al 1993). The SCA1 repeat is in the protein coding region of a novel gene, ataxin-1, and encodes a polyglutamine tract 6 to 39 repeats long in normal individuals versus 41 to 81 repeats long in affected individuals (Matilla et al 1993, Orr et al 1993, Ranum et al 1994). Predicted to be 87 kDa, ataxin-1 has no significant homology to other known proteins (Banfi et al 1993). One unusual feature of the ataxin-1 gene is that its transcript contains very large 5′ and 3′ untranslated regions. The 5′ region contains seven small exons and is subject to alternative splicing. These features suggest that the transcriptional and translational regulation of ataxin-1 is highly complicated. Untranslated regions of the ataxin-1 transcript could have biologic activity of their own, as has been shown for several 3′ untranslated regions of mRNAs (Rastinejad & Blau 1993).

Whereas the expanded SCA1 repeat is an uninterrupted sequence of CAG repeats, the normal SCA1 allele is usually interrupted midway by one to three

CAT trinucleotides (Chung et al 1993). These interruptions are thought to contribute to allele stability. Similar interruptions have been found in the FRAXA repeat (Eichler et al 1994, Fu et al 1991).

Dentatorubral-Pallidoluysian Atrophy

The sixth trinucleotide repeat disease to be identified, DRPLA is a rare neurodegenerative disorder characterized by ataxia, choreathetosis, dementia, myoclonus, and epilepsy (Naito and Oyanagi 1982, Smith et al 1958). Although DRPLA has clinically been divided into three subtypes—myoclonic-epilepsy, pseudo-Huntington's, and ataxic-choreoathetoid forms—the distinctions are not precise, and more than one form may be seen in a family. There is anticipation, with the severe early-onset cases most often paternally transmitted (Koide et al 1994, Komure et al 1995, Nagafuchi et al 1994b, Takahishi et al 1988). As with HD, DRPLA in juvenile-onset patients differs from the disease in adults: Younger patients are more likely to display progressive myoclonic-epilepsy and deteriorate more quickly.

The neuronal degeneration in DRPLA is widespread but particularly severe in the globus pallidus, dentatorubral system, and subthalamic nucleus (Luys body) (Neumann 1959, Takahishi et al 1988). Though DRPLA can resemble HD clinically, it usually is distinguished pathologically from HD by the presence of dentatorubral atrophy and the absence of significant striatal atrophy.

Because DRPLA shares clinical and genetic features with SCA1 and HD, two groups of investigators hypothesized that it, too, might be a trinucleotide repeat disorder, and analyzed a gene on chromosome 12 that had previously been shown to contain a polymorphic CAG repeat (Koide et al 1994, Nagafuchi et al 1994b). This CAG repeat proved to be expanded in all DRPLA patients, from its normal range of 7 to 25 copies to an expanded range of 49 to 88 copies (Koide et al 1994, Komure et al 1995, Nagafuchi et al 1994b). The repeat encodes a polyglutamine tract in the middle of a novel protein predicted to be 124 kDa (Nagafuchi et al 1994a). In addition to the polyglutamine tract, the predicted DRPLA gene product (also called atrophin-1) contains small homoproline and homoserine tracts upstream of the polyglutamine tract, and two stretches of arginine-glutamate dipeptides and one of alternating histidines downstream. Though the function of the DRPLA gene product is unknown, the dipeptide sequences are reminiscent of similar stretches found in a group of nuclear RNA-binding proteins and in components of the spliceosome (SR proteins).

DRPLA is most prevalent in Japan, where it has been extensively characterized. Since the discovery of the DRPLA repeat, Burke et al (1994) have determined that the Haw River Syndrome (HRS), a dominant neurodegenerative disease affecting a large African-American family in North Carolina, is

also DRPLA. Though the genetic defect is identical, HRS differs clinically from other DRPLA in that it combines features of pseudo-Huntington's and myoclonic-epilepsy and is characterized by generalized seizures without myoclonus (the two are usually seen together in DRPLA). The pathology in HRS also involves more neuraxonal dystrophy, basal ganglia calcification, and demyelination than is typically seen in most DRPLA patients (Burke et al 1994). These differences illustrate that additional genetic factors may influence the phenotypic expression of this and, perhaps, other trinucleotide repeat diseases.

Machado-Joseph Disease

One of the more common forms of spinocerebellar degeneration, MJD is often clinically indistinguishable from SCA1 (Harding 1993, Rosenberg 1992). This recently discovered expanded repeat disorder had already been assumed by many to be a CAG repeat disease because of its clinical and genetic characteristics (autosomal dominance, variable phenotype, progressive nature, and usual adult onset with anticipation). The genetic defect was identified by directly searching for candidate CAG repeat genes. Screening a brain cDNA library, Kawaguchi and colleagues (1994) isolated a cDNA that mapped to a region of chromosome 14 where MJD had already been localized. This cDNA, MJD1, contains a CAG repeat that is expanded in MJD patients and encodes a polyglutamine tract in the predicted gene product. In the limited number of normal and affected individuals screened so far, the CAG-polyglutamine repeat is 13 to 36 repeats in normal alleles and 68 to 79 repeats in disease alleles. A small hydrophilic protein of 359 amino acids, the predicted MJD1 gene product is unique among Type 1 disease genes in that its polyglutamine tract lies near the carboxyl terminus. Other than the polyglutamine tract, it contains no obvious homology to other known proteins.

Fragile X Syndrome

Named after a folate-sensitive fragile site (the *FRAXA* locus) on the X chromosome, this X-linked disorder is the most common cause of inherited mental retardation. The mental retardation can vary from mild to severe and is associated with dysmorphic features such as enlarged ears, head, and testicles (Nussbaum & Ledbetter 1989). Although pathologic studies have demonstrated hypoplasia of the cerebellar vermis and white matter changes, no distinctive biochemical or cellular abnormality has been noted in the brain or elsewhere (Reiss et al 1991, Wisniewski et al 1991).

Until the discovery of the underlying mutation in fragile X, several aspects of its inheritance pattern were puzzling, including the fact that hemizygous males may be phenotypically normal (20%), yet females carrying the mutant

allele are often affected (30%) (Sherman et al 1985). The risk of disease increases from one generation to the next, a phenomenon known as the Sherman paradox (Sherman et al 1985). This paradox is now explained by the tendency of the trinucleotide repeat to further expand with transmission (Fu et al 1991).

Cloning of the fragile site revealed that the primary genetic defect in fragile X syndrome is an expanded $(CGG)_n$ repeat (Kremer et al 1991, Oberle et al 1991, Verkerk et al 1991). The repeat is found in the 5′ untranslated region of a gene called *FMR-1* (Figure 1) (Ashley et al 1993a). In the normal population, the repeat exists as a stable polymorphism of between 6 and 52 repeats, with a mode of 30. Repeats between 35 and 55 repeats fall into a gray zone in which alleles may be stable or unstable, depending on whether AGG triplets interrupt the CGG repeat (Eichler et al 1994). In *FRAXA* families, the repeat is expanded over a broad range, from premutations of ~60–200 repeats to full mutations of several thousand repeats or more. Expansions in the premutation range usually do not cause disease but are highly prone to expand to the full mutation, resulting in the full-blown manifestations of disease. Expansion is associated with increased methylation of both the repeat and an adjacent CpG island, leading to transcriptional silencing of the gene (Bell et al 1991, Hansen et al 1992, Pieretti et al 1991, Sutcliffe et al 1992). The CGG repeat expansion and subsequent methylation likely account for the chromosomal fragility at the *FRAXA* locus.

FRAXE The fragile site associated with the CGG repeat in *FMR-1* is designated *FRAXA*. More recently Knight et al (1993) showed that a second folate-sensitive fragile site (*FRAXE*) 600 kb distal to FRAXA contained a $(GCC)_n$ repeat that is expanded in mildly retarded individuals. This second, less common site was demonstrated when cytogenetic screening uncovered a few affected families lacking the expanded *FRAXA* repeat and instead expressing the nearby *FRAXE* fragile site. The degree of mental retardation in these families seems milder than in *FRAXA* patients, although the sample size is small (Knight et al 1993). In the few *FRAXE*-positive families screened so far, the *FRAXE* GCC repeat is enlarged to greater than 200 copies in all mentally retarded males. In contrast, normal unrelated males have a trinucleotide repeat that is 6–25 copies in length. What gene the *FRAXE* trinucleotide resides in and how it causes disease are still unknown.

OTHER FRAGILE SITES Recently, two more folate-sensitive fragile sites, *FRAXF* and *FRA16A*, were cloned. The *FRAXF* site, located on the X chromosome ~1200 kb distal to *FRAXA*, is a GCC repeat that is normally 6–29 repeats but expanded to 300–500 repeats in patients expressing the fragile site (Parrish et al 1994, Ritchie et al 1994). Whether the *FraXF* site is associated

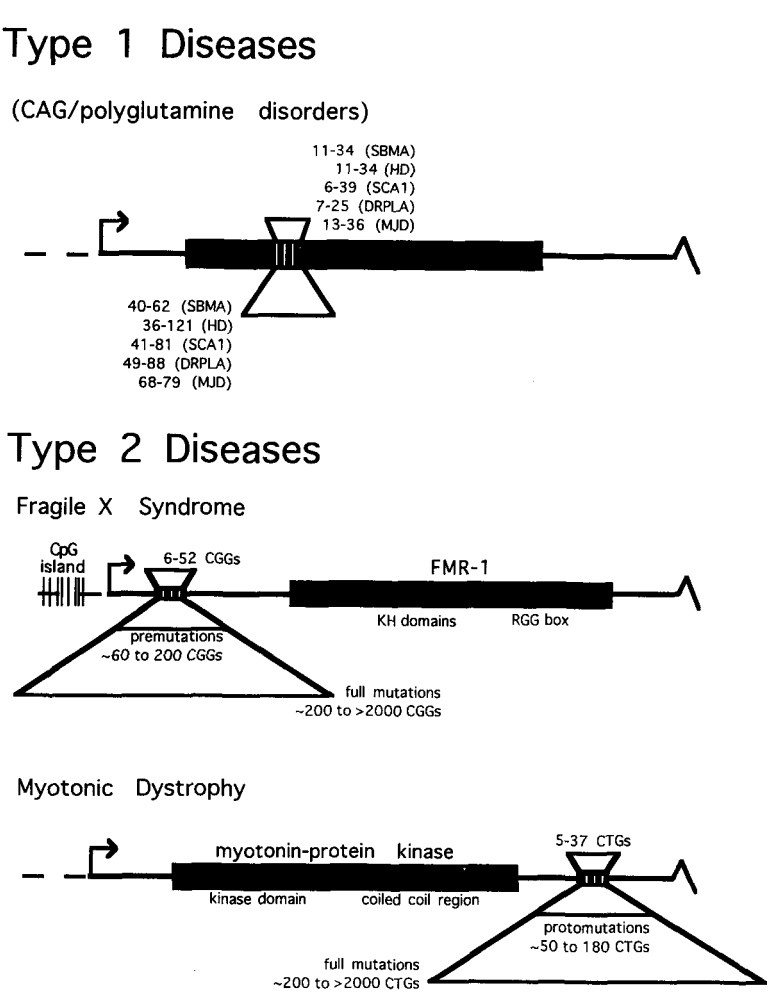

Figure 1 Location of expanded repeats in Type 1 and 2 trinucleotide repeat disorders. Shown are a schematic composite of the Type 1 disease genes and the *FMR-1* and myotonin protein kinase genes. Arrows depict transcription start sites, and shaded boxes depict protein coding regions. Striped boxes indicate the repeat; the ranges of normal and expanded repeats are indicated for each disease. (In FMR-1, the KH domains and RGG box are predicted to be RNA-binding motifs.)

with human disease is still unclear, although a few affected males have shown developmental delay.

FRA16A (chromosome 16) is the first autosomal fragile site to be cloned (Nancarrow et al 1994). As in *FRAXA, FRAXE,* and *FRAXF,* the molecular basis for the *FRA16A* site is a normally polymorphic $(CCG)_n$ repeat that is

grossly expanded (1000–2000 repeats) in individuals expressing the fragile site. When expanded, the *FRA16A* locus and an adjacent CpG island become hypermethylated like the other fragile sites. Hypermethylation seems to be a consequence of expansion (Nancarrow et al 1994). *FRA16A* and other autosomal fragile sites are not associated with any known disease, perhaps because they have not been observed in homozygous form.

The striking similarities of the four cloned fragile sites suggest that all folate-sensitive fragile sites are due to expanded repeats. If resultant hypermethylation of the repeat is itself the cause of the observed chromosomal fragility at these sites, then the sequence of the repeat must always contain the CpG dinucleotide.

Myotonic Dystrophy

An autosomal dominant disease, DM is characterized by progressive weakness and myotonia, often in association with cataracts, cardiac arrhythmias, endocrinopathy, and cognitive impairment (Harper 1989). The range in clinical severity is remarkably broad, and anticipation is common (Ashizawa et al 1992, Harley et al 1992, Harper et al 1989). Classic DM typically first appears in the second or third decades, with variable weakness and clinically evident myotonia (repetitive, rapid muscle firing after forceful contraction). Some patients, however, may be completely asymptomatic carriers or have only mild signs of disease, such as cataracts and electromyographically detectable myotonia. At the other extreme are patients, nearly always the offspring of affected women, who present with severe manifestations at birth (congenital DM). These patients are floppy as babies and usually mentally retarded. Pathologic studies of adults reveal distinct, but not specific, myopathic changes; in contrast, congenital cases show muscle fiber changes that suggest arrested development (Sarnat & Silbert 1976).

An international effort led to the identification of the defective gene on chromosome 19 (Aslandis et al 1992, Brook et al 1992, Buxton et al 1992, Fu et al 1992, Harley et al 1992, Mehadevan et al 1992). The defect is an expanded trinucleotide $(CTG)_n$ repeat in the 3′ untranslated region of a gene called myotonin-protein kinase (Figure 1). Only 5 to 37 CTGs long in normal individuals, the CTG repeat is enlarged from ~50 to greater than several thousand repeats in affected individuals. As in fragile X, smaller protomutations (~50–180 repeats) cause mild or no disease but have a high probability of expanding to much larger repeats (~200 to >3000) in the next generation. The largest repeats (>1500) are associated with congenital DM. No mutations other than the expanded repeat have been identified in patients with DM, although a few cases without detectable CTG expansion have been reported (Thornton et al 1994).

REPEAT INSTABILITY: CLINICAL ASPECTS

The discovery of expanded repeats has had an immediate clinical and scientific impact. The tests used for detecting expanded repeats are highly sensitive and specific [for example, see Kremer et al (1994)]. They permit certainty in diagnosis and eliminate the need for expensive ancillary tests (Redman et al 1993). Because the PCR-based repeat test is technically simpler and cheaper than the linkage analysis previously used for HD, requests for presymptomatic HD testing have recently increased. [Guidelines for HD testing were recently revised and may prove applicable to other trinucleotide repeat disorders (Broholm et al 1994, Hersch et al 1994).] In the near future, the ability to determine carrier status in presymptomatic individuals should facilitate trials aimed at disease prevention and may help researchers detect early biochemical changes in presymptomatic patients.

The discovery of expanded repeats has helped explain many of the unusual characteristics of these diseases (discussed individually below): variable phenotype, anticipation, parent-of-origin effects, and sporadic disease. In addition, repeat testing has helped determine the full range of clinical manifestations in each trinucleotide disorder. For example, amid the confusing nomenclature and overlapping syndromes of the hereditary ataxias, testing for the SCA1 repeat has better defined the clinical manifestations of SCA1 (Ranum et al 1994).

Repeat Size and Clinical Correlation

For each trinucleotide repeat disease, there is a correlation between increasing repeat size and disease severity. This is best illustrated for Type 1 disorders, where larger repeats generally lead to earlier onset illness. In SCA1 (Figure 2) 66% of the variability in age of onset is accounted for by repeat length (Orr et al 1993, Ranum et al 1994). In other CAG repeat disorders it is slightly lower, for example, ~50% in HD (Andrew et al 1993, Doyu et al 1992, Duyao et al 1993, Igarishi et al 1992, Koide et al 1994, LaSpada et al 1992, Nagafuchi et al 1994, Novelletto et al 1994). As this correlation implies, juvenile-onset patients usually carry the largest repeats. Longer repeats are also associated with other aspects of disease, including earlier age of death in HD (Andrew et al 1993), more rapidly progressive disease in SCA1 (Orr et al 1993, Ranum et al 1994), and expression of the progressive myoclonus epilepsy subtype in DRPLA (Koide et al 1994, Nagafuchi et al 1994b).

It is important to note that repeat length accounts for only part of the variability in disease onset. Disease severity must partly be determined by other factors, including genetic background (Burke et al 1994, Ranum et al 1994). For this reason, clinicians should not use repeat length to predict disease

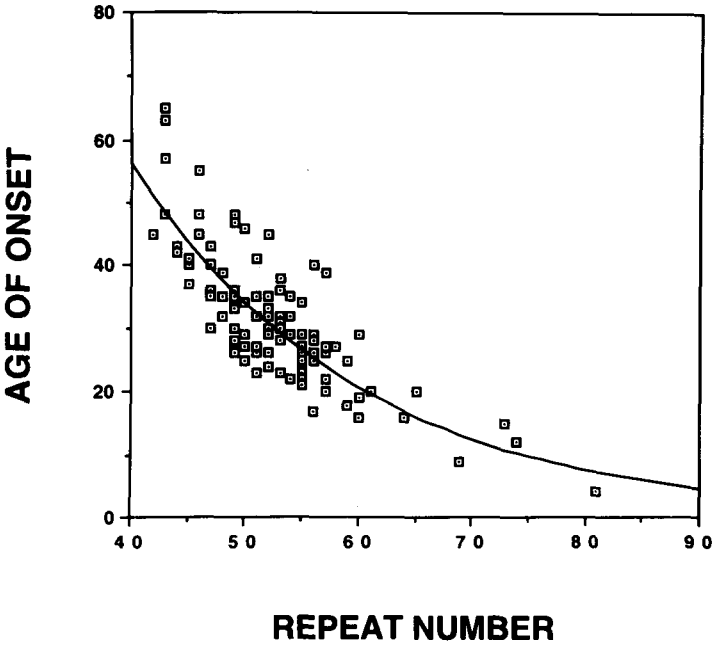

Figure 2 Inverse correlation between SCA1 repeat length and age of disease onset. In SCA1, 66% of the variability in age of onset is accounted for by repeat length. Other Type 1 diseases show a similar correlation between repeat length and age of onset. [Reprinted from Ranum et al (1994), with Dr. Ranum's permission.]

severity. At best, repeat length is only a partial predictor of age of onset, and for the most frequently observed lengths in HD (40–50 repeats), it is a poor predictor.

Type 2 disorders also show a correlation between repeat length and disease severity. In DM, for example, asymptomatic carriers typically have 60 to 120 repeats; classic adult-onset patients 200 to 400 repeats; childhood-onset patients 500 to 1500 repeats; and congenital DM 750 to >2000 repeats (Harley et al 1992, Redman et al 1993, Tsilfidis et al 1992). And in fragile X, the most profoundly retarded patients nearly always demonstrate dramatic repeat expansions, whereas smaller expansions cause milder disease (Fu et al 1991, Pieretti et al 1991). However, the correlation seen in these disorders is far from perfect. The lack of a precise correlation may be due in part to the extensive and variable somatic mosaicism seen in DM and fragile X. Repeat lengths are usually measured in peripheral blood, which may not accurately reflect repeat length in the most affected somatic tissues, such as muscle (Anvret et al 1993).

Sporadic Disease

Before direct DNA testing, patients were occasionally seen with the clinical features of HD but without a family history of disease. Testing has now confirmed that these sporadic cases, up to 3% of all HD, are due to trinucleotide expansion of the HD gene (Goldberg et al 1993, Myers et al 1993). In other Type 1 disorders the frequency of sporadic cases is unknown but is probably no greater. Sporadic HD cases arise from a small pool of intermediate size alleles (30–38 trinucleotides) that are larger than those in the general population but smaller than those in most patients. These meiotically unstable, intermediate repeats are analogous to asymptomatic premutations of FRAXA and DM. Not large enough to cause disease in the carrier, they still carry the risk of expanding into the disease range when transmitted to the next generation. Most if not all sporadic HD is paternally transmitted.

Parent-of-Origin Effect

For several Type 1 diseases (HD, DRPLA, and SCA1), severe juvenile-onset cases are usually paternally transmitted. This is due to greater repeat instability with paternal transmission (male bias) (Andrew et al 1993, Chung et al 1993, Koide et al 1994, Komure et al 1995, LaSpada et al 1992, Nagafuchi et al 1994, Novelletto et al 1994). With paternal transmission, intergenerational changes are larger (4 to 40 trinucleotides, versus typically just a few with maternal transmission), and expansions outnumber contractions (the reverse may be true for maternal transmission). Hence affected males are more likely to transmit a greatly enlarged CAG repeat of the size that typically causes juvenile-onset disease. Analysis of repeat size in sperm from HD patients suggests that this male bias is due to repeat instability during spermatogenesis (Duyao et al 1993, MacDonald et al 1993, Telenius et al 1994). Germ line instability in males also probably explains why sporadic cases of HD are paternally transmitted.

Congenital DM is nearly always maternally transmitted despite a male bias in the generation of new DM alleles (Ashizawa et al 1992, Bergoffen et al 1994, Wieringa 1994). This seems to occur, at least in part, because of constraints on the size of the repeat transmitted by males. An affected male with a very large CTG repeat—the size that might cause congenital DM—transmits a much smaller allele, usually in the premutation range. This may be due to selection against larger alleles during spermatogenesis. Since the dramatic expansions of CTG repeats appear to occur postzygotically in early embryogenesis (Jansen et al 1994, Wieringa 1994), it may instead be the case that the smaller repeats in sperm derive from embryonic tissues that did not undergo expansion. The end result is that affected females alone are able to transmit a

large repeat that with further expansion (meiotically or postzygotically) leads to congenital DM.

Similar mechanisms may help explain the fact that only maternally transmitted *FRAXA* repeats progress to full mutations. As in DM, affected FRAXA males with a full mutation contain only premutation alleles in their sperm (Reyniers et al 1993), and the dramatic *FRAXA* expansions appear to occur postzygotically early in embryogenesis (Nelson & Warren 1993, Wohrle et al 1993). Postzygotic expansion may occur exclusively on maternal *FRAXA* alleles, suggesting that a form of imprinting also contributes to the parental bias in transmission (Sutherland & Richards 1993).

REPEAT INSTABILITY: GENETIC ASPECTS

The major determinant of repeat instability appears to be repeat length itself. Normal-sized repeats have a low mutation rate that shows some length dependency [for example, a three- to fourfold increase in the mutation rate for the AR CAG repeat when going from 20–22 repeats to 28–31 repeats (Zhang et al 1994)]. However, only when repeats reach a critical size threshold, as they have in disease families, do they become highly unstable. Evidence suggests that this threshold is a stretch of roughly 35 uninterrupted trinucleotides. At the *FRAXA* locus, for example, marked instability begins when uninterrupted CGG repeat length reaches 34–38 repeats (Eichler et al 1994). This size coincides well with the upper limit of normal repeats for other trinucleotide repeat diseases (34, 34, 39, 25, 36, and 37 for SBMA, HD, SCA1, DRPLA, MJD, and DM, respectively). Moreover HD repeats of 30–38 constitute a pool of intermediate alleles that are predisposed to expand into the disease range, further suggesting that the critical threshold is in this range.

Once beyond this threshold length, trinucleotide repeats are highly mutable. For example, *FRAXA* alleles of 60–69 repeats have a 17% chance, and repeats greater than 90 a nearly 100% chance, of expanding to the full mutation when maternally transmitted (Fu et al 1991). In Type 1 disorders, where expansions seem constrained (perhaps because the repeat is translated into protein), a more modest degree of instability is observed during transmission (Andrew et al 1993, Chung et al 1993, Duyao et al 1993, Huntington's Disease Collaborative Group 1993, Koide et al 1994, LaSpada et al 1992, Nagafuchi et al 1994b, Novelletto et al 1993, Orr et al 1993, Snell et al 1993). In Type 2 diseases, a second size threshold exists (~60–70 uninterrupted repeats), beyond which repeats are prone to undergo hyperexpansion to the full mutation of several thousand repeats (Eichler et al 1994).

The existence of a similar threshold size for all trinucleotide repeat diseases suggests a common mechanism to expansion. Slipped-strand mispairing during replication has been proposed (Eichler et al 1994, Richards & Sutherland 1994)

as a mechanism of mutation that could account for several features of trinucleotide expansion: Expansions and contractions occur in a polarized manner and by multiples of three base pairs, and interruptions within the repeat seem to confer stability. Okazaki fragment slippage during replication would provide an explanation for both small changes in repeat length (as seen in Type 1 repeats and Type 2 premutations) and the enormous hyperexpansions seen in Type 2 disorders. During replication of a threshold length repeat (~35 pure repeats), slippage could occur from only one end of the Okazaki fragment (the other end being anchored outside the repeat), resulting in small to moderate expansion. During replication of longer repeats (~70 repeats), the likelihood of Okazaki fragment synthesis beginning and ending within a repeat would increase. Thus slippage could occur from both ends of the unanchored Okazaki fragment, resulting (after several rounds of replication) in hyperexpansion. DNA mismatch repair mechanisms likely play a role in correcting errors caused by slippage, since defects in mismatch repair lead to increased simple sequence repeat instability in both yeast and humans [for review, see Modrich (1994)].

Although slipped-strand mispairing offers a testable model of DNA expansion, it leaves many questions unanswered. In Type 1 disorders, for example, why is marked repeat instability confined to the male germ line? Are the dramatic expansions of Type 2 repeats truly limited to an early point in embryogenesis, and if so, why? How do we reconcile the fact that for expanded repeats, further expansions outnumber contractions, while for normal repeats, contractions are more common (Zhang et al 1994)? And why, in mice containing a human AR transgene (Bingham et al 1995), is the expanded CAG repeat stably transmitted? Taken together, these questions make it clear that factors other than repeat length must play a significant role in repeat instability. [For more thorough discussions, see Eichler et al (1994), Nelson & Warren (1993), Richards & Sutherland (1994).]

MOLECULAR MECHANISMS OF TRINUCLEOTIDE REPEAT DISEASES

Fragile X

The mechanism of disease is more clearly understood for fragile X than for other trinucleotide repeat disorders. The fragile X phenotype is due to decreased expression of the FMR-1 protein (Pieretti et al 1991). As the CGG repeat number increases in the 5′ region of the *FMR-1* gene, mRNA and protein levels decrease, and the fragile X syndrome worsens (Sutcliffe et al 1992). The degree of methylation correlates with disease severity, and hypermethylation may itself cause transcriptional silencing of the *FMR-1* gene (Bell et al 1991, McConkie-Rosell et al 1993, Nelson and Warren 1993, Pieretti et al

1991). A further argument that loss of *FMR-1* expression underlies disease is that small chromosomal deletions encompassing the *FMR-1* gene also lead to a fragile X phenotype (Gedeon et al 1992, Hirst et al 1995, Wohrle et al 1992). Transcriptional silencing seems to be the principle cause of decreased *FMR-1* expression, but translational inhibition may also play a role with the largest repeats (Feng et al 1995).

What is the role of FMR-1, the expression of which seems critical to normal brain function and development? FMR-1 is a member of a family of RNA-binding proteins (Ashley et al 1993b, Siomi et al 1993, Verheij et al 1993). It can bind its own and other human fetal brain mRNAs (Ashley et al 1993b), and alternatively spliced forms exist with distinct RNA-binding properties (Hoogeven et al 1994). The RNA-binding property of FMR-1 appears critical, as the fragile X phenotype can be generated by a point mutation that alters a highly conserved amino acid that is critical for RNA binding in related proteins (DeBoulle et al 1993). FMR-1 is widely expressed in fetal and adult tissues, most abundantly in neurons (Devys et al 1993). In the developing brain, FMR-1 is found initially in proliferating and migrating cells and later in differentiating neurons, with the highest levels in nucleus basalis magnocellularis and hippocampus (Abitbol et al 1993).

RNA-binding proteins are involved in the processing, transport, and translational control of mRNAs [see Simpson & Emeson (1996) in this volume]. Alternative splicing of transcripts helps to establish the intricate spatially and temporally restricted patterns of gene expression found in the brain. In addition, the targeting, localization, and translational control of mRNAs in subcellular domains of neurons (dendrites, for example) adds a further level of complexity to the regulation of neuronal gene expression. Which of these functions FMR-1 serves is presently unknown. Studies of *FMR-1* knockout mice, which display several phenotypic features of fragile X (Dutch-Belgian Consortium 1994), should help determine whether there is a specific class of RNAs that have processing, localization, or translation controlled by the FMR-1 protein.

The mechanism of disease in *FRAXE* may be similar to that in *FRAXA*. *FRAXE* disease severity also correlates with increasing repeat size and hypermethylation of a nearby CpG dinucleotide island (Knight et al 1993). It has been proposed that hypermethylation reduces expression of an as yet unidentified gene.

Myotonic Dystrophy

The DM gene product was given the name myotonin-protein kinase because sequence analysis showed it to have homology to members of the serine-threonine protein kinase family (Pizzuti et al 1993, Wieringa 1994). Experiments with recombinant protein have confirmed that it has kinase activity,

phosphorylating itself and other substrates at threonine and serine residues, although initial studies suggest that it differs from related serine-threonine protein kinases (Dunne et al 1994). Its amino acid sequence predicts both an α-helical coiled coil region and a hydrophobic domain near the carboxyl terminus. In muscle, the protein is predominantly localized to the neuromuscular junction (van der Ven et al 1993), although it is expressed elsewhere in muscle and in brain and heart (Jansen et al 1993). Several alternatively spliced RNAs exist that differ principally in the 5′ and 3′ regions of the transcript, leaving the kinase domain intact (Fu et al 1993). Although differential RNA processing occurs with this gene, what role it plays in the tissue-specific expression and function of myotonin-protein kinase is unknown (Pizzuti et al 1993).

Considerable focus has been directed at understanding the pathophysiology of muscle dysfunction in DM. Myotonia occurs in a number of rare inherited and acquired diseases. The underlying mechanism varies depending on the disorder, for example, chloride channel mutations in myotonia congenita and sodium channel mutations in hyperkalemic periodic paralysis (reviewed in Ptacek et al 1993). The primary electrophysiologic defect in DM is still unknown, although abnormal regulation of sodium channels has been proposed. The DM gene's kinase activity raises the possibility that in DM, myotonia may be due to altered phosphorylation of protein(s) involved in the muscle action potential (the sodium channel, for example) or excitation-contraction coupling. The normal endogenous substrates of myotonin protein kinase, however, have not yet been identified.

The muscle defects in DM cannot be due to changes in the protein itself, since the trinucleotide repeat lies in the 3′ untranslated region of the gene. Are they instead due to changes in the level of gene expression? Efforts to answer this question have been inconclusive. Studies performed with different techniques have shown either decreased or increased mRNA levels and reduced or essentially unchanged myotonin protein levels (Fu et al 1993, Novelli et al 1993, Sabouri et al 1993). These discrepancies need to be resolved before it can be determined whether altered expression plays a role in the pathogenesis of DM. The CTG expansion may also affect expression of other nearby genes, particularly in the full mutation range.

The biochemical defect in DM may occur at the RNA level instead. mRNA in the 3′ untranslated region is known to play a role in mRNA stability, cell proliferation, and differentiation (Rastinejad & Blau 1993). A greatly expanded GC-rich (CTG) repeat in the 3′ mRNA could perturb an important regulatory role for this part of the message, either by distorting its normal secondary structure or by titrating out necessary RNA-binding proteins (J Eberwine, personal communication). Such an effect could be dominant, thus accounting for DM's pattern of inheritance. An expanded repeat within the transcript could

also lead to aberrant RNA processing of the DM transcript, for which there is some experimental evidence (Krahe et al 1994).

CAG-Polyglutamine Diseases

In all five CAG repeat diseases, expanded CAG-polyglutamine tracts of about the same size cause neurodegeneration (Figure 3). Several findings support the view that expansion of CAG-polyglutamine tracts leads to a gain of function that may be toxic to neurons. First, the dominant inheritance seen in these disorders is more consistent with a gain, not loss, of function. HD in particular is a pure dominant disorder, since patients homozygous for the disease allele develop HD that is clinically indistinguishable from HD caused by a single disease allele (Wexler et al 1991). [SBMA follows an X-linked recessive pattern of inheritance, but in this case heterozygous females could be protected by lyonization (if the toxic effect is cell autonomous) or by low androgen levels (if the effect is ligand-dependent).] Second, in all these disorders the disease gene is expressed in the disease state, arguing against a simple decrease

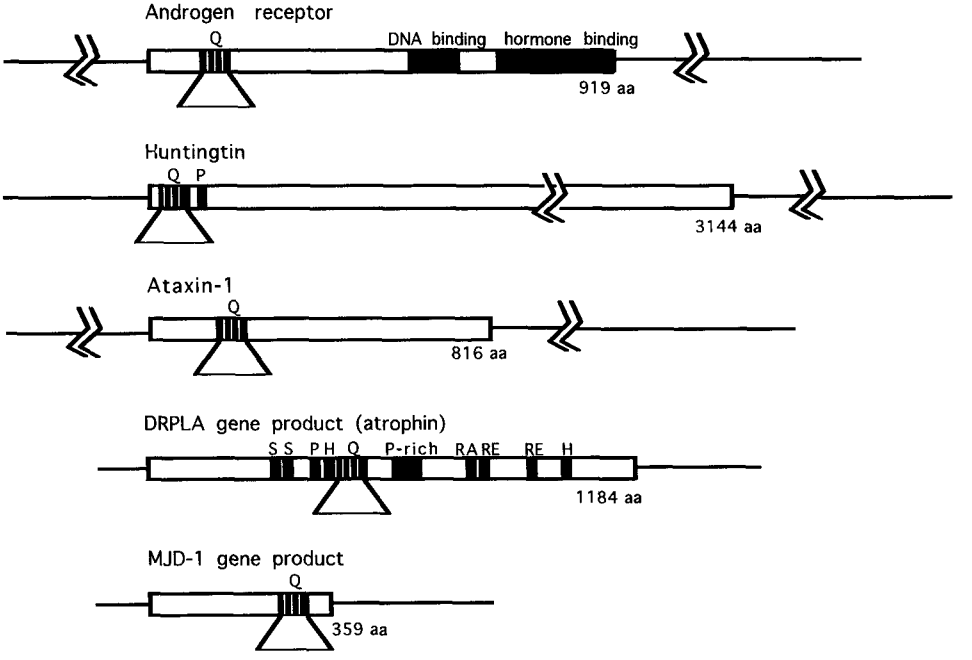

Figure 3 The Type 1 disease gene products. Shown are schematic transcripts for each Type 1 disease gene. Large open boxes indicate the protein coding domain, and smaller striped boxes the polyglutamine tracts. Shaded boxes represent regions of known function in the AR, and filled boxes represent homopolymeric or dipeptide domains (specified by the single letter amino acid code).

in gene expression as the cause of disease (Banfi et al 1994, Li et al 1993, Nagafuchi et al 1994a, Servadio et al 1995, Sharp et al 1995, Strong et al 1993, Trottier et al 1995, Warner et al 1992, Yazawa et al 1995). (MJD expression in the disease state has not yet been reported.) The fact that individuals missing one copy of the normal HD gene (because of chromosomal deletions or translocations) do not develop HD further suggests that disease is not caused by decreased gene expression. Third, the loss of function for one CAG repeat gene, the AR, leads to an entirely different disease, androgen insensitivity syndrome (testicular feminization), which does not have the neurologic manifestations of SBMA. Fourth, the recent development of a transgenic mouse model for SCA1 provides direct evidence supporting a gain of function: An expanded repeat SCA1 transgene causes cerebellar degeneration, whereas a normal repeat transgene does not (Burright et al 1995).

Any proposal to explain neurodegeneration caused by CAG-polyglutamine repeats must take into account several facts. 1. Cell death appears to be neuron specific despite the fact that the repeat-containing genes are widely expressed in nonneuronal cells (Banfi et al 1994, Li et al 1993, Nagafuchi et al 1994a, Strong et al 1993). 2. Distinct subsets of neurons die in each disorder, even though the disease gene product is expressed throughout the brain. (The androgen receptor may be an exception, since its neuronal expression is restricted to discrete regions of the brain and spinal cord.) 3. Degeneration occurs slowly, with little evidence of neuropathologic or biochemical abnormality before the clinical onset of disease. 4. The presumed neurotoxicity increases with repeat length, since longer repeats lead to earlier disease that in some cases, may progress more rapidly. 5. The diseases are caused by dissimilar genes that lack homology to one another outside of the CAG-polyglutamine repeats. There is no reason, based on sequence comparison and immunolocalization studies, to believe the encoded proteins serve similar cellular functions. 6. The gene products are all thought or known to be intracellular proteins, suggesting that the presumed toxic effect occurs intracellularly. Thus in summary, expansions in a group of dissimilar, ubiquitously expressed CAG repeat genes cause selective neurodegeneration that is dependent on the length of the CAG-polyglutamine repeat. Given the presumed common mechanism of neurodegeneration and the ubiquitous expression of CAG repeat genes, cell-specific factors must account for the death of certain neuronal populations in each disease.

The putative gain of function most likely occurs at the protein level, since the CAG repeat is translated into polyglutamine in the mature protein (this has been verified experimentally for DRPLA, HD, SBMA, and SCA1) (Servadio et al 1995, Sharp et al 1995, Trottier et al 1995, Yazawa et al 1995). However, a gain of function could also occur at the RNA level. For example, expanded CG-rich triplet repeats may form stable stem loop structures that could alter properties of the mRNA (J Eberwine, personal communication). Although the

SCA1 transgenic model supports a gain of function mechanism, whether this occurs at the protein or RNA level is unknown, and both theories warrant experimental testing.

What is the normal function of polyglutamine tracts? Increasing evidence indicates that polyglutamine domains are functionally important in regulatory proteins and transcription factors (Gerber et al 1994). Across species as evolutionarily distant as yeast, *Drosophila,* and humans, glutamine-rich domains and homopolymeric polyglutamine tracts have been found in common or ubiquitous transcription factors (e.g. TATA-binding protein, Sp1, AP1, glucocorticoid receptor), homeotic genes (e.g. *Antennapedia*), neurogenic genes (e.g. *Notch* and *mastermind*), and other developmentally important genes (e.g. embryonic polarity dorsal protein, *couch potato*) (Creusot et al 1988, Duboule et al 1987, Kao et al 1990, Mitchell & Tjian 1989, Wharton et al 1985). Mitchell & Tjian (1989) have proposed that polyglutamine tracts mediate transactivation by engaging in protein-protein interactions and thereby enabling formation of transcription activation complexes. Evidence for this hypothesis has come from deletion experiments. The yeast regulatory protein MCM1 contains a glutamine-rich region that is necessary for it to interact with the coregulatory protein alpha 1 and activate transcription (Bruhn et al 1992). When the glutamine-rich domains are deleted from Sp1, this transcription factor is also unable to form polymeric complexes and activate transcription of target genes (Pascal & Tjian 1991). More recently, studies of polyglutamine-containing fusion proteins directly demonstrated that polyglutamine stretches can activate transcription, optimally with lengths of 10–30 amino acids (Gerber et al 1994). It thus appears that polyglutamine tracts can serve as transactivation domains that probably function by modulating interactions with both homologous and heterologous transcription factors.

In fact, the only Type 1 disease gene product with a known function is a transcription factor, the AR. A member of the steroid and thyroid hormone receptor superfamily, the AR enters the nucleus upon binding androgen and activates transcription of androgen-responsive genes (Zhou et al 1994). Its polyglutamine tract is located in a large amino-terminal regulatory domain far removed from the DNA-binding and hormone-binding domains (Figure 3). The polyglutamine tract is not strictly necessary for androgen-induced gene activation. In nonneuronal cells, transcriptional activation by AR is increased when the polyglutamine tract is eliminated (Chamberlain et al 1994), and decreased when the polyglutamine tract is expanded into the disease range (Mhatre et al 1993). Whether this will also hold true in neurons or has any relationship to a toxic gain of function in neurons is unknown.

There is no direct experimental evidence to indicate that other Type 1 disease genes encode transcription factors. None of them contain definite DNA-binding motifs, although sequence analysis suggests atrophin may be an RNA-

binding protein (Nagafuchi et al 1994a) and ataxin-1 localizes to the nucleus in most neurons (Servadio et al 1995). Although huntingtin contains a sequence resembling a leucine zipper (a motif present in one group of transcription factors), recent studies indicate that in neurons huntingtin localizes to nerve terminals, not the nucleus (DiFiglia et al 1995, Sharp et al 1995, Trottier et al 1995). Regardless of the function of other Type 1 disease genes, polyglutamine stretches within these genes likely serve to modulate interactions with other cellular proteins. Perutz et al (1994) have proposed that polyglutamine tracts can self-associate in the form of a polar zipper, and recently provided direct evidence supporting this model: Incorporating a small polyglutamine tract into a normally monomeric protein caused dimers and trimers to form (Stott et al 1995).

How might expansion of a polyglutamine stretch cause neurodegeneration? There are at least three plausible mechanisms, each involving changes in protein-protein interactions. The first could be called the aberrant transregulation model. In this model, expanding the polyglutamine tract in a regulatory protein would alter, perhaps subtly, the strength of its interactions with other coregulatory proteins, thereby disrupting what is normally a highly ordered and regulated molecular cascade. For example, the AR with an expanded polyglutamine tract might display altered binding to transcription associated factors, causing aberrant transcriptional activation of one or more genes in neurons. This in turn could lead to the neuron's receiving inappropriate signals, which in some cases may result in neuronal apoptosis (DiBenedetto & Pittman 1995).

The second model could be called the protein aggregation model. In this model, expanding the polyglutamine tract would cause changes in protein-protein interactions that are not subtle and not necessarily restricted to normal cointeracting proteins. Expanded polyglutamine tracts may tend to form an extended polar zipper involving tight noncovalent interactions between proteins (Perutz et al 1994). These could be nonproductive interactions that perturb the function of any involved proteins and interfere with degradation of the aggregated proteins. For example, the expanded AR could aggregate with itself or with other proteins. This not only would deplete a potentially androgen-dependent cell of functional AR but also might lead to the accumulation of insoluble proteinaceous material that is deleterious to neurons. An attractive feature of this model is that it resembles current theories to explain two other neurodegenerative disorders, Creutzfeld Jacob disease (insoluble, protease-resistant aggregates of the prion protein) and Alzheimer's disease (aggregates of β-amyloid and microtubule-associated proteins).

The third model, proposed by Green (1993), is the transglutaminase cross-linking model. As polyglutamine tracts expand beyond the size normally found in proteins (less than 30 amino acids), they may become unusually avid

substrates for transglutaminase activity, resulting in the cross-linking of proteins within neurons. Either the cross-linked proteins themselves or the insoluble isodipeptide that is produced during subsequent proteolysis could be toxic to cells. Long-lived postmitotic cells, like neurons, would be most susceptible to the accumulation of cross-linked material. For example, the expanded AR might become cross-linked nonspecifically to itself and other proteins, effectively titrating out functional AR, perturbing nuclear and cellular processes, and leading to an accumulation of cross-linked protein or isodipeptide. Because the conformation and nuclear transport of the AR protein is controlled by androgen binding, the extent of cross-linking might be androgen dependent.

All three models can explain the repeat length dependence of the disease process. One can further speculate how each of the three could be specifically neurotoxic: In the first model one or more regulatory cascades that are neuron specific would be perturbed, and in the second and third models the long-lived postmitotic nature of neurons would make them uniquely susceptible to the accumulation of aggregated or cross-linked protein. The first model, however, would seem best at explaining the specific neuropathology seen in each CAG repeat disease; for each disease, the affected regulatory cascade(s) would be more important to cell survival in precisely those neurons that preferentially die. It is more difficult to explain how an accumulation of aggregated or cross-linked material—which should occur in most or all neurons, given the widespread CNS expression of CAG repeat genes—could lead in each disease to the death of different subsets of neurons. In order for the second and third models to explain the specific neuropathology of each disease, further disease-specific and cell-specific factors need to be invoked that result in a differential accumulation of (or sensitivity to) aggregated and cross-linked material or breakdown products, in distinct regions of the brain. No one has yet found evidence for aggregated, precipitated, or cross-linked protein in any of these diseases.

Intensive research is now underway in many laboratories to test aspects of these models. A full understanding of the molecular mechanisms at play will only come from research across a broad front, including transgenic and knock-out animal models, in vitro expression systems, yeast and bacterial systems designed to identify interacting proteins, and rigorous biochemical and immunocytochemical analyses of the normal and expanded proteins. Recent successes in creating an SCA1 transgenic model (Burright et al 1995) and in achieving targeted disruption of the HD gene [resulting in embryonic lethality (Nasir et al 1995)] are promising steps in this direction.

ACKNOWLEDGMENTS

We thank Brian Brooks, Peter Bingham, Peter Crino, and Jim Eberwine for critically reading the manuscript. H Paulson is supported by a Howard Hughes

Postdoctoral Research Fellowship for Physicians and a Neuropharmacology Research Fellowship from the American Academy of Neurology. K Fischbeck is supported by grants from the National Institutes of Health, the Muscular Dystrophy Association, the March of Dimes, and the Amyotrophic Lateral Sclerosis Association.

Literature Cited

Abitbol M, Menini C, Delezoide A-L, Rhyner T, Vekemans M, et al. 1993. Nucleus basalis magnocellularis and hippocampus are the major sites of FMR-1 expression in the human fetal brain. *Nat. Genet.* 4:147–53

Andrew SE, Goldberg YP, Kremer B, Telenius H, Theilman J, et al. 1993. The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nat. Genet.* 4:398–403

Anvret M, Ahlberg G, Grandell U, Hedberg B, Johnson K, et al. 1993. Larger expansions of the CTG repeat in muscle compared to lymphocytes from patients with myotonic dystrophy. *Hum. Mol. Genet.* 2: 1397–400

Ashizawa T, Dubel JR, Dunne PW, Dunne CJ, Fu H-Y, et al. 1992. Anticipation in myotonic dystrophy. II. Complex relationships between clinical findings and structure of the GCT repeat. *Neurology* 42:1877–83

Ashley CT, Sutcliffe JS, Kunst CB, Leiner HA, Eichler EE, et al. 1993a. Human and murine FMR-1: alternative splicing and translational initiation downstream of the CGG-repeat. *Nat. Genet.* 4:244–51

Ashley CT, Wilkinson KD, Reines D, Warren ST. 1993b. FMR-1 protein: conserved RNP family domains and selective RNA binding. *Science* 262:563–66

Aslandis C, Jansen G, Amemiya C, Shutler G, Mahadevan M, et al. 1992. Cloning of the essential myotonic dystrophy region and mapping of the putative defect. *Nature* 355: 548–51

Banfi S, Servadio A, Chung M, Kwiatkowski TJ, McCall AE, et al. 1994. Identification and characterization of the gene causing type 1 spinocerebellar ataxia. *Nat. Genet.* 7:513–20

Bates G, Lehrach H. 1994. Trinucleotide repeat expansions and human genetic disease. *BioEssays* 16:277–84

Beal MF. 1992. Does impairment of energy metabolism result in excitotoxic neuronal death in neurodegenerative diseases? *Neurology* 31:119–30

Beal MF, Brouillet E, Jenkins BG, Ferrante RJ, Kowall NW, et al. 1993. Neurochemical and histologic characterization of striatal excitotoxic lesions produced by the mitochondrial toxin 3-nitropropionic acid. *J. Neurosci.* 13: 4181–92

Bell MV, Hirst MC, Nakahori Y, MacKinnon RN, Roche A, et al. 1991. Physical mapping across the fragile X: hypermethylation and clinical expression of the fragile X syndrome. *Cell* 64:861–66

Bergoffen J, Kant J, Sladky J, McDonald-McGinn M, Zackai E, Fischbeck KH. 1994. Paternal transmission of congenital myotonic dystrophy. *J. Med. Genet.* 31:518–20

Bingham PM, Scott MO, Wang S, McPhaul MJ, Wilson EM, et al. 1995. Stability of an expanded trinucleotide repeat in the androgen receptor gene in transgenic mice. *Nat. Genet.* 9:191–96

Broholm J, Cassiman J-J, Craufurd D, Falek A, Farmer-Little C, et al. 1994. Guidelines for the molecular genetics predictive test in Huntington's disease. *Neurology* 44:1533–36

Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, et al. 1992. Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family member. *Cell* 68:799–808

Bruhn L, Hwang-Shum JJ, Sprague GF. 1992. The N-terminal 96 residues of MCM1, a regulator of cell type-specific genes in *Saccharomyces cerevisiae,* are sufficient for DNA binding, transcription activation, and interaction with alpha 1. *Mol. Cell. Biol.* 12: 3563–72

Burke JR, Wingfield MS, Lewis KE, Roses AD, Lee JE, et al. 1994. The Haw River syndrome: dentatorubral-pallidoluysian atrophy (DRPLA) in an African-American family. *Nat. Genet.* 7:521–24

Burright E, Clark HB, Servadio A, Matilla T, Feddersen RM, et al. 1995. SCA1 transgenic mice: a model for neurodegeneration caused by trinucleotide expansion. Submitted

Buxton J, Shelbourne P, Davies J, Jones C, Van Tongeren T, et al. 1992. Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature* 355: 547–48

Chamberlain NL, Driver ED, Miesfeld RL. 1994. The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. *Nucl. Acids Res.* 22:3181–86

Chung M, Ranum LPW, Duvick LA, Servadio A, Zoghbi HY, et al. 1993. Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type 1. *Nat. Genet.* 5:254–58

Creusot F, Verdiere J, Gaisne M, Slonimski PP. 1988. CYP1 (HAP1) regulator of oxygen-dependent gene expression in yeast. *J. Mol. Biol.* 204:263–76

De Boulle K, Verkerk AJMH, Reyniers E, Vits L, Hendricks J, et al. 1993. A point mutation in the *FMR-1* gene associated with fragile X mental retardation. *Nat. Genet.* 3:31–35

Devys D, Lutz Y, Rouyer N, Bellocq JP, Mandel J-L. 1993. The FMR-1 protein is cytoplasmic, most abundant in neurons, and appears normal in carriers of the fragile X premutation. *Nat. Genet.* 4:335–40

DiBenedetto AJ, Pittman RN. 1995. Death in the balance. *Perspect. Dev. Neurobiol.* In press

DiFiglia M, Sapp E, Chase K, Schwarz C, Meloni A, et al. 1995. Huntingtin is a cytoplasmic protein associated with vesicles in human and rat brain neurons. *Neuron* 14: 1075–81

Doyu M, Sobue G, Mukai E, Kachi T, Yasuda T, et al. 1992. Severity of X-linked recessive bulbospinal neuronopathy correlates with size of the tandem CAG repeat in androgen receptor gene. *Ann. Neurol.* 32:707–10

Duboule D, Haenlin M, Galliot B, Mohier E. 1987. DNA sequences homologous to the *Drosophila opa* repeat are present in murine mRNAs that are differentially expressed in fetuses and adult tissues. *Mol. Cell. Biol.* 7: 2003–6

Dunne PW, Walch ET, Epstein HF. 1994. Phosphorylation reactions of recombinant human myotonic dystrophy protein kinase and their inhibition. *Biochemistry* 33:10809–14

Dutch-Belgian Fragile X Consortium. 1994. Fmr1 knockout mice: a model to study fragile X mental retardation. *Cell* 78:23–33

Duyao MP, Ambrose C, Myers R, Novelletto A, Persichetti F, et al. 1993. Trinucleotide repeat instability and age of onset in Huntington's disease. *Nat. Genet.* 4:387–92

Eichler EE, Holden JJ, Popovich BW, Reiss AL, Snow K, et al. 1994. Length of uninterrupted CGG repeats determines instability in the FMR1 gene. *Nat. Genet.* 8:88–94

Feng Y, Zhiang F, Lokey LK, Chastain JL, Lakkis L, et al. 1995. Translational suppression by trinucleotide repeat expansion at *FMR1*. *Science* 268:731–34

Ferrante RJ, Kowall NW, Cipolloni PB, Storey E, Beal MF, et al. 1993. Excitotoxic lesions in primates as a model for Huntington's disease: histopathologic and neurochemical characterization. *Exp. Neurol.* 119:46–71

Fischbeck KH, Souders D, La Spada AR. 1991. A candidate gene for X-linked spinal muscular atrophy. *Adv. Neurol.* 56:209–13

Fu YH, Friedman DL, Richards S, Pearlman JA, Gibbs RA, et al. 1993. Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. *Science* 260:235–38

Fu YH, Kuhl DPA, Pizzuti A, Pieretti M, Sutcliffe JS, et al. 1991. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 67:1047–58

Fu YH, Pizzuti A, Fenwick RG, King J, Rajnarayan S, et al. 1992. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* 255:1256–58

Gedeon AK, Baker E, Robinson H, Partington MW, Gross B, et al. 1992. Fragile X syndrome without CCG amplification has an *FMR-1* deletion. *Nat. Genet.* 1:341–44

Gerber H-P, Seipel K, Georgiev O, Hofferer M, Hug M, et al. 1994. Transcriptional activation modulated by homopolymeric glutamine and proline stretches. *Science* 263: 808–11

Goldberg YP, Kremer B, Andrew SE, Theilmann J, Graham RK, et al. 1993. Molecular analysis of new mutations for Huntington's disease: intermediate alleles and sex of origin effects. *Nat. Genet.* 5:174–79

Goldberg YP, Telenius H, Hayden MR. 1994. The molecular genetics of Huntington's disease. *Curr. Opin. Neurol.* 7:325–32

Green H. 1993. Human genetic diseases due to codon reiteration: relationship to an evolutionary mechanism. *Cell* 74:955–56

Gusella JF, Wexler NS, Conneally PM, Naylor SL, Anderson MA, et al. 1983. A polymorphic marker genetically linked to Huntington's disease. *Nature* 306:234–38

Hansen RS, Gartler SM, Scott CR, Chen SH, Laird CD. 1992. Methylation analysis of CGG sites in the CpG island of the FMR1 gene. *Hum. Mol. Genet.* 1:571–78

Harding AE. 1993. Clinical features and classification of inherited ataxias. *Adv. Neurol.* 61:1–14

Harding AE, Thomas PK, Baraitser M, Bradbury PG, Morgan-Hughes JA, et al. 1982.

X-linked recessive bulbospinal neuropathy: a report of ten cases. *J. Neurol. Neurosurg. Psychiatry* 45:1012–19

Harley HG, Brook JD, Rundle SA, Crow S, Reardon W, et al. 1992. Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature* 355: 545–46

Harper PS. 1989. *Myotonic Dystrophy.* London/Philadelphia: Saunders. 2nd ed.

Harper PS, Morris MJ, Quarell O, eds. 1991. *Huntington's Disease.* Philadelphia: Saunders

Hersch S, Jones R, Koroshetz W, Quaid K. 1994. The neurogenetics genie: testing for the Huntington's disease mutation. *Neurology* 44:1369–73

Hirst M, Grewal P, Flannery A, Slatter R, Maher E, et al. 1995. Two new cases of FMR1 deletion associated with mental impairment. *Am. J. Hum. Genet.* 56:67–74

Hoogeveen AT, Verhey C, Bakker CE, de Graaf E, Willemsen R, et al. 1994. A tissue dependent processing and function of the FMR-1 gene product. *Am. J. Hum. Genet.* 55:A18 (Abstr.)

Huntington's Disease Collaborative Research Group. 1993. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 72:971–83

Igarashi S, Tanno Y, Onodera O, Yamazaki M, Sato S, et al. 1992. Strong correlation between the number of CAG repeats in androgen receptor genes and the clinical onset of features of spinal and bulbar muscular atrophy. *Neurology* 42:2300–2

Jansen G, Willems P, Coerwinkel M, Nillesen W, Smeets H, et al. 1994. Gonosomal mosaicism in myotonic dystrophy patients: involvement of mitotic events in $(CTG)_n$ repeat variation and selection against expansion in sperm. *Am. J. Hum. Genet.* 54:575–85

Kao CC, Lieberman PM, Schmidt MC, Zhou Q, Pei R, et al. 1990. Cloning of a transcriptionally active human TATA binding factor. *Science* 248:1646–50

Kawaguchi Y, Okamoto T, Taniwaki M, Aizawa M, Inoue M, et al. 1994. CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nat. Genet.* 8:221–28

Kennedy WR, Alter M, Sung JH. 1968. Progressive proximal spinal and bulbar muscular atrophy of late onset: a sex-linked recessive trait. *Neurology* 18:671–80

Knight SJL, Flannery AV, Hirst MC, Campbell L, Christodoulou Z, et al. 1993. Trinucleotide repeat expansion and hypermethylation of a CpG island in FRAXE mental retardation. *Cell* 74:127–34

Koide R, Ikeuchi I, Onodera O, Tanaka H, Igarishi S, et al. 1994. Unstable expansion of CAG repeat in hereditary dentatorubral-pallidoluysian atrophy (DRPLA). *Nat. Genet.* 6:9–13

Komure O, Sano A, Nishino N, Yamaguchi N, Ueno S, et al. 1995. DNA analysis in hereditary dentatorubral-pallidoluysian atrophy: correlation between CAG repeat length and phenotypic variation and the molecular basis of anticipation. *Neurology* 45:143–49

Krahe R, Ashizawa T, Narayana L, Fununage VL, Carango P, et al. 1994. Effect of the myotonic dystrophy expanded $(CTG)_n$ repeat on the transcripts of DMPK alleles. *Am. J. Hum. Genet.* 55:A227 (Abstr.)

Kremer B, Goldberg P, Andrew SE, Theilman J, Telenius H, et al. 1994. A worldwide study of the Huntington's disease mutation: the sensitivity and specificity of measuring CAG repeats. *New Engl. J. Med.* 330:1401–6

Kremer EJ, Pritchard M, Lynch M, Yu S, Holman K, et al. 1991. Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence $p(CGG)_n$. *Science* 252:1711–14

Kuhl D, Pizzuti A, Caskey CT. 1993. Triplet repeat mutations: an important etiology of genetic disease. *Curr. Opin. Neurol.* 13:55–85

LaSpada AR, Paulson HL, Fischbeck KH. 1994. Trinucleotide repeat expansion in neurological disease. *Ann. Neurol.* 36:814–22

La Spada AR, Roling D, Harding AE, Warner CL, Spiegel R, et al. 1992. Meiotic stability and genotype-phenotype correlation of the trinucleotide repeat in X-linked spinal and bulbar muscular atrophy. *Nat. Genet.* 2:301–4

La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH. 1991. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 352:77–79

Li SH, Schilling G, Young WS, Li X-J, Margolis RL, et al. 1993. Huntington's disease gene (IT15) is expressed widely in human and rat tissues. *Neuron* 11:985–93

MacDonald ME, Barnes G, Srinidi J, Duyao MP, Ambrose CM, et al. 1993. Gametic but not somatic instability of CAG repeat length in Huntington's disease. *J. Med. Genet.* 30: 982–86

Mahadevan M, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, et al. 1992. Myotonic dystrophy mutation: an unstable CTG repeat in the 3′ untranslated region of the gene. *Science* 255:1253–55

Mandel J-L. 1994. Trinucleotide diseases on the rise. *Nat. Genet.* 7:453–55

Matilla T, Volpini V, Genis D, Rosell J, Corral J, et al. 1993. Presymptomatic analysis of spinocerebellar ataxia type 1 (SCA1) via the expansion of the SCA1 CAG-repeat in a large pedigree displaying anticipation and parental male bias. *Hum. Mol. Genet.* 2: 2123–28

McConkie-Rosell A, Lachiewicz AM, Spiridigliozzi GA, Tarleton J, Schoenwald S, et al. 1993. Evidence that methylation of the FMR-1 locus is responsible for variable phenotypic expression of the fragile X syndrome. *Am. J. Hum. Genet.* 53:800–9

Mhatre AN, Trifiro MA, Kaufman M, Kazemi-Esfarjani P, Figlewicz D, et al. 1993. Reduced transcriptional regulatory competence of the androgen receptor in X-linked spinal and bulbar muscular atrophy. *Nat. Genet.* 5: 184–88

Mitchell PJ, Tjian R. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* 245: 371–78

Modrich P. 1994. Mismatch repair, genetic stability, and cancer. *Science* 266:1959–60

Myers RH, MacDonald ME, Koroshetz WJ, Duyao MP, Ambrose CM, et al. 1993. De novo expansion of a $(CAG)_n$ repeat in sporadic Huntington's disease. *Nat. Genet.* 5: 168–73

Nagafuchi S, Yanagisawa H, Ohsaki E, Shirayama T, Tadokoro K, et al. 1994a. Structure and expression of the gene responsible for the triplet repeat disorder, dentatorubral and pallidoluysian atrophy (DRPLA). *Nat. Genet.* 8:177–82

Nagafuchi S, Yanagisawa H, Sato K, Shirayama T, Ohsaki E, et al. 1994b. Dentatorubral and pallidoluysian atrophy expansion of an unstable CAG trinucleotide on chromosome 12p. *Nat. Genet.* 6:14–18

Naito H, Oyanagi S. 1982. Familial myoclonus epilepsy and choreoathetosis: hereditary dentatorubral-pallidoluysian atrophy. *Neurology* 32:798–807

Nancarrow JK, Kremer E, Holman K, Eyre H, Doggett NA, et al. 1994. Implications of FRA16A structure for the mechanism of chromosomal fragile site genesis. *Science* 264:1938–41

Nasir J, Floresco SB, O'Kusky JR, Dewart VM, Richman J, et al. 1995. Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. *Cell* 81:811–23

Nelson DL, Warren ST. 1993. Trinucleotide repeat instability: when and where? *Nat. Genet.* 4:107–8

Neumann MN. 1959. Combined degeneration of globus pallidus and dentate nucleus and their projections. *Neurology* 9:430–38

Novelletto A, Persichetti F, Sabbadini G, Mandich P, Bellone E, et al. 1994. Analysis of the trinucleotide repeat expansion in Italian families affected with Huntington's disease. *Hum. Mol. Genet.* 3:93–98

Novelli G, Gennarelli M, Zelano G, Pizzati A, Fattorini C, et al. 1993. Failure in detecting mRNA transcripts from the mutated allele in myotonic dystrophy muscle. *Biochem. Mol. Biol. Int.* 29:291–97

Nussbaum RL, Ledbetter DH. 1989. The fragile X syndrome. In *Metabolic Basis of Inherited Diseases,* ed. CR Scriver, AL Beaudet, WS Sly, D Valle, pp. 327–42. New York: McGraw-Hill

Oberle I, Rousseau F, Heitz D, Kietz C, Devys D, et al. 1991. Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science* 252:1087–102

Orr HT, Chung M, Banfi S, Kwiatkowski TJ, Servadio A, et al. 1993. Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat. Genet.* 4:221–26

Parrish JE, Oostra BA, Verkerk AJ, Richards CS, Reynolds J, et al. 1994. Isolation of a GCC repeat showing expansion in FRAXF, a fragile site distal to FRAXA and FRAXE. *Nat. Genet.* 8:229–35

Pascal E, Tjian R. 1991. Different activation domains of Sp1 govern formation of multimers and mediate transcriptional synergism. *Genes Dev.* 5:1646–56

Perutz MF, Johnson T, Suzuki M, Finch JT. 1994. Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *Proc. Natl. Acad. Sci. USA* 91:5355–58

Pieretti M, Zhang F, Fu YH, Warren ST, Oostra BA, et al. 1991. Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell* 66:817–22

Pizzuti A, Friedman DL, Caskey CT. 1993. The myotonic dystrophy gene. *Arch. Neurol.* 50: 1173–79

Ptacek LJ, Johnson KJ, Griggs RC. 1993. Genetics and physiology of the myotonic muscle disorders. *New Engl. J. Med.* 328:482–89

Ranum LP, Chung M, Banfi S, Bryer A, Schut LJ, et al. 1994. Molecular and clinical correlations in spinocerebellar ataxia type 1: evidence for familial effects on the age of onset. *Am. J. Hum. Genet.* 55:244–52

Rastinejad F, Blau HM. 1993. Genetic complementation reveals a novel regulatory role for 3′ untranslated regions in growth and differentiation. *Cell* 72:903–17

Redman JB, Fenwick RG, Fu Y-H, Pizzuti A, Caskey CT. 1993. Relationship between parental trinucleotide GCT repeat length and severity of myotonic dystrophy in offspring. *J. Am. Med. Assoc.* 269:1960–65

Reiss AL, Aylward E, Freund LS, Joshi PK, Bryan RN. 1991. Neuroanatomy of fragile X syndrome: the posterior fossa. *Ann. Neurol.* 29:26–32

Reyniers E, Vits L, De Boulle, Van Roy B, Van Velzen D, et al. 1993. The full mutation in the FMR-1 gene of male fragile X patients is absent in their sperm. *Nat. Genet.* 4:143–46

Richards RI, Sutherland GR. 1994. Simple DNA is not repeated simply. *Nat. Genet.* 6: 114–16

Ritchie RJ, Knight SJL, Hirst MC, Grewal PK, Bobrow M, et al. 1994. The cloning of FRAXF: trinucleotide repeat expansion and methylation at a third fragile site in distal Xqter. *Hum. Mol. Genet.* 3:2115–21

Rosenberg RN. 1992. Machado-Joseph disease: an autosomal dominant motor system degeneration. *Mov. Disord.* 7:193–203

Rosenberg RN. 1995. Autosomal dominant cerebellar phenotypes: the genotype has settled the issue. *Neurology* 45:1–5

Ross CA, McInnis MG, Margolis RL, Li S-H. 1993. Genes with triplet repeats: candidate mediators of neuropsychiatric disorders. *Trends Neurosci.* 16:254–60

Sabouri LA, Mahadevan MS, Narang M, Lee DSE, Surh LC, et al. 1993. Effect of the myotonic dystrophy (DM) mutation on mRNA levels of the DM gene. *Nat. Genet.* 4:233–38

Sarnat HB, Silbert SW. 1976. Maturational arrest of fetal muscle in neonatal myotonic dystrophy. *Arch. Neurol.* 33:466–74

Schut JW. 1954. Hereditary ataxia: clinical study through six generations. *Arch. Neurol. Psychiatry* 63:535–67

Servadio A, Koshy B, Armstrong D, Antalffy B, Orr HT, Soghbi HY. 1995. Expression analysis of the ataxin-1 protein in tissues from normal and spinocerebellar ataxia type 1 individuals. *Nat. Genet.* 10:94–98

Sharp AH, Loev SJ, Schilling G, Li S-H, Li X-J, et al. 1995. Widespread expression of Huntington's disease gene (IT15) protein product. *Neuron* 14:1–20

Sherman SL, Jacobs PA, Morton NE, Froster-Iskenius U, Howard-Peebler PN, et al. 1985. Further segregation analysis of the fragile X syndrome with special reference to transmitting males. *Hum. Genet.* 69:289–99

Siomi H, Siomi MC, Nussbaum RL, Dreyfuss G. 1993. The protein product of the fragile X gene, FMR-1, has characteristics of an RNA binding protein. *Cell* 74:291–98

Simpson L, Emeson RB. 1996. RNA editing. *Annu. Rev. Neurosci.* 19:In press

Smith JK, Gonda VE, Malamud N. 1958. Unusual form of cerebellar ataxia: combined dentatorubral-pallidoluysian atrophy. *Neurology* 32:798–807

Snell RG, MacMillan JC, Cheadle JP, Fenton I, Lazarou PL, et al. 1993. Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. *Nat. Genet.* 4:393–97

Sobue G, Hashizume Y, Mukai E, Hirayama M, Mitsuma T, et al. 1989. X-linked recessive bulbospinal neuronopathy: a clinicopathological study. *Brain* 112:209–32

Stott K, Blackburn JM, Butler PJG, Perutz M. 1995. Incorporation of glutamine repeats makes protein oligomerize: implications for neurodegenerative diseases. *Proc. Natl. Acad. Sci. USA*. In press

Strong TV, Tagle DA, Valdes JM, Elmer LW, Boehm K, et al. 1993. Widespread expression of the human and rat Huntington's disease gene in brain and nonneural tissues. *Nat. Genet.* 5:259–65

Sutcliffe JS, Nelson DL, Zhang F, Pieretti M, Caskey CT, et al. 1992. DNA methylation represses FMR-1 transcription in fragile X syndrome. *Hum. Mol. Genet.* 1:397–400

Sutherland GR, Richards RI. 1993. Dynamic mutations on the move. *J. Med. Genet.* 30: 978–81

Takahashi H, Okama E, Naito H, Takeda S, Nakashima S, et al. 1988. Hereditary dentatorubral-pallidoluysian atrophy: clinical and pathologic variants in a family. *Neurology* 38:1065–70

Telenius H, Kremer B, Goldberg YP, Theilman J, Andrew SE, et al. 1994. Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm. *Nat. Genet.* 6:409–14

Thornton CA, Griggs RC, Moxley RT III. 1994. Myotonic dystrophy with no trinucleotide repeat expansion. *Ann. Neurol.* 35: 269–71

Trottier Y, Devys D, Imbert G, Saudou F, An I, et al. 1995. Cellular localization of the Huntington's disease protein and discrimination of the normal and mutated form. *Nat. Genet.* 10:104–10

Tsilfidis C, MacKenzie AE, Mettler G, Barcelo J, Korneluk RG. 1992. Correlation between CTG trinucleotide repeat length and frequency of severe congenital myotonic dystrophy. *Nat. Genet.* 1:192–95

van der Ven PF, Jansen G, van Kuppevelt TH, Perryman MB, Lupa M, et al. 1993. Myotonic dystrophy kinase is a component of neuromuscular junctions. *Hum. Mol. Genet.* 2:1889–94

Verheij C, Bakker CE, de Graff E, Keulemana J, Willemsen R, et al. 1993. Characterization and localization of the FMR-1 gene product associated with fragile-X syndrome. *Nature* 363:722–24

Verkerk AJ, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DP, et al. 1991. Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 65:905–14

Vonsattel JP, Myers RH, Stevens TJ, Ferrante RJ, Bird ED, et al. 1985. Neuropathological classification of Huntington's disease. *J. Neuropath. Exp. Neurol.* 44:559–77

Warner CL, Griffen JE, Wilson JD, Jacobs LD, Murray KR, et al. 1992. X-linked spinomuscular atrophy; a kindred with associated ab-

normal androgen binding. *Neurology* 42: 2181–84
Wexler NS, Rose EA, Housman DE. 1991. Molecular approaches to hereditary diseases of the nervous system: Huntington's disease as a paradigm. *Annu. Rev. Neurosci.* 14:503–29
Wharton KA, Yedvobnick B, Finnerty VG, Artavanis-Tsakonas S. 1985. *opa:* a novel family of transcribed repeats shared by the Notch locus and other developmentally regulated loci in *D. melanogaster. Cell* 40:55–62
Wieringa B. 1994. Myotonic dystrophy reviewed: back to the future? *Hum. Mol. Genet.* 3:1–7
Wisniewski KE, Segan SM, Miezejeski CM, Serson EA, Rudelli RD. 1991. The fragile X syndrome: neurological, electrophysiological, and neuropathological abnormalities. *Am. J. Med. Genet.* 38:476–80
Wohrle D, Hennig I, Vogel W, Steinbach P. 1993. Mitotic stability of fragile X mutations in differentiated cells indicates early post-conceptional trinucleotide repeat expansion. *Nat. Genet.* 4:140–42
Wohrle D, Kotzot D, Hirst MC, Manca A, Korn B, et al. 1992. A microdeletion of less than 250 kb, including the proximal part of the *FMR-1* gene and the fragile X site, in a male with the clinical phenotype of fragile X syndrome. *Am. J. Hum. Genet.* 51:299–306
Yazawa I, Nukina N, Hashida H, Goto J, Yamada M, Kanazawa I, et al. 1995. Abnormal gene product identified in hereditary dentatorubral-pallidoluysian atrophy (DRPLA) brain. *Nat. Genet.* 10:99–103
Yu W-HA. 1989. Administration of testosterone attenuates neuronal loss following axotomy in the brain-stem motor nuclei of female rats. *J. Neurosci.* 9:3908–14
Zhang L, Leeflang EP, Yu JY, Arnheim N. 1994. Studying human mutations by sperm typing: instability of CAG trinucleotide repeats in the human androgen receptor gene. *Nat. Genet.* 7:531–35
Zhou Z-X, Wong C-I, Sar M, Wilson EM. 1994. The androgen receptor: an overview. *Recent Prog. Hormone Res.* 49:249–74