Age of onset in Huntington disease: sex specific influence of apolipoprotein E genotype and normal CAG repeat length

Patrick Kehoe, Michael Krawczak, Peter S Harper, Michael J Owen, A Lesley Jones

Abstract
Age of onset (AO) of Huntington disease (HD) is known to be correlated with the length of an expanded CAG repeat in the HD gene. Apolipoprotein E (APOE) genotype, in turn, is known to influence AO in Alzheimer disease, rendering the APOE gene a likely candidate to affect AO in other neurological diseases too. We therefore determined APOE genotype and normal CAG repeat length in the HD gene for 138 HD patients who were previously analysed with respect to CAG repeat length. Genotyping for APOE was performed blind to clinical information. In addition to highlighting the effect of the normal repeat length upon AO in maternally inherited HD and in male patients, we show that the APOE ε2/3 genotype is associated with significantly earlier AO in males than in females. Such a sex difference in AO was not apparent for any of the other APOE genotypes. Our findings suggest that subtle differences in the course of the neurodegeneration in HD may allow interacting genes to exert gender specific effects upon AO.

Keywords: Huntington disease; APOE; age of onset

Huntington disease (HD) is a late onset, autosomal dominant neurodegenerative disease associated with the expansion of a CAG repeat in the first exon of a gene on chromosome 4. The repeat number is polymorphic, with eight to 39 repeats observed in the normal population, and 36 to over 120 repeats found in the affected genes of HD patients. Age of onset (AO) of HD is variable and shows anticipation, that is, the disease tends to manifest progressively earlier from generation to generation.

The most important factor influencing AO in HD is the length of the expanded CAG repeat, which shows strong negative correlation with AO in the relatively few juvenile patients bearing long repeat arrays (>50 repeats), and less strong, albeit significant, correlation in adult onset HD patients. Moreover, previous analysis of a cohort of Welsh HD patients has shown a length effect on AO for the CAG repeat of the non-disease chromosome (that is, that from the normal parent) when paternally inherited, but not replicated in other studies. However, though variation in repeat length may explain some of the variation in AO and, together with the observed progression of expansion over generations, provides a molecular basis for anticipation, other variables must play a role, particularly in those cases where pathological repeat numbers are in the high 30s or low 40s. This insight has prompted the search in HD patients for other genes influencing AO, particularly in adult onset cases. Finding such genes is paramount since they may provide clues as to the pathology arising from the expanded repeat.

Human apolipoprotein E (ApoE) exists in three common isoforms, E2, E3, and E4, which are encoded by a gene (APOE) with three alleles, ε2, ε3, and ε4. Carriers of the ε4 allele are known to be at an increased risk of developing late onset Alzheimer disease (AD); the mechanism underlying this association is, however, still unknown. ApoE binds β-amyloid and Evans et al suggested that ApoE is involved in the initiation of β-amyloid fibril formation. Indeed, Ma et al have shown this function in vitro, with E4 exhibiting the strongest promotion of fibril formation. Carriership of ε4 is also implicated in the presentation of motor neuron disease where possession of at least one copy is associated with bulbar rather than limb onset. More recent evidence indicates that ApoE exerts an isoform specific antioxidant role. Thus, E2 gives most protection and E4 least, a finding not inconsistent with the substantially later AO of AD observed among ε2 carriers as compared with ε4 carriers.

The association of APOE genotype with several neurodegenerative disorders has been examined, with inconclusive results, but an analysis of APOE genotypes in HD patients has not been reported so far. We therefore studied a series of DNA samples from patients previously analysed with respect to their CAG repeat number and AO to see whether APOE genotype could explain a significant proportion of AO variation.

Materials and methods
A total of 138 Welsh HD patients who had had their CAG repeat numbers determined before were analysed for their APOE genotype, blind to repeat number and AO. For 121 of these patients, the sex of the transmitting parent was also known. Polymerase chain reaction and CfoI digestion were carried out as described by Wenhall et al. followed by polyacrylamide gel electrophoresis of fragments on 6% non-denaturing gels with ethidium bromide staining.

Statistical analysis was performed using the SAS/STAT® software package (version 6). Comparison of two Pearson correlation coeffi-
Table 1 Age of onset (AO) and APOE genotype in Welsh Huntington disease patients. AO is given as mean (SD), pathological repeat number is given as mean (range)

<table>
<thead>
<tr>
<th>AO</th>
<th>Pathological repeat</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>n=5</td>
<td>31.2 (6.3)</td>
<td>44.8 (12.9)</td>
</tr>
<tr>
<td>n=34</td>
<td>42.5 (10.9)</td>
<td>42.4 (11.5)</td>
</tr>
<tr>
<td>n=17</td>
<td>42.9 (11.6)</td>
<td>42.7 (13.9)</td>
</tr>
</tbody>
</table>

Model parameter estimates, r and r, from samples of sizes n, and n, was carried out by means of the z statistic,

\[
\bar{z} = \frac{[\tanh^{-1}(r_2) - \tanh^{-1}(r_1)]}{[1/(n_1-3)+1/(n_2-3)]} \tag{1}
\]

which follows a standard Gaussian distribution under the null hypothesis r = r. \(^7\) Logistic survival analysis of AO in HD patients was conducted as described by Krawczak et al. \(^8\)

Briefly, the annual hazard rate (that is, the probability of disease onset in any given year) is modelled as a logistic function of time (t), pathological repeat length (l), sex (s; 0=female, 1=male), and carriership of a given APOE genotype (c; 0=non-carrier, 1=carrier) by

\[
h(t, l, s, c) = \frac{e^{a_1 + a_2 l + a_3 s + a_4 c}}{1 + e^{a_1 + a_2 l + a_3 s + a_4 c}} \tag{2}
\]

where

\[
F(t, l, s, c) = b_0 + b_1 t + b_2 l + b_3 s + b_4 c \tag{3}
\]

is a linear combination of the parameters under study. The probability of AO= in a given patient with pathological repeat length l, sex s, and carrier status c, then equals

\[
h(a_1, l, s, c) = \prod_{k=1} F(t, l, s, c) \tag{4}
\]

which allows assessment of the significance of a given parameter for AO by means of a stepwise likelihood ratio test. Twice the log likelihood difference between any two nested models approximately follows a \(\chi^2\) distribution with degrees of freedom equal to the difference in parameter number.

Results

As previously reported, \(^4,6\) the pathologically increased CAG repeat number was found to be negatively correlated with AO of HD, irrespective of the sex of the transmitting parent (Pearson correlation coefficients: r=-0.771 maternal transmission, p<0.001; r=-0.751 paternal transmission, p<0.001). The length of the CAG repeat on non-disease chromosomes, however, was of significant influence upon AO only when paternally inherited (r=-0.304, p=0.016). For maternally derived normal CAG repeats, a similar, albeit much less pronounced, association (r=-0.152, NS) failed to attain statistical significance. The sex difference between the two coefficient correlations was not statistically significant (z=0.86, two sided p>0.3). No difference was apparent between the lengths of maternally and paternally inherited normal CAG repeats.

A strong gender specific influence on AO was observed for the ratio of pathological to normal CAG repeat number. Here, the negative correlation with AO was stronger when the normal allele was inherited from the father (r=-0.483, p<0.001) rather than from the mother (r=-0.327, p=0.014) but not in females (r=-0.039, NS). The sex difference was, however, only of borderline statistical significance (z=1.65, two sided p=0.06). Again, an inverse correlation was observed for the ratio of CAG repeat numbers (male r=-0.143, NS; female r=-0.488, p<0.001) for which the sex difference was statistically significant (z=2.17, two sided p=0.03).

APOE genotypes and mean AOs, classified by patient gender, are summarised in table 1. The APOE genotype distribution in HD patients was not significantly different from that of a control population from the same geographical area (data not shown). Only three genotypes (c,c, c,c) were present in sufficiently large numbers to allow statistical analysis (n=133). As can be inferred from table 1, genotypes c,c and c,c had no obvious effect on AO of HD. Males with genotype c,c, however, had significantly earlier onset than females of the same genotype (t=2.78, 15 df, 0.005<p<0.01) and than males with other genotypes (t=2.33, 51 df, 0.01<p<0.025). The difference in AO between c,c females and females of other APOE genotypes was of borderline statistical significance (t=1.32, 75 df, 0.05<p<0.1).

Inspection of table 1 indicates that the average pathological CAG repeat length was slightly higher among c,c males than among females of the same genotype. In order to test whether this discrepancy fully explains the observed difference in AO, data from the 17 c,c patients were subjected to logistic survival analysis (table 2). Two models were considered which both allowed for repeat length as a covariate of AO, but which differed with respect to whether the gender of a patient was considered (model I) or not (model II). Male c,c patients turned out to face an annual hazard rate that was twice as high on a logit scale than that of c,c females. Computation of twice the log likelihood difference between models I and II showed that this difference was highly

Table 2 Logistical survival analysis of age of onset in 17 HD patients with APOE genotype c,c

<table>
<thead>
<tr>
<th>Model parameter</th>
<th>Model I ML estimates</th>
<th>Model II ML estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>b_0</td>
<td>-22.200</td>
<td>-19.141</td>
</tr>
<tr>
<td>b_1</td>
<td>0.138</td>
<td>0.121</td>
</tr>
<tr>
<td>b_2</td>
<td>0.291</td>
<td>0.265</td>
</tr>
<tr>
<td>b_3</td>
<td>2.145</td>
<td>NC</td>
</tr>
<tr>
<td>ln(ML)</td>
<td>-58.367</td>
<td>-62.024</td>
</tr>
</tbody>
</table>

Model III ML estimates

| b_0 | -21.348 | -20.739 |
| b_1 | 0.154 | 0.147 |
| b_2 | 0.292 | 0.287 |
| b_3 | 1.309 | NC |
| ln(ML) | -191.482 | -193.890 |

Model IV ML estimates

| b_0 | -21.348 | -20.739 |
| b_1 | 0.154 | 0.147 |
| b_2 | 0.292 | 0.287 |
| b_3 | 1.309 | NC |
| ln(ML) | -191.482 | -193.890 |

ML: maximum likelihood; NC: not considered.
significant ($\chi^2=7.314, \text{1 df, } p=0.007$). Similarly, the annual hazard rate of male $\varepsilon2\varepsilon3$ carriers was 1.3 times higher on the logit scale than that of male HD patients with other APOE genotypes (table 3). This difference was also found to be statistically significant when tested by means of twice the log likelihood difference between the respective models, III and IV ($\chi^2=4.816, \text{1 df, } p=0.028$). Logistic survival analysis did not, however, show any significant difference between the annual hazard rates of female $\varepsilon2\varepsilon3$ carriers and non-carriers ($\chi^2=0.464, \text{1 df, } p>0.4$) which implies that their differences in AO may be explicable in terms of different pathological repeat length.

**Discussion**

In trinucleotide repeat disorders where the expanded repeat is expressed within a protein, the length of the normal allele is also critical for AO. The length of the normal allele in HD has been assessed by a number of different groups. An association of the normal repeat with AO has been reported previously in the Welsh population where an effect was found in maternally inherited HD, that is, when the normal repeat came from the father. The present study reproduces this finding and reports further data highlighting the importance of the normal repeat for AO. Undoubtedly, the phenomenon of anticipation in HD is related to the progressive expansion of the pathological CAG repeat upon transmission from parent to child, with greater expansions occurring in the paternal germline. In the 37–52 repeat range, Duyao et al found an additional sex of transmitting parent effect on AO which was independent of repeat length. It was observed that, on average, maternally inherited HD has AO two years later than paternally inherited HD with the same pathological repeat length, but no significant effect of the normal allele was apparent. Our finding of a negative correlation between AO and the length of the normal repeat in paternally derived cases, however, appears to imply an additional effect of the normal allele reducing AO.

The ratio of pathological to normal repeat length may be as important for AO as the length of the normal repeat itself; the larger the disparity between the two repeat lengths, the earlier AO (in paternally but not in maternally) inherited HD, and in female (but not in male) patients. The evidence that HD pathology proceeds at the protein level is strong and inclusions immunoreactive to huntingtin, the product of the HD gene, have been found both in mice transgenic for exon 1 of the HD gene with expanded repeats and in HD brain. Huntington interacts with molecules such as HAP1 and GAPDH in a polyglutamine tract length dependent manner in that longer repeats bind these proteins more strongly. If the disease phenotype in HD results from a greater binding affinity or higher dissociation constants preventing the release of interacting proteins from any complexes formed with the pathologically expanded polyglutamine tracts, then a longer normal polyglutamine tract (that is, a smaller ratio of normal to expanded CAG repeat) may compete better with the expanded protein, allowing fewer aberrant interactions and thus delaying AO. The observed sex specificity may reflect gender differences in the availability of various proteins for interaction. Alternatively, an imprinted modifier gene in linkage disequilibrium with HD might drive the interaction, thereby rendering the effect dependent upon sex of the transmitting parent.

While no sex difference in AO is observed for patients with APOE genotypes $\varepsilon3\varepsilon3$ or $\varepsilon3\varepsilon4$, the $\varepsilon2\varepsilon3$ genotype is associated with significantly earlier onset of HD in men than in women. Also, $\varepsilon2\varepsilon3$ males had significantly earlier onset than other male patients while for female patients this relationship appeared to be reversed. It cannot be excluded that this is because of linkage disequilibrium with a nearby gene, but APOE genotype itself is known to affect pathology in disease and traumatic injury. In AD and in head injury, for example, ApoE is associated with the deposition of $\beta$-amyloid and such deposition is more prevalent in carriers of the $\varepsilon4$ allele. Furthermore, a sex difference has been shown in prevalence and age of onset for AD: women with the $\varepsilon3\varepsilon4$ genotype have AO as early as $\varepsilon4\varepsilon4$ homozygotes while in men the $\varepsilon3\varepsilon4$ genotype gives the same (late) onset as $\varepsilon3\varepsilon3$. Carriership of $\varepsilon4$ has also been implicated in Pick disease, corticobasal degeneration, and progressive supranuclear palsy and motor neurone disease. These findings suggest that ApoE is likely to affect neuronal function in the brain over and above the formation of neurofilibrillary tangles and neuritic plaques, characteristic of AD. Expression of ApoE rises dramatically in response to brain injury and damage and it may be expressed in the damaged areas of the brains of HD patients or be associated with the observed inclusions. Although ApoE genotype is associated with late onset Alzheimer disease, the mechanism for this is unknown and thus it is difficult to assess what the mechanism behind this observed effect in HD might be.

We are grateful to the MRC for a centre grant in neuropsychiatric genetics. PK was supported by a University of Wales College of Medicine studentship; MK holds a Heisenberg grant (Ki 1093/3-2) from the Deutsche Forschungsgemeinschaft.


1093/5-2) from the Deutsche Forschungsgemeinschaft.