View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Report

Interruptions in the Expanded ATTCT Repeat of Spinocerebellar Ataxia Type 10: Repeat Purity as a Disease Modifier?

Tohru Matsuura,[1,*] Ping Fang,[1,2] Christopher E. Pearson,[3,4] Parul Jayakar,[5] Tetsuo Ashizawa,[6] Benjamin B. Roa,[1,2] and David L. Nelson[1]

[1]Department of Molecular and Human Genetics and [2]Baylor DNA Diagnostic Laboratory, Baylor College of Medicine, Houston; [3]Program of Genetics and Genomic Biology, The Hospital for Sick Children, and [4]Department of Molecular and Medical Genetics, University of Toronto, Toronto; [5]Division of Genetics and Metabolism, Miami Children's Hospital, Miami; and [6]Department of Neurology, University of Texas Medical Branch, Galveston, TX

Spinocerebellar ataxia type 10 (SCA10) is one of numerous genetic disorders that result from simple repeat expansions. SCA10 is caused by expansion of an intronic ATTCT pentanucleotide repeat tract. It is clinically characterized by progressive ataxia, seizures, and anticipation, which can vary within and between families. We report two SCA10 families showing distinct frequencies of seizures and correlations of repeat length with age at onset. One family displayed uninterrupted ATTCT expansions, whereas the other showed multiple interruptions of the repeat by nonconsensus repeat units, which differed both in the length and/or sequence of the repeat unit. Disease-causing microsatellite expansions have been assumed to be composed of uninterrupted pure repeats. Our findings for SCA10 challenge this convention and suggest that the purity of the expanded repeat element may be a disease modifier.

Unstable microsatellite repeats in disease-associated genes cause a large number of inherited neuromuscular and neurological disorders on expansion to abnormal sizes (Zoghbi and Orr 2000; Ranum and Day 2002). Expansion mutations can result in disease through a variety of mechanisms including suppression of gene expression, as in fragile X syndrome (FMR1 [MIM 309550]) and Friedreich ataxia (FRDA [MIM 229300]). Alternatively, normal protein function may be altered, as in Huntington disease (HD [MIM 143100]), Kennedy disease (SBMA [MIM 313200]), dentatorubropallidoluysian atrophy (DRPLA [MIM 125370]), and a number of spinocerebellar ataxias (SCA) caused by the expansion of CAG trinucleotide repeats coding polyglutamine tracts. A third mechanism appears to involve generation of toxic RNAs containing expanded repeats, as in myotonic dystrophy types 1 (DM1 [MIM 160900]) and 2 (DM2 [MIM 602668]) and fragile X–associated tremor/ataxia syndrome (FXTAS) (Ranum and Day 2002).

SCA10 (MIM 603516) is a rare, dominantly inherited neurodegenerative disorder characterized by a unique combination of progressive ataxia, seizure, and anticipation (Grewal et al. 1998; Matsuura et al. 1999; Zu et al. 1999). The mutation of SCA10 is an unstable and massive expansion of an ATTCT repeat in intron 9 of the SCA10 gene mapped to chromosome 22q13.3. SCA10 is the only disease known to be caused by an expansion of a pentanucleotide repeat and, in its mutant form, is among the largest expansion mutations found to date in human genetic diseases (Matsuura et al. 2000). There is a large gap between the documented normal (10–22) and mutated repeat lengths (800–4,500) associated with SCA10. The presence of "premutation" or "reduced penetrance mutation" allele ranges is not yet known for SCA10, in contrast to the polyglutamine diseases including several types of SCAs. The SCA10 ATTCT expansion has unique characteristics of instability in both somatic and germline tissues, distinct from the other repeat expansions (Matsuura et al. 2004). Phenotype and instability patterns are also variable, both within and between families (Rasmussen et al. 2001; Grewal et al. 2002); however, the underlying molecular basis is currently unknown. One possibility is that the

Received July 27, 2005; accepted for publication September 30, 2005; electronically published November 15, 2005.

Address for correspondence and reprints: Dr. David L. Nelson, Department of Molecular and Human Genetics, Room 904E Mailstop BCM225, One Baylor Plaza, Baylor College of Medicine, Houston, TX 77030. E-mail: nelson@bcm.tmc.edu

* Current affiliation: Division of Neurogenetics and Bioinformatics, Department of Advanced Medical Science, Center for Neurological Diseases and Cancer, Nagoya University Graduate School of Medicine, Nagoya, Japan.

Am. J. Hum. Genet. 2006;78:125–129. © 2005 by The American Society of Human Genetics. All rights reserved. 0002-9297/2006/7801-0013$15.00

configuration of the expanded alleles may vary with phenotype and genetic stability; however, such characterization of large expansions has been hindered by difficult cloning, propagating, and sequencing needs.

We recently identified an early-onset patient (onset at age 14 years) with ataxia who was found to have an unusual SCA10 allele of 280 ATTCT repeats, which is much smaller than expansions previously identified in SCA10 families (fig. 1*a*). In this small pedigree, the patient's asymptomatic mother has the same 280-repeat expansion, whereas her asymptomatic father has normal ATTCT alleles. This suggests that repeat lengths <800 can be pathogenic and that a 280-repeat expansion may be within a size range with reduced penetrance. Long-range PCR (LP) successfully amplified the expanded allele as well as the normal allele. The expanded allele was agarose gel–purified and was sequenced from both ends. We were surprised to find interruptions within the ATTCT repeat tract. The configuration of this small expansion was quite complex. The most proximal part (*5′* end) of the expansion is interrupted by multiple repetitive ATGCT repeats (fig. 1*b*) and the latter part (*3′* end) by ATTCTAT septanucleotide repeats. Although it was not possible to obtain the whole sequence throughout the expansion, we did note a long stretch of pure ATTCT repeats in the middle of the expanded tract. The expanded alleles from both the affected daughter and her asymptomatic mother were identical in sequence through the readable portions. We hypothesized that interruptions might play a role in disease presentation and penetrance in this pedigree.

This observation stimulated us to search for interruptions in other SCA10 patients. It is not possible, by use of current methods, to sequence through all other expansions, which, to date, exceed 800 repeats. To overcome this technical hurdle, we assessed the repeat configuration, using an ATTCT repeat-primed PCR (RP) that we previously developed to detect SCA10 expansions (Matsuura and Ashizawa 2002). Here, the reverse primer with a repeat sequence complementary to ATTCT repeats, randomly binds at multiple sites within the ATTCT repeat tract, and generates a mixture of products containing a variable number of repeats with the forward primer in the unique sequence proximal to

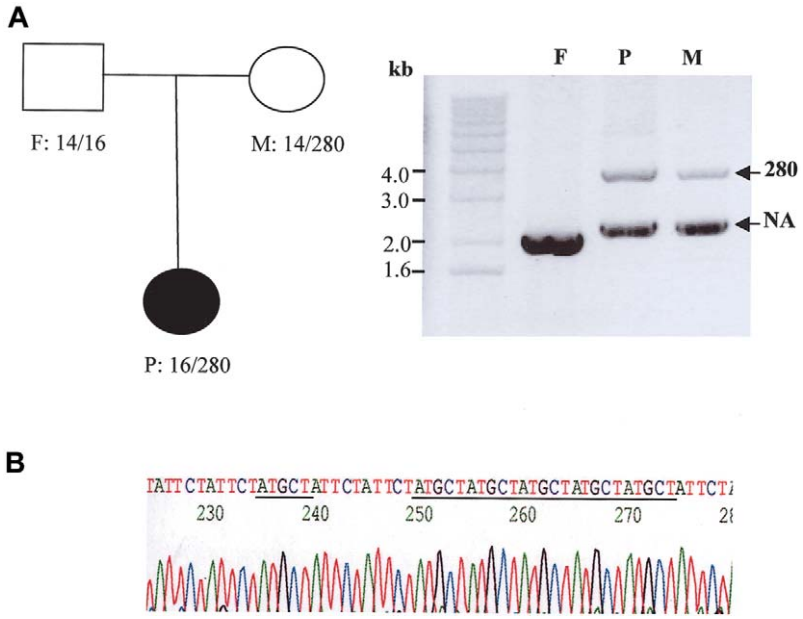

Figure 1 Small expansions in the SCA10 repeat tract. *a,* Long-range PCR of the pedigree of the patient with 280 ATTCT repeats. *Filled symbol,* Affected individual. *Open symbol,* Non-affected individual. The numbers below symbols show the numbers of SCA10 repeats. Of note, the 280-repeat allele is transmitted through her asymptomatic mother. NA = normal allele; F = father; M = mother; and P = patient. *b,* Partial sequence of the expanded 280 repeats. Direct sequencing was performed using long-range PCR products. The first part of the expansion is often interrupted by repetitive ATGCT repeats and the latter part by ATTCTAT septanucleotide repeats (not shown). Although it was impossible to obtain the whole sequence throughout the expansion, there was a long stretch of pure ATTCT repeats in the middle of the expansion. The expanded alleles from both the affected daughter and her asymptomatic mother were identical in sequence through the readable portion. The identical whole sequence is as follows ("ATGCT" and "ATTCTAT" are highlighted in bold and in bold italics, respectively): $(ATTCT)_{10}$ (**ATGCT**)$_{19}$ $(ATTCT)_{14}$ (**ATGCT**) $(ATTCT)_2$ (**ATGCT**)$_5$ $(ATTCT)_2$ (**ATGCT**) $(ATTCT)_8$ (**ATGCT**) $(ATTCT)_3$ (**ATGCT**) $(ATTCT)_{10}$ (**ATGCT**)$_2$ $(ATTCT)_5$ (**ATGCT**) $(ATTCT)_6$ (**ATGCT**) $(ATTCT)_2$ (**ATGCT**) $(ATTCT)_5$ (**ATGCT**) $(ATTCT)_4$ (**ATGCT**) (ATTCT) (**ATGCT**) $(ATTCT)_2$ (**ATGCT**)$_3$ (ATTCT) (**ATGCT**) $(ATTCT)_2$ (**ATGCT**)$_3$ (ATTCT) (**ATGCT**) $(ATTCT)_4$ (**ATGCT**)$_2$ … $(ATTCT)_{79}$ (***ATTCTAT***) $(ATTCT)_2$ (***ATTCTAT***) $(ATTCT)_3$ (***ATTCTAT***) $(ATTCT)_3$ (***ATTCTAT***) $(ATTCT)_{12}$.

the repeat. To determine whether there are interruptions in SCA10 expansions, RP products of SCA10 patients were size separated by electrophoresis at a higher resolution than is typical for diagnostic purposes. In principle, a pure uninterrupted ATTCT expansion would be predicted to show a continuous ladder of products, as the reverse primer anneals randomly within the pure ATTCT repeat tract. Interruptions in repeat purity would expectedly yield a stuttered ladder of products, with stutter locations coincident with interruptions.

We studied the two initially identified large families, since both are clinically and genetically well characterized (Grewal et al. 1998, 2002; Matsuura et al. 1999, 2000; Zu et al. 1999). All patients in family 1 showed continuous ladders, whereas those in family 2 showed a complex interruption pattern (fig. 2*a*). This pattern was similar in all the affected individuals in family 2 through three generations, although the location of interruptions was variable. The largest PCR products were gel purified, cloned, and sequenced. Expanded alleles from family 2 had two different septarepeat interruptions, ATTTTCT and ATATTCT, which are duplicated (fig. 2*b*), whereas family 1 had only pure repeats. The interrupted sequence configuration in family 2 was preserved in the affected members of three generations. We isolated no clones with >80 pentanucleotides, suggesting a limit to the capacity of *Escherichia coli* systems for propagating these sequences. The data indicate that the expanded ATTCT repeat tract is highly mutable for both length and sequence content in human patients with variable sequence changes.

Sequence interruptions in normal alleles at SCA1 (MIM 164400), SCA2 (MIM 183090), FMR1, and FRDA have been well characterized (Chung et al. 1993; Kunst and Warren 1994; Imbert et al. 1996; Pulst et al. 1996; Sanpei et al. 1996; Cossee et al. 1997; Montermini et al. 1997; Gunter et al. 1998). They appear to stabilize their repeat tracts, and the loss of interruptions is associated with instability and repeat expansion. Most normal SCA1, SCA2, and FMR1 alleles carry one or more CAT, CAA, or AGG interruption, whereas GAG-GAA hexanucleotide interruptions seen in FRDA are confined to premutations or larger normal alleles of FRDA. Interrupting sequences in a portion of large nor-

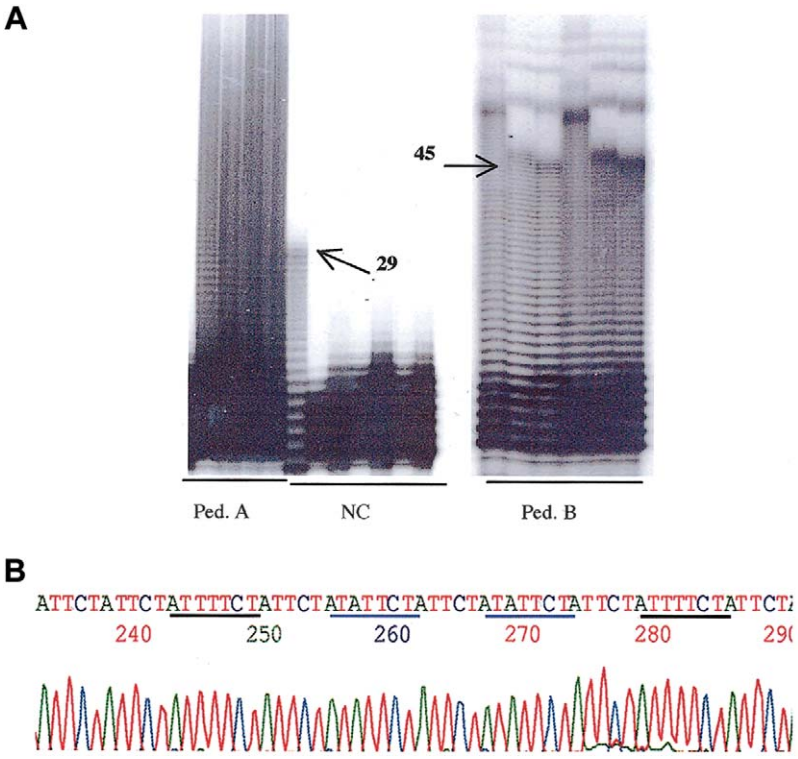

Figure 2 Different patterns in ATTCT repeat-primed PCR between SCA10 families. *a*, SCA10 Patients in family 1 showed continuous multiple ladders, whereas family 2 showed the complex interruption pattern beginning from the location of 43 repeats. NC = Normal controls. The left-end lane of NC is a recently identified normal individual with a 29 pure ATTCT repeat allele. *b*, Sequence of a representative clone from ATTCT-primed PCR products of a patient in family 2. Both septarepeat interruptions, ATTTTCT and ATATTCT, are duplicated in the repeat tract. The whole sequence of the repeat (equivalent numbers of pentamers = 76) is $(ATTCT)_{41}$ (**ATTTTCT**) (ATTCT) (**ATATTCT**) (ATTCT) (**ATATTCT**) (ATTCT) (**ATTTTCT**) $(ATTCT)_{11}$ (**ATATTCT**) $(ATTCT)_5$ (**ATATTCT**) $(ATTCT)_{10}$.

mal FRDA alleles are also thought to act as anchors to prevent further expansions. In stark contrast with these triplet-repeat diseases, SCA8 (MIM 603680) has normal alleles without interruptions and expanded alleles with interruptions (Moseley et al. 2000). In our initial study (Matsuura et al. 2000), sequence analysis of 40 SCA10 alleles from 20 normal individuals who were homozygous for alleles ranging from 11 to 16 repeats showed tandem ATTCT repeats without interruptions. Considering this situation, we reevaluated the sequencing of a wider range of normal allele sizes, focusing especially on large normal alleles (⩾17 repeats), which are found at a frequency of 7.1% (43/604 alleles) (Matsuura et al. 2000). As shown in table 1, more than half of large normal alleles have ATTGT-TTTCT or TTTCT interruption (71.1% of 38 examined), whereas all 78 normal alleles of 11–16 repeats showed an uninterrupted pentanucleotide. Interestingly, the location of all interruptions in normal alleles is confined to the second to the last (most distal) repeat. This is similar to the FRDA case, where most GAGGAA interruptions are restricted to the fifth from the last repeat (Cossee et al. 1997; Montermini et al. 1997). From these data, it is difficult to predict the consequences of ATTGT-TTTCT, TTTCT interruptions in the instability or stability of the normal length repeat tract. Interruptions have been observed in the majority of large normal alleles of ⩾17 repeats. At the same time, cloned expansion alleles from two SCA10 families have shown uninterrupted, as well as interrupted, mutant repeats.

As reported elsewhere (Grewal et al. 2002), the presence of seizures was found to be significantly different in the two families described above. Some 25% of affected members in family 1 exhibited seizure, whereas 80% in family 2 suffered from this SCA10-associated phenotype. Correlation between age at onset and repeat number also differs between the families despite clinically observed anticipation; we found a strong inverse correlation ($r^2 = 0.79$, $P = .001$) in family 1 and no significance in family 2. Although we could not apply this RP technique to amplify the 3′ end of the expansion (since the genomic sequence downstream of the ATTCT repeat is immediately followed by a long stretch of *Alu*Sg, MER11A, and MER11B repeats) to confirm that the ATTCT tracts are uninterrupted in the whole expansion in family 1, differences in the sequence interruption patterns may well explain their variable effect on phenotype and patterns of instability. As seen in SCA1, pathogenicity in the overlapping zone between normal and disease ranges is directly related to the presence/absence of interruptions (Zühlke et al. 2002) and to their locations (Sobczak and Krzyzosiak 2004). Although it will require analysis of additional families to confirm this finding, our data strongly suggest that the sequence configuration itself can be a disease modifier in this noncoding repeat expansion disease. This may be explained by the notion that these pathogenic repeat expansions act as toxic RNA species, which might have different effects dependent on sequence and/or RNA structure.

Table 1

Sequence Configuration of ATTCT Large Normal Alleles (⩾17 Repeats)

Allele Size	Pure Motif	n-AB-1	n-B-1	Total
17	6	6	4	16
18	2	9	1	12
19	1	1	3	5
20	1	1	0	2
21	0	0	1	1
22	0	0	1	1
29	1	0	0	1
Total	11	17	10	38

NOTE.—A = ATTGT and B = TTTCT.

It has been assumed that noncoding microsatellite expansions contain uninterrupted pure repeats, on the basis of the finding that loss of interruptions appears to be a precursor to repeat instability. However, large expanded repeats have not typically been cloned and sequenced; thus, there is no direct evidence for this notion. We have demonstrated that this is not the case for SCA10, and we recommend caution regarding assumptions about sequence configurations in other disorders. Indeed, interruptions may be the basis for the complex relationship between repeat length and disease severity seen in other noncoding repeat expansions, including DM1, DM2, SCA8, SCA12 (MIM 604326), and FXTAS. It is also possible that interruptions with divergent effects on phenotype may also serve to further distinguish these diseases from the polyglutamine disorders.

Acknowledgments

This work was supported by National Ataxia Foundation and National Organization for Rare Disorders grants (to T.M.); National Institutes of Health (NIH) National Institute of Child Health and Human Development grants HD38038 and HD29256 (to D.L.N.); NIH National Institute of Neurological Disorders and Stroke grant NS41547 (to T.A.); by the FRAXA Research Foundation and the BCM Mental Retardation Research Center (to D.L.N.); and by the Canadian Institutes of Health Research (CHIR) and the Fragile X Research Foundation of Canada (to C.E.P). C.E.P. is a CHIR scholar and a Canadian Genetic Disease Network scholar.

Web Resources

Accession numbers and URLs for data presented herein are as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi

.nlm.nih.gov/Omim/ (for FMR1, FRDA, HD, SBMA, DRPLA, DM1, DM2, FXTAS, SCA1, SCA2, SCA8, SCA10, and SCA12)

References

Chung MY, Ranum LP, Duvick LA, Servadio A, Zoghbi HY, Orr HT (1993) Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. Nat Genet 5:254–258

Cossee M, Schmitt M, Campuzano V, Reutenauer L, Moutou C, Mandel JL, Koenig M (1997) Evolution of the Freidreich's ataxia trinucleotide repeat expansion: founder effect and premutations. Proc Natl Acad Sci USA 94:7452–7457

Grewal RP, Achari M, Matsuura T, Durazo A, Tayag E, Zu L, Pulst SM, Ashizawa T (2002) Clinical features and ATTCT repeat expansion in spinocerebellar ataxia type 10. Arch Neurol 59:1285–1290

Grewal RP, Tayag E, Figueroa KP, Zu L, Durazo A, Nunez C, Pulst SM (1998) Clinical and genetic analysis of a distinct autosomal dominant spinocerebellar ataxia. Neurology 51:1423–1426

Gunter C, Paradee W, Crawford DC, Meadows KA, Newman J, Kunst CB, Nelson DL, Schwartz C, Murray A, Macpherson JN, Sherman SL, Warren ST (1998) Re-examination of factors associated with expansion of CGG repeats using a single nucleotide polymorphism in FMR1. Hum Mol Genet 7:1935–1946

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

Kunst CB, Warren ST (1994) Cryptic and polar variation of the fragile X repeat could result in predisposing normal alleles. Cell 77:853–861

Matsuura T, Achari M, Khajavi M, Bachinski LL, Zoghbi HY, Ashizawa T (1999) Mapping of the gene for a novel spinocerebellar ataxia with pure cerebellar signs and epilepsy. Ann Neurol 45:407–411

Matsuura T, Ashizawa T (2002) Polymerase chain reaction amplification of expanded ATTCT repeat in spinocerebellar ataxia type 10. Ann Neurol 51:271–272

Matsuura T, Fang P, Lin X, Khajavi M, Tsuji K, Rasmussen A, Grewal RP, Achari M, Alonso ME, Pulst SM, Zoghbi HY, Nelson DL, Roa BB, Ashizawa T (2004) Somatic and germline instability of the ATTCT repeat in spinocerebellar ataxia type 10. Am J Hum Genet 74:1216–1224

Matsuura T, Yamagata T, Burgess DL, Rasmussen A, Grewal RP, Watase K, Khajavi M, McCall AE, Davis CF, Zu L, Achari M, Pulst SM, Alonso E, Noebels JL, Nelson DL, Zoghbi HY, Ashizawa T (2000) Large expansion of ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. Nat Genet 26:191–194

Montermini L, Andermann E, Labuda M, Richter A, Pandolfo M, Cavalcanti F, Pianese L, Iodice L, Farina G, Monticelli A, Turano M, Filla A, De Michele G, Cocozza S (1997) The Friedreich ataxia GAA triplet repeat: premutation and normal alleles. Hum Mol Genet 6:1261–1266

Moseley ML, Schut LJ, Bird TD, Koob MD, Day JW, Ranum LP (2000) SCA8 CTG repeat: en masse contractions in sperm and intergenerational sequence changes may play a role in reduced penetrance. Hum Mol Genet 9:2125–2130

Pulst SM, Nechiporuk A, Nechiporuk T, Gispert S, Chen XN, Lopes-Cendes I, Pearlman S, Starkman S, Orozco-Diaz G, Lunkes 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

Ranum LPW, Day JW (2002) Dominantly inherited non-coding microsatellite expansion disorders. Curr Opin Genet Dev 12:266–271

Rasmussen A, Matsuura T, Ruano L, Yescas P, Ochoa A, Ashizawa T, Alonso E (2001) Clinical and genetic analysis of four Mexican families with spinocerebellar ataxia type 10 (SCA10). Ann Neurol 50:234–239

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

Sobczak K, Krzyzosiak WJ (2004) Imperfect CAG repeats form diverse structures in SCA1 transcripts. J Biol Chem 279:41563–41572

Zoghbi HY, Orr HT (2000) Glutamine repeats and neurodegeneration. Annu Rev Neurosci 23:217–230

Zu L, Figueroa KP, Grewal R, Pulst SM (1999) Mapping of a new autosomal dominant spinocerebellar ataxia to chromosome 22. Am J Hum Genet 64:594–599

Zühlke C, Dalski A, Hellenbroich Y, Bubel S, Schwinger E, Burk K (2002) Spinocerebellar ataxia type 1 (SCA1): phenotype-genotype correlation studies in intermediate alleles. Eur J Hum Genet 10:204–209