Growing triplet repeats

Sir — The discovery of expanding CAG and CGG repeats in several inherited psychomotor disorders has provided a molecular explanation of the phenomena of anticipation and other unusual inheritance patterns[1]. Trinucleotide expansions have been identified in seven disorders including Huntington's disease and at a locus on chromosome 18 (RED-1)[2]. All these diseases are associated with expansions of either CAG/CTG or CCG/CGG core sequences, but the disease mechanism is not uniform. At least three different mechanisms have been proposed[3]. We suggest that expansions in other trinucleotide motifs may also be important in the pathogenesis of inherited disorders. The human genome contains a high number of various short simple sequence repeats. Among tri–nucleotide repeats having more than 7 iterations, $(ATG)_n$, $(ATT)_n$, $(CCT)_n$, $(CGG)_n$, $(CTG)_n$, $(CTT)_n$, $(GTT)_n$, and $(TGG)_n$ have been reported in the EMBL database, and many are located within genes. The size distributions of these repeats are variable, with a few cases in each category reported to contain more than 13 repeats (Table 1).

To examine if other motifs might be expanded, we have used a method for the detection of expanded repeats directly from genomic DNA. This method, repeat expansion detection (RED)[2] has previously been used to detect CAG and CGG expansions in patient DNA as well as map the RED-1 locus[2]. Using this assay to screen 168–262 individuals for different trinucleotide motifs, we have identified four novel sequence expansions; ATG, CCT, CTT and TGG (Fig. 1). These expansions show a pattern consistent with a dominant mendelian mode of inheritance and expand or contract between generations in a fashion similar to the previously found CAG and CGG repeats. The largest ATG expansion is represented by a ligation product of 450 base pairs (bp), whereas for CCT, CTT and TGG the largest products observed are 273, 210 and 216 bp respectively (see Fig. 1). Expansion frequencies correlate reasonably well with the frequency of trinucleotide repeats in the EMBL database (Table 1). The frequency of CCT ligation products > 180 bp was 15%, similar to the size distribution of CAG repeats in the normal population[2]. ATG, CTT and TGG had expansions > 180 bp at a frequency of 1%, 10% and 7% respectively, and are also commonly reported (Table 1). We have also analysed AGT and CGT but have not found any expansions larger than 180 bp. Interestingly, the EMBL database has no entry with 7 or more iterations of these sequences in the EMBL database (Table 1). In addition we have tried studying the ATT and GTT motifs, but no ligation products were observed. This could be due to very low frequencies of longer repeats in these motifs. Alternatively, there may be technical difficulties associated with the high AT-content and the resulting decrease in T_m. (We currently lack a positive control to ensure that the RED reaction works for these motifs.)

Our observation that four additional trinucleotide motifs may expand makes it important to consider these repeat motifs in the search for genes responsible for diseases displaying anticipation. Variations in copy number of many different repeats, di- and tetra–nucleotides as well as octamers, have recently been suggested to be involved in disease mechanisms in human cancer and prion diseases[1]. This suggests that variations in repeat copy number could be of a more wides–pread biological importance and that multiple types of repeats should be considered in the search for molecular alterations leading to disease.

K. Lindblad
C. Zander
M. Schalling
Neurogenetics Unit,
Department of Molecular Medicine,
Karolinska Hospital, 171 76
Stockholm, Sweden

T. Hudson
Whitehead Institute,
Massachusetts Institute of
Technology, Cambridge,
Massachusetts 02139, USA

Table 1 Frequency of triplet repeats in the genome and expansions ≥ 180 bp detected using RED

Repeat	No. of reported sequences[a] 7–12 repeats	≥ 13 repeats	RED analysis Individuals screened	% individuals with repeats ≥180 bp
AGT	0	0	200	0
ATG	7	1	190	1
ATT	23	4	93	0[b]
CCT	18	1	262	15
CGT	0	0	182	0
CTG	57	6	168	20
CTT	7	2	172	10
GTT	18	1	187	0[b]
TGG	16	3	219	7

[a]EMBL database.
[b]Absence of single ligation product.

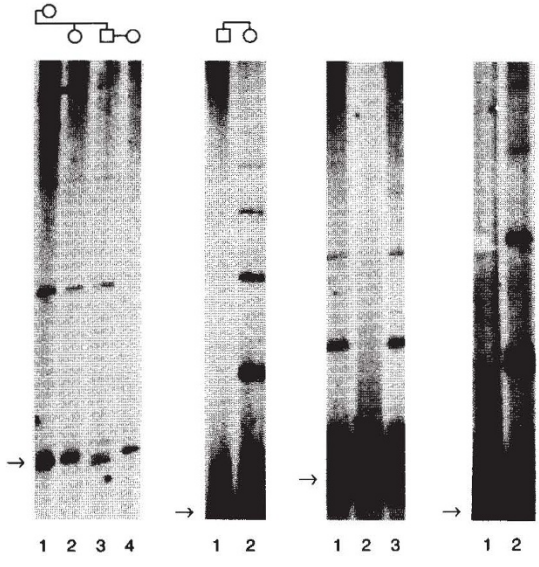

Fig. 1 Detection of ATG, CCT, CTT and TGG expansions in human genomic DNA using RED. A thermostable ligase is used to ligate repeat oligonucleotides following annealing at adjacent bases to repeat sequences in genomic DNA. Products are size separated on a denaturing polyacrylamide gel, electroblotted and hybridized to a complementary ^{32}P-labelled probe. The longest reaction product corresponds to the longest repeat size present in the genome. Reactions were run as described[2], with cycling conditions being 94 °C for 5 minutes followed by 396 cycles of either 80 °C for 30 s and 94 °C for 10 s for CCT_{13}, CTT_{14} and TGG_{12} or 75 °C for 30 s and 94 °C for 10 s for ATG_{15}. *a*, A family showing an expanded CTT repeat allele represented by formation of ligation products up to 168 bp. The arrow denotes a single ligation at 84 bp. *b*, A brother and a sister, where the sister has an ATG expansion represented by ligation products up to 450 bp. The arrow denotes a single ligation at 90 bp. *c*, Three distantly related individuals, where two individuals have an expanded CCT repeat represented by ligation products up to 234 bp. The arrow denotes a single ligation at 78 bp. *d*, Two individuals where one has a TGG expansion represented by ligation products up to 216 bp. The arrow denotes a single ligation at 72 bp.

1. Richards, R. I. & Sutherland, G. R. *Nature Genet.* **6**, 114–116 (1994).
2. Schalling, M., Hudson, T.J., Buetow, K. H. & Housman, D. E. *Nature Genet.* **4**, 135–139 (1993).
3. Warren, S.T. & Nelson, D.L. *Curr. Op. Neurobiol.* **3**, 752–759 (1993).