Trinucleotide Repeat Expansion

The human genome contains tandem repeats of trinucleotides. Normally they occur in groups of 5–35 repeats. When their number exceeds a certain threshold and they occur in a gene or close to it, they cause diseases. Once the normal, variable length has expanded, the increased number of repeats tends to increase even further when passed through the germline or during mitosis. Thus, trinucleotide expansions form a class of unstable mutations, to date observed in humans only.

A. Different types of trinucleotide repeats and their expansions

Trinucleotide repeats can be distinguished according to their localization with respect to a gene. Expansions are greater outside genes and more moderate within coding regions. In several severe neurological diseases, abnormally expanded CAG repeats are part of the gene. CAG repeats encode a series of glutamines (polyglutamine tracts). Within a normal number of repeats, which varies according to the gene involved, the gene functions normally. However, an expanded number of repeats leads to an abnormal gene product with altered function. Trinucleotide repeats also occur in noncoding regions of a gene. Fairly common types are CGG and GCC repeats. The increase in the number of these repeats can be drastic, up to 1000 or more repeats. The first stages of expansion usually do not lead to clinical signs of a disease, but they do predispose to increased expansion of the repeat in the offspring of a carrier (premutation).
B. Unstable trinucleotide repeats in different diseases

Disorders due to expansion of trinucleotide repeats can be distinguished according to the type of trinucleotide repeat, i.e., the sequence of the three nucleotides, their location with respect to the gene involved, and their clinical features. All involve the central or the peripheral nervous system. Type I trinucleotide diseases are characterized by CAG trinucleotide expansion within the coding region of different genes. The triplet CAG codes for glutamine. About 20 CAG repeats occur normally in these genes, so that about 20 glutamines occur in the gene product. In the disease state the number of glutamines is greatly increased in the protein. Hence, they are collectively referred to as polyglutamine disorders. Type II trinucleotide diseases are characterized by expansion of CTG, GAA, GCC, or CGG trinucleotides within a noncoding region of the gene involved, either at the 5′ end (GCC in fragile X syndrome type A, FRAXA), at the 3′ end (CGG in FRAXE; CTG in myotonic dystrophy), or in an intron (GAA in Friedreich ataxia).
C. Principle of laboratory diagnosis of unstable trinucleotide repeats

The laboratory diagnosis compares the sizes of the trinucleotide repeats in the two alleles of the gene examined. One can distinguish very large expansions of repeats outside coding sequences (50 to more than 1000 repeats) and moderate expansion within coding sequences (20 to 100–200). The figure shows 11 lanes, each representing one person: normal controls in lanes 1–3; confirmed patients in lanes 4–6; and a family with an affected father (lane 7), an affected son (lane 10), the unaffected mother (lane 11), and two unaffected children: a son (lane 8) and a daughter (lane 9). Size markers are shown at the left. Each lane represents a polyacrylamide gel and the \((CAG)\_n\) repeat of the Huntington locus amplified by polymerase chain reaction shown as a band of defined size. Each person shows the two alleles. In the affected persons the band representing one allele lies above the threshold in the expanded region (in practice the bands are somewhat blurred because the exact repeat size varies in DNA from different cells).
Principle of laboratory diagnosis of unstable trinucleotide repeats leading to expansion