Molecular analysis of late onset Huntington’s disease

B Kremer, F Squitteri, H Telenius, S E Andrew, J Theilmann, N Spence, Y P Goldberg, M R Hayden

Abstract
Late onset Huntington’s disease is characterised by onset of symptoms after the age of 50 and is usually associated with a milder course. We have analysed the CAG trinucleotide repeat within the HD gene in 133 late onset patients from 107 extended families. The median upper allele size for the CAG repeat was 42 with a range of 38 to 48 repeats. A significant negative correlation (r = -0.29, p = 0.001) was found between the length of repeat and age of onset for the total cohort. However, for persons with age of onset greater than 60, no significant correlation was found. In addition, no significant correlation was found between age of onset and size of the lower allele and the sex of the affected parent or grandparent. There was no preponderance of maternal descent for late onset cases in this series. This study shows that variation in repeat length only accounts for approximately 7% of the variation in age of onset for persons beyond the age of 50 and clearly shows how with increasing onset age the effect of the repeat length on this onset age seems to diminish.

Methods
PATIENT SELECTION
We have collected DNA and clinical details where possible on approximately 1100 affected persons with HD. Those patients who clearly had onset at 50 years or more and for whom DNA was readily available were chosen for study (n = 137). A total of 95 of these patients was previously included as part of our analysis of a cohort of 360 patients.

DNA ANALYSIS
Genomic DNA was isolated from leucocytes by standard extraction procedures. PCR amplification of the HD CAG repeat was performed with primers HD344 (5’ CTCAGAGTCCTCAAGTCTTC 3’) and HD482 (5’ GGCTGAGGAGGCTGAGGAGG 3’). PCR conditions were 2 mmol/l MgCl2, 50 mmol/l KCl, 20 mmol/l Tris, pH 8.4, 3.5% formamide, 15% glycerol, 200 µmol/l of each dNTP, 0.5 µmol/l of each primer, 15 mmol/l end labelled primer, and 1.25 U of Taq DNA polymerase per 25 µl reaction. Thermal cycling conditions were 95°C for three minutes, followed by 30 cycles of 94°C for one
Table 1 Clinical and genetic data of late onset cohort.

<table>
<thead>
<tr>
<th>Age of onset</th>
<th>Total group (n = 133)</th>
<th>One proband per nuclear family (n = 112)</th>
<th>One proband per extended family (n = 107)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>69</td>
<td>54</td>
<td>54</td>
</tr>
<tr>
<td>Male</td>
<td>64</td>
<td>58</td>
<td>53</td>
</tr>
<tr>
<td>Sex of affected parent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>49</td>
<td>43</td>
<td>39</td>
</tr>
<tr>
<td>Male</td>
<td>62</td>
<td>50</td>
<td>49</td>
</tr>
<tr>
<td>Sex of affected grandparent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Age of onset</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>57</td>
<td>50</td>
<td>51</td>
</tr>
<tr>
<td>Range</td>
<td>50-85</td>
<td>62</td>
<td>72°C</td>
</tr>
<tr>
<td>Upper allele size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>38-48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower allele size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>11-31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Results

SEX OF AFFECTED PARENTS OF LATE ONSET PATIENTS
A total of 137 patients from 111 separate families was included in this cohort (table 1). The sex of the affected parent was confirmed in 111 affected persons. No significant alteration in the expected sex distribution of the probands was determined. Furthermore, there was no preponderance of maternal descent for late onset patients in this cohort (table 1).

REPEAT LENGTH IN LATE ONSET PATIENTS
An allele in the range previously determined for affected patients with HD was seen in 133 of 137 persons. The median repeat length was 42 with a range of 38 to 48 (table 1, fig 1). There were four patients who did not have any detectable expansion, as all four had two alleles in the normal range but had clinical features compatible with the diagnosis of HD. Two late onset patients without repeat expansion are from families in whom at least two other affected persons have repeat expansion in the range seen in patients with HD (18/43, 17/47 and 25/50, 17/43 for lower and upper alleles in affected persons in each of these families, respectively). The presence of two normal sized alleles in these two affected subjects from two families where expansion has been shown in other affected persons is unexplained at present. These subjects may represent phenocopies and were excluded from subsequent analysis.

ASSOCIATION BETWEEN REPEAT LENGTH AND AGE AT ONSET
A significant negative correlation (n = 133, \( r = -0.29, p = 0.001 \)) was found between the length of the repeat and the age of onset for the total cohort of late onset patients (fig 1, table 2). However, for the subgroup of patients with age of onset above 60 (n = 42), no significant correlation was found between the size of the upper allele and the age of onset (n = 42, \( r = -0.05, \) NS). Similar findings were obtained when the analysis was restricted to one member per nuclear family or one member of the extended family respectively (table 2). This analysis was also conducted on affected HD members in our whole cohort with onset between 20 and 50 years and those with onset below 20. A higher correlation coefficient between CAG repeat and age at onset is evident as the age of onset decreases (table 2).

No significant correlation could be found between age at onset and the size of the lower allele, nor between the age of onset and sex of affected parent or grandparent. Furthermore, neither the sex of the affected parent or grand-

Figure 1 Age of onset v CAG repeat length in 133 late onset patients (\( r = -0.29, p = 0.001 \), for log transformed age of onset). Note that overlapping values are represented as a single dot.
Molecular analysis

Parent–child difference in age of onset

Figure 2. Parent-child difference in age of onset vs parent-child difference in CAG repeat length for 15 late onset patients and their children. The regression line and its 95% confidence interval are indicated, showing a non-significant association.

Sib–sib difference in age of onset

Figure 3. Sib-sib difference in age of onset vs sib-sib difference in CAG repeat length for 23 late onset patients and their sibs. Data of the younger sib are subtracted from data of the older sib. The regression line and its 95% confidence interval are indicated, showing a non-significant association.

culminates

RELATIONSHIP BETWEEN REPEAT LENGTH AND AGE OF ONSET IN FIRST DEGREE AFFECTED RELATIVES

There were 16 parent/child pairs for study from 15 nuclear families in whom the parent had late onset and from whom clinical data and DNA were available. In all instances, there was no difference in allele size of greater than six repeats. However, there were marked variations in age of onset (up to 33 years) in these offspring compared to their parents (fig 2). In eight parent–child pairs in whom the difference in age of onset was 15 years or greater, the transmitting parent was the father in seven instances (fig 2). No significant correlation between repeat length differences and variation in age at onset were seen in these parent–child pairs (n = 16, r = −0.16, NS) (fig 2). Therefore, while there are obvious differences in age of onset between the parent and the child, the repeat length difference did not appear to account for these variations in the age of onset.

There were 20 families available for sib–sib analysis. These included 17 with data available on two sibs and three with data on three affected sibs. Therefore 23 affected sib pairs were analysed. CAG repeat length varied by five repeats or fewer in the sibs (fig 3). In 17 sib pairs, differences in age at onset were less than 10 years with a familial tendency to late age of onset in these sibs (fig 4). However, in six sibships, differences in age of onset were between 10 and 29 years with CAG repeat length differences of less than five.

Discussion

In this study we have shown that there is a significant association between repeat length and age of onset of patients who first manifest signs and symptoms of HD after the age of 50. However, for the 42 patients with age of onset after the age of 60 in the study, no significant association could be found between age at onset and size of the upper allele.

We have previously reported on the correlations between repeat length and age of onset in a predominantly adult cohort of patients with HD and in a separate cohort of persons with juvenile onset. In that cohort of adult patients with HD (mean age of onset 41.5 years) the repeat length accounted for approximately 50% of the variation in age of onset. However, for the cohort with juvenile onset HD, the length of the repeat accounted for approximately 73% of the variation in age of onset. This trend is borne out in this study, with the variation in repeat length accounting for only approximately 7% of the variation in age of onset in the whole cohort beyond the age of 50, without any significant influence on the age of onset in persons beyond the age of 60. Thus, it is clear from these three studies that the contribution of repeat length to age of onset is greatest for those with earliest onset.
repeat length has already been associated with age of onset;\textsuperscript{16-19} it is most likely that the size of the repeat in some way directly affects the gene product resulting in a more severe phenotype. Earlier onset could reflect a more pronounced effect of the abnormal gene product on neuronal cells. The relative contribution of other factors which might contribute to the pathogenesis of HD would decrease with increased CAG size.

In this study, we show no preponderance of maternal transmission in persons with late onset HD. There have been at least three previous studies which adopted a similar methodology in terms of ascertainment of patients and have resulted in conflicting findings (table 3). Interestingly, if one combines the data from all studies, while there is a small preponderance of affected mothers for late onset cases, this does not reach any statistical significance. There are numerous potential pitfalls in such studies which in part may account for the conflicting findings. A bias might be introduced by the omission of many cases where the sex of the affected parent was not known. The sex of the affected parent of the late onset cases was not known in approximately one third of the cases in the study by Myers et al,\textsuperscript{12} but only in 14% in the report of Farrer and Connelly\textsuperscript{13} and in 17% in this study. This bias would, however, only be in effect if the unknown sex of the transmitting parent was not random. Therefore, this is unlikely to be a major factor accounting for these different findings. Another potential bias might be counting a parent more than once if there are multiple affected sibs. However, when this particular issue was looked at specifically in the data set of this study, this did not appear to influence the results. A third factor that should be considered is whether these different study groups were strictly compatible in terms of age of onset. If preponderance of maternal descent for HD was mainly seen in those with latest age at onset, inclusion of a large number of very late onset cases would bias these results. To explore this, we examined our cohort of 42 with onset after 60 years. In this group, the number of affected mothers was 15 and affected fathers 17, while the sex of the transmitting parent was unknown in 10 persons. Even in this selected group there is no evidence for preponderance of maternal descent.

Our study included 16 parent/child pairs for whom DNA analysis and also age of onset data were available. Similarly, data were available for 23 pairs of affected sibs. In some families there is familial aggregation of late onset cases both within nuclear families and in extended pedigrees (fig 4). This supports findings from

Table 3  Sex of affected parents of late onset cases.

<table>
<thead>
<tr>
<th></th>
<th>Father</th>
<th>Mother</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myers et al\textsuperscript{12}</td>
<td>13</td>
<td>32</td>
</tr>
<tr>
<td>Farrer and Connelly\textsuperscript{13}</td>
<td>35</td>
<td>56</td>
</tr>
<tr>
<td>Adams\textsuperscript{1}</td>
<td>63</td>
<td>53</td>
</tr>
<tr>
<td>This study</td>
<td>62</td>
<td>49</td>
</tr>
<tr>
<td>Total</td>
<td>173</td>
<td>190</td>
</tr>
</tbody>
</table>
Molecular analysis of late onset Huntington’s disease

We thank members of our research group for their helpful comments. This work was supported by the MRC (Canada), the Canadian Genetic Disease Network, and the Huntington Society of Canada. YPG is a Sandoz postdoctoral fellow. BK and HT are MRC postdoctoral fellows. FS is supported by the Instituto Scientific Sanatrix. MRH is an established investigator for the BC Children’s Hospital.

1 Hayden MR. Huntington’s chorea. New York: Springer-Verlag, 1981.

This and previous studies,23–24 that have suggested there is familial aggregation of age of onset both within nuclear and extended families for both early and late onset illness. These factors can be taken into account as broad principles when counselling at risk family members of affected persons. These data strongly indicate that there are familial factors present in these families, besides CAG repeat size, influencing the onset and severity of the illness.

Previous studies have appropriately commented on the difficulties in the designation of age of onset in affected persons (summarised in reference 1, pp 45–9). This may be particularly apparent for late onset cases where the disease may be milder in presentation. This factor may play some role in reducing the strength of the association between CAG repeat length and age at onset in late onset cases. Recent studies17–19 determining the relationship between trinucleotide repeat length and age of onset clearly may underestimate the relationship between such variables in view of the fact that the estimates of age of onset may show considerable variation around the true age of onset. Nevertheless, it is evident from this study that as the age of onset increases, the impact of repeat length on this clinical onset appears to show less influence. Factors other than the CAG repeat length are likely to play a more important role in determination of age of onset for persons with onset beyond the age of 50. These might include other genetic factors as well as possible environmental factors that are influencing the expression of the abnormal gene. Understanding these factors will not doubt provide valuable insights into the pathogenesis of this devastating disorder.