Original Paper

Eur J Hum Genet 1994;2:44–50

Limited Expansion of the (CAG)n Repeat of the Huntington Gene: A Premutation (?)

Eric Legius[a]
Harry Cuppens[a]
Herman Dierick[a]
Karl Van Zandt[a]
René Dom[b]
Jean-Pierre Fryns[a]
Gerry Evers-Kiebooms[a]
Marleen Decruyenaere[a]
Koen Demyttenaere[c]
Peter Marynen[a]
Jean-Jacques Cassiman[a]

[a] Center for Human Genetics, University of Leuven, and Departments of
[b] Neurology and
[c] Psychiatry, University Hospital Gasthuisberg, Leuven, Belgium

Key Words
Huntington disease
Unstable repeat
Premutation
Autosomal dominant disorder
Presymptomatic testing
DNA diagnosis
Chromosome 4

Abstract

Huntington's disease (HD) is an autosomal dominant disorder with choreic movements, psychiatric manifestations and cognitive dysfunction. Recently the IT15 gene on chromosome 4p has been identified containing an unstable and expanded trinucleotide repeat in patients with HD. We report on the characteristics of this repeat in 248 individuals from 41 Belgian HD families. The length of the expanded repeat was defined precisely and reproducibly on an ALF sequencer and correlated well with the age of onset ($r = -0.72$). Paternal transmission of the expanded repeat resulted on average in a significantly longer repeat length (+2.79 repeats) than maternal transmission (–0.29 repeats). (CAG)n repeat of a premutation (?) size was observed in this population with subsequent expansion in the disease range. Presymptomatic or prenatal testing using only linked markers may be problematic in these cases.

Introduction

Huntington's disease (HD) is an autosomal dominant neurodegenerative disorder characterized by presenile dementia, motor disturbance and psychiatric manifestations. It affects about 1 in 10,000 individuals of European origin and symptoms usually appear in the fourth to fifth decade [1, 2]. There is good correlation for age of onset between individuals from the same family, but paternal transmission of the disorder is often associated with earlier onset of symptoms [3]. The genetic defect causing HD has been mapped to chromosome band 4p16.3 by linkage analysis, making presymptomatic testing for this disorder possible in an affected family setting [4, 5]. Recently a new gene, IT15, was isolated

Received: June 23, 1993
Revision received: December 6, 1993
Accepted: December 13, 1993

Prof. J.J. Cassiman
Center for Human Genetics
Campus Gasthuisberg
Herestraat 49
B–3000 Leuven (Belgium)

1018–4813/94/
0021–0044$5.00/0

containing an expanded and unstable trinucleotide repeat ((CAG)n) in HD individuals [6]. We report on the expansion and transmission of this repeat in 248 individuals from 41 families. This new information will be important for presymptomatic testing and prenatal diagnosis.

Patients and Methods

Clinical information and genomic DNA were obtained from 248 individuals belonging to 41 families with HD as part of a presymptomatic testing programme at our center. From this group a total of 103 individuals were diagnosed with HD, and 38 asymptomatic individuals had entered the presymptomatic testing program. Age of onset information for HD was collected as part of a retrospective study and we estimate the maximum error for age of onset to be 2 years. All individuals concerned by this study had consented at the time of blood sampling to make this DNA also available for research in the future, under the condition that the results of such tests would not be made available to them. To protect the rights of the individuals (patients and controls) all samples were coded and provided for research only with the essential clinical information. As a result a new counselling session and a new blood sample will have to be taken from any individual who wants to confirm the result of the predictive testing by this new assay.

Polymerase chain reaction (PCR) amplification of the (CAG)n repeat in IT15 was performed on each sample using genomic primers [6], one of which was conjugated with fluorescein isothiocyanate (FITC). A 50 µl PCR reaction consisted of 500 ng genomic DNA, 10 pmol of each primer, 200 µ*M* each dNTP, 1.2 m*M* $MgCl_2$, 10% dimethylsulfoxide (DMSO), 16.6 m*M* $(NH_4)_2SO_4$, 67 m*M* Tris-HCl (pH 8.8), 10 m*M* β-mercaptoethanol, 6.7 µ*M* EDTA, 0.85 µg BSA and 1.25 units of Taq polymerase (Cetus). A total of 40 cycles were performed at the following temperatures: 95 °C for 1 min, 60 °C for 1 min and 72 °C for 2 min after an initial denaturation of 3 min at 95 °C. Denatured PCR products were electrophoresed on an automatic sequencing device (ALF, Pharmacia), sized according to a reference sequence, and the corresponding length of the (CAG)n repeat was calculated, knowing that a PCR fragment of 247 bp contains 18 CAG repeats [6]. DNA samples from the 8 individuals from table 3 were also analyzed with a different primer set immediately flanking the (CAG)n repeat (sense primer: 5′-ATG AAG GCC TTC GAG TCC CTC AAG TCC TTC-3′ and anti-sense primer: 5′-CGG CGG CGG CGG TGG CGG CTG T-3′.

Table 1. The different observed (CAG)n alleles in the IT15 gene are shown for a group of 65 founders without HD

(CAG)n	Number	Frequency	SE
12	2	0.0154	0.0108
13	1	0.0077	0.0077
15	5	0.0385	0.0169
16	3	0.0231	0.0132
17	12	0.0923	0.0254
18	16	0.1231	0.0288
19	15	0.1154	0.0280
20	14	0.1077	0.0272
21	20	0.1538	0.0316
22	21	0.1615	0.0323
23	5	0.0385	0.0169
24	5	0.0385	0.0169
25	3	0.0231	0.0132
26	3	0.0231	0.0132
27	5	0.0385	0.0169

The first column shows the number (n) of (CAG)n repeats, the second column the number of times a certain repeat length was observed, the third column lists the allele frequencies and the fourth column the standard error (SE) for the respective allele frequencies. Allele frequencies and standard errors were calculated with the USERM1 module of the MENDEL computer programme, and represent maximum likelihood estimates [13].

Results

In a sample of 65 founders (unrelated spouses), without HD, we observed 15 different alleles (table 1) with a heterozygote fraction of 0.893 and a polymorphism information content (PIC) value of 0.883. Affected individuals from every HD family showed an increase in (CAG)n repeat length ranging

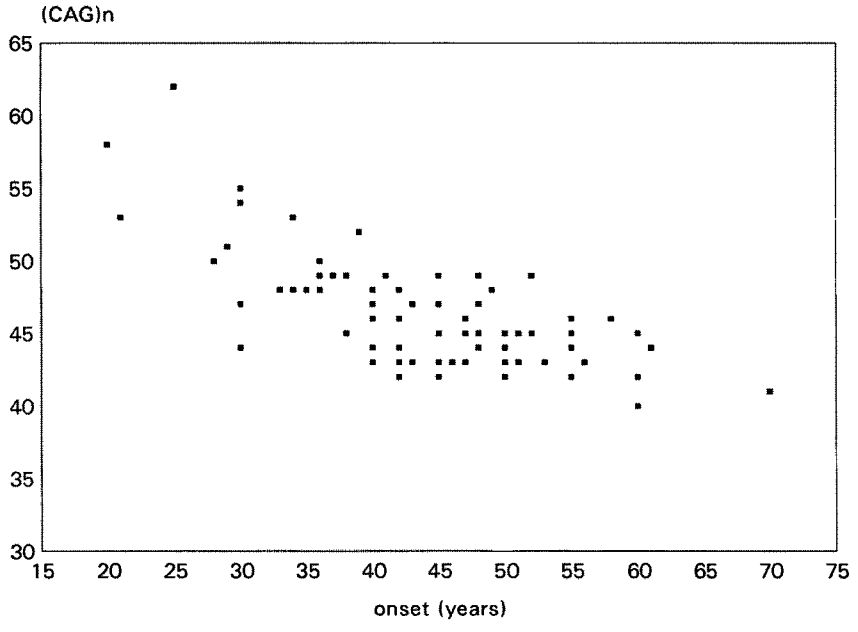

Fig. 1. Relation between age of onset of symptoms and size of expanded (CAG)n repeat in 84 individuals with HD. Regression: Y = 124 – 1.73X, where Y is age of onset in years and X is the number of CAG repeats.

CI

from 42 to 62 on the HD allele (102 HD alleles).

The length of the expanded (CAG)n repeat in IT15 correlated with age of onset of HD (fig. 1) for a group of 84 HD patients, with Pearson's correlation coefficient r = –0.72 (95% CI $-0.81 < r < -0.60$) and $r^2 = 0.49$. We did not find an additional correlation between age of onset and (CAG)n repeat length on the normal IT15 allele. For individuals with a repeat unit length in the 42 to 50 size range the correlation was lower (r = –0.46). Results were available for a total of 21 parent-to-child transmissions of an expanded (CAG)n repeat (table 2a). Paternal transmission resulted in a significantly larger average repeat length in offspring of 2.79 repeats (SE 0.74) as compared with an average decrease of 0.29 repeats in offspring (SE 0.36) in case of maternal transmission (Wilcoxon two sample test, p = 0.0018). For several pedigrees no DNA was available from the transmitting parent. For those families we examined the maximum difference in repeat length between sibs from the same sibship (table 2b). Paternal transmission resulted in a significantly larger mean difference in repeat length per sibship than maternal transmission (table 2b, Wilcoxon two sample test, p = 0.027). The maximum repeat length in a sibship also correlated with the maximum difference in repeat length between sibs (Pearson's correlation coefficient $r = 0.56$, 95% CI $0.25 < r < 0.76$).

In these 41 HD pedigrees, 8 individuals carried an expanded (CAG)n repeat smaller than 42 repeats (table 3). None of these indi-

Table 2. a Difference in repeat length between parent and offspring shown for mother-child and father-child transmissions, and **b** maximal difference in repeat length between sibs of the same sibship is listed for sibships with maternal transmission vs. sibships with paternal transmission

viduals

a Generation shift	Maternal (7)	Paternal (14)
−2 repeats	1	0
−1	1	0
0	4	2
1	1	4
2	0	2
3	0	3
6	0	2
10	0	1
Mean	−0.29	2.79
(SE)	(0.36)	(0.74)

b Maximal sibship difference (n)	Maternal (16)	Paternal (14)
0 repeats	7	2
1	4	3
2	2	2
3	2	3
4	1	0
5	0	1
6	0	1
8	0	1
9	0	1
Mean	1.12	3.14
(SE)	(0.33)	(0.76)

a Difference in repeat length between parent and offspring (generation shift) is shown for mother-child (maternal, n = 7) and for father-child transmission (paternal, n = 14). The mean generation shift is given with standard error (SE), for the two modes of transmission. The differences in generation shift between maternal and paternal transmission are significant (p = 0.0018, Wilcoxon two sample test, Bartlett's test for homogeneity of variance, p = 0.009).

b The maximal difference in repeat length between sibs of the same sibship is listed for sibships with maternal transmission (n = 16), versus sibships with paternal transmission (n = 14). The mean difference in repeat length is calculated as well as the standard error (SE). The means for the two modes of transmission are significantly different as estimated with the Wilcoxon two sample test (p = 0.027). T testing is not applicable because Bartlett's test for homogeneity of variance suggests significantly different variances for the two samples (p = 0.005).

viduals had a definite diagnosis of HD, but 3 were tentatively labeled as probably early HD, and individual 5.1 showed aspecific mental deterioration from the age of 80 years on.

It is known that the region of the (CAG)n repeat contains 2 CCG repeats that can be polymorphic [7]. To find out if this influenced the findings in the 8 individuals described in table 3, we analyzed their DNA samples also with PCR primers immediately flanking the (CAG)n repeat. The results of these tests are also listed in table 3 ((CAG)n;B). It is clear that there might be some variation in the CCG repeats but these 8 individuals clearly have repeats expanded outside the normal range but without the symptoms of typical HD.

Discussion

Analysis of the (CAG)n repeat in the IT15 gene in these 40 families supports the idea of genetic homogeneity in HD, because an expanded (CAG)n repeat was observed in every HD family tested. Analysis of the genetic transmission of this repeat also explains several known genetic characteristics of HD. The

Table 3. Summary of clinical data for 8 individuals with an expanded repeat smaller than 42 is shown

Individual	(CAG)n		Clinical status	Onset HD years	Clinical manifestations
	A	B			
5.1	39	38	NL	NA	no chorea, ambulatory, nursing home at the age of 90 years because of mental deterioration from 80 years on
16.3	39	37	NL	NA	normal mental and physical health at the age of 67 years
17.2	39	36	NL	NA	normal mental and physical health, lives independently at the age of 90 years
18.5	39	36	NL	NA	normal mental and physical health, normal professional activities at the age of 45 years
39.4	39	37	NL	NA	normal mental and physical health, lives independently at the age of 69 years
8.4	40	38	HD?	66	normal mental health at the age of 69 years, lives independently, minor motor symptoms starting at the age of 66 years
21.4	40	38	HD?	61	mental deterioration from 60 years on, but no chorea
16.5	41	39	HD?	70	choreic movements starting at the age of 70 years, lives independently at the age of 73 years

Individuals are listed according to their respective pedigree, and 5.1 indicates individual 1 belonging to pedigree 5. NA = Not applicable; NL = normal; HD? = probable Huntington's disease. (CAG)nA refers to the PCR test with primers amplifying also the flanking CCG repeat, and (CAG)nB refers to the PCR with primers immediately flanking the (CAG)n repeat and excluding the CCG repeat.

repeat size itself correlates with age of onset of HD (fig. 1) and 49% of the variance in age of onset is explained by variance in repeat length ($r^2 = 0.49$), at least for the observed range in repeat lengths in HD patients. However, the correlation is weaker in the 42 to 50 repeat unit length range, and the length of the expanded repeat cannot be used to predict age of onset. This is also observed in other studies [8–10]. The observed expansion in repeat size in case of paternal transmission and the nearly unchanged repeat size in case of maternal transmission (table 2a, b), may explain why paternal transmission of HD is associated on average with an earlier age of onset than maternal transmission, and why the parent-offspring age-of-onset correlation is stronger in case of maternal transmission [3]. The difference in repeat length between sibs is larger in case of a longer repeat, suggesting that a longer expansion of the (CAG)n repeat is more unstable in meiosis than a shorter expansion. This phenomenon may be responsible for the very long expansions (>100 repeats) observed in some cases of juvenile HD [6] after paternal transmission.

The most important finding in this study however is the presence of 8 individuals positive for the haplotype segregating with HD in their family, but without symptoms or only

atypical symptoms of HD at an advanced age. These 8 individuals show an expanded (CAG)n repeat length between 39 and 41 (table 3) as measured with the primers published by the Huntington's Disease Collaborative Research Group [6]. The expanded repeats have a length between 36 and 39 (CAG)n repeats when measured with primers immediately flanking the CAG repeat in IT15 [7]. These preliminary data would indicate that, at least in the examined population using the described technique, there might be a continuum in the number of expanded (CAG)n repeats in the IT15 gene with apparent normal mental and physical health, over repeats associated with minor motor symptoms or aspecific mental deterioration beginning after the age of 60, to longer repeats with increasing severity of symptoms and decreasing age of onset of typical HD. Affected relatives of individuals 16.3, 16.5, 18.5 and 39.4 (table 3) showed expanded (CAG)n repeats of 42, 43 and 44. Although the HD parents were not available for study in these cases, there is indirect evidence for a contraction of the repeat to a premutation size in individuals listed in table 3. On the other hand, we observed expansion from 34 (CAG)n repeats to 44 repeats after paternal transmission. The father with 34 (CAG)n repeats was asymptomatic. This indicates that a repeat length of 34 can be unstable and might expand in the disease range of repeat length after transmission. The expanded (CAG)n repeat in the 8 individuals listed in table 3 compares best with asymptomatic premutations observed in the fragile-X syndrome and in myotonic dystrophy [11]. It is tempting to speculate that expanded (CAG)n repeats in the range of 40 might also be responsible for other forms of hereditary chorea without dementia [12].

From these data it appears that precise measurement of the (CAG)n repeat may become very important for accurate genetic counselling in cases of presymptomatic or prenatal testing for HD. Presymptomatic testing with linked markers only will not detect a decrease in repeat length and might result in a very high risk for developing HD, whereas in some individuals with a (CAG)n repeat expansion of 39 or lower, the risk for developing HD within a normal life span may be close to zero. On the other hand, children from a normal 70-year-old individual belonging to a HD family might still be at risk for developing HD in case the repeat expanded and this may require DNA testing. The (CAG)n repeat test will also allow to test sporadic cases of HD, and presymptomatic testing will be possible for children of such an individual. However, we are still convinced that all presymptomatic testing should be done as part of a presymptomatic testing programme, and 'urgent' PCR-based presymptomatic testing outside a testing programme should be avoided because of the possible negative psychological impact for the tested individual and partner. A collaborative effort should be set up to assess the precise meaning of our observations on borderline expansions of the (CAG)n repeat in HD. In the mean time the (CAG)n test should be used with caution and other diseases, which share some characteristics with Huntington's chorea, should be studied.

Acknowledgements

The authors would like to express their gratitude to the HD families and to the Belgian HD League for the fruitful collaboration. The expertise and essential assistance of Andrea Boogaerts and Trees Cloostermans in gathering family information and samples is gratefully acknowledged. We would also like to thank Katrien Van den Bosch, Fabienne Rans and Mariëlle Braeken for expert technical assistance and R. Vlietinck for helpful discussion of the statistical aspects. This work was supported by the Interuniversity Network for Fundamental Research sponsored by the Belgian Government (1991–1995).

References

1 Martin JB, Gusella JF: Huntington's disease: Pathogenesis and management. N Engl J Med 1986;315: 1267–1276.

2 Harper PS: The epidemiology of Huntington's disease. Hum Genet 1992;89:365–376.

3 Myers RH, Madden JJ, Teague JL, Falek A: Factors related to onset age of Huntington's disease. Am J Hum Genet 1982;34:481–488.

4 Gusella JF, Wexler NS, Conneally PM, Naylor SL, Anderson MA, Tanzi RE, Watkins PC, Ottina K, Wallace MR, Sakaguchi AY, Young AB, Shoulson I, Bonilla E, Martin JB: A polymorphic DNA marker genetically linked to Huntington's disease. Nature 1983;306:234–238.

5 Wasmuth JJ, Hewitt J, Smith B, Allard D, Haines JL, Skarecky D, Partlow E, Hayden MR: A highly polymorphic locus very tightly linked to the Huntington's disease gene. Nature 1988;332:734–736.

6 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;72:1–20.

7 Rubinstein DC, Barton DE, Davison BCC, Ferguson-Smith MA: Analysis of the Huntington 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;2:1713–1715.

8 Snell RG, MacMillan JC, Cheadle JP, Fenton I, Lazarou LP, Davies P, MacDonald ME, Gusella JF, Harper PS, Shaw DJ: Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. Nature Genet 1993;4: 393–397.

9 Duyao M, Ambrose C, Myers R, Novelletto A, Persichetti F, Frontali M, Folstein S, Ross C, Franz M, Abbott M, Gray J, Conneally P, Young A, Penney J, Hollingsworth Z, Shoulson I, Lazzarini A, Falek A, Koroshetz W, Sax D, Bird E, Vonsattel J, Bonilla E, Alvir J, Bichham Conde J, Cha J-H, Dure L, Gomez F, Ramos M, Sanchez-Ramos J, Snodgrass S, de Young M, Wexler N, Moscowitz C, Penchaszadeh G, MacFarlane H, Anderson M, Jenkins B, Srinidhi J, Barnes G, Gusella J, MacDonald M: Trinucleotide repeat length instability and age of onset in Hungtington's disease. Nature Genet 1993;4:387–392.

10 Andrew SE, Goldberg YP, Kremer B, Telenius H, Theilmann J, Adam S, Starr E, Squitieri F, Lin B, Kalchman MA, Graham RK, Hayden MR: The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. Nature Genet 1993;4:398–403.

11 Caskey CT, Pizzuti A, Fu Y-H, Fenwick RG, Nelson DL: Triplet repeat mutations in human disease. Science 1992;256:784–789.

12 Behan PO, Bone I: Hereditary chorea without dementia. J Neurol Neurosurg Psychiatry 1977;40:687–691.

13 Lange K, Weeks D, Boehnke M: Programs for pedigree analysis: Mendel, Fisher, and dGene. Genet Epidemiol 1988;50:471–472.