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

Christophe Pécheux, 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)3(CCT)2, 48% (CCG)5(CCT)3, and 2% (CCG)6(CCT)4. In contrast (CCG)5(CCT)3 was found in 85% of the HD alleles which represents significant linkage disequilibrium with the HD mutation.

Huntington's disease (HD) is associated with the expansion of a CAG triplet repeat (37 to 140) in the 5' coding region of the IT15 transcript. 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, 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. The 176 bp allele is the most frequently observed in normal subjects and HD patients and corresponds to the sequence originally published. 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. The normal and abnormal alleles from heterozygous HD patients were then separated on a 1:5% low melting agarose gel and directly sequenced either automatically or manually. Polymorphisms both in the CCG triplet number ((CCG)n) or in the CCT triplet number ((CCT)m) adjacent to the CAG repeat were observed. In normal chromosomes, two major alleles (CCG5(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. In contrast, the HD chromosomes carried allele 1 in 85% of the cases confirming the linkage disequilibrium already described (x2=13, p<0.001). However, in contrast to Barron et al., 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 (CCG5(CCT)2 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-

---

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

<table>
<thead>
<tr>
<th>Allele</th>
<th>Normal alleles</th>
<th>HD alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>%</td>
</tr>
<tr>
<td>1</td>
<td>(CAG)CCA CAG CCG CCA (CCG)3 (CCT)2 CAG CCT</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>(CAG)CCA CAG CCG CCA (CCG)5 (CCT)2 CAG CCT</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>(CAG)CCA CAG CCG CCA (CCG)10 (CCT)2 CAG CCT</td>
<td>20</td>
</tr>
</tbody>
</table>
The amplification 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.

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