Genetic polymorphisms adjacent to the CAG repeat influence clinical features at onset in Huntington’s disease

I Vuillaume, P Vermersch, A Destée, H Petit, B Sablonnière

Abstract
Objectives—To evaluate possible influences of CCG and Δ2642 glutamic acid polymorphisms adjacent to the (CAG)n trinucleotide repeat in Huntington’s disease gene IT15 on some clinical features (age and symptoms) at onset.

Methods—84 patients and a control group of 68 unaffected relatives were studied. Patients all belonged to a group of affected persons tested for molecular confirmation of Huntington’s disease. The length of the CAG repeat sequence in the IT15 gene and the adjacent CCG and Δ2642 polymorphisms were determined by quantitative polymerase chain reaction.

Results—Two intragenic polymorphisms were studied: (CCG)n and Δ2642 glutamic acid. Patients were classified firstly according to the size of the CCG rich segment adjacent to the CAG repeat into genotype groups CCG 7/7, 7/8, 7/9, 7/10, and 10/10 and then according to Δ2642 polymorphism into genotype groups A/A (absence of the Δ2642 deletion), A/B, and B/B (presence of the Δ2642 deletion in respectively one and two alleles). The presence of Δ2642 mutation was associated with a significant decrease in age at onset, although there was no significant increase in CAG size. A good correlation was found between the (CAG)n, trinucleotide repeat size and the age at onset in patients with genotype AA (r²=0.72). Within patients of the A/B genotype group however, a significant correlation was found but with a drop of the r² value to 0.44. No association was found between age at onset and the CCG polymorphism. Although an increased percentage of patients within the A/A genotype group had a neurological onset, we found no overall significant association between CCG or Δ2642 polymorphisms and the nature of symptoms at onset.

Conclusions—The Δ2642 glutamic acid polymorphism did not affect CAG repeat size nor the nature of symptoms at onset but seems to influence the age at onset in patients with Huntington’s disease.

Keywords: Huntington’s disease; clinical features; genotype; polymorphisms

Huntington’s disease is an autosomal dominant neurodegenerative disorder affecting about 1 in 10 000 people in European populations. It usually presents in adult life with chorea, psychiatric manifestations, and cognitive impairment, leading to progressive dementia.1 2 This disorder nevertheless shows variation in age and in clinical features at onset. Onset symptoms are neurological, psychiatric/ cognitive, or combined (neurological and psychiatric/cognitive) in respectively 46% to 59%, 23% to 36%, and 18% to 30% of cases.3 4 5 The gene responsible for Huntington’s disease was mapped to the tip of the short arm of chromosome 4 in 1983.6 The genetic defect associated with the disease was identified in 1993 in a novel gene (IT15) containing a trinucleotide repeat (CAG)n that is expanded in patients with Huntington’s disease.7 Whereas the number of CAG repeats in IT15 gene ranges from nine to 35 in healthy persons, it ranges from 36 to 121 in patients with Huntington’s disease.8 The expansion of the (CAG)n repeat thus proved to be a highly specific marker for the diagnosis of Huntington’s disease.9

After cloning of the Huntington’s disease mutation, two genetic polymorphisms were identified close to the CAG tract. The first one was a CCG rich segment downstream to the (CAG)n stretch and the second one was the Δ2642 glutamic acid polymorphism concerning a deletion of three nucleotides at codon positions 2642–2645.10–15 Both presented polymorphic frequencies on normal and affected chromosomes. Moreover, both were independently associated with differences in CAG repeat length on normal chromosomes, as shown by haplotype analysis. In these studies, a strong linkage disequilibrium was found between the Huntington’s disease mutation and alleles at both polymorphic regions: CCG rich length alleles were underrepresented whereas Δ2642 was overrepresented in Huntington’s disease chromosomes.

Numerous attempts to determine correlations between the Huntington’s disease repeat length and clinical features have been undertaken. A strong negative correlation (r=−0.70) between age at onset and CAG repeat size was reported overall, the number of CAG repeats in Huntington’s disease chromosomes accounting for 50% of the variation of age at onset.10–14 This correlation could be even stronger as shown recently by Brinkman et al.15 As for symptoms at onset, no particular clinical feature correlated with the CAG repeat size.10 16 21 22 As the variation in age at onset is partially explained by the trinucleotide repeat length, it seemed logical to investigate whether...
adjacent polymorphisms (for example, the CCG rich segment and Δ2642 mutation) also account in part for the clinical heterogeneity of the disease. We therefore examined in 84 independent patients with Huntington's disease the correlation between the age and the nature of symptoms at onset with the CAG repeat size or with any of the genotypes determined by either CCG or Δ2642 polymorphisms.

**Table 1** Distribution of genotypes determined by CCG or Δ2642 glutamic acid polymorphisms in controls and patients with Huntington's disease

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Control group</th>
<th>HD group</th>
<th>p Value overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCG 7/7</td>
<td>19</td>
<td>31</td>
<td>0.10</td>
</tr>
<tr>
<td>CCG 7/8</td>
<td>2</td>
<td>3</td>
<td>0.65</td>
</tr>
<tr>
<td>CCG 7/9</td>
<td>5</td>
<td>1</td>
<td>0.20</td>
</tr>
<tr>
<td>CCG 7/10</td>
<td>37</td>
<td>28</td>
<td>0.33</td>
</tr>
<tr>
<td>CCG 10/10</td>
<td>5</td>
<td>2</td>
<td>0.12</td>
</tr>
<tr>
<td>Total</td>
<td>68</td>
<td>84</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

A/B 0 0 1 4 4.7
A/B 11 16.2 34 40.5
A/A 57 83.8 46 54.8
Total 68 100 84 100 <0.005

**Results**

**DISTRIBUTION OF GENOTYPES DETERMINED BY THE CCG OR THE Δ2642 POLYMORPHISMS**

The CCG segment of seven repeats was overrepresented in Huntington's disease chromosomes (up to 95% vs. 60.3% in normal chromosomes). The unbalanced distribution of linkage disequilibrium between CCG rich segments and the Huntington's disease mutation was highly significant (data not shown). Genotype assessment in controls and patients led to five different groups (table1) according to the number of CCG repeats. The distribution of those genotype groups (CCG 7/7, 7/8, 7/9, 7/10, and 10/10) showed a significant difference among patients and controls (p<0.005). Genotype CCG 7/7 was present in 60.7% of patients with Huntington's disease whereas only in 27.9% of controls. On the other hand, genotype 7/10 was more frequent in controls (34.4%) than in patients (33.3%). Genotype frequency for the Δ2642 glutamic acid polymorphism was also assessed on controls and patients. Two different genotype groups were identified: group A/A characterised by the absence of the deletion, group A/B and B/B determined by the presence of the deletion on respectively one and two alleles. The distribution was significantly different between patients and controls (p<0.001). Absence of the Δ2642 residue (genotype A/B) was found in 16.2% of controls leading to an allele frequency of 8% considering that no homozygote (genotype B/B) was found for this deletion. However, the Δ2642 glutamic acid polymorphism was overrepresented among patients with Huntington's disease, being present in 45.2% of our Huntington's disease cohort. In this group, we found four patients homozygous for the deletion, therefore calculating a frequency of 25.0% for the B allele in patients with Huntington's disease.

**RELATION BETWEEN PATIENT GENOTYPE (CCG OR Δ2642 POLYMORPHISMS) AND AGE AT ONSET**

The CCG repeat size associated with normal and affected chromosomes showed two well separated distributions. The number of CAG repeats varied between 12 and 32 (mean 19.5)
in the normal allele, and between 39 and 76 (mean 45.8) in the Huntington’s disease allele. In our cohort of patients, age at onset varied between 10 and 70 years (mean 43.0 (SD 13.8)). Age at onset for 63 of them varied between 21 and 59 years, eight patients were juvenile cases and 11 patients had a late onset. The now long established inverse correlation between age at onset and CAG repeat size in patients was confirmed in our study ($r=-0.757, \ p<10^{-4}$). To evaluate the effect of either the CCG rich or the Δ2642 glutamic acid polymorphisms, we compared the CAG repeat number and the age at onset in patients carrying different genotypes. Mean age at onset among the different genotype groups according to the CCG polymorphisms varied between 42.8 and 43.1 years, showing no significant difference (table 2). Patients bearing the Δ2642 glutamic acid mutation (A/B+B/B genotype group) had an earlier onset of 6.5 years than those belonging to the A/A genotype group (table 2). In this case, the difference in mean age at onset was significant ($p<0.05$). In these two genotype groups, log of age at onset correlated significantly with the CAG stretch ($r=-0.848$ and $r=-0.664$ in the A/A and in the A/B+B/B genotype group respectively). However, there was no significant difference between the size of the CAG repeat in these same two genotype groups, suggesting that the differences in mean age at onset could be explained by an indirect influence of Δ2642 mutation. The difference in age at onset between the A/A and A/B+B/B genotype groups cannot be simply explained by an apparent bias in the distribution of large CAG repeat expansions. In fact, we found a similar distribution of patients with a CAG repeat expansion larger than 50 CAG repeats between the two groups (13.3% vs 20%, $\chi^2=0.98, \ p=0.45$, data not shown). Moreover, the contribution of the CAG repeat length in the variation of age at onset did not exceed 44% ($r^2=0.44$) in patients carrying the Δ2642 muta-

### Table 2: Distribution of age at onset and CAG repeat length and their correlations in each genotype group

<table>
<thead>
<tr>
<th>Genotype</th>
<th>n</th>
<th>Age at onset</th>
<th>CAG</th>
<th>Correlation age at onset/CAG</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCG 7/7</td>
<td>51</td>
<td>42.8 (12.8)</td>
<td>45.6 (5.4)</td>
<td>$-0.744, 0.55$</td>
</tr>
<tr>
<td>CCG 7/X*</td>
<td>31</td>
<td>43.1 (15.0)</td>
<td>45.9 (7.3)</td>
<td>$-0.759, 0.57$</td>
</tr>
<tr>
<td>A/A</td>
<td>45</td>
<td>45.5 (12.2)</td>
<td>45.0 (6.4)</td>
<td>$-0.848, 0.72$</td>
</tr>
<tr>
<td>A/B + B/B</td>
<td>37</td>
<td>39.0 (14.8)</td>
<td>46.5 (5.8)</td>
<td>$-0.664, 0.44$</td>
</tr>
<tr>
<td>P</td>
<td>&lt;0.05</td>
<td>&lt;0.05</td>
<td>&lt;0.05</td>
<td>&lt;0.05</td>
</tr>
</tbody>
</table>

* X contained either 8, 9, or 10 CCG repeats.

In our cohort of patients, age at onset varied (mean 45.8) in the Huntington’s disease allele. The now long established inverse correlation between age at onset and CAG repeat size in patients was confirmed in our study ($r=-0.757, \ p<10^{-4}$). To evaluate the effect of either the CCG rich or the Δ2642 glutamic acid polymorphisms, we compared the CAG repeat number and the age at onset in patients carrying different genotypes. Mean age at onset among the different genotype groups according to the CCG polymorphisms varied between 42.8 and 43.1 years, showing no significant difference (table 2). Patients bearing the Δ2642 glutamic acid mutation (A/B+B/B genotype group) had an earlier onset of 6.5 years than those belonging to the A/A genotype group (table 2). In this case, the difference in mean age at onset was significant ($p<0.05$). In these two genotype groups, log of age at onset correlated significantly with the CAG stretch ($r=-0.848$ and $r=-0.664$ in the A/A and in the A/B+B/B genotype group respectively). However, there was no significant difference between the size of the CAG repeat in these same two genotype groups, suggesting that the differences in mean age at onset could be explained by an indirect influence of Δ2642 mutation. The difference in age at onset between the A/A and A/B+B/B genotype groups cannot be simply explained by an apparent bias in the distribution of large CAG repeat expansions. In fact, we found a similar distribution of patients with a CAG repeat expansion larger than 50 CAG repeats between the two groups (13.3% vs 20%, $\chi^2=0.98, \ p=0.45$, data not shown). Moreover, the contribution of the CAG repeat length in the variation of age at onset did not exceed 44% ($r^2=0.44$) in patients carrying the Δ2642 muta-

### Table 3: Relation between patient genotype and nature of symptoms at onset

<table>
<thead>
<tr>
<th>Patient genotype</th>
<th>CCG 7/7</th>
<th>CCG 7/X</th>
<th>A/A</th>
<th>A/B + B/B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurological</td>
<td>41.2</td>
<td>34.5</td>
<td>46.7</td>
<td>29.7</td>
</tr>
<tr>
<td>Psychiatric/cognitive</td>
<td>37.4</td>
<td>24.1</td>
<td>22.2</td>
<td>27.0</td>
</tr>
<tr>
<td>Combined*</td>
<td>31.4</td>
<td>41.4</td>
<td>31.1</td>
<td>43.3</td>
</tr>
<tr>
<td>% of patients</td>
<td>0.70</td>
<td>0.45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Association of two or three types of symptoms at onset (neurological / psychiatric / cognitive) within the first year of onset.

### Discussion

Analysis of the distribution of genotypes determined by the CCG or Δ2642 polymorphisms in our cohort of 84 patients with Huntington’s disease and 68 controls is similar to previous published data. Thus despite the impossibility of assessing the phase of the Δ2642 polymorphism in our set of patients, the genotype frequency for this polymorphism shows similar deduced allele frequencies in the Huntington’s disease gene to those values reported by Almqvist and others, assuming the fact that a strong disequilibrium exists between Δ2642 and the Huntington’s disease mutation. Regarding the CCG rich segment distribution, our results confirmed what was previously published by other investigators, showing that expanded (CAG) repeats are preferentially associated with a CCG repeat of seven (96% of patients have a CCG of seven repeats) whereas it reached 72% in patients not carrying the deletion ($r=0.72$). On the other hand, the contribution of the CAG repeat size in the variation of age at onset was similar among the two genotype groups CCG 7/7 and CCG 7/X (X containing either 8, 9, or 10 CCG repeats) determined by the CCG polymorphism. The inverse correlation between CAG size and age at onset was significant with an $r=0.55$ for patients in genotype group CCG 7/7 and an $r=0.57$ for patients in genotype group CCG 7/X underlying the absence of any apparent effect of the CCG polymorphism on the correlation between CAG size and age at onset.

### Relation between Patient Genotype (CCG or Δ2642 Polymorphisms) and the Nature of Symptoms at Onset

An assessment was made of the relation between the (CAG)n expansion length and the presentation of clinical features within the first year after onset. The cohort was therefore divided into those who either had neurological symptoms, psychiatric or cognitive manifestations, and into those who had combined symptoms (neurological and psychiatric, or neurological and cognitive, or cognitive and psychiatric symptoms). There was no association between the CAG repeat length and a particular clinical presentation at onset (data not shown).

The influence of either the CCG or the Δ2642 glutamic acid polymorphism on the nature of symptoms at onset was also examined. There was no overall statistical relation ($p>0.05$) between the clinical presentation at onset and the different genotypes determined by the CCG rich segment on the one hand, and the Δ2642 glutamic acid polymorphism on the other hand (table 3). Nevertheless, it was noteworthy that 46.7% of patients with genotype A/B or B/B had a neurological onset with no other associated symptom. The distribution of genotypes including the various CCG polymorphisms was similar among patients presenting either psychiatric, cognitive, or combined symptoms.
Genetic polymorphisms and clinical features in Huntington’s disease
761

is the major factor for the reduction in age at
ments. As the expansion of CAG repeat length
sion on the Huntington’s disease chromosome
on the Huntington’s disease. This finding suggests that
function of the Huntington’s disease gene is
influenced not only by the CAG repeat expansion
on the Huntington’s disease chromosome
but also by possible cis-acting specific
As the expansion of CAG repeat length is the major factor for the reduction in age at
onset, presence of the Δ2642 mutation may add to the severity of the disease, carriers of
this mutation tending to develop the disease earlier. This finding agrees with the opinion
that age at onset of Huntington’s disease may depend on different factors including the
influence of the normal allele and non-allelic modifying genes. Genetic components
such as genetic imprinting and aging genes could be evoked. Other factors including envi-
ronmental factors may also play a part in the
pHENOSCOPE. No correlation was found between nature of symptoms at onset and (CAG)n trinucleotide expansion,
concurring with results of other studies. No
significant association was found between a specific type of symptom at onset
(neurological, psychiatric, or cognitive) and the genetic polymorphisms. However, strikingly,
a larger number of patients (46.7%) with
genotype A/A than those with genotype A/B or
B/B (29.7%) had neurological features without
any psychiatric or cognitive symptoms at onset. This difference was not significant but
could indicate a tendency towards a possible association between the presence of genotype
A/A and the predominance of neurological features at onset. Genetic variants acting cis or
to the expanded CAG repeat could then influence the phenotypic presentation. The
assessment of the Δ2642 polymorphisms disclosed the potential indirect influence of
this genetic polymorphism in the onset of the disease. Whether or not this polymorphism
may represent an independent predictor of the
age at onset would need further studies of the
respective contribution of normal and expanded alleles (deleted or not), on the pheno-
type; requiring haplotype studies. It could also
be relevant to analyse the influence of the
 genetic polymorphisms on the mode of
progression of the disease, as was already
reported for the influence on the age at
onset. This genotypic study of patients
based on the Δ2642 and CCG-rich segment
analysis may contribute to our understanding of the
 genotype-phenotype relations in Hunt-
ington’s disease.

We are grateful to Guy Bocquillon, Patrick Devos, Christiane
Marzys, and Carole Verlez for their skilful technical assistance
and to Isabelle Delalande for her invaluable help with accumu-
lation of clinical data. We are also grateful to A Duhamel (stat-
istician from the CERIM, CHRU de Lille) for help in statistical
valuation. This work was supported by grants from the CHRU of
Lille.

1 Hayden MR. Huntington’s chorea. New-York: Springer
Verlag 1981.
3 Polstein JE. Huntington’s disease: a disorder of families.
in 510 patients with Huntington’s disease. J Med Genet
5 Claes S, Van Zand K, Legius E, et al. Correlations between
triplet repeat expansion and clinical features in Hunt-
dNA marker genetically linked to Huntington’s disease.
7 The Huntington’s Disease Collaborative Research Group. A
novel gene containing a trinucleotide repeat that is
expanded and unstable on Huntington’s disease chromo-
264:1401–6.