

## Original articles

## A molecular investigation of true dominance in Huntington's disease

Yolanda Narain, Andreas Wytttenbach, Julia Rankin, Robert A Furlong, David C Rubinsztein

## Abstract

**Huntington's disease (HD) is thought to show true dominance, since subjects with two mutant alleles have been reported to have similar ages at onset of disease compared to heterozygous sibs. We have investigated this phenomenon using a cell culture model. Protein aggregate formation was used as an indicator for pathology, as intraneuronal huntingtin inclusions are associated with pathology in vitro and in vivo. We showed that cytoplasmic and nuclear aggregates are formed by constructs comprising part of exon 1 of huntingtin with 41, 51, 66, or 72 CAG repeats, in a rate that correlates with repeat number. No inclusions were seen with 21 CAG repeat constructs. Mutant and wild type huntingtin fragments can be sequestered into inclusions seeded by a mutant huntingtin. Wild type huntingtin did not enhance or interfere with protein aggregation. The rate of protein aggregation was dose dependent for all mutant constructs tested. These experiments suggested a model for the dominance observed in HD; the decrease in the age at onset of a mutant homozygote may be small compared to the variance in the age at onset for that specific repeat number in heterozygotes. Our experiments also provide a model, which may explain the different repeat size ranges seen in patients and healthy controls for the different polyglutamine diseases.**

(J Med Genet 1999;36:739-746)

Keywords: Huntington's disease; huntingtin; CAG repeats; true dominance

Huntington's disease (HD) is an autosomal dominant neurodegenerative condition associated with abnormal movements, cognitive deterioration, and psychiatric symptoms. The disease typically presents in middle age. However, it can present at any age from early childhood to advanced age.<sup>1</sup> The causative mutation is a (CAG)<sub>n</sub> repeat expansion close to the 5' end of the coding part of the transcript, which is translated into an abnormally long polyglutamine tract in the huntingtin protein. Dis-

ease chromosomes have 36 or more uninterrupted CAG codons, while normal chromosomes are polymorphic in length with 35 or fewer CAG repeats. The age at onset of disease is inversely correlated with the number of repeats on disease chromosomes.<sup>1</sup>

Huntington's disease is a member of a family of neurodegenerative diseases caused by CAG/polyglutamine expansions, which include spinobulbar muscular atrophy, spinocerebellar ataxias (SCA) types 1, 2, 3, 6, and 7, and dentatorubral-pallidoluysian atrophy (DRPLA).<sup>2-4</sup> These diseases are probably caused by similar pathological processes, since they share a number of features. In all cases, disease severity is inversely correlated with repeat number. Genetic and transgenic experiments suggest that the polyglutamine expansion mutation causes disease by conferring a novel deleterious function on the mutant protein.<sup>5-8</sup> Recently, the discovery of intraneuronal protein aggregates (inclusions) containing the mutant proteins has provided a pathological common denominator in many of these diseases.<sup>3 9 10</sup>

To date, three articles have suggested that HD patients with two mutant alleles have a clinical presentation similar to their sibs with only one mutant allele.<sup>11-13</sup> This phenomenon of true dominance is almost unique in human genetics. The two early reports relied on inferences from linkage data.<sup>12 13</sup> Recently, Durr *et al*<sup>11</sup> reported the first cases with molecular confirmation, where an HD patient with 42 and 46 CAG repeat alleles had a similar presentation to his brother with 48 and 17 CAG repeat alleles.<sup>11</sup> In contrast to HD, patients with two mutant alleles in the SCA1, SCA2, SCA3, SCA6, and DRPLA genes have an earlier age at onset compared to subjects with one mutant allele.<sup>14-19</sup>

We have tried to find a molecular explanation for the apparent true dominance in HD, as it may provide insights into the disease process. A cell based model is a powerful tool to address this issue, since cells can be transfected with various doses and combinations of constructs with different CAG repeats driven by the same promoter. Many cell lines form inclusions in vitro when transfected with constructs expressing expanded polyglutamine tracts.<sup>3 20-23</sup> In

Department of Medical Genetics, Cambridge Institute for Medical Research, Wellcome/MRC Building, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2XY, UK  
Y Narain  
A Wytttenbach  
J Rankin  
R A Furlong  
D C Rubinsztein

Correspondence to: Dr Rubinsztein.

Revised version received 7 July 1999  
Accepted for publication 8 July 1999

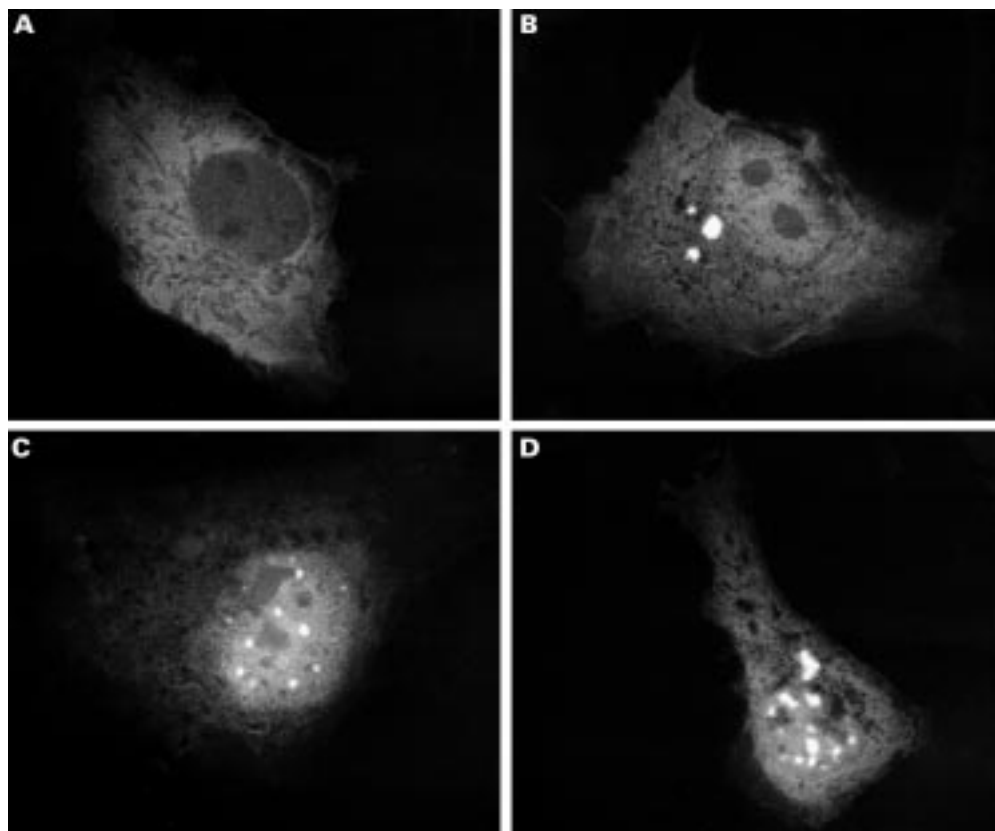
HD, DRPLA, and SCA3, these phenotypes only occur in cells expressing a fragment of the disease gene containing the repeats, and there are increasing data suggesting that an initial cleavage process generating a toxic fragment may be a critical step in the pathological process in some polyglutamine diseases.<sup>3 21-23</sup>

A causal role for inclusions in HD pathology has been suggested, as they appear before the signs of disease in a transgenic mouse expressing exon 1 of the HD gene with expanded repeats.<sup>10</sup> In this mouse model, the density and number of aggregates in the neuropil are highly correlated with the development of neurological symptoms.<sup>24</sup> In addition, the numbers of these inclusions in the cortex of HD patients correlates with CAG repeat number/age at onset.<sup>9</sup> Inclusion formation in vitro correlates with susceptibility to cell death.<sup>21 22</sup> Cell death occurred at similar rates in vitro, when huntingtin inclusions were targeted to the nucleus or the cytoplasm.<sup>25</sup> There is, however, some debate as to whether these inclusions cause the disease or are an epiphenomenon. Saudou *et al*<sup>26</sup> reported a dissociation between cell death and inclusion formation in primary cultures following various perturbations. These findings were not straightforward and may be compatible with a pathogenic role for huntingtin polymerisation.<sup>2 4</sup> Klement *et al*<sup>27</sup> deleted the self-association domain of a SCA1 transgene with mutant expanded repeats, which prevented inclusion formation, but nevertheless resulted in a SCA1-like phenotype in trans-

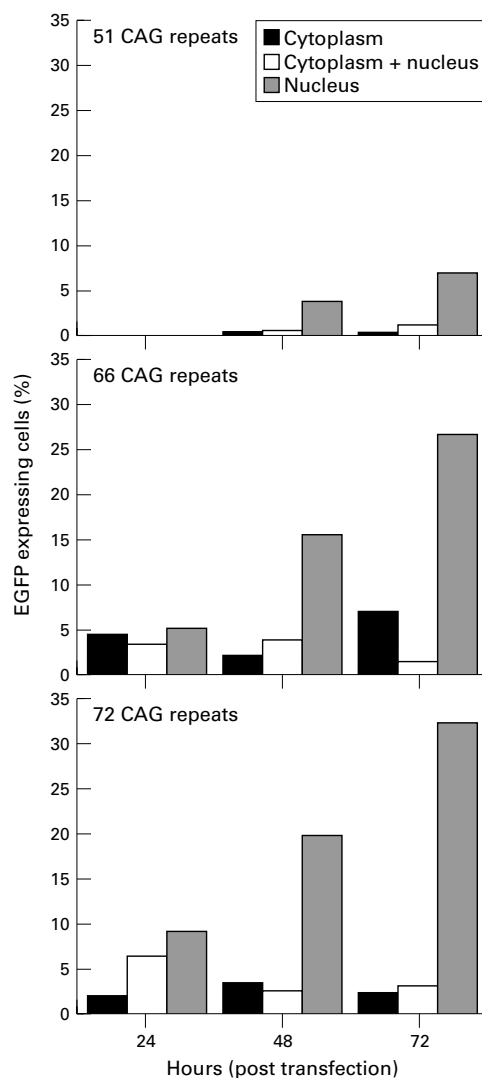
genic mice. However, it is not clear if this phenotype was the result of the deletion altering the protein structure,<sup>2</sup> since no data were presented for this deletion in association with the wild type allele. Furthermore, these mice do not show the progressive disease seen in mice expressing SCA1 transgenes with expanded repeats, including the self-association domain.<sup>28</sup> Since the latter mice develop inclusions, these aggregates may be necessary for disease progression. Even if inclusions are an epiphenomenon, in vitro and in vivo data suggest that inclusion formation correlates with the disease process and serves as a good surrogate for cell damage in vitro.

The true dominance observed in HD could be explained by an interaction between mutant and wild type HD gene products. Snell *et al*<sup>30</sup> reported a significant effect of the normal allele on the age at onset of HD in maternally inherited disease. This finding was compatible with earlier results of Farrer *et al*,<sup>31</sup> who reported a notable tendency of sibs to share the same age at onset if they inherited the same allele at D4S10, a locus linked to the HD gene. A recent report has shown that wild type ataxin-3 is recruited into aggregates formed by mutant ataxin-3.<sup>29</sup> Thus, we have tested if a similar phenomenon occurs in HD in vitro, and whether coexpression of wild type huntingtin enhances inclusion formation by the mutant form.

Additionally, we examined the dependency of inclusion formation on the expression levels



**Figure 1** COS-7 cells transfected with pEGFP-(CAG)<sub>21</sub> (A). Inclusions formed by pEGFP-(CAG)<sub>72</sub> in the cytoplasm only (B), in the nucleus only (C), and in the cytoplasm and the nucleus (D). COS7 cells were transfected with fragments of exon 1 of huntingtin containing 21 or 72 CAG repeats cloned downstream of GFP in pEGFP-C1. No inclusions were present with pEGFP-(CAG)<sub>21</sub>.



**Figure 2** Time course experiment. We used 1  $\mu$ g of pEGFP-(CAG)<sub>41</sub>-(CAG)<sub>51</sub>-(CAG)<sub>66</sub> or -(CAG)<sub>72</sub>. Cells were fixed 24, 48, or 72 hours after transfection. The proportions of GFP expressing cells with cytoplasmic inclusions (found only in the cytoplasm), nuclear inclusions (found only in the nucleus area), or cytoplasmic + nuclear inclusions are indicated (see Methods for details). At 24 hours, <1% of GFP expressing cells had inclusions with pEGFP-(CAG)<sub>51</sub>. Since very few inclusions were formed with pEGFP-(CAG)<sub>41</sub> (2% of GFP expressing cells had inclusions after 48 hours, 2.5% of GFP expressing cells had inclusions after 72 hours), these were not separated according to location.

of alleles with various CAG repeat numbers. Although transgenic experiments suggest a correlation between transgene expression and disease severity for SCA1, SCA3, and HD,<sup>10 32-35</sup> this association has not previously been investigated in cell culture models.

## Methods

### CONSTRUCTS

#### HD pEGFP-C1 constructs

Genomic DNA from subjects with 21, 41, 51, or 72 CAG repeats in the HD gene was used as a template for amplification of codons 8 to 57 inclusive of HD exon 1 (numbering applies to wild type protein with 23 glutamine residues). We used primers HDF-5' ATG AAG GCC TTC GAG TCC CTC AAG TCC TTC 3' and

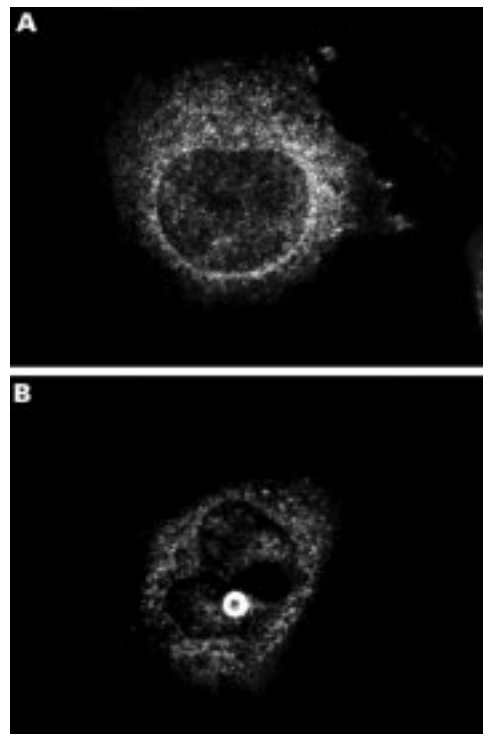
HU3-5' GGC GGC TGA GGA AGC TGA GAA 3'. Aliquots of these PCR products were reamplified with primers HM1 (5' G CGC AGA TCT ATG AAG GCC TTC GAG TCC CTC AAG TCC TTC 3') and HM3 (5' G CGC GAA TTC GGC GGC TGA GGA AGC TGA GGA 3'), which have *Bgl*II and *Eco*RI restriction sites incorporated (underlined) to allow cloning into pEGFP-C1 (4.2 kb) (Clontech, Palo Alto, USA).

#### HD pHM6 constructs

The HD pEGFP constructs were used as templates for amplification with primers HDF-HindIII (5' A CTG AAG CTT ATG AAG GCC TTC GAG TCC CTC AAG TCC TTC 3') and HM3, to enable cloning into the *Hind*III and *Eco*RI sites of pHM6 (5.4 kb) (Roche Diagnostics). All constructs were validated by sequencing. DNA concentrations were determined spectrophotometrically and were consistent with concentrations derived from gel estimates. Both pEGFP and pHM6 have the cytomegalovirus immediate early promoter sequence.

#### CELL CULTURE AND TRANSIENT TRANSFECTION

Green monkey kidney cells (COS-7) were maintained in Dulbecco's modified Eagle's medium (DMEM, Sigma, Dorset, UK) supplemented with 100 IU/ml penicillin and streptomycin, 2 mmol/l L-glutamine, 1 mmol/l sodium pyruvate, and 10% fetal bovine serum (FBS) (Life Technologies, Paisley, UK) at 37°C, 5% CO<sub>2</sub>.



**Figure 3** COS-7 cells transfected with (A) pHM6-(CAG)<sub>21</sub> and (B) pHM6-(CAG)<sub>72</sub>. Note the ring-like shape of some of the inclusions formed by pHM6-(CAG)<sub>72</sub>; 1  $\mu$ g of each construct was used and cells were stained after 24 hours. They were visualised using a mouse monoclonal anti-HA antibody and a goat anti-mouse secondary antibody conjugated to Texas Red. No inclusions were present for pHM6-(CAG)<sub>21</sub>.

COS-7 cells were grown to 60-80% confluence on sterile coverslips in six well plates for 24 hours before transfection. Cells were treated with LipofectAMINE Reagent™ (Life Technologies, Paisley, UK) at 10  $\mu$ l/well and plasmid DNA at 1.6  $\mu$ g/well for five hours in serum free medium. At 5 hours the medium was supplemented with 20% FBS. In the time course experiment we used 1  $\mu$ g of plasmid DNA per well.

At the end of each experiment, the cells were washed with PBS and fixed with 4% paraformaldehyde in 1  $\times$  PBS for 30 minutes; the coverslips were mounted in Citifluor medium (Citifluor, London, UK) containing DAPI (3  $\mu$ g/ml) and examined by fluorescence microscopy. Transfection efficiencies were similar for the pEGFP-C1 constructs with different CAG repeat lengths. The total proportion of GFP expressing cells with inclusions was determined by counting a minimum of 220 GFP expressing cells. Inclusions were defined by fluorescence microscopy as green aggregates that were not seen in cells expressing the empty pEGFP vector. A minimum of 50 cells with inclusions were counted to determine the proportions of inclusion containing cells with only nuclear, nuclear and cytoplasmic, and only cytoplasmic inclusions. Inclusions were scored as nuclear if they were completely surrounded by DAPI staining. Cells were counted in random fields with slides blinded to the observer, using the high magnification objective (100  $\times$ ). Statistics were performed using Statview for Windows (version 4.5).

#### IMMUNOCYTOCHEMISTRY

Cells transfected with pEGFP vectors can be analysed by fluorescent microscopy without using immunocytochemical staining procedures. Cells transfected with pHM6 vectors were washed twice with 1  $\times$  phosphate buffered saline (PBS), then fixed with ice cold methanol for 5-10 minutes. Cells were then washed three times with 1  $\times$  PBS and incubated in blocking buffer (5% fetal bovine serum in 1  $\times$  PBS) for 10 minutes at room temperature. Cells were incubated with primary antibody (monoclonal mouse anti-HA.11 monoclonal (BabCO, Richmond, USA) at 1:100 for one hour at room temperature, then washed three times with 1  $\times$  PBS followed by incubation with goat anti-mouse IgG conjugated to Texas Red (Molecular Probes, Eugene, US) at 1:100 for one hour at room temperature. Cells were finally washed with 1  $\times$  PBS three times and air dried and mounted in antifadent (Citifluor, London, UK) supplemented with DAPI (3  $\mu$ g/ml). Confocal microscopy was used to study colocalisation of protein products of pEGFP- and pHM6-huntingtin constructs.

#### Results

##### HD EXON 1 CONSTRUCTS CONTAINING EXPANDED REPEATS FORM INCLUSIONS IN COS-7 CELLS

COS-7 cells were transfected for 24, 48, and 72 hours with constructs comprising the N-terminus of huntingtin (amino acids 8 to 57, numbered as if the protein has 23 glutamine residues) with 21, 41, 49 or 51, 66, and 72 uninterrupted CAG repeats. A frag-

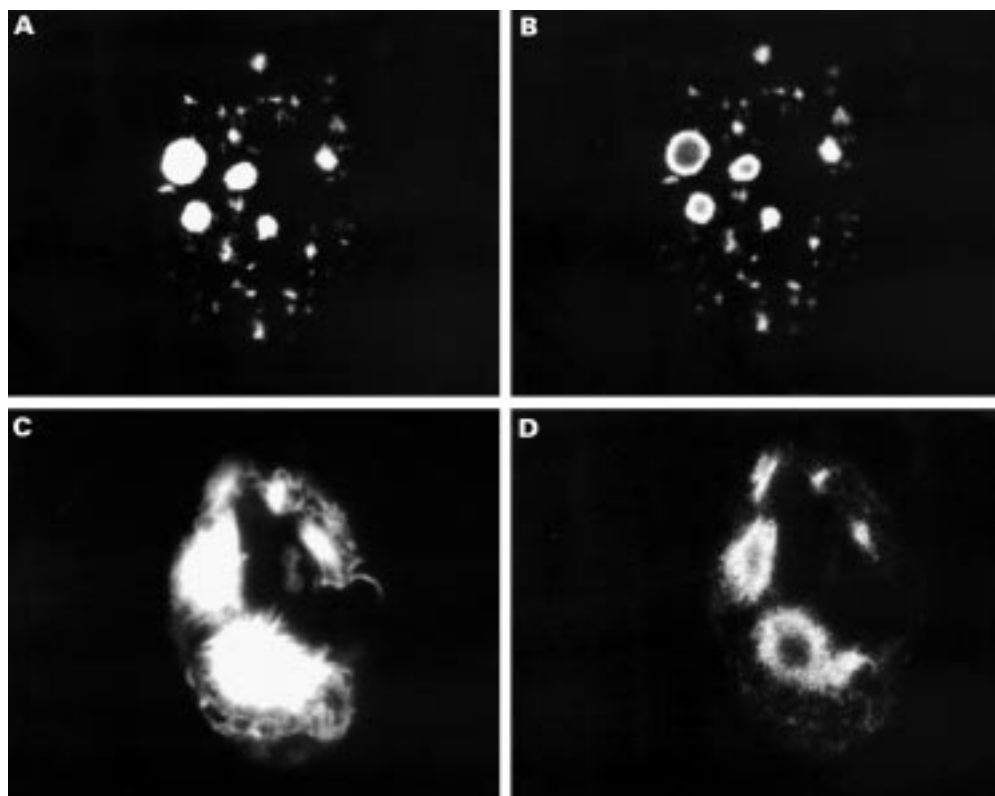
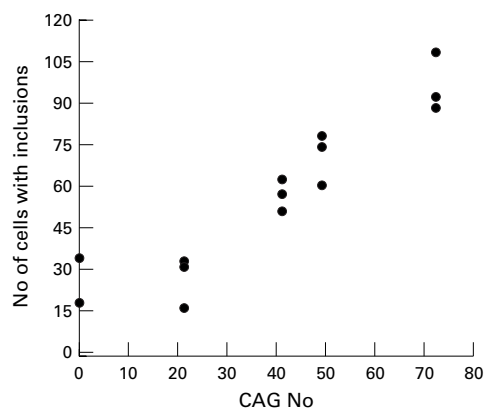


Figure 4 Colocalisation of proteins formed by pEGFP-(CAG)<sub>21</sub> with inclusions seeded by pHM6-(CAG)<sub>21</sub> (A, B) and pHM6-(CAG)<sub>21</sub> with inclusions seeded by pEGFP-(CAG)<sub>21</sub> (C, D) after 24 hours. Panels A and C show the GFP signals and B and D show staining against the HA tag of pHM6 labelled with Texas Red; 1  $\mu$ g of each construct was used for transfection and cells were fixed and stained after 24 hours. These figures have been deliberately taken with a low background to facilitate assessment of colocalisation.



**Figure 5** Cotransfection of COS-7 cells with 0.4  $\mu$ g pEGFP-(CAG)<sub>72</sub> and 1.2  $\mu$ g of either pHM6-(CAG)<sub>21</sub>, -(CAG)<sub>41</sub>, -(CAG)<sub>49</sub>, -(CAG)<sub>72</sub>, or empty pHM6 for 24 hours. The numbers of GFP expressing cells with GFP inclusions are shown. A total of 220 GFP expressing cells were counted for three independent experiments. The numbers of inclusions in cells cotransfected with pEGFP-(CAG)<sub>72</sub> and empty pHM6 are indicated at 0 CAG repeats, where there are two points with 18 inclusions. GFP inclusion formation was dependent on CAG repeat number in pHM6-huntingtin constructs ( $p < 0.0001$ ,  $R^2 = 0.919$ , regression for 41-72 repeats).

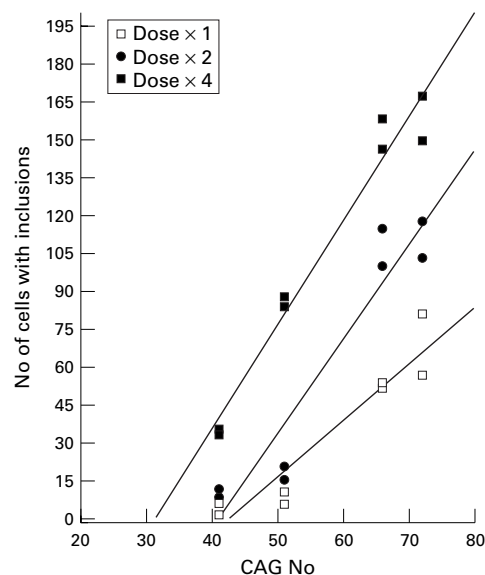
ment of huntingtin was used, as previous groups have shown that such fragments form inclusions in transient transfection experiments, while this does not occur with much larger fragments or the full length protein with expanded repeats.<sup>21-23</sup> These fragments were either cloned downstream of enhanced green fluorescent protein (EGFP) in pEGFP-C1 or in pHM6, which has a haemagglutinin (HA) tag at the N-terminus. We will refer to the constructs as pEGFP-(CAG)<sub>21</sub>, -(CAG)<sub>41</sub>, -(CAG)<sub>51</sub>, -(CAG)<sub>66</sub>, -(CAG)<sub>72</sub> and pHM6-(CAG)<sub>21</sub>, -(CAG)<sub>41</sub>, -(CAG)<sub>49</sub>, -(CAG)<sub>72</sub>. No inclusions were detected after 24, 48, and 72 hours in cells transfected with either pEGFP-(CAG)<sub>21</sub> or pHM6-(CAG)<sub>21</sub>.

Vectors containing 41 or more CAG repeats formed nuclear or cytoplasmic inclusions or both. Fig 1 shows examples of cells with only nuclear, only cytoplasmic, and both nuclear and cytoplasmic inclusions formed by pEGFP-(CAG)<sub>72</sub>. The proportions of GFP expressing cells with inclusions increased with time after transfection and were positively correlated with CAG repeat number (fig 2). We have observed similar relationships in PC12 and SH-SY5Y cells (data not shown). An increasing proportion of inclusions was found in the nucleus as a function of time post-transfection for pEGFP-(CAG)<sub>41</sub>, -(CAG)<sub>51</sub>, -(CAG)<sub>66</sub> and -(CAG)<sub>72</sub>. Since nuclear and cytoplasmic inclusions have been associated with a similar toxicity<sup>25</sup> in a cellular model of HD, we determined the total proportion of GFP expressing cells with inclusions in subsequent analyses. The fluorescence images of inclusions formed by the two set of vectors looked different; the larger inclusions formed by pHM6 vectors often have a ring shaped appearance (fig 3) in contrast to the inclusions formed by pEGFP vectors, which fluoresce uniformly.

#### HUNTINGTIN CONSTRUCTS WITH PATHOLOGICAL REPEATS SEQUESTER EXON 1 FRAGMENTS WITH WILD TYPE OR MUTANT REPEAT LENGTHS

Cells were cotransfected with pHM6-(CAG)<sub>72</sub> and pEGFP-(CAG)<sub>21</sub> or pEGFP-(CAG)<sub>72</sub>. Inclusions formed by pHM6-(CAG)<sub>72</sub> sequestered both pEGFP-(CAG)<sub>21</sub> and pEGFP-(CAG)<sub>72</sub> (fig 4A, B). Alternatively, cells were cotransfected with pEGFP-(CAG)<sub>72</sub> and pHM6-(CAG)<sub>21</sub> or pHM6-(CAG)<sub>72</sub>. The inclusions formed by pEGFP-(CAG)<sub>72</sub> sequestered both pHM6-(CAG)<sub>21</sub> and pHM6-(CAG)<sub>72</sub> (fig 4C, D). Cells transfected with a 1:1 ratio of pEGFP-(CAG)<sub>72</sub> and pHM6-(CAG)<sub>21</sub> showed inclusions labelled with both GFP and Texas Red in less than 50% GFP inclusion containing cells. However, when we counted 20 cells transfected with a 1:1 ratio of pEGFP-(CAG)<sub>21</sub> and pHM6-(CAG)<sub>72</sub>, all inclusions showed colocalisation of both constructs. Confocal colour analysis confirmed colocalisation (data not shown) and no GFP signal was detected coming through the red filter. No GFP inclusions were seen in cells cotransfected with pHM6-(CAG)<sub>72</sub> and empty pEGFP vector or emerlin tagged EGFP (data not shown).

Since pEGFP-(CAG)<sub>21</sub> is efficiently sequestered into inclusions seeded by an excess of pHM6-(CAG)<sub>72</sub>, we examined if inclusions are formed at similar rates in "heterozygote" and "homozygote" HD cells. Cells were cotransfected with 0.4  $\mu$ g of pEGFP-(CAG)<sub>72</sub> and 1.2  $\mu$ g of either pHM6-(CAG)<sub>21</sub>, -(CAG)<sub>41</sub>, -(CAG)<sub>49</sub>, -(CAG)<sub>72</sub>, or empty pHM6. A three-fold excess of the pHM6-huntingtin constructs versus pEGFP-(CAG)<sub>72</sub> was used to ensure



**Figure 6** COS-7 cells transfected with three different doses of pEGFP-(CAG)<sub>41</sub>, -(CAG)<sub>51</sub>, -(CAG)<sub>66</sub> or -(CAG)<sub>72</sub> for 72 hours. For each construct we transfected cells with three doses (dose  $\times$  1: 0.4  $\mu$ g, dose  $\times$  2: 0.8  $\mu$ g, and dose  $\times$  4: 1.6  $\mu$ g) of pEGFP-(CAG)<sub>41</sub>, -(CAG)<sub>51</sub>, -(CAG)<sub>66</sub> and -(CAG)<sub>72</sub>. The total amount of DNA transfected was maintained constant (2  $\mu$ g) by making up the difference with pEGFP empty vector. At all doses there was a strong linear correlation between CAG repeat length and inclusion formation: dose  $\times$  1:  $p < 0.0001$ ,  $R^2 = 0.896$ ; dose  $\times$  2:  $p = 0.0002$ ,  $R^2 = 0.910$ ; dose  $\times$  4:  $p < 0.0001$ ,  $R^2 = 0.967$ . Inclusion formation was dose dependent ( $p < 0.0001$ , multiple regression taking CAG repeats into account).

that practically all cells expressing GFP also expressed pHM6. The proportion of GFP expressing cells with inclusions was positively correlated with pHM6-CAG repeat length ( $p < 0.0001$ ,  $R^2 = 0.919$ ) (fig 5).

It was notable that similar proportions of inclusions were observed when cotransfecting pEGFP-(CAG)<sub>72</sub> with either empty pHM6 or pHM6-(CAG)<sub>21</sub> (fig 5). This experiment also suggests a strong dose dependence on inclusion formation in vitro, since cells transfected with pEGFP-(CAG)<sub>72</sub> and empty pHM6 make fewer inclusions (consistent with their lower mutant huntingtin expression) than cells transfected with pEGFP-(CAG)<sub>72</sub> and pHM6-(CAG)<sub>72</sub>.

#### HUNTINGTIN INCLUSION FORMATION IS DOSE DEPENDENT AT ALL CAG REPEAT LENGTHS

In order to investigate the relationship between the expression level of mutant huntingtin exon 1 on inclusion formation in vitro at different CAG repeat lengths, we transfected cells with three doses (dose  $\times$  1 (0.4  $\mu$ g), dose  $\times$  2 (0.8  $\mu$ g), and dose  $\times$  4 (1.6  $\mu$ g) of pEGFP-(CAG)<sub>413</sub>, -(CAG)<sub>513</sub>, -(CAG)<sub>663</sub>, and -(CAG)<sub>72</sub> (fig 6). The total amount of DNA transfected in each case was maintained constant (2  $\mu$ g), making up the difference with pEGFP empty vector. (We have previously shown that this approach, where one modifies the concentration of the target vector in a fixed amount of total transfected DNA, gives reproducible expression, which correlates with target vector concentration<sup>36</sup>.) At all doses, there was a strong linear correlation between CAG repeat length and the proportion of GFP expressing (that is, transfected) cells with inclusions (fig 6) (dose 1:  $p < 0.0001$ ,  $R^2 = 0.896$ ; dose 2:  $p = 0.0002$ ,  $R^2 = 0.910$ ; dose 4:  $p < 0.0001$ ,  $R^2 = 0.967$ ). The dose dependency of the huntingtin repeat construct on inclusion formation was not saturated even at 72 CAG repeats.

#### Discussion

Three previous reports have suggested that HD may show true dominance, since subjects with two mutant alleles had similar ages of onset compared to their heterozygous sibs<sup>11-13</sup> (see Introduction). Despite these limited data, the validity of their conclusions has not been widely challenged. We have investigated this phenomenon using a cellular model. The interpretation of our data assumes that inclusion formation is an appropriate marker for cell pathology. The use of inclusions as a marker is appropriate as long as their rate of formation correlates with the pathological process, even if inclusions are merely innocent bystanders or represent a protective response. This correlation is supported by studies in HD brains,<sup>9</sup> transgenic mice,<sup>24</sup> and in cell culture.<sup>20-23</sup> We are not aware of any study which has attempted to investigate the molecular basis for the "true dominance" reported in HD, and reconcile this phenomenon with the dose dependency seen in other polyglutamine diseases and in transgenic models of SCA1, SCA3, and HD.<sup>10 32-35</sup>

Our model shows that the rate of inclusion formation is CAG repeat dependent, like

previous models of polyglutamine diseases.<sup>20-22</sup> We observed an increase in the proportion of inclusion containing cells with only nuclear inclusions with time. While this is consistent with the suggestion of Lunke and Mandel<sup>37</sup> that cytoplasmic inclusions may be the precursors of nuclear inclusions, it is important to bear in mind that our studies have only examined snapshots in time in a population of cells. This issue needs further clarification using longitudinal studies of single cells.

While pEGFP-(CAG)<sub>21</sub> was found in virtually all inclusions seeded by pHM6-(CAG)<sub>72</sub>, pHM6-(CAG)<sub>21</sub> was detected in <50% of inclusions seeded by pEGFP-(CAG)<sub>72</sub>. In addition, the inclusions found by pHM6 constructs were ring shaped, compared to the uniformly fluorescent EGFP inclusions. This could be because the HA epitope may occasionally be buried in inclusions preventing antibody recognition. Alternatively, the strong fluorescent signal of EGFP may allow greater sensitivity of recognition. Thus, despite their different appearances, the pEGFP- and pHM6-huntingtin inclusions may have similar morphologies.

#### INCLUSION FORMATION IS NOT ENHANCED BY THE NORMAL GENE PRODUCT

Since pEGFP-(CAG)<sub>21</sub> was efficiently recruited into inclusions formed by pHM6-(CAG)<sub>72</sub>, we examined the rates of GFP inclusion formation in "heterozygote" cells (expressing pEGFP-(CAG)<sub>72</sub> and pHM6-(CAG)<sub>21</sub>) and "homozygote" cells (expressing pEGFP-(CAG)<sub>72</sub> and pHM6-(CAG)<sub>413</sub>, -(CAG)<sub>493</sub>, and -(CAG)<sub>72</sub>). No previous studies have tested if wild type huntingtin enhances inclusion formation by mutant huntingtin. The overall rate of inclusions formed by the pEGFP-(CAG)<sub>72</sub> construct was neither enhanced nor reduced by the presence of cotransfected pHM6-(CAG)<sub>21</sub>, compared to the empty vector. The absence of any effect is compatible with our failure to replicate the previous report,<sup>31</sup> suggesting an effect of the wild type allele on the age at onset of HD.<sup>38</sup> In addition, the proportion of inclusion containing cells depended on the CAG repeat length of the cotransfected mutant pHM6 construct. Consequently, one would expect more inclusions in HD homozygotes than in heterozygotes.

#### INCLUSION FORMATION CORRELATES WITH EXPRESSION LEVELS OF THE MUTANT ALLELE

Inclusion formation in cells transfected with pEGFP-huntingtin constructs was dependent on the expression levels of all mutant CAG lengths tested ( $p < 0.0001$ ; multiple regression analysis, taking CAG repeats into account). This suggests that our system was not saturated at expression levels many fold higher than physiological levels. These data from a cell model complement a recent study, which showed that inclusion formation in vitro experiments using purified recombinant huntingtin N-terminal peptides was dose dependent.<sup>39</sup> Thus, one would anticipate a dose dependency in vivo, unless there were a rate

limiting step upstream of inclusion formation in disease pathogenesis. For instance, if the huntingtin cleavage reaction (see Introduction) were saturated, then the rate of toxic fragment production may be similar for both homozygotes and heterozygotes. However, transgenic mice expressing full length huntingtin cDNA with expanded repeats show an earlier onset of symptoms in homozygotes (which express five times the endogenous levels of huntingtin), compared to heterozygotes.<sup>33</sup> A relationship between increased mutant gene expression and earlier onset of disease was also recently reported in a YAC transgenic model of HD, which expresses the full length gene.<sup>35</sup> In the polyglutamine disease SCA3, which does not show true dominance,<sup>16,19</sup> the SCA3 protein may also require cleavage to form a toxic fragment.<sup>23</sup>

We suggest the following reasoning to reconcile the available molecular and transgenic data with the true dominance observed clinically. The repeat size range that is likely to be observed in a homozygote is constrained by the repeat sizes that will allow his or her parents to reach reproductive age. Most HD patients will have 40–48 CAG repeats, which generally does not change when maternally inherited, but tends to increase by about four repeats when paternally inherited.<sup>40</sup> Within the range of repeat sizes expected in an HD homozygote, the decrease in the age at onset conferred by a double dose of the mutant gene may be small, relative to the variance in the age at onset expected for each repeat size. Thus, a typical HD homozygote with two identically sized alleles may have mean ages at onset only 5–10 years earlier than a heterozygote with a similar allele. This decrease will be difficult to detect because the 95% confidence intervals for age at onset for a given repeat size in heterozygotes range from 20–30 years<sup>41</sup> and the age at onset of a disease with an insidious onset is difficult to pinpoint. These factors are important to consider when assessing publications comprising only three papers,<sup>11–13</sup> where CAG repeat data are confined to one homozygote and a heterozygous sib.<sup>11</sup> In this report, the homozygote's two mutant alleles were both smaller than the mutation in his sib. Our proposed dosage effect may not be noticed, unless studied systematically in large numbers of cases documented with CAG repeat numbers.

Another mystery is why the different CAG/polyglutamine diseases have different normal and disease size ranges,<sup>3,4</sup> but similar shaped curves of repeat size versus age at onset. This may be related to the protein context in which the expanded polyglutamines are found, and also to the expression levels, as suggested by the different dose response curves in fig 6. If all CAG/polyglutamine diseases (with the possible exception of SCA6) have shared molecular mechanisms, variation in expression of the relevant genes in the target tissues could effectively alter the smallest repeat size sufficient to cause disease. If one were to extrapolate the curves in fig 6, these would intercept the X axis at 31 repeats for the 1.6 µg dose, 41 repeats for the 0.8 µg dose, and 43 repeats for the 0.4 µg

dose. If this in vitro model is a reflection of the in vivo processes, then the X axis intercepts (equivalent to zero inclusions) would relate to the lowest repeat size causing disease at the relevant expression level at a common time (equivalent to a similar age).

In conclusion, if this cell model reflects the in vivo situation, then the apparent true dominance of HD cannot be accounted for by a role of the normal allele in inclusion formation. There is a dose dependency for inclusion formation in vitro at all CAG lengths up to 72 repeats. The apparent dominance in HD patients could be because of the small effect of doubling the dose of the mutant gene, compared to the variance in age at onset for heterozygotes.

Drs Jenny Carmichael and Jina Swartz are thanked for technical assistance. We are grateful to The Rehabilitation and Medical Research Trust, Friends of Peterhouse, The University of Cambridge Department of Pathology, and The Isaac Newton Trust for supporting this work. AW is grateful to the Swiss National Science Foundation and the Hereditary Disease Foundation for Research Fellowships, RAF is a Peterhouse Senior Research Fellow in Neuroscience, and DCR is a Glaxo Wellcome Research Fellow.

- Ross CA, Hayden MR. Huntington's disease In: Rubinsztein DC, Hayden MR, eds. *Analysis of triplet repeat disorders*. Oxford: Bios Scientific, 1998:169–208.
- Perutz MF. Glutamine repeats and neurodegenerative diseases: molecular aspects. *Trends Biochem Sci* 1999;24:58–63.
- Paulson HL. Protein fate in neurodegenerative proteinopathies: polyglutamine diseases join the (mis) fold. *Am J Hum Genet* 1999;64:339–45.
- Rubinsztein DC, Wyttenbach A, Rankin J. Intracellular inclusions, pathological markers in diseases caused by expanded polyglutamine tracts? *J Med Genet* 1999;36:265–70.
- Duyao MP, Auerbach AB, Ryan A, et al. Inactivation of the mouse Huntington's disease gene homolog Hdh. *Science* 1995;269:407–10.
- Nasir J, Floresco SB, O'Kusky JR, et al. Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. *Cell* 1995;81:811–23.
- White JK, Auerbach W, Duyao MP, et al. Huntingtin is required for neurogenesis and is not impaired by the Huntington's disease CAG expansion. *Nat Genet* 1997;17:404–10.
- Zeitlin S, Liu JP, Chapman DL, Papaioannou VE, Efstratiadis A. Increased apoptosis and early embryonic lethality in mice nullizygous for the Huntington's disease gene homologue. *Nat Genet* 1995;11:155–63.
- Becher MW, Kotzok JA, Sharp AH, et al. Intracellular neuronal inclusions in Huntington's disease and dentatorubral and pallidolysian atrophy: correlation between the density of inclusions and IT15 CAG triplet repeat length. *Neurobiol Dis* 1997;4:1–11.
- Davies SW, Turmaine M, Cozens BA, et al. Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell* 1997;90:537–48.
- Durr A, Hahn-Barm V, Brice A, Pecheux C, Dode C, Feingold J. Homozygosity in Huntington's disease. *J Med Genet* 1999;36:172–3.
- Myers RH, Leavitt J, Farrer LA, et al. Homozygote for Huntington disease. *Am J Hum Genet* 1989;37:511–23.
- Wexler NS, Young AB, Tanzi RE, et al. Homozygotes for Huntington disease. *Nature* 1987;326:194–7.
- Goldfarb LG, Vasconcelos O, Platonov FA, et al. Unstable triplet repeat and phenotype variability of spinocerebellar ataxia type I. *Ann Neurol* 1996;39:500–6.
- Matsumura R, Futamura N, Fujimoto Y, et al. Spinocerebellar ataxia type 6. Molecular and clinical features of 35 Japanese patients including one homozygous for the CAG repeat expansion. *Neurology* 1997;49:1238–43.
- Lerer I, Merims D, Abeliovich D, Zlotgora J, Gadoth N. Machado-Joseph disease: correlation between the clinical features, the CAG repeat length and homozygosity for the mutation. *Eur J Hum Genet* 1996;4:3–7.
- Sanpei K, Takano H, Igarashi S, et al. Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. *Nat Genet* 1996;14:277–84.
- Sato K, Kashihara K, Okada S, et al. Does homozygosity advance the onset of dentato-rubral-pallidolysian atrophy? *Neurology* 1995;45:1934–6.
- Sobue G, Doyu M, Nakao N, et al. Homozygosity for Machado-Joseph disease gene enhances phenotypic severity. *J Neurol Neurosurg Psychiatry* 1996;60:354–6.
- Cooper JK, Schilling G, Peters MF, et al. Truncated N-terminal fragments of huntingtin with expanded

- glutamine repeats form nuclear and cytoplasmic aggregates in cell culture. *Hum Mol Genet* 1998;7:783-90.
- 21 Igarashi S, Koide R, Shimohata T, et al. Suppression of aggregate formation and apoptosis by transglutaminase inhibitors in cells expressing truncated DRPLA protein with an expanded polyglutamine stretch. *Nat Genet* 1998;18:111-17.
  - 22 Martindale D, Hackam A, Wiczorek A, et al. Length of huntingtin and its polyglutamine tract influences localization and frequency of intracellular aggregates. *Nat Genet* 1998;18:150-4.
  - 23 Paulson HL, Perez MK, Trotter Y, et al. Intranuclear inclusions of expanded polyglutamine protein in spinocerebellar ataxia type 3. *Neuron* 1997;19:333-44.
  - 24 Li H, Li SH, Cheng AL, Mangiarini L, Bates GP, Li XJ. Ultrastructural localisation and progressive formation of neuropil aggregates in Huntington's disease transgenic mice. *Hum Mol Genet* 1999;8:1227-36.
  - 25 Hackam AS, Singaraja R, Zhang T, Gan L, Hayden MR. In vitro evidence for both the nucleus and cytoplasm as subcellular sites of pathogenesis in Huntington's disease. *Hum Mol Genet* 1999;8:25-33.
  - 26 Saudou F, Finkbeiner S, Devys D, Greenberg ME. Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear inclusions. *Cell* 1998;95:55-66.
  - 27 Klement IA, Skinner PJ, Kaytor MD, et al. Ataxin-1 nuclear localisation and aggregation: role in polyglutamine-induced disease in SCA1 transgenic mice. *Cell* 1998;95:41-53.
  - 28 Orr HT, Skinner PA, Klement CJ, Cummings CJ, Zoghbi HJ. The role of ataxin-1 nuclear expression and aggregates in SCA1 pathogenesis. *Am J Hum Genet* 1998;63(suppl):A8.
  - 29 Perez MK, Paulson HL, Pendse SJ, Saionz SJ, Bonini NM, Pittman RN. Recruitment and the role of nuclear localisation in polyglutamine-mediated aggregation. *J Cell Biol* 1998;143:1457-70.
  - 30 Snell RG, MacMillan JC, Cheadle JP, et al. Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. *Nat Genet* 1993;4:393-7.
  - 31 Farrer LA, Cupples LA, Wiater P, Conneally PM, Gusella JF, Myers RH. The normal HD allele, or a closely linked gene, influences the age at onset of Huntington disease. *Am J Hum Genet* 1993;53:125-30.
  - 32 Burright EN, Clark HB, Servadio A, et al. SCA1 transgenic mice: a model for neurodegeneration caused by an expanded CAG trinucleotide repeat. *Cell* 1995;82:937-48.
  - 33 Reddy PH, Williams M, Charles V, et al. Behavioural abnormalities and selective neuronal loss in HD transgenic mice expressing mutated full-length HD cDNA. *Nat Genet* 1998;20:198-202.
  - 34 Warrick JM, Paulson HL, Gray-Board GL, et al. Expanded polyglutamine protein forms nuclear inclusions and causes neural degeneration in *Drosophila*. *Cell* 1998;93:939-49.
  - 35 Hodgson JG, Agopyan N, Gutekunst CA, et al. A YAC mouse model for Huntington's disease with full-length mutant huntingtin, cytoplasmic toxicity, and selective striatal neurodegeneration. *Neuron* 1999;23:181-92.
  - 36 Coles R, Caswell R, Rubinsztein DC. Functional analysis of the Huntington's disease (HD) gene promoter. *Hum Mol Genet* 1998;7:791-800.
  - 37 Lunke A, Mandel JL. A cellular model that recapitulates major pathogenic steps of Huntington's disease. *Hum Mol Genet* 1998;7:1355-61.
  - 38 Rubinsztein DC, Leggo J, Chiano M, et al. Genotypes at the GluR6 kinate receptor locus are associated with variation in the age of onset of Huntington disease. *Proc Natl Acad Sci USA* 1997;94:3872-6.
  - 39 Scherzinger E, Sittler A, Schweiger K, et al. Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology. *Proc Natl Acad Sci USA* 1999;96:4604-9.
  - 40 Ranen NG, Stine OC, Abbott MH, et al. Anticipation and instability of IT-15 (CAG)<sub>N</sub> repeats in parent-offspring pairs with Huntington disease. *Am J Hum Genet* 1995;57:593-602.
  - 41 Brinkman RR, Mezei MM, Theilmann J, Almqvist E, Hayden MR. The likelihood of being affected with Huntington disease by a particular age, for a specific CAG size. *Am J Hum Genet* 1997;60:1202-10.