Identification and characterization of CGG triplet repeat containing genes: Potential candidates for neurological disorders.

Wufang Fan, Greg Lennon, Harvey W. Mohrenweiser, Evan E. Eichler.
Human Genome Center, Lawrence Livermore National Laboratory, Livermore, CA, 94551.

Expanding CGG trinucleotide repeats have been implicated in the molecular etiology of three neuropsychiatric disorders, namely; the Fragile X syndrome, FRAX(E) mental retardation and Jacobsen syndrome. In each case the molecular lesion was shown to result from instability and hyperexpansion of a CGG trinucleotide repeat located in the 5' UTR of the respective gene. In an effort to identify candidate disease genes which may be involved in such diseases, two approaches were undertaken to isolate and characterize additional CGG repeat-containing genes. 165,800 arrayed I.M.A.G.E clones representing ten different human tissue types were screened using end-labelled CGG repeat as probe. 832 positive clones (0.4% of all clones) were identified. Clone distribution among the various cDNA libraries revealed that libraries derived from adult brain tissue were enriched 5-10 fold for CGG -repeat containing transcripts (417/832 or 50% of all clones). Sequence information was available for nearly 1/3 of all clones (236/832 or 28%). Of these, 72 cDNA clones showed no homology upon BLAST sequence similarity searches of existing databases. 55/164 unique loci were identified which showed homology or identity with previously identified genes. Interestingly, a significant proportion of these unique loci (18/55) reveal homology to genes involved in transcription/translational regulation or RNA binding. In order to identify additional cDNA clones with potentially longer tracts of CGG repeats, a second approach using the GENETRAPPER TM cDNA Positive Selection System was employed. A significant enrichment (150-200X) of CGG-repeat containing cDNA clones was obtained from a SuperScriptTM Human brain cDNA library. Hybridization data indicates that approximately 2% of all clones were considered to contain CGG repetive tracts greater than 10 CGG repeats. Unique loci with long repeats from both the I.M.A.G.E collection and GENETRAPPER source are currently being sequenced and mapped. The results of this analysis should provide a useful resource of candidiate genes for consideration of other anticipatory neuropsychiatric disorders.
*Performed under auspices of the US DOE by LLNL; contract No. W-7405-ENG-48.