Analysis of triplet repeats in the huntingtin gene in Japanese families affected with Huntington’s disease

Naoki Masuda, Jun Goto, Nobuhiro Murayama, Masahiko Watanabe, Ikuko Kondo, Ichiro Kanazawa

Abstract
Huntington’s disease (HD) is a progressive neurodegenerative disorder which is clinically characterised by chorea, cognitive decline, and emotional disturbance; it is inherited in an autosomal dominant manner. An unstable and abnormally expanded CAG repeat in the huntingtin (HTT) gene has been shown to cause the disease by the Huntington’s Disease Collaborative Research Group. The frequency of HD is approximately 3 to 8 per 100,000 in most western populations; however, the frequency is less than one-tenth in Japan.13

There is another triplet sequence, a CCG repeat, immediately 3’ adjacent to the CAG repeat in the huntingtin gene. The triplet sequence is also polymorphic, alleles of seven or 10 repeats are predominant in populations, and strong linkage disequilibrium between the former allele and HD has been shown in white populations.9,13 As one hypothesis which could explain this observation it has been proposed that the allele of seven repeats could be critical in Huntington’s disease.

Materials and methods
FAMILY STUDIES
The HD families analysed had been examined by neurologists or psychiatrists or both. Information on clinical status, sex, age at onset, and pedigree structure was recorded for all HD families. Ascertainment of a patient was based on characteristic clinical features and family history. The age of onset was defined as the age at which choreic movements or psychiatric impairment were first observed.

DNA ISOLATION
Genomic DNA was isolated from either lymphoblastoid cell lines or the buffy coat layer of fresh blood by standard extraction methods.13

PCR ANALYSIS
PCR amplification of the triplet repeat sequence in the huntingtin gene was performed using the primers HD-1, HD-2, HD-3, and HD-4F, which have been described previously.1,13 The primers HD-1 (ATGAGGACGCTGAGCTCTCAAGTCCT), HD-2 (AAACTTCACGCTCGGTGCA-GCGCTCCCTAG) produce a product containing the CAG repeat and the GC rich region, which includes the polymorphic CCG repeat. The primers HD-3 (GGCGTTGCGGCT-GTTGCTGCTGCTGC) and HD-4F (GCAGCGAGACGACAAAGCCGCAC-CGCC) are complementary and are at the boundary between the CAG repeat and the CCG repeat; the primers HD-1 and HD-3 amplify only the CAG repeat sequence, and the primers HD-2 and HD-4F amplify the GC rich region. These three PCR amplifications were performed for all DNA samples analysed. The amplified alleles were separated on a 5% polyacrylamide gel containing 7 mol/l urea. Allele sizes were determined by comparison with the sizes of the amplified products of pHE5A and pHE5Na and M13mp18 sequence ladders. pHE5A and pHE5Na are the plasmids constructed as follows. Genomic DNA from a
Huntington's disease patient was amplified by PCR using primers HD-1 and HD-2. The amplified products were separated on an agarose gel and the expanded allele product and the normal sized product were purified and subcloned separately into the Smal site of pBluescriptII SK+; designated pHE5A and pHE5Na, respectively. The triplet repeat sequence in pHE5A and pHE5Na were of the form TTC(CAG)$_n$CAA and TTC(CAG)$_n$CCGCA(CCG)$_n$CCT, respectively.

**HAPLOTYPING**

The haplotype of each chromosome studied was determined by comparison of the lengths of products of the three sets of PCR amplification, and was deduced and confirmed by phase determination where two or three generations were available for molecular analysis.

**Results**

**DISTRIBUTION OF CAG REPEAT NUMBER**

Our linkage studies showed that Japanese HD families were tightly linked to chromosome 4p16.3.$^{67}$ All 20 HD families previously studied were analysed for determination of number of the CAG repeat. A typical autoradiograph of one of the families is shown in fig 1. It was confirmed that all affected subjects with HD show expansion of the CAG repeat (data not shown).

Additional Japanese HD families that were not studied for linkage analysis were also analysed. Molecular data were obtained from 92 affected subjects from 58 unrelated HD families. A control group consisted of 185 unrelated non-HD chromosomes which were identified in the families studied. The distribution of normal and abnormal CAG repeat number is shown in fig 2. The normal range varied from seven to 29 repeats (mode 17 repeats) and the observed heterozygosity and PIC in the normal Japanese population are 0.742 and 0.706, respectively. The HD range is from 37 to 95 and the mode is 42. The peaks of the normal and HD ranges are well separated and the two distributions do not overlap.

**AGE OF ONSET AND CAG REPEAT LENGTH**

Fig 3 shows the relationship between age of onset and the CAG repeat size in both alleles in affected subjects. A significant negative correlation ($r = -0.709$, $p<0.001$) was evident for the relationship between the age of onset and the number of longer CAG repeats. The regression curve was derived according to the formula (CAG repeat number) = $112 - 17.72 	imes \ln$ (age of onset).

Data on age of onset were obtained from 22 father–child pairs and five mother–child pairs in this cohort. The Wilcoxon signed rank test showed that children with paternal transmission developed symptoms significantly earlier than their fathers ($p = 0.0001$), whereas differences between mother–child pairs were not significant ($p = 0.124$). The mean difference between father–child pairs was 20–23 (SE 2.69) years, and 12–20 (SE 6.29) years between mother–child pairs. Molecular data of CAG expansion were obtained from 20 father–child pairs and 14 mother–child pairs. The difference in CAG repeat size on transmission of the HD chromosome was calculated and the results are shown in fig 4. On paternal transmission the repeat size increased by 4.05 (SE 1.98) repeats on average, whereas the size decreased by 0.36 (SE 0.86) repeats on maternal transmission. The Wilcoxon signed rank test showed that children with paternal transmission had significantly larger expansions than their father ($p = 0.015$), whereas expansion from affected mothers to their children was not significant ($p = 0.685$). This study contained three children with the rigid form of HD and in all cases their HD mutations were inherited from their father.

**CCG REPEAT POLYMORPHISM AND HUNTINGTON'S DISEASE**

The CCG repeat sequence immediately 3' adjacent to the CAG repeat is polymorphic in white populations.$^{811}$ The frequencies of the CCG repeat alleles in the normal Japanese population is shown in the table. Alleles of seven and 10 repeats occur in the majority of
Analysis of triplet repeats in the huntingtin gene in Japanese families affected with Huntington's disease

Figure 3 Relationship between the number of CAG repeats and the age of onset. Diamonds = expanded allele, crosses = normal allele. The broken line is the regression curve which is of the formula (CAG repeat number) = 112 - 17·72 x ln (age of onset).

Figure 4 Variation of CAG repeat size on transmission of Huntington's disease chromosomes.

Figure 5 Haplotype of the CAG and CCG repeat polymorphism in the normal Japanese population. Hatched bar = (CCG)6 allele. Filled bar = (CCG)7 allele. Open bar = (CCG)10 allele.

The observed allele frequencies for the CCG repeat polymorphism in the HD population are significantly different from those obtained from the normal population ($\chi^2 = 39.39$, df = 2, $p = 0.0001$). This result supports strong linkage disequilibrium between HD and the CCG allele with 10 repeats in the Japanese population.

Haplotype analysis of the CAG and CCG repeat polymorphism in the normal Japanese population

Fig 5 shows the distributions of CAG repeat length of normal chromosomes which have a different CCG repeat allele. The mean CAG repeat size is 17·45 (SE 0·13) for (CCG)6 chromosomes and 17·65 (SE 0·13) for (CCG)10 chromosomes. The latter is slightly larger than the former, although the difference is not significantly statistically ($t$ value = 0·7, $p = 0·4839$) and the two distributions are indistinguishable ($\chi^2 = 15·78$, df = 12, $p = 0·2013$).

The primers HD-1 and -2, which have been previously described,1 amplify the sequence containing both CAG and CCG repeats and the products show combined polymorphism. Heterozygosity and PIC for the CCG repeat polymorphism in the Japanese population is 0·483 and 0·372, respectively.

Recent studies have shown that more than 90% of HD chromosomes of western descent have the seven repeat allele and strong linkage disequilibrium between this allele and HD is evident.9-11 We first analysed the 20 families which had been previously used for linkage analysis. In only three families, the seven repeat allele cosegregated with the HD mutation; however, the 10 repeat allele cosegregated in 17 families. Fig 1 shows a family in which the 10 repeat allele unambiguously cosegregates with the expansion of the CAG repeat. Additional Japanese HD chromosomes were analysed and haplotypes were identified for a total of 58 unrelated Japanese HD chromosomes (table). Forty-nine HD chromosomes (85·5%) cosegregated with the 10 repeat allele, whereas only nine (15·5%) carried the seven repeat allele. The observed allele frequencies for the CCG repeat polymorphism in the HD population are significantly different from those obtained from the normal population ($\chi^2 = 39.39$, df = 2, $p = 0.0001$). This result supports strong linkage disequilibrium between HD and the CCG allele with 10 repeats in the Japanese population.

Discussion

Expansion of the CAG Repeat

Analysis of the CAG repeat showed that the range of expansion of the repeat was between 37 and 95 (mean = 42), that of normal chromosomes was between 17 and 29, and alleles
containing repeats ranging from 30 to 36 were
not found. The distribution of the expanded
repeat of HD was similar to that for North
American and European populations described
previously.\textsuperscript{14-22} It is difficult to define precisely
the smallest size of the CAG repeat expansion
which causes the disorder. In most studies
numbers around 37 were reported as the lowest
numbers for HD of European descent: 35 for
the Scottish population,\textsuperscript{20} 36 for American,\textsuperscript{19}
37 for Italian,\textsuperscript{23} 38 for Canadian,\textsuperscript{16,19} 39 for
German,\textsuperscript{17,19} 39 for Danish,\textsuperscript{18} and 41 for Russian.\textsuperscript{22}

A significant correlation was found for the
relationship between the age of onset and the
CAG repeat expansion. The longer the CAG
repeat, the younger the age of onset. A similar
relationship was described for the HD popu-
lation of white descent.\textsuperscript{14-22} The inverse
relationship is clear for juvenile onset HD, but
for adult onset cases any given number of
repeats can be associated with an age of onset
that varies by decades.

We obtained molecular data from 20 HD
father–child pairs and 14 mother–child pairs
affected with HD. A significant increase of
CAG repeat size on paternal transmission of
HD and greater instability of paternally trans-
mitted repeats has been shown in western HD
families.\textsuperscript{14,20} Analysis of Japanese affected par-
ent-offspring pairs in this study also showed
that on paternal transmission there appeared
to be a younger age of onset together with
the longer CAG repeat, but not on maternal
transmission. Anticipation corresponds to ex-
pansion of the CAG repeat and it occurs in
paternal transmission.

The above features of the CAG repeat are
mostly similar to those of western populations.

\textbf{THE CCG REPEAT POLYMORPHISM}

The CCG repeat polymorphism is composed
of the two predominant alleles (CCG)\textsubscript{7} and
(CCG)\textsubscript{10} and some other minor alleles in Ja-
apanese or western populations. Frequencies of
both major alleles in the Japanese population
were not significantly different from those in
western populations which have been described
previously.\textsuperscript{9-11} Other minor alleles are rare in
Japan, compared with data from western or
black people,\textsuperscript{9-11,22} and we identified only one
chromosome with the (CCG)\textsubscript{7} allele in this
study.

The prevalence of HD varies considerably
among different ethnic groups. The frequency
of the disease is very low in Japan and less
than one-tenth of the prevalence in western
countries. It was suggested that the low fre-
cuency of HD reflects the origins of the HD
mutation. The original HD gene was con-
sidered to be western European in origin and
subsequently was thought to have spread to
Japan by emigration.\textsuperscript{2} Haplotype analysis of
HD chromosomes provides a clue to the origins
of the HD gene. It is interesting to note that
there is a striking discrepancy between Japanese
and western HD, in terms of a CCG repeat
allele showing strong association with the CAG
repeat expansion. In fact, strong linkage dis-
equilibrium is found between Japanese HD
chromosomes and (CCG)\textsubscript{10}, whereas western HD
chromosomes are strongly associated with
(CCG)\textsubscript{7}.\textsuperscript{9-11} The most plausible explanation for
this discrepancy is that HD mutations in these
two races originated from different ancestral
lineages; the western HD mutation is mostly
derived from a chromosome containing
(CCG)\textsubscript{7} in western Europe, while Japanese HD
is mostly from a (CCG)\textsubscript{10} associated chro-
mosome in the ancestors of the Japanese popu-
lation. Although the mutation rate for HD
was estimated to be extremely low,\textsuperscript{23} recent
molecular genetic findings in HD suggest that
new mutation events appear to occur more
frequently than considered previously.\textsuperscript{23} Two
groups described several new mutations for
HD that were shown by molecular analysis of
expansion of a CAG repeat.\textsuperscript{24,25} In addition,
analysis of HD chromosomes with multiallelic
markers was reported to show many different
haplotypes, suggesting a variety of independent
HD mutations.\textsuperscript{26} Therefore, in addition to the
hypothesis of spread of the HD gene from
western Europe as the origin of Japanese HD, it
is reasonable to postulate that a new mutational
event in the Japanese ancestral population
might contribute to the present HD patients
in Japan.

In addition to the founder effect, we can
consider other hypotheses to account for the
disequilibrium of the CCG repeat poly-
morphism. First, one could assume that a cis-
acting element links to the CCG repeat and
predisposes to expansion of the CAG repeat.
The CCG repeat sequence itself can be con-
sidered to have an effect on the CAG repeat
expansion. Although it has been proposed that
the (CCG)\textsubscript{7} allele could be critical in HD,\textsuperscript{10}
Japanese HD chromosomes associate strongly
with the (CCG)\textsubscript{10} allele. Therefore, it appears
unlikely that a particular allele of the CCG
repeat polymorphism might be involved in the
mechanism of the disease. Another possibility
is that relatively longer CAG repeat alleles in
the normal range could be prone to mutate
to expansion. The mean length of the CAG
repeat of chromosomes containing (CCG)\textsubscript{10}
was slightly longer than that of chromosomes
carrying (CCG)\textsubscript{7},\textsuperscript{14} although the difference was
not significant. Longer alleles of the CAG re-
peat polymorphism, which constitute a very
small fraction in our cohort, could be mostly
associated with an allele of (CCG)\textsubscript{10} in Japan
and might constitute a reservoir providing pre-
mutations. Finally, one could speculate that
the threshold of CAG repeat size to predispose
to expansion may vary in a chromosome hav-
ing a different sized allele of the CCG poly-
morphism. The threshold in a chromosome
carrying (CCG)\textsubscript{7} might be smaller than in a
(CCG)\textsubscript{7} chromosome. However, the dis-
tributions of CAG repeat size appear to be
similar for both (CCG)\textsubscript{10} and (CCG)\textsubscript{7} chro-
mosomes (data not shown), though we have
little molecular data about HD chromosomes
carrying (CCG)\textsubscript{7}.

We wish to thank the members of the HD families and their
physicians for their cooperation. This work was supported by
Grant-in-Aid for Scientific Research on Priority Areas and
for General Scientific Research (No 04404041), Ministry of
Education, Science, and Culture, Japan.
Analysis of triplet repeats in the huntingtin gene in Japanese families affected with Huntington's disease


