# Genetics of Childhood Disorders: IX. Triplet Repeat Disorders

RUSSELL L. MARGOLIS, M.D., AND CHRISTOPHER A. ROSS, M.D., PH.D.

Trinucleotide, or triplet, repeats consist of 3 nucleotides consecutively repeated (e.g., CAG CAG CAG CAG) within a region of DNA. All possible combinations of nucleotides are known to exist as triplet repeats, although some are more common than others. These repeated sequences are found both within gene boundaries and in the large stretches of DNA that lie between genes. If the triplet repeats lie within a gene, they may be found within the flanking upstream promoter region, within exons, or within introns. If they lie within exons, they may be present in the sequence that will be translated into protein. In that case, the repeat encodes a series of identical amino acids. The triplets may also occur at the 5′ or 3′ untranslated portion of the transcript. The different regions in which triplet repeats may lie are summarized in Figure 1.

Thousands of trinucleotide repeats exist throughout the human genome. Many are the same length in all individuals, while others are of variable length. The variable, or polymorphic, repeats are almost always transmitted without change in length from one generation to the next. However, some do change in length when passed on, and when that occurs, the gene is often disrupted.

This mutational type, first discovered in 1991, was termed a dynamic or expansion mutation. Over the past 8 years, more than a dozen diseases caused by trinucleotide repeat expansions have been identified. The discovery of this class of mutations has led to great excitement in the field of genetics, since expansion mutations often have properties contrary to the conventional wisdom of mendelian genetics.

The currently known expansion mutation disorders fall into 3 general groups (Fig. 1). Eight diseases are caused by expanded CAG repeats encoding the amino acid glutamine. The best known and most thoroughly characterized member of this group is Huntington disease. The other members include spinocerebellar ataxia types 1, 2, 3, 6, and 7; dentatorubral-pallidoluysian atrophy; and spinal and bulbar muscular atrophy. The repeated expansion occurs in different genes for each of these 8 disorders. However, for this group of disorders, the number of repeats at which the expansion is sufficient to cause disease is similar, typically about 40 triplets.

Each disease within this first group is characterized by neurodegeneration in a particular set of cortical and subcortical brain regions. The pathology appears to result from neurotoxic properties of the excessively long stretches of glutamine residues encoded by the CAG repeat expansions. The clinical manifestations of these diseases, as exemplified by Huntington disease, may include abnormalities of voluntary and involuntary movement; dementia; affective, psychotic, or obsessive-compulsive symptoms; apathy; irritability; and other less specific personality changes.

A second group of illnesses caused by relatively short repeat expansions consists of mostly developmental disorders. The

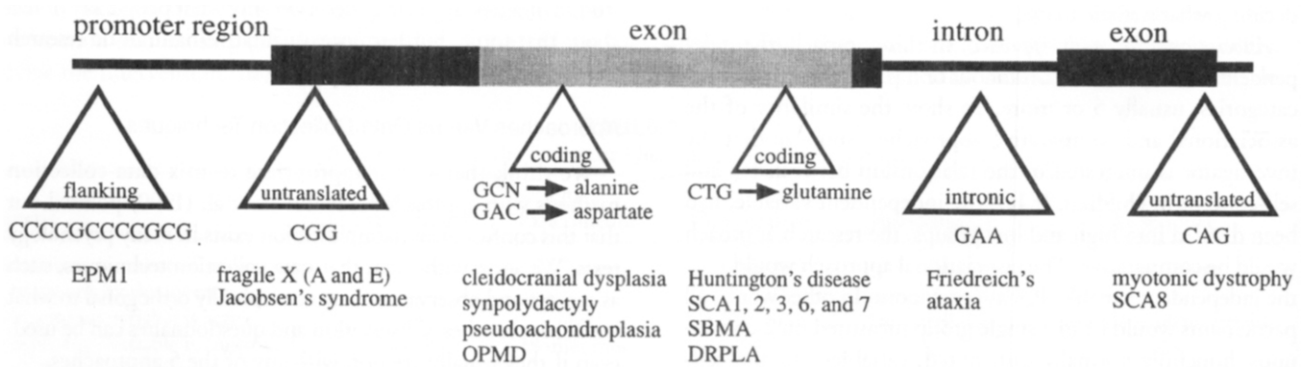


**Fig. 1** Genetic locations of repeat expansions. The relative locations of known repeat expansions are portrayed on a prototypical gene. Expansions have been discovered in all regions of a gene. They are found in the upstream flanking region, as in episodic myoclonic epilepsy. They may occur in exons or in introns. Those that are present within exons may occur within the upstream untranslated region, as is found in fragile X syndrome; they may occur within the protein coding region, as is the case with Huntington disease; or they may occur within the downstream untranslated region, as is seen with myotonic dystrophy. There is one example of an intronic repeat expansion: Friedreich ataxia. The size of the triangles reflects the size of the triplet repeat expansion. EPM1 = episodic myoclonic epilepsy type 1; OPMD = oculopharyngeal muscular dystrophy; SCA = spinocerebellar ataxia; SBMA = spinal and bulbar muscular atrophy; DRPLA = dentatorubral-pallidoluysian atrophy.

number of repeats in some of these illnesses is as few as 1 or 2 additional triplets, resulting in proteins with only a few more amino acid residues than is normally found. However, these additional extra amino acids are sufficient to disrupt the normal function of the protein. One example, synpolydactyly, is characterized by abnormal skeletal patterning. It is caused by 7 to 10 additional GCN triplets (in which N is any of the 4 base pairs). These additional codons will add the amino acid alanine into the protein sequence of the transcription factor HOXD13.

Cleidocranial dysplasia is another disorder of skeletal development that is caused by the presence of approximately 10 GCN triplets. In this case, 10 alanines are added to the protein CBFA1 (core-binding factor A1) that encodes a subunit of another regulatory transcription factor. Oculopharyngeal muscular dystrophy is caused by expansion of a GCG repeat that places additional alanine residues in the *PABP2* (polyA binding protein 2) gene. It is interesting that this disease can be either autosomal recessive or autosomal dominant, depending on the repeat length. Finally, the normal repeat of the cartilage oligomeric matrix protein contains 5 GAC triplets encoding the amino acid aspartate. If the repeat contracts by a single triplet or expands by 1 or 2 triplets, the result is 1 of 2 related developmental disorders of the skeleton, pseudoachondroplasia or multiple epiphyseal dysplasia.

A third and more heterogeneous group of disorders results from repeat expansions outside of the protein coding region. These expansions are characterized by their large size, in some cases repeating hundreds or even thousands of times. The prototypical disease of this type is fragile X syndrome, caused by the expansion of a CGG repeat in the 5′ untranslated region of the gene *FMR1* (fragile X mental retardation 1). There remains debate as to exactly how the large repeat leads to disease. One proposed mechanism is that the expansion becomes chemically modified through a process termed methylation and that this interferes with normal transcription machinery. The net effect of the large expansion is a decrease or absence of any transcription from the *FMR1* gene, with an accompanying decrease in functional protein.

The *FMR1* gene is on the X chromosome. Males with the mutation lack a second, normal copy of the gene. These children have a more severe phenotype than females who have a normal *FMR1* gene on their second X chromosome. Affected males typically have dysmorphic facial features, gonadal hypertrophy, mental retardation, and psychiatric symptoms, including some of the signs and symptoms typically seen in autism. Females with one normal and one expanded repeat have a milder phenotype and typically show milder cognitive, affective, and social difficulties. In both males and females, there is a rough correlation between the length of the repeat and the severity of the illness. A second triplet repeat disorder that causes mental retardation was recently discovered nearby on the X chromosome and is caused by the expansion of a GCC repeat in its 5′ untranslated region.

Several other disorders in this third group have expansions outside the protein-coding region. A 3′ untranslated CTG repeat causes myotonic dystrophy. With extreme expansions, myotonic dystrophy can be a life-threatening disorder with prominent cognitive impairment. Friedreich ataxia is a recessive disorder with childhood onset and multiple abnormalities including prominent cerebellar signs. It is usually the result of a long GAA repeat expansion in an intron of the gene *frataxin*. An expansion of a dodecamer repeat (CCCCGCCCCGCG) in the promoter region of the gene *cystatin B* is the most common genetic mutation causing episodic myoclonic epilepsy type 1. This illness is characterized by its childhood onset, seizures, myoclonus, dementia, and affective symptoms. Expansion of the 5′ untranslated CCG repeat in the proto-oncogene *CBL2* can result in deletion of the terminal arm of chromosome 11, giving rise to a contiguous gene syndrome with multiple developmental abnormalities. Spinocerebellar ataxia type 8, unlike the other spinocerebellar ataxias in which the repeat lies within the protein-coding region, results from an CTG expansion in the 3′ untranslated region.

The discovery of triplet repeat mutations solved a longstanding enigma in genetics. From early in this century, it was noticed that some diseases have an earlier age of onset or a more severe phenotype with each successive generation in an affected family. As there was no biological explanation for this phenomenon, known as anticipation, many considered it an artifact of subject ascertainment. Anticipation in some of these disorders can be dramatic. In myotonic dystrophy, for example, repeats just above the threshold for illness may result only in cataracts late in life. Within 2 or 3 generations, the progeny of such individuals may have remarkably long repeats, resulting in fatal congenital disease.

It is now clear that anticipation is not an artifact and can be explained by the unique properties of expansion mutations. The molecular basis for anticipation is that most repeats long enough to cause disease are unstable and have a tendency to get longer with each successive generation. The longer the expansions are, the earlier symptoms arise or the more severe these symptoms become. Two mechanisms have been proposed to explain these findings. In some disorders, the expansion produces more of a toxic gene product. This is likely to occur in Huntington and related diseases. For other repeat disorders, the expansion results in less transcription from the affected gene, and as a consequence, less of the functional protein is produced. Fragile X is an illness that exemplifies this type of mechanism. Finally, for some disorders, such as myotonic dystrophy, the expansion occurs in the 3′ untranslated region and the mechanisms to explain the increase in symptoms are yet to be discovered.

Other unusual genetic phenomena have been observed in triplet repeat disorders. All are attributable to the variability in repeat length that is found among affected individuals. Examples include monozygotic twins who are discordant for a particular illness, presumably as a consequence of postzygotic changes in repeat length. Others include skipped generations (a consequence of incomplete penetrance), disappearance of disease from a branch of a family (reversion of a repeat to a length below the disease threshold), and seemingly sporadic cases of a disease (expansion of a repeat from below to above disease threshold).

What is the relevance of repeat expansions to child and adolescent psychiatry? As was mentioned earlier, many of the known repeat expansion disorders present as congenital or childhood disorders that often have prominent neurological as well as psychiatric features. Fragile X, the most common familial cause of mental retardation, is certainly a striking example of this, as is juvenile Huntington disease. Understanding the unusual genetics of these disorders is an important factor in establishing an accurate diagnosis and prognosis and is essential in providing counseling to affected families. Both Huntington disease and fragile X syndrome will be reviewed in the next 2 columns in this *Journal.*

Repeat expansion disorders may also be of relevance for other psychiatric disorders. Most interestingly, anticipation has been detected in bipolar disorder and schizophrenia. In autism, anticipation per se has been less thoroughly addressed, but it appears that the parents of some affected children have milder symptoms, suggesting an increase in phenotypic severity over successive generations. Interpretation of these findings, however, remains difficult as the age-of-onset estimates are subject to multiple biases. Even if anticipation is present, it may be explained by several biological mechanisms in addition to repeat expansion. The question of whether or not trinucleotide expansions contribute to the genetic susceptibility to psychiatric disorders will be answered only after a systematic search for such mutations is conducted.

## WEB SITES OF INTEREST

http://www.familyvillage.wisc.edu/lib_frgx.htm
http://www.interlog.com/~rlaycock/what.html
http://www.ahsc.arizona.edu/~msrgsn/gd/gdvol10d.htm

## ADDITIONAL READINGS

Hagerman RJ, Cronister A (1996), *Fragile X Syndrome: Diagnosis, Treatment, and Research,* 2nd ed. Baltimore: Johns Hopkins University Press

McInnis MG (1996), Anticipation: an old idea in new genes. *Am J Hum Genet* 59:973–979

McInnis MG, Margolis RL (1998), Anticipation, triplet repeats, and psychiatric disorders. In: *Analysis of Triplet Repeat Disorders,* Rubinsztein DC, Hayden MR, eds. Oxford, England: Bios Scientific Press, pp 401–411

Ross CA (1995), When more is less: pathogenesis of glutamine repeat neurodegenerative diseases. *Neuron* 15:493–496

Ross CA, Margolis RL, Rosenblatt A, Ranen NG, Becher MW, Aylward E (1997), Huntington disease and the related disorder, dentatorubral-pallidoluysian atrophy (DRPLA). *Medicine (Baltimore)* 76:305–338

Verkerk A, Pieretti M, Sutcliffe J et al. (1991), Identification of gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 65:905–914

Wells RD, Warren ST (1998), *Genetic Instabilities and Hereditary Neurological Diseases.* San Diego: Academic Press

---

*Accepted June 15, 1999.*

*Dr. Margolis is Associate Professor, Department of Psychiatry, and Dr. Ross is Professor, Department of Psychiatry and Neuroscience, Johns Hopkins University School of Medicine, Baltimore.*

*Correspondence to Dr. Lombroso, Child Study Center, Yale University School of Medicine, 230 South Frontage Road, New Haven, CT 06520; e-mail: paul.lombroso@yale.edu.*

*To read all the columns in this series, visit the Web site at http://info.med.yale.edu/chldstdy/plomdevelop/*