

Short reports

Sequence analysis of the CCG polymorphic region adjacent to the CAG triplet repeat of the HD gene in normal and HD chromosomes

Christophe Pêcheux, Jean-François Mouret, Alexandra Dürr, Yves Agid, Josué Feingold, Alexis Brice, Catherine Dodé, Jean-Claude Kaplan

Abstract

The CAG expansion responsible for Huntington's disease (HD) is followed by an adjacent polymorphic CCG repeat region which may interfere with a PCR based diagnosis. We have sequenced this region in 52 unrelated HD patients, from both normal and HD chromosomes. Fifty percent of the normal alleles were (CCG)₇(CCT)₂, 48% (CCG)₁₀(CCT)₂, and 2% (CCG)₇(CCT)₃. In contrast (CCG)₇(CCT)₂ was found in 85% of the HD alleles which represents significant linkage disequilibrium with the HD mutation.

(J Med Genet 1995;32:399-400)

Huntington's disease (HD)¹ is associated with the expansion of a CAG triplet repeat (37 to 140) in the 5' coding region of the IT15 transcript.² The CAG repeat region is also polymorphic in normal alleles (11 to 34 repeats). In the first PCR assay described to estimate the number of CAG repeats,² primers encompassing both the CAG and a 3' adjacent CCG repeat region were used. However, PCR assays have recently shown that the adjacent CCG rich region is also polymorphic, and five alleles of 170, 176, 179, 182, and 185 bp have been reported.^{3,4} The 176 bp allele is the most frequently observed in normal subjects and HD patients and corresponds to the sequence originally published.² We have performed sequence analysis of the polymorphic region adjacent to the CAG repeat on both chromosomes in 52 unrelated random HD patients. This was possible because the normal and HD alleles are easily distinguished after electrophoresis of the PCR products encompassing the whole region. The PCR amplification was performed

using primers and experimental conditions which have been previously described.^{2,5} The normal and abnormal alleles from heterozygous HD patients were then separated on a 1.5% low melting agarose gel and directly sequenced either automatically⁵ or manually. Polymorphisms both in the CCG triplet number ((CCG)₇ or (CCG)₁₀) and in the CCT triplet number ((CCT)₂ or (CCT)₃) adjacent to the CCG repeat were observed. In normal chromosomes, two major alleles (CCG)₇(CCT)₂ (allele 1) and (CCG)₁₀(CCT)₂ (allele 3) were found (figure, table 1) corresponding respectively to the 176 bp and the 185 bp fragments observed in the previous studies.^{3,4} In contrast, the HD chromosomes carried allele 1 in 85% of the cases confirming the linkage disequilibrium already described ($\chi^2=13$, $p<0.001$).^{3,6} However, in contrast to Barron *et al*,⁴ we found that allele 3 can be observed together with the expanded CAG repeat that causes HD (11% of the cases in our series). Allele 2 is characterised by both (CCG)₇(CCT)₃ and a G→A transition in the penultimate CAG of the uninterrupted CAG stretch (table 1).

Since HD is likely to result from the abnormal size of the Gln stretch encoded by both the CAG and CAA codons, it is noteworthy that considering only the number of the uninterrupted CAG repeat in allele 2 leads to a four residue underestimation of the Gln stretch. Alleles 1, 2, and 3 were observed with the same frequencies in females and males. Table 2 shows the distribution of alleles 1, 2, and 3 in relation to the size of the CAG repeat in normal and HD chromosomes. We have compared, in non-HD chromosomes, the mean number of CAG repeats associated with allele 1 (19.9) and allele 3 (17.05) using the *t* test. The difference is significant ($p<0.01$). In con-

Laboratoire de Biochimie Génétique, CHU Cochin Port-Royal, Université René Descartes Paris V, 123 Boulevard Port-Royal, 75014 Paris, France
C Pêcheux
C Dodé
J-C Kaplan

INSERM U289, Hôpital de la Salpêtrière, 47 Boulevard de l'Hôpital, 75651 Paris Cedex 13, France
A Dürr
Y Agid
A Brice

Applied Biosystems, BP 50086, 13 Rue de la Perdrix, 95948 Roissy Cedex, France
J-F Mouret

INSERM U155, Château de Longchamp, Bois de Boulogne, 75016 Paris, France
J Feingold

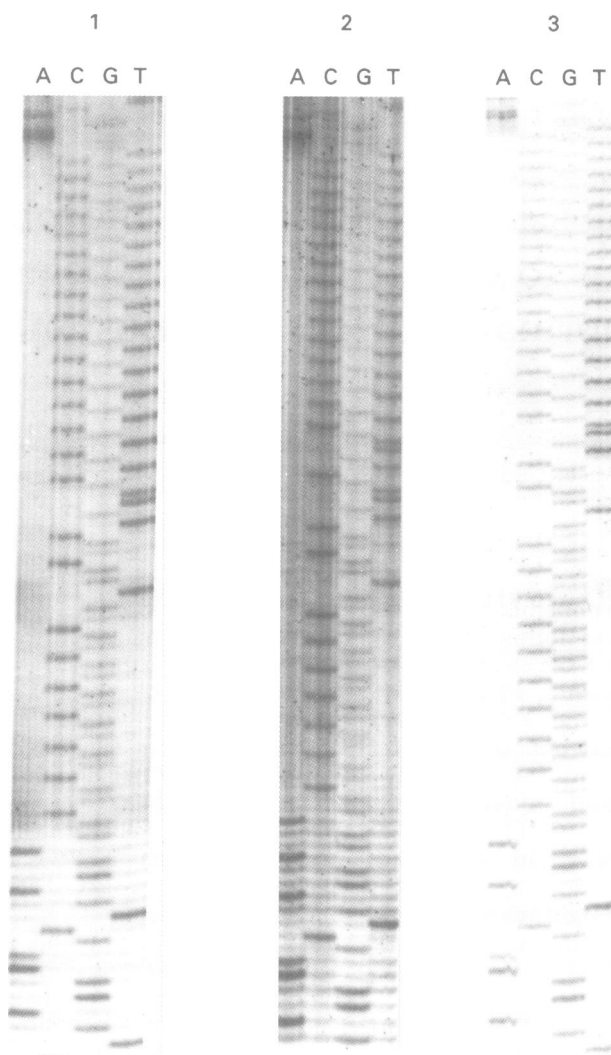
Correspondence to: Dr Dodé.

Received 23 September 1994
Revised version accepted for publication 2 December 1994

Table 1 Allele frequencies for polymorphisms in the CCG rich region on HD and non-HD chromosomes

Allele		Normal alleles		HD alleles	
		No	%	No	%
1	(CAG) _n CAA CAG CCG CCA (CCG) ₇ (CCT) ₂ CAG CTT	21	50	44	85
2	(CAG) _n CAA CAG CAA CAG CCG CCA (CCG) ₇ (CCT) ₃ CAG CTT	1	2	2	4
3	(CAG) _n CAA CAG CCG CCA (CCG) ₁₀ (CCT) ₂ CAG CTT	20	48	6	11

Allele



Direct DNA sequencing of the CCG rich region of the purified alleles 1, 2, and 3. The amplified DNA was electrophoresed on a 1.5% Nusieve agarose gel. The band containing the relevant DNA was cut out and heated at 65°C for 10 minutes with the same volume of water. DNA was extracted successively by phenol and chloroform and then ethanol precipitated and resuspended in 10 µl of Tris 10 mmol/l, EDTA 1 mmol/l. Sequencing reactions were performed using the sequenase kit version 2 (United States Biochemical Corporation) with the following modifications. Denaturation was done by heating 5 µl of amplified DNA (estimated to be equivalent to 2 pmol/l of single stranded DNA), 10% dimethylsulphoxide, and 2 pmol/l of primer 5' AACTCAGGGTCGGTGACG CCGCTCCTCAG 3' in a boiling water bath for 10 minutes. Annealing was then performed by direct chilling in dry ice. Both the labelling reaction with (³⁵S)dATP (37 TBq/mmol) and extension by T7 polymerase were done as recommended by the manufacturer. Sequencing gels (6% acrylamide/8 mol/l urea) were run at 1.5 kV.

trast the difference observed on HD chromosomes is not significant. In our series we did not observe any difference in the age of onset of the disease between allele 1 and allele 3 for a given number of CAGs. Our results establish that not only the CCG but also the CCT triplet number account for the size vari-

Table 2 Distribution of alleles 1, 2, and 3 in HD and non-HD chromosomes compared to the CAG repeat number

No of CAG	Allele 1	Allele 2	Allele 3
<i>Normal alleles</i>			
10			1
15	2		5
16	1		1
17	3	1	3
18	4		5
19	1		2
20	3		2
21	1		1
22	1		
23	1		
24	2		
27	1		
29	1		
<i>HD alleles</i>			
37	2		
39	6		
40	3		
41	5	1	
42	4		1
43	1		
44	5	1	1
45	2		
46	4		1
47	1		
48			1
49	1		
50	2		
53	1		
54	1		
57	1		
60	2		
65			2
70	1		
80	1		
140	1		

ation of the CCG rich region in both normal and HD chromosomes. As the sequence based method described here does not require a family study, it is more reliable than previously published protocols.³

We thank Assistance Publique, Hôpitaux de Paris for financial support.

- 1 Harper PS. *Huntington's disease*. London: WB Saunders, 1991.
- 2 The Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 1993;72:971-83.
- 3 Rubinsztein DC, Barton DE, Davison BCC, Ferguson-Smith MA. Analysis of the huntingtin gene reveals a trinucleotide length polymorphism in the region of the gene that contains two CCG rich stretches and a correlation between decreased age of onset of Huntington's disease and CAG repeat number. *Hum Mol Genet* 1993;10:1713-15.
- 4 Barron LH, Rae A, Holloway S, Brock DJH, Warner JP. A single allele from the polymorphic CCG rich sequence immediately 3' to the unstable CAG trinucleotide in the IT15 cDNA shows almost complete disequilibrium with Huntington's disease chromosomes in the Scottish population. *Hum Mol Genet* 1994;3:173-5.
- 5 Dodé C, Dürr A, Pêcheux C, et al. Huntington's disease in French families: CAG repeat expansion and linkage disequilibrium analysis. *CR Acad Sci Paris* 1993;316:1374-80.
- 6 Andrew SE, Goldberg YP, Theilmann J, Zeisler J, Hayden MR. A CCG repeat polymorphism adjacent to the CAG repeat in the Huntington disease gene: implications for diagnostic accuracy and predictive testing. *Hum Mol Genet* 1994;3:65-7.