

ORIGINAL INVESTIGATION

Quasar Saleem · Shweta Choudhry · Mitali Mukerji
Leena Bashyam · Madakasira V. Padma
Ambar Chakravarthy · Mool Chand Maheshwari
Satish Jain · Samir K. Brahmachari

Molecular analysis of autosomal dominant hereditary ataxias in the Indian population: high frequency of SCA2 and evidence for a common founder mutation

Received: 22 November 1999 / Accepted: 22 December 1999 / Published online: 3 February 2000

© Springer-Verlag 2000

Abstract Expansion of CTG/CAG trinucleotide repeats has been shown to cause a number of autosomal dominant cerebellar ataxias (ADCA) such as SCA1, SCA2, SCA3/MJD, SCA6, SCA7, SCA8 and DRPLA. There is a wide variation in the clinical phenotype and prevalence of these ataxias in different populations. An analysis of ataxias in 42 Indian families indicates that SCA2 is the most frequent amongst all the ADCAs we have studied. In the SCA2 families, together with an intergenerational increase in repeat size, a horizontal increase with the birth order of the offspring was also observed, indicating an important role for parental age in repeat instability. This was strengthened by the detection of a pair of dizygotic twins with expanded alleles showing the same repeat number. Haplotype analysis indicates the presence of a common founder chromosome for the expanded allele in the Indian population. Polymorphism of CAG repeats in 135 normal individuals at the SCA loci studied showed similarity to the Caucasian population but was significantly different from the Japanese population.

Introduction

The autosomal dominant cerebellar ataxias (ADCA) are a clinically and genetically heterogeneous group of neurodegenerative disorders that lead to generalized failure of coordination and particularly affect gait, speech and limbs. Various spinocerebellar ataxias, such as SCA1, SCA2, SCA3/MJD, SCA6, SCA7, and DRPLA, have been shown to be caused by the expansion of CTG/CAG trinucleotide repeats in the coding region of genes responsible for these diseases (Wells and Warren 1998). However, in the case of SCA8, the expansion is observed in the 3' untranslated region of the gene (Koob et al. 1999).

The CAG repeats at different spinocerebellar ataxia (SCA) loci are highly polymorphic in normal individuals. Once the repeat number crosses a particular threshold, specific to an individual locus, instability is observed, leading to the manifestation of disease. These disorders also exhibit the phenomenon of anticipation, wherein there is an increase in severity in successive generations with a reduction in the age of onset. This phenomenon has been explained by the increase in repeat number when transmitted from parent to offspring (Wells and Warren 1998).

While the exact mechanism responsible for repeat expansion is still unknown, a number of hypotheses such as strand slippage, hairpin mediated slippage and recombination have been proposed (Brahmachari et al. 1995; McMurray 1995; Wells and Warren 1998). The length of uninterrupted stretches of CAG repeats and the point mutations within the repeats have been shown to play a role in the instability process (Chung et al. 1993). Screening in the populations for such polymorphism in the repeats at various loci involved in ADCAs may provide a novel approach towards understanding the mechanisms responsible for repeat expansion. Molecular analyses of several of these SCAs have been carried out in a number of populations (Leggo et al. 1997; Lopes-Cendes et al. 1997; Silveira et al. 1998; Takano et al. 1998; Jin et al. 1999), however there are no studies pertaining to the frequencies of

Q. Saleem and S. Choudhry have contributed equally to this work.

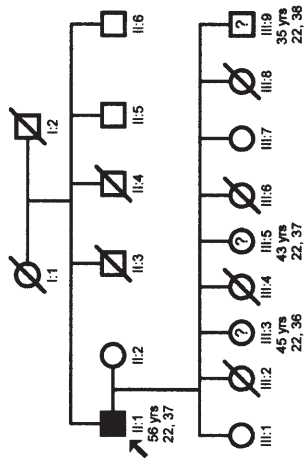
Q. Saleem · S.K. Brahmachari
Molecular Biophysics Unit,
Indian Institute of Science, Bangalore, India

S. Choudhry · M. Mukerji · L. Bashyam · S.K. Brahmachari (✉)
Functional Genomics Unit,
Centre for Biochemical Technology (CSIR),
Delhi University Campus, Mall Road, Delhi, 110 007, India
e-mail: skb@cvt.res.in,
Tel.: +91-11-7416489, Fax: +91-11-7257471

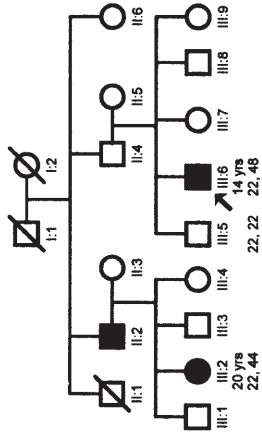
M.V. Padma · M.C. Maheshwari · S. Jain
Neuroscience Centre,
All India Institute of Medical Sciences, New Delhi, India

A. Chakravarthy
Department of Neurology,
Vivekananda Institute of Medical Sciences, Calcutta, India

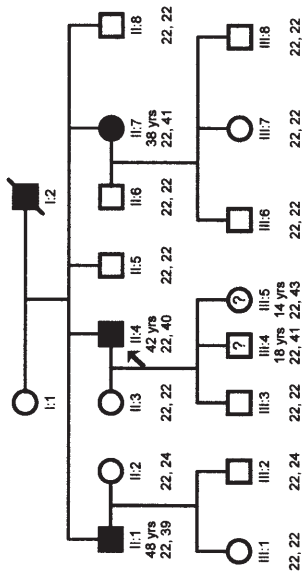
AT010



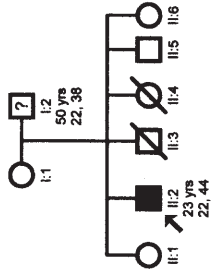
AT009



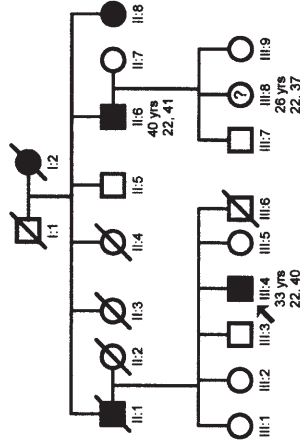
AT006



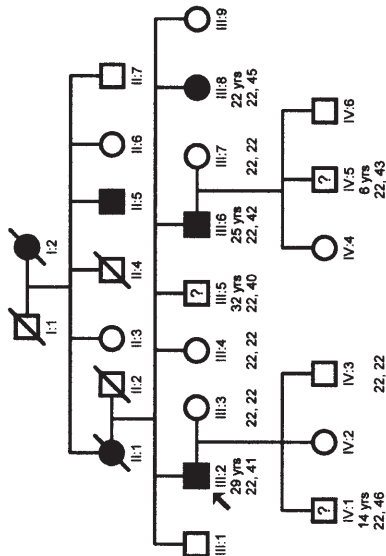
AT029



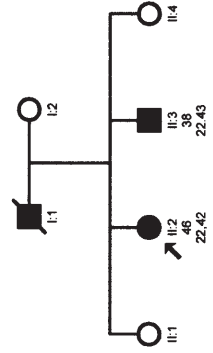
AT024



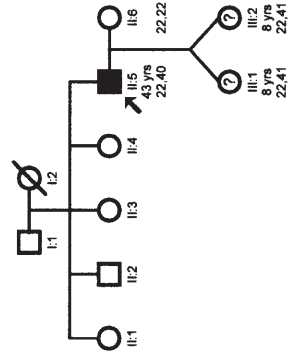
AT017



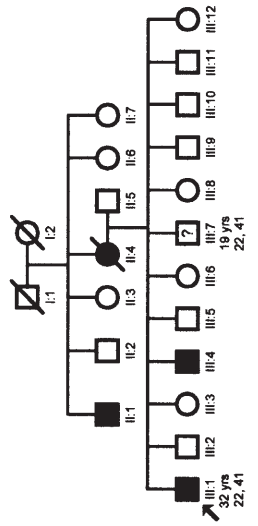
ATV001

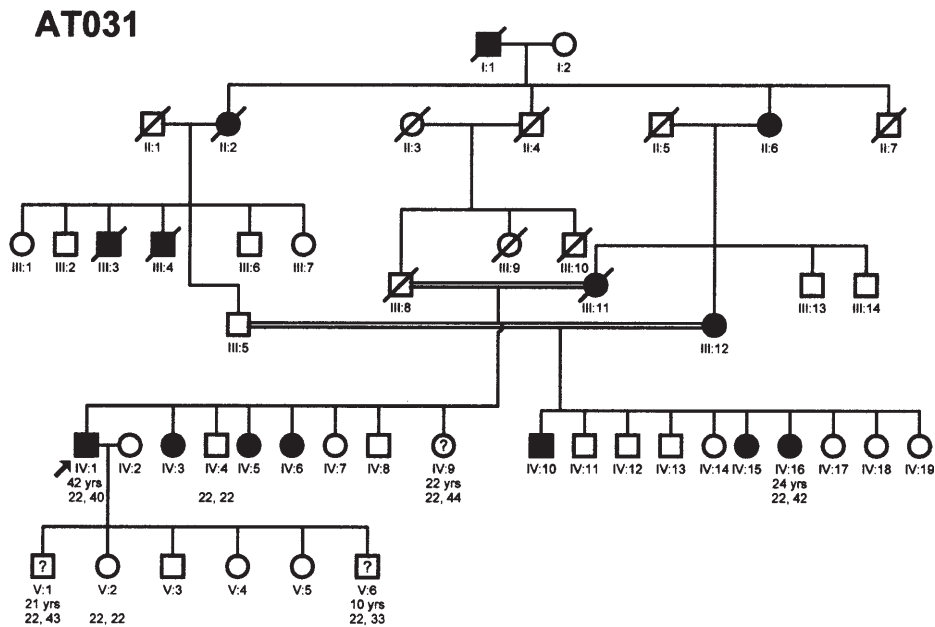


AT038



AT027





▲ **Fig. 1** Pedigree of SCA2 affected families. Numbers below symbols indicate age at onset and repeat number at the SCA2 locus. Filled symbols indicate clinically affected individuals and symbols with a question mark indicate individuals with CAG expansion who have not yet developed clinical symptoms

these ataxias in the Indian population. To identify the various types of ADCAs in the population we have carried out an extensive study of patients suffering from this class of disorders. We have also estimated the CAG repeat polymorphism at the various disease loci in the normal Indian population as the nature of the distribution of repeats has been shown to influence disease prevalence (Takano et al. 1998).

Materials and methods

The study was carried out on 42 families comprising 122 members including the patients and family members. Members of 35 families, mostly from northern India, were diagnosed for ataxia at the Neuroscience Centre, All India Institute of Medical Sciences, New Delhi, which is a major tertiary case referral centre in northern India. The patients were examined by a team of neurologists and evaluated in detail after admission to the hospital. Besides clinical evaluation, all patients underwent a detailed neurophysiological examination (nerve conduction and evoked potentials) and imaging of the brain (CT or MRI scan). All patients presented with a progressive form of cerebellar ataxia often but not always associated with pyramidal signs, extra pyramidal signs and ocular findings. The ADCA type was diagnosed by the standard criteria (Harding 1993; Junck and Fink 1996). The remaining 7 families were identified for ataxia at the Vivekannanda Institute for Medical Sciences, Calcutta and were from eastern India. Samples were also collected from 135 unrelated normal individuals of northern Indian origin. Informed consent was obtained from all patients and normal individuals. Pedigree charts were created using Cyrillic version 2.1 (Cherwell Scientific, Oxford, UK).

DNA was isolated from peripheral blood leukocytes using a modification of the salting-out procedure (Miller et al. 1988). Repeat sizes were estimated at the SCA1, SCA2, MJD/SCA3, SCA6, SCA7, SCA8 and DRPLA loci by PCR using previously published

primers (Orr et al. 1993; Kawaguchi et al. 1994; Koide et al. 1994; Pulst et al. 1996; Zuchenko et al. 1997; Gouw et al. 1998; Koob et al. 1999). One of the primers was fluorescently labelled with either a HEX, TET or FAM fluorophore. The size of the repeat in the fluorescently labelled PCR product was determined by Gene Scan analysis using an ABI Prism 377 automated DNA sequencer (Perkin Elmer, Foster City, USA). As an additional screening procedure, patient samples were also checked for expansion at the Friedreich's ataxia, spinobulbar muscular atrophy (SBMA) and Huntington's disease loci.

The distribution of CAG repeats at the SCA1, SCA2, SCA3/MJD, SCA6 SCA7, SCA8 and DRPLA loci was determined in 135 unrelated normal individuals. Sequencing of DNA from a few normal individuals and all the available members of the SCA2, SCA1 and SCA3/MJD positive families was carried out to confirm the repeat size as well as the nature of the interruptions. Sequencing was carried out using dye terminator chemistry on an ABI Prism 377 automated DNA sequencer.

Haplotype analysis of the SCA2 positive families consisting of affected and unaffected members was carried out using microsatellite markers D12S1332, D12S1333 and D12S1672. These markers span a region around the SCA2 CAG repeat in the following order: centromere – D12S1332 (350 kb) – D12S1672 (200 kb) – D12S1333 – telomere with D12S1132 and D12S1333 flanking the CAG repeat. D12S1672 is in the first intron of the SCA2 gene and is 20 kb centromeric to the CAG repeat stretch, which is in the first exon of the gene (Pulst et al. 1996; Sahba et al. 1998). The frequency of these markers was also determined in normal chromosomes derived from unaffected SCA2 family members and normal control individuals from the same geographic origin as the patients. The phase of the markers was determined in the normal chromosomes segregating in the SCA2 affected pedigrees. In all the analyses, the CEPH family member 1347-02 was used as a control for verification of the repeat size.

Results

Molecular analysis of ataxia families

Among the 42 families studied, three families showed expansions at the Friedreich's ataxia locus and were excluded from further analysis. In the remaining 39 kin-

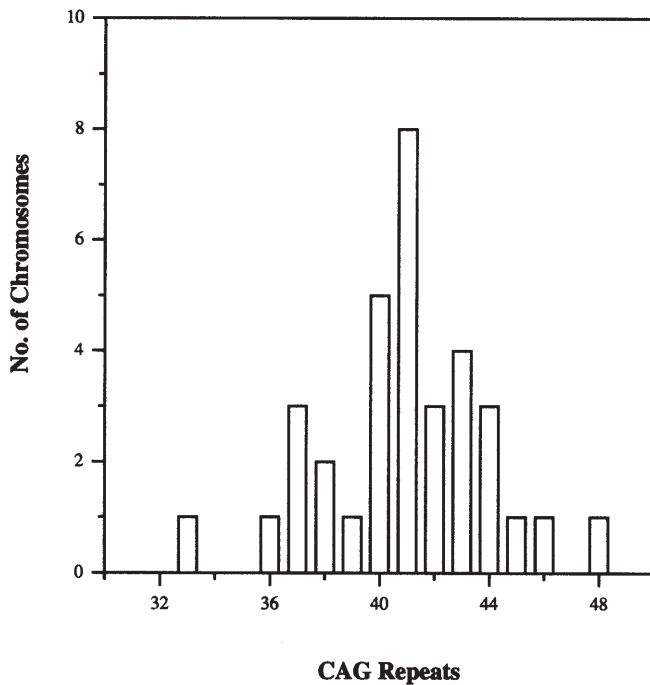


Fig.2 Distribution of CAG repeats in expanded alleles at the SCA2 locus in 34 individuals. The 33-repeat allele has arisen by a contraction from a diseased individual having 40 repeats and was asymptomatic at the time of the study

dreds, there was an expansion at the SCA2 locus in 10 pedigrees, at SCA1 in 3 and at SCA3/MJD in 2 pedigrees. Both the SCA3/MJD pedigrees and one SCA2 family were from eastern India and the remaining affected individuals were from northern India. Genotyping of the rest of the families did not show expansion at any of the other loci analysed including SCA6, SCA7, SCA8 and DRPLA. In the case of the SCA1 positive individuals, the size of the repeats in the expanded allele ranged from 48–57. In the four positive MJD individuals from two pedigrees, repeat sizes ranged from 69–75. All the expanded alleles were sequenced and found to contain uninterrupted stretches of CAG repeats. In all the cases, the repeat sizes were well within the reported disease range. Since our sample had a maximum frequency of SCA2, a more detailed study of the affected and unaffected relatives of the SCA2 families is presented below (Fig. 1).

Analysis of SCA2 families

All the patients who were confirmed for SCA2 had gait ataxia and cerebellar atrophy, with 86% having slow saccades and dysarthria. The other clinical signs were pyramidal signs (66%), peripheral neuropathy (66%), face fasciculations (53%) and broken pursuits (53%). Nystagmus and extrapyramidal signs were not observed in any of the cases.

Expanded alleles at the SCA2 locus were found in 34 individuals who were affected and asymptomatic at-risk

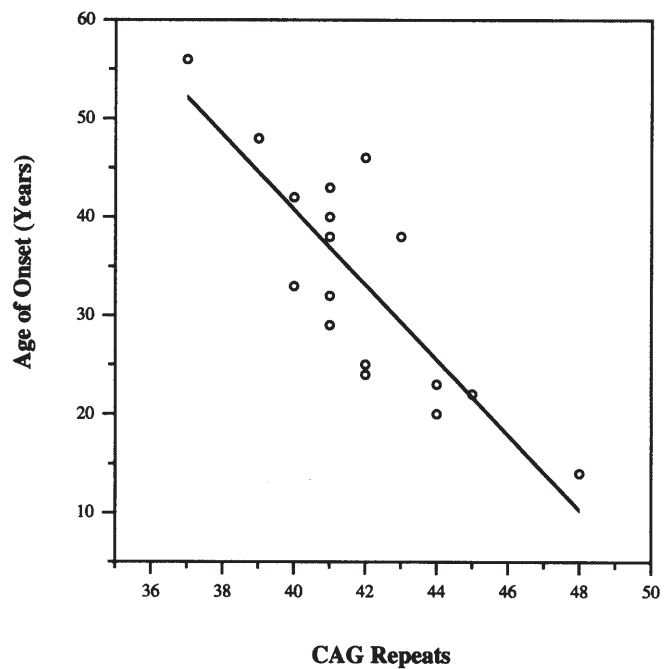


Fig.3 Correlation between number of CAG repeats in expanded alleles at the SCA2 locus and age of disease onset ($r=-0.74$)

individuals. The repeat size in the expanded allele ranged from 36 to 49 (Fig. 2). One individual with an expanded allele of 33 repeats was also observed in an SCA2 pedigree (Fig. 1, AT 031:V6). This individual who was 10 years old at the time of the study was asymptomatic, however his young age does not exclude a later disease onset. Haplotype analysis confirmed that this allele had arisen by contraction of an allele with 40 repeats during transmission from the affected father. All the expanded alleles (including the 33-repeat allele) were sequenced and were found to contain no interruptions. In 11 cases analysed, transmission was paternal and variation in size was from -7 to +5 repeats. In AT 031:V6 and AT024:III-8 contractions of 7 repeats and 4 repeats, respectively, were observed (Fig. 1). A significant linear inverse correlation characteristic of anticipation was observed between repeat size and age at onset, ($r=-0.74$, $P<0.005$) as shown in Fig. 3. Together with a vertical increase in repeat length, we observed a horizontal increase in CAG repeat number in a majority of cases, i.e. an increase in repeat size with the birth order of the offspring (Table 1). We had a total of 17 affected sib pair combinations, of which 14 showed higher repeat numbers in the younger sib as compared to the elder sib, a trend which was statistically significant ($\chi^2=5.65$, $P<0.05$). In the family AT038 with affected dizygotic sibs, there was no difference in the size of the repeat in the twins. The zygosity of the sibs was confirmed by haplotyping. Horizontal expansion was not seen in these twins, indicating that the age of the parents might indeed influence the size of the transmitted expanded allele.

Table 1 Sib pair combinations with difference in repeat sizes showing that the majority of younger sibs had larger repeat sizes than their elder sibs

Serial number	Family number	Elder sib (repeat size)	Younger sib (repeat size)	Difference in repeat size (younger sib-elder sib)	Transmission
1	AT 006	II-1 (39)	II-4(40)	1	Paternal
2		II-1 (39)	II-7 (41)	2	Paternal
3		II-4 (40)	II-7 (41)	1	Paternal
4		III-4 (41)	III-5 (43)	2	Paternal
5	AT 010	III-3 (36)	III-5 (37)	1	Paternal
6		III-3 (36)	III-9 (38)	2	Paternal
7		III-5 (37)	III-9 (38)	1	Paternal
8	AT 017	III-2 (41)	III-5 (40)	-1	Maternal
9		III-2 (41)	III-6 (42)	1	Maternal
10		III-2 (41)	III-8 (45)	4	Maternal
11		III-5 (40)	III-6 (42)	2	Maternal
12		III-5 (40)	III-8 (45)	5	Maternal
13		III-6 (42)	III-8 (45)	3	Maternal
14	AT027	III-1 (41)	III-7 (41)	0	Maternal
15	AT031	IV-1 (40)	IV-9 (44)	4	Paternal
16		V-1 (43)	V-6 (33)	-10	Paternal
17	ATV001	II-2 (42)	II-3 (43)	1	Paternal
18 ^a	AT038	IV-1 (41)	IV-2 (41)	0	Paternal

^aIndicates that in this case the siblings were dizygotic twins and were not used in statistical analyses

Table 2 Markers associated with expanded alleles in different families. D12S1332: allele 8=208 bp, 5=202 bp, 6=204 bp; D12S1672: 3=281 bp, 4=283 bp; D12S1333:1=225bp

Family number	D12S1332	D12S1672	D12S1333
AT006	8	3	1
AT009	8	3	1
AT027	8	3	1
AT038	8	3	1
AT024	5	3	1
AT031	5	3	1
AT010	6	3	1
AT029	5	4	1
AT017	5	4	1

Haplotype analysis of SCA2 families

Haplotype analysis was carried out in 9 SCA2 pedigrees using the microsatellite markers D12S1332, D12S1333 and D12S1672 (Table 2). In the remaining SCA2 family, the DNA was insufficient to carry out haplotype analysis. All the 9 pedigrees were from northern India. The allele 1

of D12S1333 was present on the chromosome containing the expanded repeat in all the SCA2 positive families. The alleles 3 and 4 of the marker D12S1672 and alleles 5 and 8 of D12S1332 were the most commonly linked to the expanded alleles in the patients. The truncated haplotype 3-1 for markers at D12S1333 and D12S1672 was associated with SCA2 expansion in all but two families, AT029 and AT017. In these two exceptions the haplotype was 4-1.

The frequency of the alleles associated with the expanded chromosomes was also estimated in the normal chromosomes derived from the same geographical location (Table 3). The allele 1 in the case of D12S1333 and the allele 3 of D12S1672 were in complete linkage disequilibrium with the SCA2 expanded allele ($P < 0.001$). In the case of D12S1332, the frequencies of alleles 5 and 8 in the normal and the expanded alleles indicate that the association of allele 8 with the expanded allele is much more significant ($P < 0.001$). The haplotype 8-3-1 and the truncated haplotype 3-1 were in strong linkage disequilibrium with the expanded SCA2 allele ($P < 0.001$). This haplotype was not present in chromosomes from normal individuals in which the phase could be determined.

Table 3 Linkage disequilibrium between the microsatellite markers and the SCA2

Locus	Allele	Patients (frequency)	Controls (frequency)	χ^2	Probability
D12S1332	5	4/9 (0.44)	66/200 (0.33)	0.506	0.476
	6	1/9 (0.11)	24/200 (0.12)	0.006	0.93
	8	4/9 (0.44)	14/200 (0.07)	15.3	<0.0001
D12S1672	3	7/9 (0.77)	30/197 (0.15)	22.8	<0.0001
	4	2/9 (0.22)	9/197 (0.05)	5.3	0.02
D12S1333	1	9/9 (1.00)	26/198 (0.13)	46.2	<0.0001
Haplotype	8-3-1	4/9 (0.44)	0/61 (0.00)	28.7	<0.0001
	5-3-1	2/9 (0.22)	1/61 (0.02)	8.1	0.004
	5-4-1	2/9 (0.22)	0/61 (0.00)	13.7	0.0002
	6-3-1	1/9 (0.11)	0/61 (0.00)	6.86	<0.0001
Truncated haplotype	-3-1	7/9 (0.77)	1/61 (0.02)	44.916	<0.0001

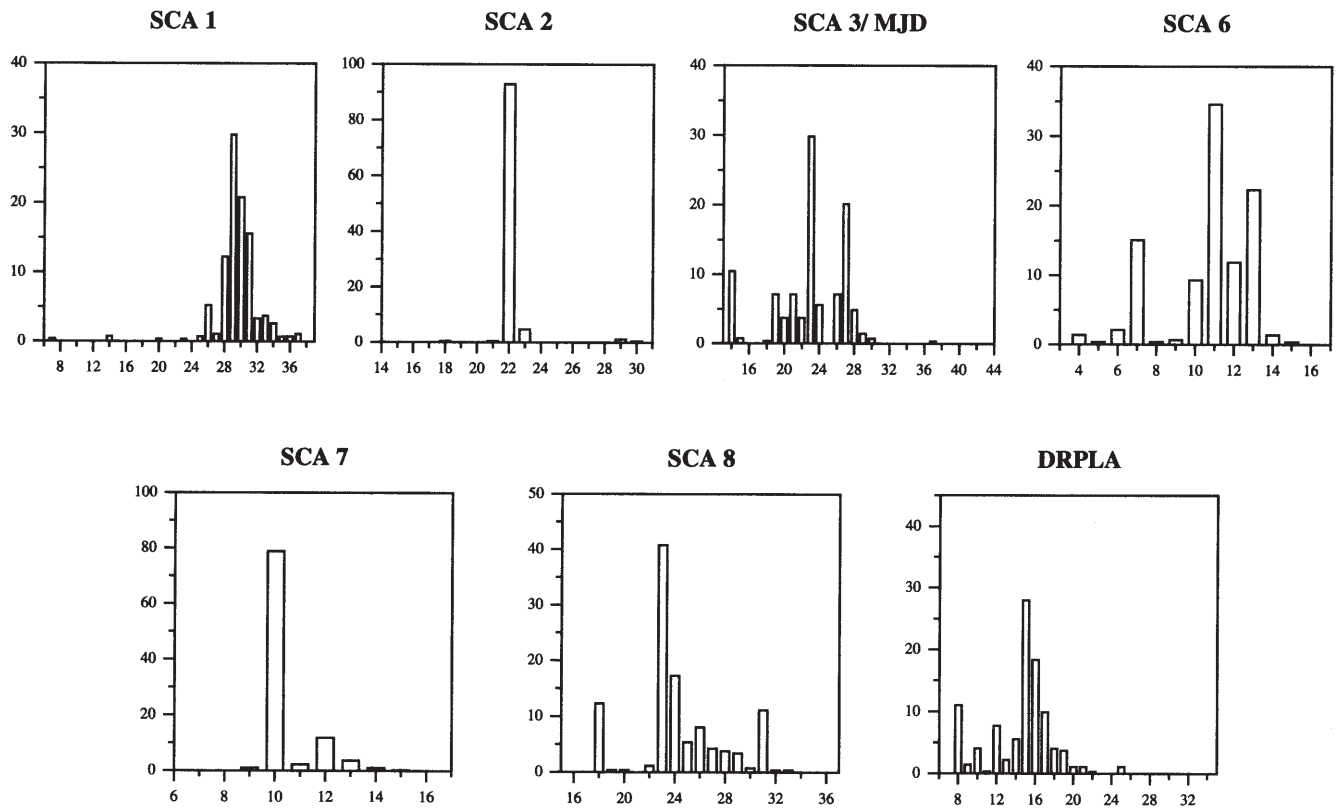


Fig. 4 Distribution of CAG repeats at various SCA loci in 270 normal chromosomes from unrelated individuals. *X axis* represents CAG repeat number and *Y axis* the percentage frequency

Distribution of CAG repeats at various SCA loci in normal individuals

Figure 4 shows the frequency distribution of the trinucleotide repeats at the SCA1, SCA2, SCA3/MJD, SCA6, SCA7, SCA8 and DRPLA loci of 270 normal chromosomes predominantly from the northern part of India. The distribution of the normal alleles of SCA1, SCA2, SCA3, SCA6 and DRPLA resembles the Caucasian population to a great extent and differs from the Japanese (Takano et al. 1998). In the case of SCA1, normal alleles with 7–37 CAG repeats were present, with 90% of them having repeats from 27 to 37. All the normal alleles of SCA1 which were sequenced had CAT interruptions within the CAG repeat stretch. In the case of SCA2, though 91% of the alleles had repeats of less than 23, large normal (LN) alleles were also observed. All the normal alleles sequenced contained 1, 2 or 3 interruptions of CAA except one LN allele with 29 repeats which did not contain any interruption. This is the first report of an SCA2 allele of this size without interruptions observed in the normal range. In the case of SCA7, the heterozygosity was comparatively low, with 80% of the alleles having 10 repeats. A large normal allele carrying 18 repeats was observed at this locus which has not been reported previously. At the SCA8 locus, a stretch of CTA repeats is contiguous with the CTG repeat, both of which are polymorphic (Koob et al. 1999). The repeat

Table 4 Frequencies of large normal alleles at the SCA loci in Indian, Japanese and Caucasian populations. Data from Japanese and Caucasian populations are taken from Takano et al. (1998)

SCA locus	Repeat size	Indian	Caucasian	Japanese
SCA1	>30	0.27	0.26	0.09
	>31	0.12	0.16	0.04
SCA2	>22	0.06	0.12	0.01
	>23	0.02	0.03	0.01
SCA3/MJD	>27	0.07	0.09	0.21
	>28	0.03	0.04	0.09
SCA6	>13	0.017	0.04	0.20
	>14	0.0035	0.00	0.08
DRPLA	>17	0.11	0.04	0.26
	>18	0.07	0.06	0.13

numbers estimated were, therefore, a composite of the CTA and CTG repeats. At this locus, the repeats varied from 18 to 33 with 90% of the alleles having less than 27 repeats. LN alleles at the SCA8 locus ranging from 33 to 92 repeats, which have been reported previously, have not been observed in the samples studied. Frequency of LN alleles and a comparison with the Caucasian and Japanese populations is given in Table 4. The range representing LN alleles were determined as per the criterion described by Takano et al. (1998).

Discussion

Forty-two Indian families diagnosed for spinocerebellar ataxia have been analysed for expansions at the SCA1,

SCA2, SCA3/MJD, SCA6, SCA7, SCA8 and DRPLA loci. Of these, three families showed expansions at the Friedreich's ataxia locus and were excluded from further analysis. In our study, we observed the maximum frequency for SCA2 (10/39), followed by SCA1 (3/39) and MJD (2/39). The remaining families (24/39) were negative for expansion at all the loci studied.

The expanded repeat range and the clinical features of SCA2 patients were similar to those described earlier (Imbert et al. 1996; Pulst et al. 1996; Sanpei et al. 1996; Cancel et al. 1997; Geschwind et al. 1997; Giunti et al. 1998; Wadia et al. 1998). Anticipation was observed in the SCA2 pedigrees (Fig. 3). The SCA2 families analysed exhibited a number of interesting features, the most striking of which was the horizontal increase in repeat number in the majority of sib pairs. Our results indicate that as the age of the parent increases there is an increased instability (predominately expansion) during transmission to the offspring. Single sperm analysis for various SCAs has indicated that there is considerable size variation in the repeat number for the expanded allele (Kunst et al. 1997; Chong et al. 1995; Chong et al. 1997; Leeflang et al. 1995; Takiyama et al. 1997; Zhang et al. 1995). In the case of Huntington's disease, single sperm analysis was carried out in a patient at an interval of 2 years (Leeflang et al. 1998). It was shown that though the overall distribution of the expanded alleles was the same, a few larger alleles not seen previously were observed in the sample taken after 2 years. In the case of SCA2, if instability occurs during spermatogenesis then the observed increase in repeat size in siblings would indicate that there is a progressive temporal increase in repeat size in the germline. This is further strengthened by our observation that in the dizygotic affected twins, where the age of the transmitting parent is the same for both sibs, there was no difference in the size of the expanded alleles. It is interesting to note that in the case of some SCA1 transgenic mice the magnitude of repeat instability during maternal transmission increased as the mice grew older (Kaytor et al. 1997). However, in all these cases the instability was characterized by a preponderance of contractions. Thus, age of the transmitting parent seems to play a role in repeat instability in both male and female germ cells. In case of Friedreich's ataxia, a direct correlation has been shown between repeat instability in the offspring and parental age at birth, with a preference for expansions during maternal transmissions and the inverse in paternal transmissions (De Michele et al. 1998). Although this phenomena of horizontal increase in repeat size has to be validated in a larger number of SCA2 families, the parental age bias is in favour of a pre-zygotic occurrence of instability. In the two cases where the female was the transmitting parent, we find both an increase and a decrease in repeat length with the birth order of the offspring. Analysis needs to be carried out to determine whether the sex of the transmitting parent has an effect on this phenomenon. It is important to analyse more maternal transmissions and the variation in size of the expanded repeat at the SCA2 locus in single sperm over an interval of time.

It is noteworthy that the frequency distribution of the expanded alleles also shows a resemblance to earlier reports (Giunti et al. 1998; Riess et al. 1997). A modal allele around the range of 40 repeats is seen both in our study (Fig. 2) and the earlier reports. This similarity might underlie a common mechanism resulting in some alleles being over-represented in the distribution of expanded alleles. Another feature which we observe in the families is an increase in the repeat by one unit during transmission to offspring. Since it is difficult for one repeat unit to form a hairpin, strand slippage rather than hairpin-mediated slippage may also be involved in the generation of small expansions.

The truncated haplotype 3-1 (D12S1672-D12S1333) found in the SCA2 pedigrees was far more common in the patients than in the normal population, suggesting the presence of a common ancestral founder (Table 3). This is quite likely since all the SCA2 families analysed were from neighbouring geographic regions (Punjab, Haryana and Uttar Pradesh, three adjacent states of northern India). Two of the families had the truncated haplotype 4-1. The allele 4 differs from 3 by one repeat unit. It is probable that the allele 4 might have arisen *de novo* due to a slippage event on a pre-existing 3-1 haplotype. The alleles 5, 6 and 8 for D12S1332 were present on the chromosomes containing the SCA2 expansions. Since this marker is comparatively more distant from the SCA2 locus, it is possible that a rare recombination could have resulted in different haplotypes. The significantly lower frequency of the 8 allele (D12S11332) in the normal chromosomes indicates that the 8-3-1 haplotype could be the founder for the SCA2 mutation and 5-3-1 could have arisen by recombination since the frequency of the latter is very high in the normal population. In earlier studies, haplotype analysis of SCA2 families from different geographic origins revealed many different haplotypes suggesting the presence of multiple founders (Giunti et al. 1998; Didierjean et al. 1999). However, a common haplotype was seen in patients from countries which were geographically close (Didierjean et al. 1999). The haplotype in our analysis resembles two SCA2 pedigrees found in France (Didierjean et al. 1999). The similarity in haplotypes between our patients and the geographically distant French patients could indicate that either the SCA2 mutation is very ancient or these haplotypes actually represent an 'at risk' chromosome.

In different populations, the disease prevalence of these ataxias have been shown to be associated with the presence of LN alleles at the respective loci (Goldberg et al. 1993; Myers et al. 1993; Squitieri et al. 1994; Takano et al. 1998). We have estimated the frequency of various LN alleles at the different SCA loci (Fig. 4) and compared this with the frequencies reported for the Caucasian and the Japanese populations (Takano et al. 1998), shown in Table 4. As data was not available for SCA7 and SCA8 in these populations, such a comparison could not be carried out. For all the loci, the frequencies of LN alleles were significantly different from the Japanese population (Table 4). However, when compared to the Caucasian

population, all of the loci were similar with two exceptions. In the case of DRPLA, the Indian population had a significantly higher frequency of LN alleles >17 repeats ($\chi^2=6.74$, $P=0.009$) which was, however, less than that seen in the Japanese population. DRPLA has been shown to occur almost solely in the Japanese, but some Caucasian families with the disease have also been reported (Kondo 1998). It would be interesting to examine a larger sample of Indian ataxia patients to see whether there is truly a higher prevalence of the disease than that previously reported for the Caucasian population. In the case of SCA2, repeats >22 were present at a significantly lower frequency than that of the Caucasian population ($\chi^2=6.86$, $P=0.009$). SCA2 is unique among the loci studied, in that it has a near unimodal distribution of normal alleles in all the populations studied. No difference in LN allele frequency was seen between the Japanese and Caucasian populations in spite of the marked difference in prevalence of this disease in the two groups (Takano et al. 1998). It is likely that in the case of SCA2, the LN allele frequency is not the best indicator of disease prevalence. As SCA2 repeats also possess CAA interruptions, the frequency of large uninterrupted alleles might be more informative in this regard. The MJD cases we have identified were all from eastern India and none were detected in the larger sample from northern India. It is important to note that the normal samples we have studied were derived from northern India. It is probable that the prevalence of these ataxias might vary in the different regions of India. Given the large ethnic variation in the Indian population, a larger study of this kind is important both for estimating the frequency of known SCAs and also to search for families which could represent novel disease-causing mutations.

Acknowledgements The authors would like to thank Ms. P. Anuradha, Ms. Gurjit, Ms. Laxmi, Ms. R. Jaya and Ms. Sanghamitra Roy for technical support. We also thank Dr. Vani Brahmachari and Dr. Partha P. Majumder for critical evaluation of the manuscript and Dr. Chandrika B. Rao for help with statistical analyses. Financial support from the Dept. of Biotechnology, Govt. of India on Programme on Functional Genomics to S.K.B. is duly acknowledged.

References

- Brahmachari SK, Meera G, Sarkar PS, Balagurumoorthy P, Tripathi J, Raghavan S, Shaligram U, Pataskar S (1995) Simple repetitive sequences in the genome: structure and functional significance. *Electrophoresis* 16:1705–1714
- Cancel G, Durr A, Didierjean O, Imbert G, Burk K, Lezin A, Belal S, Benomar A, Abada-Bendib M, Vial C, Guimaraes J, Chneiweiss H, Stevanin G, Yvert G, Abbas N, Saudou F, Lebre AS, Yahyaoui M, Hentati F, Vernant JC, Klockgether T, Mandel JL, Agid Y, Brice A (1997) Molecular and clinical correlations in spinocerebellar ataxia 2: a study of 32 families. *Hum Mol Genet* 6:709–715
- Chong SS, McCall AE, Cota J, Subramony SH, Orr HT, Hughes MR, Zoghbi HY (1995) Gametic and somatic tissue-specific heterogeneity of the expanded SCA1 CAG repeat in spinocerebellar ataxia type 1. *Nat Genet* 10:344–350
- Chong SS, Almqvist E, Telenius H, LaTray L, Nichol K, Bourdelat-Parks B, Goldberg YP, Haddad BR, Richards F, Sillence D, Greenberg CR, Ives E, Van den Engh G, Hughes MR, Hayden MR (1997) Contribution of DNA sequence and CAG size to mutation frequencies of intermediate alleles for Huntington disease: evidence from single sperm analyses. *Hum Mol Genet* 6:301–309
- Chung M-Y, Ranum LPW, Duvick LA, Servadio A, Zoghbi HY, Orr HT (1993) Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type 1. *Nat Genet* 5:254–258
- De Michele G, Cavalcanti F, Criscuolo C, Pianese L, Monticelli A, Filla A, Coccozza S (1998) Parental gender, age at birth and expansion length influence GAA repeat intergenerational instability in the X25 gene: pedigree studies and analysis of sperm from patients with Friedreich's ataxia. *Hum Mol Genet* 7:1901–1906
- Didierjean O, Cancel G, Stevanin G, Durr A, Burk K, Benomar A, Lezin A, Belal S, Abada-Bendib M, Klockgether T, Brice A (1999) Linkage disequilibrium at the SCA2 locus. *J Med Genet* 36:415–417
- Geschwind DH, Perlman S, Figueroa CP, Treiman LJ, Pulst SM (1997) The prevalence and wide clinical spectrum of the spinocerebellar ataxia type 2 trinucleotide repeat in patients with autosomal dominant cerebellar ataxia. *Am J Hum Genet* 60:842–850
- Giunti P, Sabbadini G, Sweeney MG, Davis MB, Veneziano L, Mantuano E, Federico A, Plasmati R, Frontali M, Wood NW (1998) The role of the SCA2 trinucleotide repeat expansion in 89 autosomal dominant cerebellar ataxia families. Frequency, clinical and genetic correlates. *Brain* 121:459–467
- Goldberg YP, Kremer B, Andrew SE, Theilmann J, Graham RK, Squitieri F, Telenius H, Adam S, Sajoo A, Starr E, et al (1993) Molecular analysis of new mutations for Huntington's disease: intermediate alleles and sex of origin effects. *Nat Genet* 5:174–179
- Gouw LG, Castaneda MA, McKenna CK, Digre KB, Pulst SM, Perlman S, Lee MS, Gomez C, Fischbeck K, Gagnon D, Storey E, Bird T, Jeri FR, Ptacek LJ (1998) Analysis of the dynamic mutation in the SCA7 gene shows marked parental effects on CAG repeat transmission. *Hum Mol Genet* 7:525–532
- Harding AE (1993) Clinical features and classification of inherited ataxias. *Adv Neurol* 61:1–14
- Imbert G, Saudou F, Yvert G, Devys D, Trottier Y, Garnier JM, Weber C, Mandel JL, Cancel G, Abbas N, Durr A, Didierjean O, Stevanin G, Agid Y, Brice A (1996) Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. *Nat Genet* 14:285–291
- Jin DK, Oh MR, Song SM, Koh SW, Lee M, Kim GM, Lee WY, Chung CS, Lee KH, Im JH, Lee MJ, Kim JW, Lee MS (1999) Frequency of spinocerebellar ataxia types 1,2,3,6,7 and dentatorubral pallidolusian atrophy mutations in Korean patients with spinocerebellar ataxia. *J Neurol* 246:207–210
- Junck L, Fink JK (1996) Machado-Joseph disease and SCA3: the genotype meets the phenotype. *Neurology* 46:4–8
- Kawaguchi Y, Okamoto T, Taniwaki M, Aizawa M, Inoue M, Katayama S, Kawakami H, Nakamura S, Nishimura M, Akiyuchi I, et al (1994) CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nat Genet* 8:221–228
- Kaytor MD, Burright EN, Duvick LA, Zoghbi HY, Orr HT (1997) Increased trinucleotide repeat instability with advanced maternal age. *Hum Mol Genet* 6:2135–2139
- Koide R, Ikeuchi T, Onodera O, Tanaka H, Igarashi S, Endo K, Takahashi H, Kondo R, Ishikawa A, Hayashi T, et al (1994) Unstable expansion of CAG repeat in hereditary dentatorubral-pallidolusian atrophy (DRPLA). *Nat Genet* 6:9–13
- Kondo I (1998) Clinical aspects of DRPLA. In: Wells RD, Warren ST (eds) *Genetic instabilities and hereditary neurological diseases*. Academic Press, San Diego, pp 197–211

- Koob MD, Moseley ML, Schut LJ, Benzow KA, Bird TD, Day JW, Ranum LP (1999) An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). *Nat Genet* 21:379–384
- Kunst CB, Leeflang EP, Iber JC, Arnheim N, Warren ST (1997) The effect of FMR1 CGG repeat interruptions on mutation frequency as measured by sperm typing. *J Med Genet* 34:627–631
- Leeflang EP, Zhang L, Tavare S, Hubert R, Srinidhi J, MacDonald ME, Myers RH, Young M de, Wexler NS, Gusella JF, et al (1995) Single sperm analysis of the trinucleotide repeats in the Huntington's disease gene: quantification of the mutation frequency spectrum. *Hum Mol Genet* 4:1519–1526
- Leeflang EP, Tavare S, Marjoram P, Grewal R, Neal COS, Arnheim N (1998) Human germline mutation analysis by single genome PCR: application to dynamic mutations. In: Wells RD, Warren ST (eds) *Genetic instabilities and hereditary neurological diseases*. Academic Press, San Diego, pp 543–558
- Leggo J, Dalton A, Morrison PJ, Dodge A, Connarty M, Kotze MJ, Rubinsztein DC (1997) Analysis of spinocerebellar ataxia types 1, 2, 3, and 6, dentatorubral-pallidoluysian atrophy, and Friedreich's ataxia genes in spinocerebellar ataxia patients in the UK. *J Med Genet* 34:982–985
- Lopes-Cendes I, Teive HG, Calcagnotto ME, Da Costa JC, Cardoso F, Viana E, Maciel JA, Radvany J, Arruda WO, Trevisol-Bittencourt PC, Rosa Neto P, Silveira I, Steiner CE, Pinto Junior W, Santos AS, Correa Neto Y, Werneck LC, Araujo AQ, Carakushansky G, Mello LR, Jardim LB, Rouleau GA (1997) Frequency of the different mutations causing spinocerebellar ataxia (SCA1, SCA2, MJD/SCA3 and DRPLA) in a large group of Brazilian patients. *Arq Neuropsiquiatr* 55: 519–529
- McMurray CT (1995) Mechanisms of DNA expansion. *Chromosoma* 104:2–13
- Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16:1215
- Myers RH, MacDonald ME, Koroshetz WJ, Duyao MP, Ambrose CM, Taylor SA, Barnes G, Srinidhi J, Lin CS, Whaley WL, et al (1993) De novo expansion of a (CAG)_n repeat in sporadic Huntington's disease. *Nat Genet* 5:168–173
- Orr HT, Chung MY, Banfi S, Kwiatkowski TJ Jr, Servadio A, Beaudet AL, McCall AE, Duvick LA, Ranum LP, Zoghbi HY (1993) Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat Genet* 4:221–226
- Pulst SM, Nechiporuk A, Nechiporuk T, Gispert S, Chen XN, Lopes-Cendes I, Pearlman S, Starkman S, Orozco-Diaz G, Lunke A, DeJong P, Rouleau GA, Auburger G, Korenberg JR, Figueroa C, Sahba S (1996) Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nat Genet* 14:269–276
- Riess O, Laccone F, Gispert S, Schols L, Zuhlke C, Vieira-Saecker AM, Herlt S, Wessel K, Epplen J, Weber B, Kreuz F, Chahrokh-Zadeh S, Meindl A, Lunke A, Aguiar J, Macek Jr M, Krebsova A, Macek Sr M, Burk K, Tinschert S, Schreyer I, Pulst SM, Auburger G (1997) SCA2 trinucleotide expansions in German SCA patients. *Neurogenetics* 1:59–64
- Sahba S, Nechiporuk A, Figueroa KP, Nechiporuk T, Pulst SM (1998) Genomic structure of the human gene for spinocerebellar ataxia type 2 (SCA2) on chromosome 12q24.1. *Genomics* 47:359–364
- Sanpei K, Takano H, Igarashi S, Sato T, Oyake M, Sasaki H, Wakisaka A, Tashiro K, Ishida Y, Ikeuchi T, Koide R, Saito M, Sato A, Tanaka T, Hanyu S, Takiyama Y, Nishizawa M, Shimizu N, Nomura Y, Segawa M, Iwabuchi K, Eguchi I, Tanaka H, Takahashi H, Tsuji S (1996) Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. *Nat Genet* 14:277–284
- Silveira I, Coutinho P, Maciel P, Gaspar C, Hayes S, Dias A, Guimaraes J, Loureiro L, Sequeiros J, Rouleau GA (1998) Analysis of SCA1, DRPLA, MJD, SCA2, and SCA6 CAG repeats in 48 Portuguese ataxia families. *Am J Med Genet* 81:134–138
- Squitieri F, Andrew SE, Goldberg YP, Kremer B, Spence N, Zeisler J, Nichol K, Theilmann J, Greenberg J, Goto J, et al (1994) DNA haplotype analysis of Huntington disease reveals clues to the origins and mechanisms of CAG expansion and reasons for geographic variations of prevalence. *Hum Mol Genet* 3:2103–2114
- Takano H, Cancel G, Ikeuchi T, Lorenzetti D, Mawad R, Stevanin G, Didierjean O, Durr A, Oyake M, Shimohata T, Sasaki R, Koide R, Igarashi S, Hayashi S, Takiyama Y, Nishizawa M, Tanaka H, Zoghbi H, Brice A, Tsuji S (1998) Close associations between prevalences of dominantly inherited spinocerebellar ataxias with CAG-repeat expansions and frequencies of large normal CAG alleles in Japanese and Caucasian populations. *Am J Hum Genet* 63:1060–1066
- Takiyama Y, Sakoe K, Soutome M, Namekawa M, Ogawa T, Nakano I, Igarashi S, Oyake M, Tanaka H, Tsuji S, Nishizawa M (1997) Single sperm analysis of the CAG repeats in the gene for Machado-Joseph disease (MJD1): evidence for non-Mendelian transmission of the MJD1 gene and for the effect of the intragenic CGG/GGG polymorphism on the intergenerational instability. *Hum Mol Genet* 6:1063–1068
- Wadia N, Pang J, Desai J, Mankodi A, Desai M, Chamberlain S (1998) A clinicogenetic analysis of six Indian spinocerebellar ataxia (SCA2) pedigrees. The significance of slow saccades in diagnosis. *Brain* 121:2341–2355
- Wells RD, Warren ST (eds) (1998) *Genetic instabilities and hereditary neurological diseases*. Academic Press, San Diego
- Zhang L, Fischbeck KH, Arnheim N (1995) CAG repeat length variation in sperm from a patient with Kennedy's disease. *Hum Mol Genet* 4:303–305
- Zhuchenko O, Bailey J, Bonnen P, Ashizawa T, Stockton DW, Amos C, Dobyns WB, Subramony SH, Zoghbi HY, Lee CC (1997) Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the alpha 1A-voltage-dependent calcium channel. *Nat Genet* 15:62–69