Annu. Rev. Genetics. 1995. 29:703–28

TRINUCLEOTIDE REPEAT EXPANSION AND HUMAN DISEASE

Claude T. Ashley, Jr., and Stephen T. Warren
Howard Hughes Medical Institute and Departments of Biochemistry and Pediatrics, Emory University School of Medicine, Atlanta, Georgia 30322

KEY WORDS: trinucleotide repeat, genome instability, genetic anticipation, mental retardation, neurodegenerative disease, dynamic mutations, fragile sites

ABSTRACT

Eleven human loci, responsible for nine diseases, exhibit an unprecedented form of mutation: the expansion of trinucleotide repeats. Normally polymorphic CGG/CCG or CAG/CTG repeats (means of ~20 triplets) are found enlarged to either 2–3 or 10–1000 times normal lengths. The smaller expansions are found within genes coding for polyglutamine and are associated with neurodegenerative diseases. The larger expansions are most commonly associated with chromosomal fragile sites and can be found isolated or in the untranslated regions of genes. Expanded alleles of all these loci exhibit remarkable meiotic instability, frequently lengthening upon transmission. Since the abnormal repeat length can be correlated with incomplete penetrance and/or variable expressivity, the elongation in subsequent generations explains the genetic anticipation in these disorders.

CONTENTS

INTRODUCTION

The remarkable fidelity of DNA replication and the predictable transmission of Mendelian traits sometimes appear to attenuate the fluidity of the genome. In humans, abundant polymorphic loci with multiple alleles, particularly those involving micro- or minisatellites, attest to the lubricious nature of certain sequences (8). However, the potential for genome instability has not been more apparent than with the recent discoveries of trinucleotide repeat expansion mutations.

Since 1991, 11 human loci, responsible for 9 genetic diseases, have been identified in which a normally polymorphic trinucleotide repeat undergoes a mutational change whereby the repeat length expands, sometimes quite remarkably (Table 1). This dynamic mutation is unlike conventional mutations because the expanded repeat can undergo further change (usually continued expansion) with each subsequent generation. The magnitude of the expansion is, in part, related to the parental repeat length. For those expansions that result in a phenotype, the subsequent expansion of the repeat in descending generations can lead to non-Mendelian inheritance patterns due to incomplete penetrance and/or variable expressivity, which can generally be directly correlated with repeat length. Therefore the discovery of trinucleotide repeat expansions has uncovered a previously unrecognized form of mutational change and has provided a molecular explanation for some unusual inheritance patterns.

Loci that undergo trinucleotide repeat expansion can currently be placed within two groups based upon the involved repeat; either the CGG repeat (equally a CCG, the 5′ sequence on the other strand) or the CAG repeat (alternatively CTG, again of the other strand). All four designations are used in the literature depending largely upon which strand is coding, if in the context of a gene. In addition, depending on the designation of the first base of a repeat tract, a GCC repeat, largely equivalent to a CCG or CGG repeat, may be reported. Alternatively, it has been argued that an alphabetical repeat nomenclature (CCG or AGC, for each repeat) be used (171). However, the repeat nomenclature used below reflects the predominant usage among investigators.

The CGG/CCG repeat loci involve noncoding (but sometimes transcribed) triplets and, when expanded, are associated with chromosomal fragile sites. The CAG/CTG repeats are most often coding and associated, when expanded, with neurodegenerative diseases. The exception is the 3′ untranslated CTG repeat of myotonic dystrophy. The CTG repeat associated with myotonic dystrophy, like the CGG/CCG fragile site repeats, can undergo massive expansions from normal lengths of ~30 repeats to several thousand when fully expanded, while the coding CAG repeats undergo a more modest two- to threefold increase over a normal length of ~20 repeats.

Table 1 Trinucleotide repeat expansions in humans

Locus	Disorder	Repeat Alleles: Normal	Repeat Alleles: Mutant	Repeat location	Functional consequence of expansion
FMR1 (FRAXA)	Fragile X syndrome	$(CGG)_{6-52}$	$(CGG)_{60-200}$ (premutation) $(CGG)_{230\rightarrow1000}$ (full)	5'-UTR	Loss-of-function Fragile site
FRAXE	Fragile XE mental retardation	$(GCC)_{7-35}$	$(GCC)_{130-150}$ (premutation) $(GCC)_{230-750}$ (full)	ND[a]	Fragile site Loss-of-function ?
FRAXF	None	$(GCC)_{6-29}$	$(GCC)_{300-1000}$	ND	Fragile site
FRA16A	None	$(CCG)_{16-49}$	$(CCG)_{1000-1900}$	ND	Fragile site
FRA11B (CBL2)	Predisposition toward Jacobsen (11q-) syndrome in offspring	$(CGG)_{11}$	$(CGG)_{80}$ (premutation) $(CGG)_{100\rightarrow1000}$ (full)	5'UTR	Fragile site
AR	Spinal and bulbar muscular atrophy	$(CAG)_{11-33}$	$(CAG)_{38-66}$	Coding	Gain and partial loss-of-function
SCA1	Spinocerebellar ataxia Type 1	$(CAG)_{6-39}$	$(CAG)_{41-81}$	Coding	Gain of function
Hdh (IT15)	Huntington's disease	$(CAG)_{10-35}$	$(CAG)_{36-121}$	Coding	Gain of function
B37 (DRPLA)	Dentatoribral-pallidoluysian atrophy Haw River syndrome (phenotypic variant)	$(CAG)_{7-25}$	$(CAG)_{49-75}$ $(CAG)_{63-68}$	Coding	Gain of function
MJD1 (SCA3)	Machado-Joseph disease	$(CAG)_{12-37}$	$(CAG)_{61-84}$	Coding	Gain of function
DMPK	Myotonic dystrophy	$(CTG)_{5-37}$	$(CTG)_{50\rightarrow3000}$	3'UTR	Processing of *DMPK* Message abnormal?

[a] ND = not determined

CHROMOSOMAL FRAGILE SITES AND CGG/CCG REPEATS

Fragile X Syndrome (FMR1/FRAXA)

Chromosomal fragile sites are cytogenetically visible, nonstaining gaps on metaphase chromosomes. The human fragile site at Xq27.3, termed *FRAXA* (FRAgile site, X chromosome, A site), is associated with fragile X syndrome, the most frequent cause of inherited mental retardation (190–192). Patients with fragile X syndrome display the triad of mental retardation, ranging from mild to severe, characteristically long and narrow facies with large ears and frontal prominence, and postpubescent macroorchidism (59).

The inheritance pattern of fragile X syndrome is complex but is consistent with an X-linked dominant with reduced penetrance. In a study of 206 families segregating fragile X syndrome, reduced penetrance of 80% in males and 30% in females was found (153, 154). In addition, Sherman et al (153, 154) observed that penetrance depended upon one's position within the pedigree, with a general trend of increasing penetrance in successive generations. While the term genetic anticipation, which refers to increasing severity or decreasing age of onset of disease in successive generations, might be applied to such observations today, these peculiarities in the inheritance pattern of fragile X syndrome were frequently referred to as the Sherman paradox.

Events preceding and culminating in the identification of the gene responsible for fragile X syndrome, referred to as fragile X mental retardation 1 (*FMR1*), have recently been reviewed (190). Within the *FMR1* gene, an unusual CGG repeat displaying marked expansion in fragile X patients was identified (94, 126, 183, 198). This trinucleotide repeat is highly polymorphic within the normal population, ranging from 6–52 copies in normal individuals of various ethnic origins (49). The repeat is cryptic with interspersed AGG triplets occurring most often at repeat positions 10 and 20 (38, 95, 183). In affected individuals, expansions ranging from 230→1000, referred to as "full mutations", were seen (44, 49). Premutation alleles, alleles with repeat sizes between 60–230 repeats that have a high probability of expansion in subsequent offspring of carrier females, have also been identified in the phenotypically normal transmitting males and their carrier daughters (49). The premutation length is directly related to the risk of full expansion, and therefore penetrance, thus resolving the Sherman paradox (49). Since carrier males only transmit premutation alleles regardless of somatic tissue repeat length, the absence of penetrance in obligate carrier daughters has been explained (136, 194). Reductions in *FMR1* repeat sizes have been occasionally observed, primarily upon male transmissions of premutations (44), whereas subsequent female transmissions result almost exclusively in repeat expansions. Somatic

mosaicism of the expanded *FMR1* trinucleotide repeat is common, with both premutation and multiple full mutation bands being observed in many affected males (49). This has been interpreted as evidence for a postzygotic window of massive repeat expansion.

Concomitant with *FMR1* trinucleotide repeat expansion into the full mutation range is methylation of both the CGG repeat as well as the nearby CpG island that correlates with the absence of *FMR1* gene expression (73, 131, 167). Normal alleles and premutation alleles remain unmethylated, consistent with findings of normal amounts of *FMR1* mRNA and the encoded protein, FMRP (42). Within the range of 200–300 repeats, methylation is the most sensitive indicator of penetrance (104, 113), although ribosome stalling on the expanded repeat within mRNA can lead to diminished protein levels (43). Confirmation that the absence of *FMR1* gene expression is responsible for the fragile X syndrome came with the identification of multiple patients with *FMR1* deletions (133 and references therein) and a single point mutation (29), all in CGG repeat expansion-negative individuals. Also, an *Fmr1* knockout mouse exhibits features of fragile X syndrome, including enlarged testes, cognitive impairment, and hyperactivity (28).

Consistent with the absence of new mutations resulting in fragile X syndrome (159), marked linkage disequilibrium suggests a small pool of founder chromosomes (129, 138). Haplotyping of normal and fragile X chromosomes revealed distinct and unequal allelic frequencies among these two classes of chromosomes (95). Through sequence analysis of normal *FMR1* repeats, it has been shown that haplotypes normally containing repeats that are more nearly pure (less cryptic) tend to be overrepresented on fragile X chromosomes, suggesting that long perfect repeats are the predisposed, ancestral alleles (71, 95, 161). Eichler et al (38) have shown that observable intergenerational instability correlates with uninterrupted repeats, 34–38 repeats in length. In addition, Kunst & Warren (95) have identified a distinct polarity of variation in CGG repeats, with all variation in length occurring on the 3′ end of the repeat. Based on this observation, a model of trinucleotide repeat expansion due to slippage occurring during lagging strand synthesis of DNA replication has been proposed (95).

The *FMR1* gene consists of 17 exons spanning 38 kb of Xq27.3 (40). Expression of the gene is widespread but not ubiquitous, with highest amounts of expression in brain, testes, and ovaries (2, 68, 69). Alternative splicing producing multiple protein isoforms of 39–90 kD, some with novel carboxy termini, has been observed (10, 89, 181, 182). Immunohistochemical analyses have revealed that FMRP is predominately cytoplasmic, although nuclear staining has occasionally been observed (32, 180).

Two distinct ribonucleoprotein (RNP)-family motifs frequently found in proteins that interact with RNA have been identified within the FMRP amino

acid sequence (11, 156). RNA-binding activity of FMRP toward RNA homopolymers and messenger RNA molecules has been documented, with FMRP interacting specifically with approximately 4% of messages present in human fetal brain (11, 157). Protein products containing the lone *FMR1* missense mutation, Ile367Asn, display decreased levels of RNA-binding activity in vitro (155, 181). Although the intracellular role of FMRP remains to be elucidated, the recent association of FMRP with ribosomal subunits (DE Eberhart, Y Feng, F Zhang & ST Warren, submitted) suggests a possible role in translation.

Fragile X E (FRAXE)

When the diagnosis of fragile X syndrome incorporated testing for the *FMR1* trinucleotide repeat expansion, a small group of families that were cytogenetically positive for the fragile site but negative for *FMR1* expansion was identified (14, 31). Analysis of such families identified a second folate-sensitive fragile site, termed *FRAXE*, located 600 kb telomeric to *FRAXA* (*FMR1*) (46, 90, 168). A highly polymorphic GCC repeat ranging from 7–35 repeats, with a mean of 15–16 repeats, was identified in normal individuals (4, 90, 91). Repeats of 230–750 were observed in *FRAXE*-positive individuals, and more modest repeats of 130–150 were found in carrier females who did not express *FRAXE* (90). Reminiscent of fragile X syndrome, trinucleotide repeat expansions at the *FRAXE* locus of >150 repeats were correlated with abnormal methylation of a CpG island located approximately 1 kb upstream of the repeat, perhaps influencing a nearby gene (90). It was observed that major *FRAXE* GCC repeat expansions occur more frequently upon maternal transmission, although, in contrast to *FMR1*, transmission of a large repeat from a male to his daughter has been documented (60).

Some individuals displaying *FRAXE* expression display mild mental impairment, with IQ scores in the range of 60–80 (60, 91, 117). Analogous to *FMR1*, phenotypic expression has been correlated with the degree of methylation of the nearby CpG island, as individuals with large GCC expansions at *FRAXE*, but only partial methylation, are phenotypically normal (117). Penetrance in *FRAXE* families is most often observed in individuals displaying >230 repeats (90), similar in size to the *FMR1* full mutation. Recently, deletions in the *FRAXE* region have been found in two unrelated individuals with developmental and speech delay, suggesting that the responsible gene for the *FRAXE* phenotype may reside in a 200-kb interval (51).

Fragile X F (FRAXF)

A third folate-sensitive fragile site in the region of Xq27-Xq28 has recently been identified. This fragile site (*FRAXF*) was mapped at least 600 kb distal

to *FRAXE* (70, 130). A GCC trinucleotide repeat was subsequently identified at the *FRAXF* locus (130). Normal repeat lengths ranged from 6–29 (130), while individuals expressing the *FRAXF* fragile site displayed GCC repeat expansions of 300–1000 repeats together with concomitant methylation of the repeat region (130, 140). Since *FRAXF* families without any noticeable phenotype exist (141, 168, 184), it is possible that no genes reside near *FRAXF*, or that nearby methylation effects are without consequence.

Fragile 16 A (FRA16A)

Differential methylation was identified in individuals expressing the autosomal, folate-sensitive fragile site at 16p13.1 relative to controls, and corresponding YACs revealed a highly polymorphic CCG repeat at this locus (124). Individuals expressing the *FRA16A* fragile site have large amplifications of the CGG repeat, ranging from 1000–1900 triplets (124). Repeat sizes of normal Caucasian controls ranged from 16–49 triplets, with a mean of 22 (124). Similar to *FMR1*, normal alleles with long uninterrupted repeats are believed to be predisposed toward instability (123). Instability has been observed in maternal transmissions of the *FRA16A* expansion and, unlike full expansions at *FMR1*, paternal transmission of the expansion occurs, but with little change in repeat length (123). Importantly, Nancarrow et al (123) showed that while methylation of the region was present in *FRA16A* individuals, no methylation was observed in normal controls. These findings confirmed previous predictions at *FMR1* that methylation seems to occur as a consequence rather than a cause of CGG repeat expansion (167, 169). No phenotype has yet been associated with the *FRA16A* locus, although the lack of homozygous individuals could account for this observation.

Fragile 11 B (FRA11B)

The folate-sensitive fragile site *FRA11B*, known to reside at 11q23.3 (170), has been characterized and linked to Jacobsen syndrome (82, 83). Jacobsen syndrome, involving dysmorphic features and severe mental retardation, is a chromosomal deletion syndrome with deletions in 11q (78, 150). The *FRA11B* fragile site and the breakpoints in a Jacobsen patient, whose mother expressed *FRA11B*, were recently colocalized (82, 83) near the protooncogene *CBL2* (148, 177), a gene previously shown to contain a CGG repeat in its 5′ untranslated region (17, 139). Jones et al (82) documented expansion of this CGG repeat from the common repeat length of 11 triplets in normal chromosomes to repeat sizes >100 triplets in *FRA11B*-expressing individuals. Evidence of somatic mosaicism with multiple bands on Southern blots corresponding to a repeat range from 80 to >1000 has been observed (82). In another family, a cytogenetically normal mother with 80 repeats had three *FRA11B*-expressing

offspring all with repeat lengths in excess of 100; this suggests the existence of premutations similar to that described for *FMR1*. Also similar to *FMR1*, there is evidence for abnormal methylation in the vicinity of the expanded *FRA11B* repeat (82).

In three of eight families displaying 11q23 deletions and Jacobsen syndrome, expansion of the GCC repeat of *CBL2* has been documented, consistent with the possibility that the inherent instability of the expanded *CBL2* GCC repeat could predispose toward deletions resulting in Jacobsen syndrome (82). Therefore, these data resurrect the possibility that fragile sites may influence nonrandom chromosome rearrangements, particularly those involving cancer (199). It is not known, in this instance, if *CBL2* is inactivated by abnormal *FRA11B* methylation. However, since *CBL2* is a dominant oncogene rather than a tumor suppressor, hemizygous expression of *CBL2* may not predispose toward cancer.

NEURODEGENERATIVE DISORDERS AND CAG/CTG REPEATS

Spinal and Bulbar Muscular Atrophy (SBMA)

Spinal and bulbar muscular atrophy (SBMA), also known as Kennedy's disease (88), is an X-linked recessive, adult-onset motor neuropathy (97). Affected individuals typically present in the 3rd to 5th decade with motor neuron signs such as proximal muscle weakness and fasciculations, as well as signs of bulbar muscle dysfunction such as dysphagia (96). In addition, affected males often display signs of mild androgen insensitivity, including gynecomastia, reduced fertility, and testicular atrophy (96). Autopsies of affected individuals revealed diffuse lower motor neuron degeneration as well as loss of brainstem motor nuclei corresponding to the trigeminal, facial, and hypoglossal nerves (162).

In 1991, the SBMA disorder was colocalized to Xq12 with the androgen receptor (*AR*) gene (45). The *AR* gene contains a trinucleotide repeat within the first exon that encodes a polyglutamine tract. This normally polymorphic CAG repeat of 11–33 repeats undergoes expansion to 38–66 repeats in SBMA patients (16). Repeats within the normal range were meiotically stable in all transmissions, whereas instability was observed in 27% of transmissions of expanded CAG alleles (96). Both contractions and expansions of the *AR* repeat have been observed, with expansions occurring in >75% of cases (15, 96). Analysis of normal sperm has revealed a CAG repeat mutation rate of 6.7×10^{-3}; however, sperm of an affected individual (47 repeats) showed 60% expansions and 15% contractions (202, 203). Strong correlations between repeat length and age of onset (35, 75, 96) have been documented.

The effect of polyglutamine lengthening in the *AR* protein is currently

unknown. Although gynecomastia, testicular atrophy, and sterility as well as decreased androgen binding are consistent with androgen insensitivity, these patients are phenotypic males. This is in marked contrast to XY individuals with testicular feminization who are phenotypic females because of the absence or inactivity of the *AR* protein (108). Furthermore, the absence of SBMA-type neurodegenerative changes in individuals with testicular feminization or severe androgen insensitivity suggests a novel activity of the *AR* with an expanded polyglutamine tract that is not present in normals or individuals with null alleles. Since the *AR* is a transcription factor, one hypothesis is that polyglutamine expansion alters gene expression, thus leading to neuron degeneration (114). Altered transcriptional activation by *AR* proteins with expanded repeats has been found and interpreted to suggest that the polyglutamine expansion leads both to a partial loss-of-function, leading to incomplete androgen insensitivity, as well as a gain-of-function that is selectively motor neurotoxic (86, 114).

Spinocerebellar Ataxia, Type 1 (SCA1)

Spinocerebellar atrophy, type 1 (*SCA1*), is an autosomal dominant neurodegenerative disease characterized by progressive limb and gait ataxia, dysarthria, dysphagia, opthalmoparesis, muscle wasting, and peripheral neuropathy (143). Neuropathologic features include loss of Purkinje cells of the cerebellar cortex, atrophy of the brainstem, and neuronal loss in the cerebellum, brainstem, and spinocerebellar tracts (204). The range of onset can be from 4–70 years within a single family, with juvenile onset cases being observed in the later generations indicative of genetic anticipation (205).

Orr et al (127) identified a CAG repeat at the *SCA1* critical region on chromosome 6. Repeat sizes within the normal population ranged from 6–39 repeats, while repeats of 41–81 were observed in affected patients (112). An inverse correlation between repeat length and age of onset has been documented, as well as a strong direct correlation between repeat length and severity of disease (81, 134). All cases of juvenile onset *SCA1* occur through paternal transmissions of expanded repeats in excess of 58 repeats (127). Chung et al (27) reported that in 16 paternal transmissions observed, 63% resulted in CAG expansion in the next generation, whereas in 28 maternal transmissions, 69% of the CAG repeats either decreased or remained the same size.

The CAG repeat of *SCA1* is a cryptic repeat in the normal population, with two interrupting CAT triplets (CAG-repeat positions 13 and 15) being observed in 98% of normal alleles analyzed (27). However, no interruptions are present within the expanded alleles, suggesting that loss of the interrupting CAT triplets could account for the instability observed in *SCA1* families.

The *SCA1* gene is comprised of nine exons covering 450 kb of the human

genome (12). The transcript is 10,660 nucleotides in length with expression in all tissues analyzed (12, 127). A 2448-nucleotide open reading frame encodes a predicted 87-kD polypeptide of unknown function, referred to as ataxin-1, that contains a polymorphic polyglutamine tract interrupted by two histidine residues (corresponding to the CAT interruptions of the CAG repeat) (12). The remainder of the message is a 936-base 5′ untranslated region (UTR) and a uniquely large 3′ UTR of 7277 nucleotides (12). Immunohistochemical analyses detect both wild-type and mutant ataxin-1 in extracts of cultured cells from *SCA1* patients, indicating that proteins with expanded CAG repeats are expressed within the cell (151). Additionally, subcellular localization studies have revealed that ataxin-1 exhibits predominately nuclear staining in neuronal cells, cytoplasmic staining in skeletal muscle, and both cytoplasmic and nuclear staining in Purkinje cells (151). Burright et al (EN Burright, HB Clark, A Servadio, T Matilla, RM Feddersen, WS Yunis, LA Duvick, HY Zoghbi & HT Orr, submitted) have found that mice containing a transgene with 82 repeats in ataxin-1 experience Purkinje cell degeneration and ataxia, confirming the dominant nature of the expanded allele.

Huntington's Disease (HD)

Huntington's disease (HD) is an autosomal dominant movement disorder affecting approximately 1 in 10,000 Caucasians (65). The disorder is characterized by neuronal loss in the caudate and putamen, primarily of medium spiny neurons (58, 67). Characteristic features include progressive cognitive decline and chorea, which typically become manifest in midlife but which can occur anytime between ages 2 and 90 (58).

An unstable CAG repeat in HD patients, encoding a polyglutamine tract in the *IT15* gene, was reported in 1993 (57). Repeat lengths in affected individuals from numerous ethnic groups ranged from 36–121 (av 44–45), expanded from normal repeat lengths of 10–35 (av 18–19) (6, 36, 93, 160). Both repeat expansions and reductions were observed, with expansions being much more frequent in the paternal lineage, consistent with predominant paternal transmission in juvenile onset cases (3, 58, 160, 206). Strong correlations in age of onset or disease progression and repeat length have been obtained (6, 76, 105).

Repeat sizes within the range of 30–38 are exceedingly rare, (93, 135). The identification of multiple families with apparently sporadic HD in which an expanded HD chromosome was inherited from an unaffected individual with a repeat size of 30–38 in length has lead to speculation that an intermediate class of alleles analogous to the premutation alleles of fragile X syndrome may exist (53, 54, 118). Expansions of intermediate alleles resulting in de novo expression of HD have been confined exclusively to paternal transmissions

(53, 54, 118). Consistent with paternal instability, marked heterogeneity in repeat size has been shown in sperm of affected HD males (36, 106, 174).

The human HD gene (*IT15*) encompasses 180–200 kb and consists of 67 exons (5). *IT15* expression is found throughout the brain as well as in other somatic tissues (57, 101, 166). The 348-kD protein, referred to as huntingtin, is of unknown function and is predominately cytoplasmic (33, 72, 152, 175). Expression of both normal and expanded *IT15* transcripts and protein has been documented (84, 164), indicating an enlargement of the huntingtin protein by the expanded polyglutamine tract. Disruption of the murine Hdh gene produces embryonic lethality in homozygotes (37, 125).

Dentatorubral-Pallidoluysian Atrophy (DRPLA)

Autosomal dominant, dentatorubral-pallidoluysian atrophy (DRPLA) occurs with an estimated prevalence of 1 in 1 million in Japan (115), with rare occurrences in Europe and North America (132, 188, 189). DRPLA is named for the neuropathological findings of atrophy of the dentate nucleus of the cerebellum and its projections to the red nucleus (rubro) combined with atrophy of the lateral segment of the globus pallidus and its projections to subthalamic nuclei (of Luys) (121, 158, 172). DRPLA is similar to HD and is characterized by cerebellar ataxia, choreoathetosis, and dementia, although patients < 20 years old show progressive myoclonic epilepsy and mental retardation (92, 172).

Koide et al (92) and Nagafuchi et al (120), through screening of databases for CAG repeat-containing sequences, each independently identified a CAG-containing sequence, known as *B37*, originally identified by Li et al (100) and mapped to human chromosome 12, where DRPLA localized. CAG repeat sizes range from 7–25 pure repeats in the normal population, with expansion to 49–75 repeats in affected individuals (92, 120). Consistent with other trinucleotide repeat diseases, larger expanded repeats have been correlated with more severe phenotypes and/or earlier age of disease onset (92, 120). Somatic variation in CAG repeat length between various brain structures has recently been documented (178). Repeat expansions are more frequently observed in paternal transmissions (92, 120, 147).

The gene for DRPLA encodes a 4.5-kb transcript expressed in all tissues, with highest levels in brain, ovary, testis, and prostate (119). The predicted polypeptide is 1184 amino acids in length and contains a polyglutamine tract corresponding to the CAG repeat (119). Antibodies against the predicted DRPLA protein detect an enlarged protein product in the brains of DRPLA patients (197), consistent with the expanded polyglutamine tract.

HAW RIVER SYNDROME Haw River syndrome (HRS) is an autosomal dominant neurodegenerative disorder found in a single African-American family

from Haw River, North Carolina (41). Phenotypic manifestations and neuropathologic findings are similar to DRPLA, but HRS differs from DRPLA in that myoclonus is not observed in the HRS family (22). Burke et al (22) showed that the DRPLA trinucleotide repeat was expanded in members of this HRS family. Therefore, the same CAG repeat expansion accounted for both DRPLA and HRS, with 63–68 triplets found in HRS, similar to DRPLA (22). Although not an allelic variant of DRPLA, HRS may represent the influence of different alleles of a modifying gene(s).

A comparison of normal DRPLA repeat alleles in African-American, American-Caucasian, and Japanese populations has revealed that normal alleles in excess of 19 repeats are most common (7.4%) in the Japanese population, as compared to 1% and 0% for the African-American and Caucasian, respectively, implying that the presence of larger CAG alleles might account for the increased prevalence of disease in the Japanese and rarely in African-Americans populations (21).

Machado-Joseph Disease (MJD/SCA3)

The disease most recently added to the rapidly growing list of disorders caused by unstable trinucleotide repeats is Machado-Joseph disease (MJD, now also known as spinocerebellar ataxia type 3; *SCA3*). Initially described in individuals of Portuguese-Azorean descent (122, 144), MJD is now recognized as a common cause of spinocerebellar degeneration in individuals of most ethnic origins (85, 142). Degenerative changes have been noted in neurons of the substantia nigra, dentate nucleus of the cerebellum, neurons of the spinocerebellar and pyramidal tracts of the spinal cord, and peripheral neurons (142, 163). Clinically, these changes in both the central and peripheral nervous systems are observed as cerebellar ataxia, extrapyramidal features such as dystonia-rigidity reminiscent of parkinsonism, peripheral neuropathy, facial and lingual fasciculation, and external ophthalmoplegia (142, 163).

Kawaguchi et al (85) have recently identified a gene mapping to 14q32.1 within the MJD linkage interval that contains a CAG trinucleotide repeat observed to be unstable in MJD families. The gene, termed *MJD1*, encodes a predicted 359-amino acid protein of unknown function. The unstable CAG repeat ranges from 12–37 repeats in normal controls with expansion to 61–84 repeats in MJD patients (107, 111). Expansion of the CAG repeat within both maternal and paternal lineages has been found, with the greater increase in repeat length being noted in paternal transmissions (111). An interruption in the CAG repeat at the 5′ end of the repeat has been observed in all patients and controls, with all variation in repeat size occurring on the 3′ end of the repeat (85).

Myotonic Dystrophy (DM)

Autosomal dominant myotonic dystrophy (DM) is the most common form of adult muscular dystrophy (18). Clinical features of the DM disorder include myotonia, progressive muscular weakness and wasting, cardiac conduction abnormalities, and cataracts, with frontal balding and testicular atrophy in affected males (18). The phenotype is quite variable, ranging from only cataracts or balding in mildly affected individuals with adult-onset DM to severe hypotonia, muscular atrophy, and mental retardation in congenital DM (176). Genetic anticipation in DM families had been clearly documented with the milder phenotype occurring in earlier generations (64, 74).

DNA fragments in the region of human chromosome 19q13.3 were identified that displayed length variation in DM patients (18, 23, 50, 62, 109). This length variation was mapped to a normally polymorphic CTG trinucleotide repeat that exists as 5–37 copies in the normal Caucasian population, with expansions from 50 to >3000 in affected patients (18, 50, 99, 109). The extent of trinucleotide repeat expansion is in direct correlation with disease severity (18, 23, 99, 109). Individuals with intermediate-sized repeats, between 50–80, display very mild phenotypes, if any (20, 25). The distribution of normal CTG repeat lengths in the DM gene of Caucasians is trimodal, with peaks at 5 (35–40%), 11–17 (50%), and 19–30 (10%) (77, 201). The normal DM repeats in Southern African populations range from 5–22 repeats, and this reduction of large normal repeats has been correlated with marked decreases in DM prevalence in this population (55). In the Caucasian population, linkage disequilibrium between a nearby 1-kb *Alu* insertion/deletion polymorphism and DM has been established (77, 98), and it has been suggested that the class of 19–30 repeats might represent a pool of predisposed normal alleles (77).

Individuals displaying congenital DM and their mothers display larger repeat amplifications than noncongenital DM affecteds (63, 176). Although the overall trend in CTG instability is toward amplification, reductions in the expanded CTG repeat in paternal transmissions have been described (1, 9, 19, 30). Two "reverse" mutations have been reported in which expanded alleles in affected male individuals contracted to 19 and 24 repeats in phenotypically normal offspring (19).

Tissue-specific somatic mosaicism of CTG repeat length has been reported, in favor of a postzygotic window of trinucleotide repeat expansion in DM (7, 80, 116, 196, 200). Repeat sizes in muscle tissue exceed that in peripheral lymphocytes in DM patients (7, 200). Age-dependent mosaicism in DM affecteds has also been described, with increasing variation in somatic cell repeat size skewed toward expansion rather than contraction in DM individuals who have been followed over time (196, 200). Variation in repeat sizes in sperm included both expansions and contractions, with the tendency toward expan-

sion observed only in those individuals with small or intermediate-sized repeats (80). Contractions of the DM repeat in sperm into the normal range were observed, whereas expansions >1000 repeats have not been reported, accounting for both the occurrence of reverse mutations in paternal transmissions as well as the predominately maternal transmission of congenital cases of DM, respectively (80).

The aberrantly expanded CTG repeat of DM resides in the 3′ untranslated region of a serine/threonine protein kinase referred to as DM protein kinase (*DMPK*) (18, 50, 109). The DM gene consists of 15 exons spanning approximately 14 kb (110). The gene is expressed most highly in skeletal and cardiac muscle (79). Immunohistochemical studies have localized a 69-kD DM protein to the neuromuscular junction of skeletal muscle as well as to intercalated disks in cardiac muscle (179, 193). The effect of trinucleotide repeat expansion within the DM gene on expression of *DMPK* remains controversial. Although Carango et al (24) reported decreased expression from expanded alleles, conflicting reports of both decreased (48) as well as increased (146) expression of DM mRNA exist. Consistent with down-regulation of the abnormal DM allele, expanded CTG repeats of the DM gene have recently been demonstrated to be sites of preferential nucleosome assembly, with increasing repeat lengths directly correlated with more vigorous nucleosome assembly (186, 187). Altered chromatin structure associated with DM trinucleotide repeat expansion has also been documented, consistent with the apparently disproportionate levels of DM allele expression (128). More recently, unusual nuclear sequestration of DM messages with expanded CTG repeats has been reported from both muscle biopsies and fibroblasts of affected individuals (173). Wang et al (185) observed decreased accumulation of the entire poly A+ RNA pool in muscle biopsy specimens from DM affecteds, suggesting that the expanded CTG repeat could represent a *trans*-factor that has a dominant-negative effect on multiple cellular RNAs.

CONCLUSIONS

General Features of Expansion Mutations

While there are no obvious intrinsic reasons for trinucleotides to participate in expansions (compared, for example, to di- or tetranucleotides), the commonality of all 11 loci being CG-rich trinucleotide repeats certainly suggests a bias toward this type of repeat. Also without obvious explanation is the limitation of this form of mutation to the human species. However, studies of expansion-prone human loci in other species suggest a tendency in nonhumans toward smaller repeats or cryptic repeats, both of which would presumably stabilize the triplet repeat (10, 13, 39, 102). It has been proposed that these

smaller, cryptic repeats tend to increase in length and purity through evolution, suggesting that humans are intrinsically at higher risk for this form of mutation (39, 145). However, the failure of human transgenes containing long, previously unstable trinucleotide repeats to exhibit instability in the mouse (16) could indicate species differences in DNA replication/repair that influence repeat instability. Alternatively, such transgenic data could be explained by the absence of critical *cis*-acting elements from the cDNA constructs used. If certain *cis*-acting sequences are critical for expansion, the orthogenous loci may lack this critical genomic organization.

The expansion mutations characterized to date can be categorized by those undergoing short expansions (2–3 times normal repeat length) versus those that undergo long expansions (>10 normal repeat length). All coding CAG repeat loci fall within the former classification while those involving fragile sites and the myotonic dystrophy locus are long repeat expansions. However, the latter group do exhibit short expansion phenomena. For example, the premutation alleles at the *FMR1* locus behave similarly to the short expansions of the coding CAG repeat loci. Since evidence suggests that the long expansions are postmeiotic events, occurring in early embryogenesis (80, 116, 195), a congruent possibility is that all the loci exhibit a meiotic short expansion mutational mechanism, which at certain loci, may undergo continued expansion mitotically in the early embryo. This notion is attractive because it is consistent with data from both long and short expansion loci indicating that purity of the repeat (27, 95), promotes instability (seen particularly in pure tracts of >38 trinucleotides) (39), and it also accounts for the mosaicism common with the long expansions. Therefore, there may be a similar meiotic mechanism operating at all these loci where long perfect repeat tracts are inherently unstable. Mechanisms have been proposed suggesting that such reiterated tracts could lead to lagging strand instability, since 38 trinucleotides approximate the length of an Okazaki fragment (95, 137). Alternatively, it has been proposed that the trinucleotide repeats form unusual DNA structures that confer instability (26, 47).

Also consistent among most of these disorders is a parent-of-origin effect. While most of the CAG repeat loci show a clear paternal bias for expansion, fragile X syndrome and DM exhibit a maternal origin for large expansion. While the former is likely a meiotic event, the latter may involve a maternal imprint.

Consequences of Trinucleotide Repeat Expansion

The consequence of trinucleotide repeat expansion appears to be based upon the specific repeat, the length of the expansion, and the location of the repeat within the genome. The CGG/CCG repeats, which, unlike the CAG repeats,

contain a methylatable CpG dinucleotide, acquire heavy DNA methylation once the repeat length exceeds approximately 200 trinucleotides. This methylation appears to be in response to the long repeat since there is evidence that expansion precedes methylation (167) or is normally not observed at a given locus except in the presence of an expanded repeat (82, 124). How expanded CGG/CCG repeats acquire DNA methylation is unclear. However, long reiterated sequences may be viewed as foreign by the cell, which then may attempt to inactivate the repeat by methylation, similar to the acquisition of methylation by integrated viral genomes (34). Once the lengthy CGG/CCG repeats become methylated, they appear to replicate later in the cell cycle (61), which may therefore be the molecular basis of the cytogenetic fragile sites. The long, methylated CGG/CCG repeats also influence gene expression, as is the case with *FMR1*, and can therefore impart a loss-of-function phenotype if in the vicinity of a gene. There is evidence of yet uncharacterized human repeat loci that can frequently expand without apparent phenotypic consequence (149), suggesting, as have others (139), that many more examples of trinucleotide repeat expansion will be found.

The consequences of the CAG repeat expansions appear to be at the level of the protein. These expansions are thought to confer a gain or altered function rather than a loss-of-function. The polyglutamine tracts encoded by the normal CAG repeat, which double or triple in length in proteins encoded by the expanded alleles, have been speculated to become substrate for transglutaminases and cross-link to lysyl donor proteins (56), oligomerize with themselves or other polyglutamine-containing proteins (165), or alter transcriptional activation for the nuclear proteins (52). Irrespective of the ultimate mechanism, all coding CAG repeat disorders share a remarkable feature of neuronal degeneration and, therefore, whatever the change induced by the expanded polyglutamine, it likely will influence programmed cell death to some degree.

The consequence of the CTG expansion in myotonic dystrophy is less clear. RNA processing has generally been implicated, even in conflicting studies, and is somewhat appealing given the location of the repeat within the 3′ UTR and the emerging importance of 3′ UTRs in RNA processing (87). However, the presence of another gene near the *DMPK* repeat (79), which could also contribute to the phenotype, confounds the understanding of the consequence of the CTG expansion.

One feature of the trinucleotide repeat expansions that is clearly apparent is the presence of genetic anticipation in affected pedigrees (66). The correlation of repeat expansion with the decreasing age of onset in coding CAG expansions, with the increasing penetrance in fragile X syndrome, and with the increasing disease severity in myotonic dystrophy, is striking. All of these phenomena, which can be grouped as genetic anticipation, influence or modify to varying degrees the expected Mendelian inheritance patterns and had been

previously explained, when not dismissed outright as ascertainment errors (66), as the results of complex interactions of genes and environment. It is now apparent that a single mutant allele, varying from generation to generation or even among sibs, can explain, to a large degree, genetic anticipation. Thus, other disorders of complex inheritance, such as psychiatric disease (103), thought to result from numerous interacting genetic and environmental influences, can now be reconsidered as perhaps being the result of only a limited number of interacting influences; one possibility being a trinucleotide repeat expansion. Therefore, while it is remarkable that since 1991, 11 loci that mutate by this previously unprecedented trinucleotide repeat expansion have been found in humans, it will be of even greater interest to determine to what degree human variation can be attributed to this form of genome fluidity.

ACKNOWLEDGMENTS

We thank Kerstin Lindblad for reviewing the manuscript and Janelle Clark for assistance in preparation. STW is an investigator with the Howard Hughes Medical Institute.

Literature Cited

1. Abeliovich D, Lerer I, Pashut-Lavon I, Shmueli E, Raas-Rothschild A, Frydman M. 1993. Negative expansion of the myotonic dystrophy unstable sequence. *Am. J. Hum. Genet.* 52:1175–81
2. Abitbol M, Menini C, Delezoide A-L, Rhyner T, Vekemans M, Mallet J. 1993. Nucleus basalis magnolaris and hippocampus are the major sites of *FMR-1* expression in the human fetal brain. *Nat. Genet.* 4:147–52
3. Albin RL, Tagle DA. 1995. Genetics and molecular biology of Huntington's disease. *Trends Neurosci.* 18:11–14
4. Allingham-Hawkins DJ, Ray PN. 1995. *FRAXE* expansion is not a common etiological factor among developmentally delayed males. *Am. J. Hum. Genet.* 56:72–76
5. Ambrose CM, Duyao MP, Barnes G, Bates GP, Lin CS, et al. 1994. Structure and expression of the Huntington's disease gene: evidence against simple inactivation due to an expanded CAG repeat. *Somat. Cell Mol. Genet.* 20:27–38
6. Andrew SE, Goldberg YP, Kremer B, Telenius H, Theilmann J, et al. 1993. The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nat. Genet.* 4:398–403
7. Anvret M, Ahlberg G, Grandell U, Hedberg B, Johnson K, Edstrom L. 1993. Larger expansions of the CTG repeat in muscle compared to lymphocytes from patients with myotonic dystrophy. *Hum. Mol. Genet.* 2:1397–400
8. Armour JAL, Jeffreys AJ. 1992. Biology and applications of human minisatellite loci. *Curr. Opin. Genet. Dev.* 2:850–56
9. Ashizawa T, Anvret M, Baiget M, Barcelo JM, Brunner H, et al. 1994. Characteristics of intergenerational contractions of the CTG repeat in myotonic dystrophy. *Am. J. Hum. Genet.* 54:414–23
10. Ashley CT, Sutcliffe JS, Kunst CB, Leiner HA, Eichler EE, et al. 1993. Human and murine *FMR-1*: alternative splicing and translational initiation downstream of the CGG-repeat. *Nat. Genet.* 4:244–51
11. Ashley CT, Wilkinson KD, Reines D,

Warren ST. 1993. *FMR1* protein: conserved RNP family domains and selective RNA binding. *Science* 262:563–66
12. Banfi S, Servadio A, Chung M-y, Kwiatkowski TJ Jr, McCall AE, et al. 1994. Identification and characterization of the gene causing type 1 spinocerebellar ataxia. *Nat. Genet.* 7:513–20
13. Baxendale S, Abdulla S, Elgar G, Buck D, Berks M, et al. 1995. Comparative sequence analysis of the human and pufferfish Huntington's disease genes. *Nat. Genet.* 10:67–76
14. 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
15. Biancalana V, Serville F, Pommier J, Hanauer A, Mandel JL. 1992. Moderate instability of the trinucleotide repeat in spino bulbar muscular atrophy. *Hum. Mol. Genet.* 1:255–58
16. 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
17. Blake TJ, Shapiro M, Morse HC, Langdon WY. 1991. The sequence of the human and mouse *c-cbl* proto-oncogenes show *v-cbl* was generated by a large truncation encompassing a proline-rich domain and a leucine zipper-like motif. *Oncogene* 6:653–57
18. 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
19. Brunner HG, Jansen G, Nillesen W, Nelen MR, de Die CEM, et al. 1993. Reverse mutation in myotonic dystrophy. *N. Engl. J. Med.* 328:476–80
20. Brunner HG, Nillesen W, van Oost BA, Jansen G, Wieringa B, et al. 1992. Presymptomatic diagnosis of myotonic dystrophy. *J. Med. Genet.* 29:780–84
21. Burke JR, Ikeuchi T, Koide R, Tsuji S, Yamada M, et al. 1994. Dentatorubral-pallidoluysian atrophy and Haw River syndrome. *Lancet* 344:1711–12
22. Burke JR, Wingfield MS, Lewis KE, Roses AD, Lee JE, et al. 1994. The Haw River syndrome: dentatorubropallidoluysian atrophy in an African American family. *Nat. Genet.* 7:521–24
23. 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* 355:547–51
24. Carango P, Noble EN, Marks HG, Funanage VL. 1993. Absence of myotonic dystrophy protein kinase (*DMPK*) mRNA as a result of triplet repeat expansion in myotonic dystrophy. *Genomics* 18:340–48
25. Carey N, Johnson K, Nokelainen P, Peltonen L, Savontaus M-L, et al. 1994. Meiotic drive at the myotonic dystrophy locus. *Nat. Genet.* 6:117–18
26. Chen X, Mariappan SVS, Catasti P, Ratliff R, Moyzis RK, et al. 1995. Hairpens are formed by the single DNA strands of the fragile X triplet repeats: structure and biological implications. *Proc. Natl. Acad. Sci. USA* 92:5199–203
27. Chung M-y, Ranum LPW, Duvick LA, Servadio A, Zoghbi HY, Orr HT. 1993. Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. *Nat. Genet.* 5:254–58
28. Consortium D-BFX. 1994. *FMR1* knockout mice: a model to study fragile X mental retardation. *Cell* 78:23–33
29. De Boulle K, Verkerk AJMH, Reyniers E, Vits L, Hendrickx J, et al. 1993. A point mutation in the *FMR-1* gene associated with fragile X mental retardation. *Nat. Genet.* 3(1):31–35
30. de Die-Smulders CEM, Howeler CJ, Mirandolle JF, Brunner HG, Hovers V, et al. 1994. Anticipation resulting in elimination of the myotonic dystrophy gene: a follow up study of one extended family. *J. Med. Genet.* 31:595–601
31. Dennis N, Curtis G, Macpherson J, Jacobs P, Heitz D. 1992. Two families with Xq27.3 fragility, no detectable insert in the *FMR-1* gene, mild mental impairment and absence of Martin-Bell phenotype. *Am. J. Med. Genet.* 43:232–36
32. Devys D, Lutz Y, Rouyer N, Bellocq J-P, 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
33. 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
34. Doerfler W. 1992. DNA methylation: eukaryotic defense against the transcription of foreign genes? *Microb. Pathol.* 12:1–8
35. Doyu M, Sobue G, Mukai E, Kachi T, Yasuda T, et al. 1992. Severity of X-linked recessive bulbospinal neuro-

nopathy correlates with size of the tandem CAG repeat in androgen receptor gene. *Ann. Neurol.* 32:707–10
36. Duyao M, Ambrose C, Myers R, Novelletto A, Persichetti F, et al. 1993. Trinucleotide repeat length instability and age of onset in Huntington's disease. *Nat. Genet.* 4:387–92
37. Duyao MP, Auerbach AB, Ryan A, Persichetti F, Barnes GT, et al. 1995. Inactivation of the mouse Huntington's disease gene homolog Hdh. *Science* 269: 407–10
38. Eichler EE, Holden JJA, 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
39. Eichler EE, Kunst CB, Lugenbeel KA, Ryder OA, Davison D, et al. 1995. Evolution of the cryptic *FMR1* CGG repeat. *Nat. Genet.* In press
40. Eichler EE, Richards S, Gibbs RA, Nelson DL. 1993. Fine structure of the human *FMR1* gene. *Hum. Mol. Genet.* 2:1147–53
41. Farmer TW, Wingfield MS, Lynch SA, Vogel FS, Hulette C, et al. 1989. Ataxia, chorea, seizures, and dementia: pathologic features of a newly defined familial disorder. *Arch. Neurol.* 46:774–79
42. Feng Y, Lakkis L, Devys D, Warren ST. 1995. Quantitative comparison of *FMR1* gene expression in normal and premutation alleles. *Am. J. Hum. Genet.* 56:106–13
43. Feng Y, Zhang F, Lokey LK, Chastain JL, Lakkis L, et al. 1995. Translational suppression by trinucleotide repeat expansion at *FMR1*. *Science* 268:731–34
44. Fisch GS, Snow K, Thibodeau SN, Chalifaux M, Holden JJA, et al. 1995. The fragile X premutation in carriers and its effect on mutation size in offspring. *Am. J. Hum. Genet.* 56:1147–55
45. Fischbeck KH, Souders D, La Spada AR. 1991. A candidate gene for X-linked spinal muscular atrophy. *Adv. Neurol.* 56:209–14
46. Flynn GA, Hirst MC, Knight SJL, Macpherson JN, Barber JCK, et al. 1993. Identification of the *FRAXE* fragile site in two families ascertained for X-linked mental retardation. *J. Med. Genet.* 30:97–100
47. Fry M, Loeb LA. 1994. The fragile X syndrome d(CGG)n nucleotide repeats form a stable tetrahelical structure. *Proc. Natl. Acad. Sci. USA* 91:4950–54
48. Fu Y-H, 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
49. Fu Y-H, Kuhl DP, 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
50. Fu Y-H, Pizzuti A, Fenwock RG, King J, Rajnarayan S, et al. 1992. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* 255: 1256–58
51. Gedeon AK, Keinanen M, Ades LC, Kaarlainen H, Gecz J, et al. 1995. Overlapping submicroscopic deletions in Xq28 in two unrelated boys with developmental disorders: identification of a gene near *FRAXE*. *Am. J. Hum. Genet.* 56:907–14
52. 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
53. Goldberg YP, Andrew SE, Theilmann J, Kremer B, Squitieri F, et al. 1993. Familial predisposition to recurrent mutations causing Huntington's disease: genetic risk to sibs of sporadic cases. *J. Med. Genet.* 30:987–90
54. 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
55. Goldman A, Ramsay M, Jenkins T. 1994. Absence of myotonic dystrophy in southern African Negroids is associated with a significantly lower number of CTG trinucleotide repeats. *J. Med. Genet.* 31:37–40
56. Green H. 1993. Human genetic diseases due to codon reiteration; relationship to an evolutionary mechanism. *Cell* 74: 955–56
57. Group THDCR. 1993. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 72:971–83
58. Gusella JF, MacDonald ME, Ambrose CM, Duyao MP. 1993. Molecular genetics of Huntington's disease. *Arch. Neurol.* 50:1157–63
59. Hagerman RJ, Silverman AC, eds. 1991. *Fragile X Syndrome, Diagnosis, Treatment and Research.* Baltimore: Johns Hopkins Univ. Press
60. Hamel BCJ, Smits APT, de Graaff E, Smeets DFCM, Schoute F, et al. 1994. Segregation of *FRAXE* in a large family: clinical, psychometric, cytogenetic, and

molecular data. *Am. J. Hum. Genet.* 55:923–31
61. Hansen RS, Gartler SM, Scott CR, Chen S-H, Laird CD. 1992. Methylation analysis of CGG sites in the CpG island of the human *FMR1* gene. *Hum. Mol. Genet.* 1(8):571–78
62. 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
63. Harley HG, Rundle SA, MacMillan JC, Myring J, Brook JD, et al. 1993. Size of the unstable CTG repeat sequence in relation to phenotype and parental transmission in myotonic dystrophy. *Am. J. Hum. Genet.* 52:1164–74
64. Harper PS. 1989. Myotonic dystrophy and related disorders. In *Principles and Practice of Medical Genetics*, ed. A Emery, D Rimoin, 2:579–97. London: WB Saunders
65. Harper PS. 1991. The epidemiology of Huntington's disease. In *Huntington's Disease*, ed. PS Harper, pp. 251–80. London: Saunders
66. Harper PS, Harley HG, Reardon W, Shaw DJ. 1992. Anticipation in myotonic dystrophy: new light on an old problem. *Am. J. Hum. Genet.* 51:10–16
67. Hedreen JC, Peyser CE, Folstein SE, Ross CA. 1991. Neuronal loss in layers V and VI of cerebral cortex in Huntington's disease. *Neurosci. Lett.* 133: 257–61
68. Hergersberg M, Matsuo K, Gassman M, Schaffner W, Luscher B, et al. 1995. Tissue-specific expression of a *FMR1*/B-galactosidase fusion gene in transgenic mice. *Hum. Mol. Genet.* 4: 359–66
69. Hinds HL, Ashley CT, Nelson DL, Warren ST, Housman DE, Schalling M. 1993. Tissue specific expression of *FMR1* provides evidence for a functional role in fragile X syndrome. *Nat. Genet.* 3:36–43
70. Hirst MC, Barnicoat A, Flynn G, Wang Q, Daker M, et al. 1993. The identification of a third fragile site, *FRAXF*, in Xq27-Xq28 distal to both *FRAXA* and *FRAXE*. *Hum. Mol. Genet.* 2:197–200
71. Hirst MC, Grewal PK, Davies KE. 1994. Precursor arrays for triplet repeat expansion at the fragile X locus. *Hum. Mol. Genet.* 3:1553–60
72. Hoogeveen AT, Willemsen R, Meyer N, de Rooij KE, Roos RAC, et al. 1993. Characterization and localization of the Huntington disease gene product. *Hum. Mol. Genet.* 2:2069–73
73. Hornstra IK, Nelson DL, Warren ST, Yang TP. 1993. High resolution methylation analysis of the *FMR1* gene trinucleotide repeat region in fragile X syndrome. *Hum. Mol. Genet.* 2(10): 1659–65
74. Howeler C, Busch HFM, Geraedts JPM, Niermeijer MF, Staal A. 1989. Anticipation in myotonic dystrophy: fact or fiction. *Brain* 112:779–97
75. 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 clinical onset of features of spinal and bulbar muscular atrophy. *Neurology* 42: 2300–2
76. Illarioshkin SN, Igarashi S, Onodera O, Markova ED, Nikolskaya NN, et al. 1994. Trinucleotide repeat length and rate of progression of Huntington's disease. *Ann. Neurol.* 36:630–35
77. Imbert G, Kretz C, Johnson K, Mandel J-L. 1993. Origin of the expansion mutation in myotonic dystrophy. *Nat. Genet.* 4:72–76
78. Jacobsen P, Hauge M, Henningsen K, Hobolth N, Mikkelsen M, Philip J. 1973. An (11;21) translocation in four generations with chromosome 11 abnormalities in the offspring. *Hum. Hered.* 23: 568–85
79. Jansen G, Mahadevan M, Amemiya C, Workskamp N, Segers B, et al. 1992. Characterization of the myotonic dystrophy region predicts multiple protein isoform-encoding mRNAs. *Nat. Genet.* 1:261–66
80. 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 extreme expansion in sperm. *Am. J. Hum. Genet.* 54:575–85
81. Jodice C, Malaspina P, Persichetti F, Novelletto A, Sparado M, et al. 1994. Effect of trinucleotide repeat length and parental sex on phenotypic variation in spinocerebellar ataxia I. *Am. J. Hum. Genet.* 54:959–65
82. Jones C, Penny L, Mattina S, Yu E, Baker L, et al. 1995. The fragile site *FRA11B*, in protooncogene *CBL2*, is associated with a chromosome deletion syndrome. *Science*. In press
83. Jones C, Slijepcevic P, Marsh S, Baker E, Langdon WY, et al. 1994. Physical linkage of the fragile site *FRA11B* and a Jacobsen syndrome chromosome deletion breakpoint in 11q23.3. *Hum. Mol. Genet.* 3:2123–30
84. Jou Y-S, Myers RM. 1995. Evidence from antibody studies that the CAG

repeat in the Huntington's disease gene is expressed in the protein. *Hum. Mol. Genet.* 4:465–69
85. 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
86. Kazemi-Esfarjani P, Trifiro MA, Pinsky L. 1995. Evidence for a repressive function of the long polyglutamine tract in the human androgen receptor: possible pathogenetic relevance for the (CAG)n-expanded neuronopathies. *Hum. Mol. Genet.* 4:523–27
87. Keller W. 1995. No end yet to messenger RNA 3′ processing. *Cell* 81:829–32
88. 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
89. Khandjian EW, Fortin A, Thibodeau A, Tremblay S, Cote F, et al. 1995. A heterogeneous set of *FMR1* proteins is widely distributed in mouse tissues and is modulated in cell culture. *Hum. Mol. Genet.* 4:783–89
90. Knight SJL, Flannery AV, Hirst MC, Campbell L, Christodoulou Z, et al. 1993. Trinucleotide repeat amplification and hypermethylation of a CpG island in *FRAXE* mental retardation. *Cell* 74: 127–34
91. Knight SJL, Voelckel MA, Hirst MC, Flannery AV, Moncla A, Davies KE. 1994. Triplet repeat expansion at the *FRAXE* locus and X-linked mild mental handicap. *Am. J. Hum. Genet.* 55: 81–86
92. Koide R, Ikeuchi T, Onodera O, Tanaka H, Igarashi S, et al. 1994. Unstable expansion of CAG repeat in hereditary dentatorubral-pallidoluysian atrophy (DRPLA). *Nat. Genet.* 6:9–13
93. Kremer B, Goldberg P, Andrew SE, Theilmann J, Telenius H, et al. 1994. A worldwide study of the Huntington's disease mutation. *N. Engl. J. Med.* 330: 1401–6
94. 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(CCG)n. *Science* 252:1711–14
95. Kunst CB, Warren ST. 1994. Cryptic and polar variation of the fragile X repeat could result in predisposing normal alleles. *Cell* 77:853–61
96. La Spada AR, Roling DB, 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
97. 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
98. Lavedan C, Hofmann-Radvanyi H, Boileau C, Bonaiti-Pellie C, Savoy D, et al. 1994. French myotonic dystrophy families show expansion of a CTG repeat in complete linkage disequilibrium with an intragenic 1 kb insertion. *J. Med. Genet.* 31:33–36
99. Lavedan C, Hofmann-Radvanyi H, Rabes JP, Roume J, Junien C. 1993. Different sex-dependent constraints in CTG length variation as explanation for congenital myotonic dystrophy. *Lancet* 341: 236–37
100. Li S-H, McInnis MG, Margolis RL, Antonarakis SE, Ross CA. 1993. Novel triplet repeat containing genes in human brain: cloning, expression, and length polymorphisms. *Genomics* 16: 572–79
101. Li S-H, Schilling G, Young WS III, Li X-J, Margolis RL, et al. 1993. Huntington's disease gene (*IT15*) is widely expressed in human and rat tissues. *Neuron* 11:985–93
102. Lin B, Nasir J, MacDonald H, Hutchinson G, Graham RK, et al. 1994. Sequence of the murine huntington disease gene: evidence for conservation, and polymorphism in a triplet (CCG) repeat alternate splicing. *Hum. Mol. Genet.* 3: 85–92
103. Lindblad K, Nylander P-O, De Bruyn A, Sourey D, Zander C, et al. 1995. Detection of expanded CAG repeats in bipolar affective disorder using the repeat expansion detection (RED) method. *Neurobiol. Dis.* In press
104. Loesch DZ, Huggins R, Hay DA, Gedeon AK, Mulley JC, Sutherland GR. 1993. Genotype-phenotype relationships in fragile X syndrome: a family study. *Am. J. Hum. Genet.* 53:1064–73
105. Lucotte G, Turpin JC, Riess O, Epplen JT, Siedlaczk I, et al. 1995. Confidence intervals for predicted age of onset, given the size of (CAG)n repeat, in Huntington's disease. *Hum. Genet.* 95: 231–32
106. MacDonald ME, Barnes G, Srinidhi 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
107. Maciel P, Gaspar C, DeStefano AL, Silveira I, Coutinho P, et al. 1995. Correlation between CAG repeat length and

clinical features in Machado-Joseph disease. *Am. J. Hum. Genet.* 57:54–61
108. MacLean HE, Choi W-T, Rekaris G, Warne GL, Zalac JD. 1995. Abnormal androgen receptor binding affinity in subjects with Kennedy's disease (spinal and bulbar muscular atrophy). *J. Clin. Endocrinol. Metab.* 80:508–16
109. 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
110. Mahadevan MS, Amemiya C, Jansen G, Sabourin L, Baird S, et al. 1993. Structure and genomic sequence of the myotonic dystrophy (DM kinase) gene. *Hum. Mol. Genet.* 2:299–304
111. Maruyama H, Nakamura S, Matsuyama Z, Sakai T, Doyu M, et al. 1995. Molecular features of the CAG repeats and clinical manifestation of Machado-Joseph disease. *Hum. Mol. Genet.* 4: 807–12
112. 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
113. 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
114. 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
115. Miwa S. 1994. Triplet repeats strike again. *Nat. Genet.* 6:3–4
116. Monckton DG, Wong L-JC, Ashizawa T, Caskey CT. 1995. Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses. *Hum. Mol. Genet.* 4:1–8
117. Mulley JC, Yu S, Loesch DZ, Hay DA, Donnelly A, et al. 1995. *FRAXE* and mental retardation. *J. Med. Genet.* 32: 162–69
118. Myers RH, MacDonald ME, Korshetz 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
119. Nagafuchi S, Yanagisawa H, Oshaki E, Shirayama T, Tadokoro K, et al. 1994. Structure and expression of the gene responsible for the triplet repeat disorder, dentatorubral and pallidoluysian atrophy (DRPLA). *Nat. Genet.* 8:177–82
120. Nagafuchi S, Yanagisawa H, Sato K, Shirayama T, Ohsaki E, et al. 1994. Dentatorubral and pallidoluysian atrophy expansion of an unstable CAG trinucleotide on chromosome 12p. *Nat. Genet.* 6:14–18
121. Naito H, Oyanagi S. 1982. Familial myoclonus epilepsy and choreoathetosis; hereditary dentatorubral-pallidoluysian atrophy. *Neurology* 32: 798–807
122. Nakano KK, Dawson DM, Spence A. 1972. Machado-disease: a hereditary ataxia in Portuguese immigrants to Massachusetts. *Neurology* 22:49–55
123. Nancarrow JK, Holman K, Mangelsdorf M, Hori T, Denton M, et al. 1995. Molecular basis of p(CCG)n repeat instability at the *FRA16A* fragile site locus. *Hum. Mol. Genet.* 4:367–72
124. 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
125. Nasir J, Floresco SB, O'Kusky JR, Diewert VM, Richman JM, 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
126. Oberlé I, Rousseau F, Heitz D, Kretz C, Devys D, et al. 1991. Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science* 252:1097–102
127. 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
128. Otten AD, Tapscott SJ. 1995. Triplet repeat expansion in myotonic dystrophy alters the adjacent chromatin structure. *Proc. Natl. Acad. Sci. USA* 92:5465–69
129. Oudet C, Mornet E, Serre JL, Thomas F, Lentes-Zengerling SL, et al. 1993. Linkage disequilibrium between the fragile X mutation and two closely linked CA repeats suggests that fragile X chromosomes are derived from a small number of founder chromosomes. *Am. J. Hum. Genet.* 52:297–304
130. Parrish JE, Oostra BA, Verkerk AJMH, 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
131. 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
132. Potter NT, Meyer MA, Zimmerman AW, Eisenstadt ML, Anderson IJ. 1995. Molecular and clinical findings in a family with dentatorubral-pallidoluysian atrophy. *Ann. Neurol.* 37: 273–77
133. Quan F, Zonana J, Gunter K, Peterson KL, Magenis RE, Popovich BW. 1995. An atypical case of fragile X syndrome caused by a deletion that includes the *FMR1* gene. *Am. J. Hum. Genet.* 56: 1042–51
134. Ranum LPW, Chung M-y, Banfi S, Bryer A, Schut LJ, et al. 1994. Molecular and clinical correlations in spinocerebellar ataxia type I: evidence for familial effects on the age at onset. *Am. J. Hum. Genet.* 55:244–52
135. Read A. 1993. Huntington's disease: testing the test. *Nat. Genet.* 4:329–30
136. Reyniers E, Vits L, De Boulle K, 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
137. Richards RI. 1994. Simple repeat DNA is not replicated simply. *Nat. Genet.* 6:114–16
138. Richards RI, Holman K, Friend K, Kremer E, Hillen D, et al. 1992. Evidence of founder chromosomes in fragile X syndrome. *Nat. Genet.* 1:257–60
139. Riggins GJ, Lokey LK, Chastain JL, Leiner HA, Sherman SL, et al. 1992. Human genes containing polymorphic trinucleotide repeats. *Nat. Genet.* 2:186–91
140. 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
141. Romain DR, Chapman CJ. 1992. Fragile site Xq27.3 in a family without mental retardation. *Clin. Genet.* 41:33–35
142. Rosenberg RN. 1992. Machado-Joseph disease: an autosomal dominant motor system degeneration. *Mov. Disord.* 7: 193–203
143. Rosenberg RN. 1995. Autosomal dominant cerebellar phenotypes: the genotype has settled the issue. *Neurology* 45:1–5
144. Rosenberg RN, Nyhan WL, Bay C. 1976. Autosomal dominant striatonigral degeneration: a clinical, pathologic, and biochemical study of a new genetic disorder. *Neurology* 26:703–14
145. Rubinsztein DC, Amos W, Leggo J, Goodburn S, Ramesar RS, et al. 1994. Mutational bias provides a model for the evolution of Huntington's disease and predicts a general increase in disease prevalence. *Nat. Genet.* 7:525–30
146. Sabourin LA, Mahadevan MS, Narang M, Lee DSC, Surh LC, Korneluk RG. 1993. Effect of the myotonic dystrophy (DM) mutation on mRNA levels of the *DM* gene. *Nat. Genet.* 4(3):233–38
147. Sano A, Yamauchi N, Kakimoto Y, Komure O, Kawai J, et al. 1994. Anticipation in hereditary dentatorubral-pallidoluysian atrophy. *Hum. Genet.* 93: 699–702
148. Savage PD, Shapiro M, Langdon WY, Geurts van Kessel AD, Seuanez HN, et al. 1991. Relationship of the human protooncogene *CBL2* on 11q23 to the t(4;11), t(11;22), t(11;14) breakpoints. *Cytogenet. Cell Genet.* 56:112–15
149. Schalling M, Hudson TJ, Buetow KH, Housman DE. 1993. Direct detection of novel expanded trinucleotide repeats in the human genome. *Nat. Genet.* 4:135–39
150. Schinzel A, auf der Maur P, Moser H. 1977. Partial deletion of long arm of chromosome 11[del(11)(q23)]: Jacobsen syndrome. *J. Med. Genet.* 14:438–44
151. Servadio A, Koshy B, Armstrong D, Antalffy B, Orr HT, Zoghbi HT. 1995. Expression analysis of the ataxin-1 protein in tissues from normal and spinocerebellar ataxia type 1 individuals. *Nat. Genet.* 10:94–98
152. 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: 1065–74
153. Sherman SL, Jacobs PA, Morton NE, Froster-Iskenius U, Howard-Peebles PN, et al. 1985. Further segregation analysis of the fragile X syndrome with special reference to transmitting males. *Hum. Genet.* 69:289–99
154. Sherman SL, Morton NE, Jacobs PA, Turner G. 1984. The marker (X) syndrome: a cytogenetic and genetic analysis. *Ann. Hum. Genet.* 48:21–37
155. Siomi H, Choi M, Siomi NC, Nussbaum RL, Dreyfuss G. 1994. Essential role for the KH domains in RNA binding: impaired RNA binding by a mutation in the KH domain of *FMR1* causes fragile X syndrome. *Cell* 77:33–39
156. Siomi H, Matunis MJ, Michael WM, Dreyfuss G. 1993. The pre-mRNA bind-

ing K protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res.* 21(5):1193–98
157. Siomi H, Siomi MC, Nussbaum RL, Dreyfuss G. 1993. The protein product of the fragile X gene, *FMR1*, has characteristics of an RNA binding protein. *Cell* 74:291–98
158. Smith JK, Gonda VE, Malamud N. 1958. Combined dentato-rubral and pallido-Luysian degeneration. *Neurology* 8:205–9
159. Smits APT, Dreesen JCFM, Post JG, Smeets DFCM, de Die-Smulders C, et al. 1993. The fragile X syndrome: no evidence for any recent mutations. *J. Med. Genet.* 30:94–96
160. Snell RG, MacMillan JC, Cheadle JP, Fenton I, Lazarou LP, et al. 1993. Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. *Nat. Genet.* 4: 393–97
161. Snow K, Tester DJ, Kruckeberg KE, Schaid DJ, Thibodeau SN. 1994. Sequence analysis of the fragile X trinucleotide repeat: implications for the origin of the fragile X mutation. *Hum. Mol. Genet.* 3:1543–51
162. Sobue G, Hashizume Y, Mukai E, Hirayama M, Mitsuma T, Takahashi A. 1989. X-linked recessive bulbospinal neuronopathy. *Brain* 112:209–32
163. St. George-Hyslop P, Rogaeva E, Huterer J, Tsuda T, Santos J, et al. 1994. Machado-Joseph disease in pedigrees of Azorean descent is linked to chromosome 14. *Am. J. Hum. Genet.* 55:120–25
164. Stine OC, Li S-H, Pleasant N, Wagster MV, Hedreen JC, Ross CA. 1995. Expression of the mutant allele of *IT-15* (the *HD* gene) in striatum and cortex of Huntington's disease patients. *Hum. Mol. Genet.* 4:15–18
165. 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* 92:6509–13
166. 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 nonneuronal tissues. *Nat. Genet.* 5:259–65
167. 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
168. Sutherland GR, Baker E. 1992. Characterization of a new rare fragile site easily confused with the fragile X. *Hum. Mol. Genet.* 1:111–13
169. Sutherland GR, Gedeon A, Kornman L, Donnelly A, Byard RW, et al. 1991. Prenatal diagnosis of fragile X syndrome by direct detection of the unstable DNA sequence. *N. Engl. J. Med.* 325: 1720–22
170. Sutherland GR, Jacky PB, Baker E, Manuel A. 1983. Heritable fragile sites on human chromosomes. X. New folate-sensitive fragile sites: 6p23, 9p21, 9q32, 11q23. *Am. J. Hum. Genet.* 35: 432–37
171. Sutherland GR, Richards RI. 1995. Simple tandem DNA repeats and human genetic disease. *Proc. Natl. Acad. Sci. USA* 92:3636–41
172. Takahashi H, Ohama E, Naito H, Takeda S, Nakashima S, et al. 1988. Hereditary dentatorubral-pallidoluysian atrophy: clinical and pathological variants in a family. *Neurology* 38:1065–70
173. Taneja KL, McCurrach M, Schalling M, Housman D, Singer RH. 1995. Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J. Cell Biol.* 128:995–1002
174. Telenius H, Almqvist E, Kremer B, Spence N, Squitieri F, et al. 1995. Somatic mosaicism in sperm is associated with intergenerational (CAG)n changes in Huntington disease. *Hum. Mol. Genet.* 4:189–95
175. 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
176. 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
177. Tunnacliffe A, Mcguire RS. 1990. A physical linkage group in human chromosome band 11q23 covering a region implicated in leukocyte neoplasia. *Genomics* 8:447–53
178. Ueno S-i, Kondoh K, Kotani Y, Komure O, Kuno S, et al. 1995. Somatic mosaicism of CAG repeat in dentatorubral-pallidoluysian atrophy (DRPLA). *Hum. Mol. Genet.* 4:663–66
179. van der Ven PFM, Jansen G, van Kuppevelt THMSM, Perryman MB, Lupa M, et al. 1993. Myotonic dystrophy kinase is a component of neuromuscular junctions. *Hum. Mol. Genet.* 2: 1889–94

180. Verheij C, Bakker CE, deGraaff E, Keulemans J, Willemson R, et al. 1993. Characterization and localization of the *FMR-1* gene product associated with fragile X syndrome. *Nature* 363: 722–24
181. Verheij C, de Graaff E, Bakker CE, Willemson R, Willems PJ, et al. 1995. Characterization of *FMR1* proteins isolated from different tissues. *Hum. Mol. Genet.* 4:895–901
182. Verkerk AJMH, de Graaff E, De Boulle K, Eichler EE, Konecki DS, et al. 1993. Alternative splicing in the fragile X gene *FMR1*. *Hum. Mol. Genet.* 2(4): 399–404
183. Verkerk AJMH, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DPA, 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
184. Voullaire LE, Webb GC, Leversha MA. 1987. Chromosome deletion at 11q23 in an abnormal child from a family with inherited fragility at 11q23. *Hum. Genet.* 76:202–4
185. Wang J, Pegoraro E, Menegazzo E, Gennarelli M, Hoop RC, et al. 1995. Myotonic dystrophy: evidence for a possible dominant-negative RNA mutation. *Hum. Mol. Genet.* 4:599–606
186. Wang Y-H, Amirhaeri S, Kang S, Wells RD, Griffith JD. 1994. Preferential nucleosome assembly at DNA triplet repeats from the myotonic dystrophy gene. *Science* 265:669–71
187. Wang Y-H, Griffith J. 1995. Expanded CTG triplet blocks from the myotonic dystrophy gene create the strongest known natural nucleosome positioning elements. *Genomics* 25:570–73
188. Warner TT, Williams L, Harding AE. 1994. DRPLA in Europe. *Nat. Genet.* 6:225
189. Warner TT, Williams LD, Walker RWH, Flinter F, Robb SA, et al. 1995. A clinical and molecular genetic study of dentatorubropallidoluysian atrophy in four European families. *Ann. Neurol.* 37:452–59
190. Warren ST, Ashley CT. 1995. Triplet repeat expansion mutations: the example of fragile X syndrome. *Annu. Rev. Neurosci.* 18:77–99
191. Warren ST, Nelson DL. 1993. Trinucleotide repeat expansions in neurologic disease. *Curr. Opin. Neurobiol.* 3:752–59
192. Warren ST, Nelson DL. 1994. Advances in molecular analysis of fragile X syndrome. *J. Am. Med. Assoc.* 271:536–42
193. Whiting EJ, Waring JD, Tamai K, Somerville MJ, Hincke M, et al. 1995. Characterization of myotonic dystrophy kinase (*DMK*) protein in human and rodent muscle and central nervous tissue. *Hum. Mol. Genet.* 4:1063–72
194. Willems PJ, Van Roy B, De Boulle K, Vits L, Reyniers E, et al. 1992. Segregation of the fragile X mutation from an affected male to his normal daughter. *Hum. Mol. Genet.* 1(7):511–17
195. 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
196. Wong L-JC, Ashizawa T, Monckton DG, Caskey CT, Richards CS. 1995. Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent. *Am. J. Hum. Genet.* 56:114–22
197. Yazawa I, Nukina N, Hashida H, Goto J, Yamada M, Kanazawa I. 1995. Abnormal gene product identified in hereditary dentatorubral-pallidoluysian atrophy (DRPLA) brain. *Nat. Genet.* 10:99–103
198. Yu S, Pritchard M, Kremer E, Lynch M, Nancarrow J, et al. 1991. Fragile X genotype characterized by an unstable region of DNA. *Science* 252:1179–81
199. Yunis JJ, Soreng AL. 1984. Constitutive fragile sites and cancer. *Science* 226: 1199–204
200. Zatz M, Passos-Bueno MR, Cerqueira A, Marie SK, Vainzof M, Pavanello RCM. 1995. Analysis of the CTG repeat in skeletal muscle of young and adult myotonic dystrophy patients: when does the expansion occur. *Hum. Mol. Genet.* 4:401–6
201. Zerylnick C, Torroni A, Sherman SL, Warren ST. 1995. Normal variation at the myotonic dystrophy locus in global human populations. *Am. J. Hum. Genet.* 56:123–30
202. Zhang L, Fischbeck KH, Arnheim N. 1995. CAG repeat length variation in sperm from a patient with Kennedy's disease. *Hum. Mol. Genet.* 4:303–5
203. 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
204. Zoghbi HY. 1993. The spinocerebellar degenerations. In *Current Neurology*, ed. SH Appel, pp. 87–110. St. Louis: Mosby-Year Book
205. Zoghbi HY, Pollack MS, Lyons LA,

Ferrell RE, Daiger SP, Beaudet AL. 1988. Spinocerebellar ataxia: variable age of onset and linkage to human leukocyte antigen in a large kindred. *Ann. Neurol.* 25:580–84

206. Zuhlke C, Riess O, Bockel B, Lange H, Thies U. 1993. Mitotic stability and meiotic variability of the (CAG)n repeat in the Huntington disease gene. *Hum. Mol. Genet.* 2:2063–67

Annual Review of Genet.
Volume 29 (1995)

CONTENTS