

Detection of Large Pathogenic Expansions in *FRDA1*, *SCA10*, and *SCA12* Genes Using a Simple Fluorescent Repeat-Primed PCR Assay

Claudia Cagnoli,* Chiara Michielotto,*
Tohru Matsuura,† Tetsuo Ashizawa,‡
Russell L. Margolis,§ Susan E. Holmes,§
Cinzia Gellera,¶ Nicola Migone,* and
Alfredo Brusco*

From the Dipartimento di Genetica Biologia e Biochimica,*
Università degli Studi di Torino and Azienda Ospedaliera San
Giovanni Battista di Torino, S.C. Genetica Medica, Torino, Italy;
the Department of Molecular and Human Genetics,† Baylor
College of Medicine, Houston, Texas; the Department of
Neurology,‡ The University of Texas Medical Branch, Galveston,
Texas; the Department of Psychiatry,§ Johns Hopkins University
School of Medicine, Baltimore, Maryland; and the Divisione di
Biochimica e Genetica,¶ Istituto Neurologico "Carlo Besta",
Milano, Italy

At least 18 human genetic diseases are caused by expansion of short tandem repeats. Here we describe a successful application of a fluorescent PCR method for the detection of expanded repeats in *FRDA1*, *SCA10*, and *SCA12* genes. Although this test cannot give a precise estimate of the size of the expansion, it is robust, reliable, and inexpensive, and can be used to screen large series of patients. It proved useful for confirming the presence of large expansions in the Friedreich ataxia gene following an ambiguous result of long-range PCR, as well as rapid pre-screening for large repeat expansions associated with Friedreich ataxia and *SCA10* and the shorter repeat expansions associated with *SCA12*. (*J Mol Diagn* 2004, 6:96–100)

At least 18 neurological and muscular diseases have been shown so far to derive from expansion of short tandem repeats located either in coding or non-coding gene sequences.¹ The pathogenic expansions in coding regions are mostly due to CAG tri-nucleotides and rarely exceed 100 triplets, whereas the expansions of repeats located in introns or in 3' or 5' untranslated gene regions (UTRs) may be much longer. Indeed, in the genes for myotonic dystrophy (*DM1*, CTG repeat in 3'UTR, and *DM2*, CCTG repeat in intron 1), Friedreich ataxia (*FRDA1*, GAA repeat in intron 1), and spinocerebellar ataxia type 10 (*SCA10*, ATTCT repeat in intron 9), the pathological expansion varies from hundreds to thousands of repeating units. The most striking example occurs in *SCA10*,

with an upper limit of ~4500 pentamers, corresponding to a total expansion of 22.5 kb.² Such large expansions cannot be detected by standard PCR protocols.

Routine testing for *FRDA1* expansions is usually performed by Southern blot analysis or long-range PCR.^{3–5} Though the latter technique is generally preferred over the more cumbersome Southern blot analysis, it may generate faint bands within the pathological range that may or may not reflect true expansions. Interpretation of these bands requires further tests such as Southern analysis of the PCR products with (GAA)_{10–12} oligonucleotide probes, especially in healthy carriers.

Looking for alternative methods, Warner et al⁶ proposed a repeat-primed fluorescent PCR to detect large CTG-repeats in myotonic dystrophy. Here we describe the successful application of the fluorescent PCR protocol to routine testing of three diseases characterized (Friedreich ataxia, and spinocerebellar ataxia type 10) or potentially characterized (spinocerebellar ataxia type 12) by large repeat expansions.

In developing our analysis we considered the polymorphism of the normal and expanded repeats in the three genes. Approximately 83% of the GAA-repeat alleles in the *FRDA1* gene have 6 to 12 triplets (the 9 GAA-variant alone accounts for ~50% of the alleles in this range) and ~17% show 16 to 29 triplets; non-pathogenic alleles with 34 to 36 repeats (often interrupted by at least one "not perfect" triplet) have occasionally been described.^{5,7,8} Pathogenic alleles have been reported to vary from 66 to 1200 GAA repeats, but alleles <200 repeats are very rare.⁹ Normal *SCA10* alleles range from 10 to 22 with a peak between 13 and 14 pentaplets, while the pathogenic alleles are between 280 to 4500.^{2,10} *SCA12* normal alleles are between 7 to 31 CAGs with a main 10 repeats allele accounting for nearly 50% of *SCA12* genes. The pathological range starts from 55 up to 78 CAGs, although a 45 repeat allele in a gray zone has been described.^{11–13} We used these parameters to establish cri-

Supported in part by the "Associazione Emma & Ernesto Rulfo per la Genetica Medica" and National Institutes of Health grants NS41547 (T.A.) and NS042930 (R.L.M. and S.E.H.). Dr. C. Cagnoli has been recipient of a fellowship from the "Associazione Italiana Sindromi Atassiche - Piemonte e Valle d'Aosta".

Accepted for publication December 2, 2003.

Address reprint requests to Dr. Alfredo Brusco, Dipartimento di Genetica Biologia e Biochimica, Università degli Studi di Torino, via Santena 19, 10126, Torino, Italy. E-mail: alfredo.brusco@unito.it.

Table 1. Genotypes of Patients/Carriers and Normal Allele Range in Studied Healthy Controls

Friedreich ataxia (GAA repeats)			SCA10 (ATTCT repeats)	SCA12 (CAG repeats)
Patients (n = 43)		Carriers (n = 28)	Patients (n = 10)	Patients (n = 6)
200/200	160/830	9/100	14/1380	13/66
220/220	660/830	7/170	14/1760	14/67
300/330	300/900	9/260	14/2240	10/72
260/360	850/1050	8/270	14/2340	10/73
270/370	230/1100	12/270	14/2540	15/74
380/380	230/1100	17/300	12/2780	10/76
300/400	430/1100	9/320	13/2780	
330/400	650/1100	10/360	12/2820	
200/430	760/1100	9/430	13/3500	
270/430	1000/1100	10/430	12/3700	
350/450	830/1200	9/500		
410/450	500/1200	10/500		
160/530	330/1250	9/520		
330/560	950/1250	9/600		
360/600	260/1430	10/600		
160/600		9/730		
160/660		9/730		
360/660		9/750		
650/650		10/760		
660/700		20/760		
700/730		10/830		
150/750		9/930		
430/760		7/950		
600/760		8/980		
700/760		8/980		
170/770		9/1050		
250/800		8/1100		
800/800		9/1100		
Allele range in controls				
(n = 152)		(n = 50)		(n = 50)
6–26		11–21		8–21

teria for distinguishing among normal and expanded repeats.

Materials and Methods

Repeat-Primed PCR

Our PCR assay is based on the method of Warner et al.⁶ a fluorescent primer is designed in a locus-specific region upstream of the unstable repeat of interest. The companion reverse primer (on the complementary strand) consists of 5 to 8 units of the repeat and a 5' tail that is used as an anchor for a second reverse primer, which prevents progressive shortening of the PCR products during subsequent cycles. PCRs were performed with the following general conditions: 200 to 1000 ng of genomic DNA, 800 mmol/L of locus-specific (forward) and 40 mmol/L of repeat-specific (reverse) primer and 800 mmol/L of the "common" flag primer (5'-tacgcatccagtttgagacg), 200 μ mol/L dNTPs, 1.5 mmol/L MgCl₂, and 1U of TaqGold in 1X specific buffer (Applera Italia, Monza, Italy).

Locus-specific PCR conditions were: *FRDA1* (primers: 5'-HEX-gggattggtgccagtgctaaagtttag; 5'-tacgcatccagtttgagacg ttcttcttcttcttcttcttctt), 3% dimethyl sulfoxide (DMSO) or 1 mol/L betaine, with the following cycling

parameters: initial denaturation 7 minutes at 95°, 40 cycles consisting in 45 seconds at 95°, 1 minute at 60°, 3 minutes at 72°, and a final extension of 10 minutes at 72°. *SCA10* (primers: 5'-6-FAM-cagatggcagaatgataaactcaa; 5'-tacgcatccagtttgagacg agaatagaatagaatagaat), with the following cycling parameters: initial denaturation 7 minutes at 95°, 40 cycles consisting in 45 seconds at 95°, 1 minute at 55°, 3 minutes at 72°, and a final extension of 10 minutes at 72°. *SCA12* (primers: 5'-6-FAM-tgctgggaaagagtcgtg; 5'-tacgcatccagtttgagacg ctgctgctgctgctg), 15% DMSO, with the following cycling parameters: initial denaturation 9 minutes at 95°, 40 cycles consisting in 45 seconds at 95°, 1 minute at 56°, 4 minutes at 72° +10 seconds each cycle, and 10 minutes final extension at 72°.

Genomic DNA and Patients

Genomic DNA was extracted from blood samples using the QIAamp DNA mini blood kit (QIAGEN, Milan, Italy). Our controls for *FRDA1* were 152 healthy subjects, 43 patients with known expansions on both alleles, and 28 obligate carriers. The presence and size of the expansion were investigated in controls, carriers, and patients by a long-range PCR assay described by Monrós et al.¹⁴ For *SCA10* and *SCA12* we analyzed 50 healthy subjects as controls, and 10 *SCA10* and 6 *SCA12* patients with known

expansions. All were tested by PCR and Southern blot, whenever necessary, for SCA10,² or by PCR for SCA12.¹¹ The patient and control genotypes at the three repeated regions are shown in Table 1. Each patient and control was tested at least twice, and results were blindly analyzed by a different operator.

Analysis of Fluorescent PCR Fragments and Interpretation of Data

Analysis of fluorescent PCR products was performed with an ABI-Prism 377 automatic sequencer using a 36-cm gel (5.5% Long ranger acrylamide mix (BMA-Cambrex, Italy), 7 mol/L urea, 1X TBE), and a ROX-GS500 internal standard marker (Applera). Data were examined using Genescan 3.1 software. Samples were considered reliable in the range of 200 to 7000 fluorescence units (FU).

To quantitatively measure the type of skewness of the peaks created by our repeat-primer assay, the size (abscissa) and height (ordinate) of the peaks were plotted using Microsoft Excel 2000 software. The peaks used were selected in the 230- to 330-bp range for the FRDA1 locus and in the 133- to 300-bp range for the SCA10 locus, with a minimum height of 20 FU. The program calculates a regression line and a square regression coefficient (R^2), which gives a parameter to measure the slope of the peaks.

Results

Friedreich Ataxia

Representative examples of fluorescent PCR profiles of genotypes from two controls (one homozygote, 9/9, and one heterozygote, 9/26), one healthy carrier of a pathogenic repeat (9/100) and two FA patients (200/200 and 300/900) are shown in Figure 1 (panels 1 to 5).

In all cases and controls tested, the contour profile presented multiple spikes due to the possible alternative annealing sites of the reverse primer. Instead of the expected 3-bp periodicity, the spikes appeared regularly spaced by 1-bp, likely produced by a polymerase artifact. In controls, the pattern given by normal alleles abruptly end at the right tail (see Figure 1, panels 1 and 2), and the number of triplets of the larger allele could be estimated from the size of the main peak at the right side of the panel. The peculiarity of the pattern given by expanded alleles in carriers and affected patients is shown by the shape at the right side of the tail, where the PCR product exhibited a characteristic array of spikes that gradually end (in Figure 1, compare panels 1 and 2 and 3 and 5, and note the conformation marked by the hyphenated arrow). Since peaks rapidly disappear reaching a fluorescence close to zero (baseline), it is impossible to estimate the dimension of the expanded alleles.

A parameter to measure how the peaks decrease is the R^2 coefficient calculated from a regression line obtained by plotting the height *versus* size of the peaks (see Figure 2 as example). Our results clearly show that this approach is always able to distinguish normal subjects

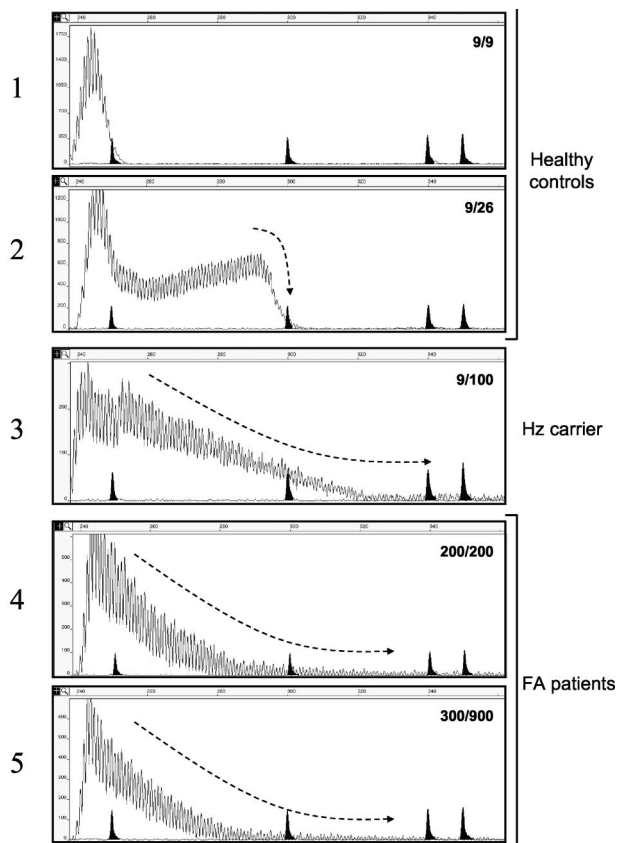


Figure 1. Fluorescent repeat-primed PCR analysis of the Friedreich ataxia locus. The peak profiles of healthy controls with two alleles in the normal size range (panels 1 and 2) are distinct from the profiles found in subjects with one (Hz carrier, panel 3) or two pathogenic alleles (FA patients, panels 4 and 5). The genotype of *FRDA1* GAA triplets is shown inside each panel. Marker peaks of 250, 300, 340, and 350 bp are shadowed. Size in bp is also reported on abscissa, while ordinate shows an arbitrary fluorescence intensity scale. The **hyphenated arrow** marks the right tail shape of the peak traces.

from those carrying an expansion, although in most cases it was impossible to establish if one or two expanded alleles were present (Table 2). This is also apparent from the examples reported in Figure 1 (compare panel 3 and panels 4 and 5).

Patients/carriers were reliably discriminated from controls over a wide range of expansions (see Table 1). We did not test subjects with expanded alleles at the uppermost end of the normal range (32 to 34 GAAs) and the lowermost end of the expanded range (66 GAAs). These variants are uncommon; fewer than 1% of normal alleles have more than 26 triplets and pathogenic alleles carrying fewer than 150 triplets are extremely rare.^{5,7,8} However, since we could accurately estimate the length of longer normal alleles, we expect that even the extremely rare variants with normal alleles above 32 triplets would be correctly classified.

SCA10

Standard PCR protocols cannot amplify the long pathogenic alleles of SCA10 (280 to 4500 repeating pentameric units), but only the short repeats contained in normal

alleles, which ranged from 11 to 21 pentaplets in our 50 healthy controls. Two of us¹⁵ recently set up a repeat-primed PCR similar to the one here described, using radioactive-labeled primers. Here we report the modification of this technique. Using fluorescent primers, we were able to visualize a short series of discrete peaks with a 5-bp periodicity in both cases and controls.

In all 10 patients tested (heterozygous for an expanded allele with 1380 to 3500 repeats) the fluorescent PCR profile was clearly distinguishable from that found in the control cases by the presence of a smeared ladder of peaks, longer (to the right) than observed with normal alleles, with a characteristic 5-bp periodicity that progressively decreased in intensity to eventually fuse with the base line (Figure 3, bottom). Also the R^2 value of the regression line, calculated as in FA, provided a useful quantitative parameter for establishing the presence of an expansion (see Table 2). In genotypes lacking pathogenic alleles, the number of repeats present in the largest allele could be calculated by adding four to the total number of peaks (Figure 3, top, the first peak reflects the number of repeats in the primer). The smeared ladder adjacent to the normal allele does not provide the size of the pathogenic allele; quantitation requires Southern analysis.

SCA12

Pathogenic expansions in the unstable CAG repeat located in an upstream region of *PPP2R2B* have been identified in only a few families. The size range of pathogenic alleles so far described is 55 to 78 repeats.^{11,12} The PCR described in the first report of a *SCA12* family¹¹ could easily detect the size of the expanded allele. Given the paucity of data on the size of pathogenic alleles, we decided to set up the repeat-primed PCR for a routine *SCA12* screening. The test was validated in six different *SCA12* patients carrying 66 to 76 repeats, and in 50 healthy subjects (range of normal repeats 8 to 21). As shown in Figure 3, the expanded allele was clearly recognized as additional, bell-shaped peaks consistent in size with the previously estimated number of CAG repeats.

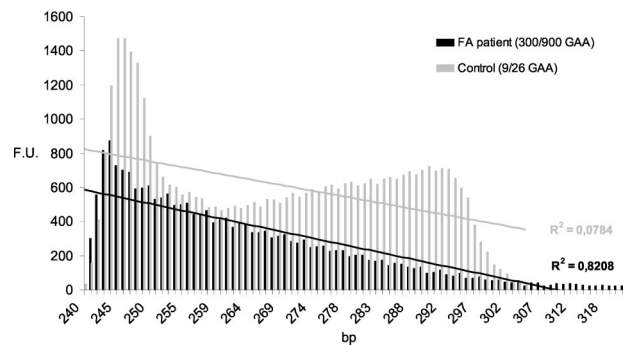


Figure 2. Measurement of the slope of the peaks profile. Height (F.U.) and size (bp) of repeat-primed PCR peaks can be plotted to calculate a regression line and an R^2 coefficient (see Materials and Methods). A control subject and a patient corresponding to panels 2 and 5 in Figure 1 are displayed as example. The R^2 values were always able to discriminate a carrier of one or two expansions *versus* a normal control (see Table 2).

Discussion

The fluorescent repeat-primed PCR in our hands was free from false-positive or false-negative results, confirming that this technique is robust, sensitive, specific, and reliable. However, we did not succeed in using this method to detect *SCA8* repeats, suggesting that local conformational constraints might prevent the amplification of certain repeats. Our method is particularly valuable in routine screening of large populations; we used this test to determine the prevalence of *SCA10* among Italian *SCA* families (Cagnoli C, Michielotto C, Matsuura T, Ashizawa T, Margolis RL, Holmes SE, Gellera C, Migone N, Brusco A, manuscript in preparation) and we believe it could be useful for screening Friedreich ataxia among patients with a generic diagnosis of ataxia. Indeed, supposedly idiopathic cases of cerebellar ataxia have occasionally been revealed as atypical Friedreich ataxia.¹⁶ This is of particular interest since *SCA* patients are often not tested for *FRDA1* expansions in a routine, clinical setting, in part due to the lack of a simple and inexpensive pre-screening test.

Table 2. Data Analysis of the *FRDA1* and *SCA10* Repeat-Primer PCR Results

	<i>FRDA1</i> gene		
	Number of expanded alleles		
	0	1	2
R^2 mean \pm S.D.	0.116 \pm 0.105	0.797 \pm 0.091	0.859 \pm 0.046
Minimum	0.001	0.595	0.753
Maximum	0.331	0.952	0.947
Tested subjects	30	26	43
	<i>SCA10</i> gene		
	Number of expanded alleles		
	0	1	
R^2 mean \pm S.D.	0.131 \pm 0.099	0.612 \pm 0.102	
Minimum	0.003	0.515	
Maximum	0.287	0.833	
Tested subjects	15	10	

R^2 , square linear regression coefficient; S.D., standard deviation.

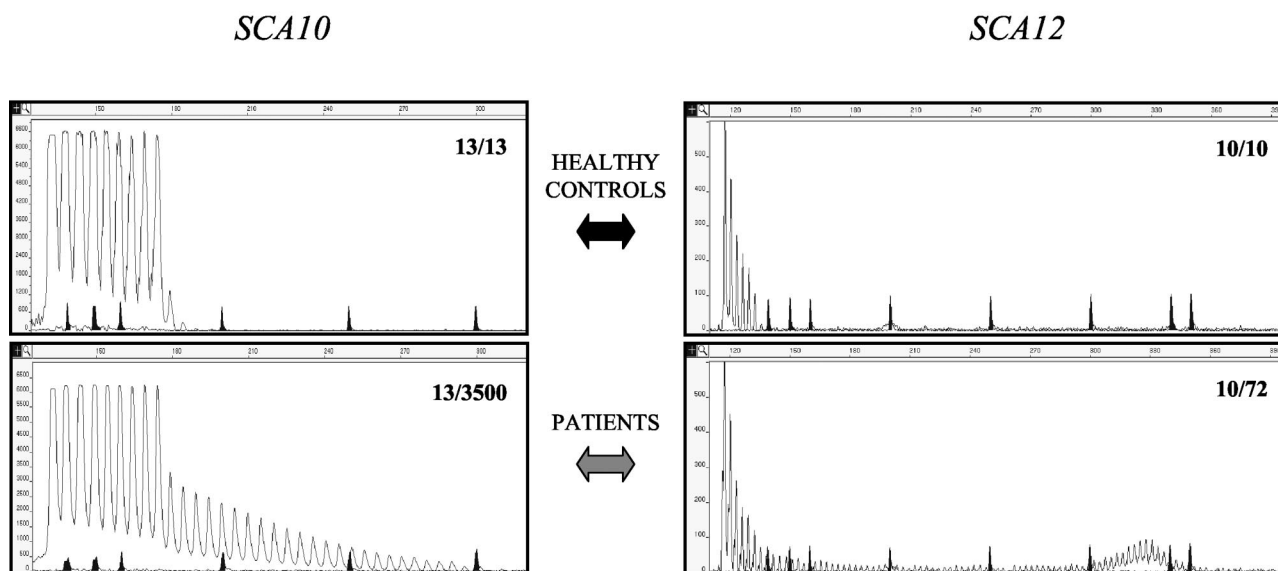


Figure 3. Fluorescent repeat-primed PCR analysis in the *SCA10* (left) and *SCA12* (right) loci. **Top panels** show healthy control subjects, while **bottom panels** display patients carrying one expanded allele. The genotype at the two loci expressed in number of repeats is reported inside each **panel**. Marker peaks of 139, 150, 160, 200, 250, 300, 340, and 350 bp are shadowed. Size in bp is also reported on abscissa, while ordinate shows an arbitrary fluorescence intensity scale.

Finally, we have shown that the fluorescent repeat-primed PCR, originally developed for the analysis of myotonic dystrophy type I, is a valuable tool for the identification of large expansions in Friedreich ataxia, spinocerebellar ataxia type 10, and SCA12, diseases associated with unstable repeats in non-coding regions that are potentially susceptible to large expansions that can be missed by typical long-range PCR. We suggest that repeat-primer PCR can be used clinically as a rapid pre-screening test in the diagnostic protocols of FA, SCA10, and SCA12, followed by other techniques, such as long-range PCR or Southern blotting, for positive cases in which the size of the expansion is critical or in which heterozygotes must be distinguished from homozygotes.

Acknowledgments

We thank M. De Marchi for critical reading of the manuscript.

References

- Ranum LPW, Day JW: Dominantly inherited, non-coding microsatellite expansion disorders. *Curr Opin Genet Dev* 2002, 12:266–271
- 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: Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. *Nat Genet* 2000, 26:191–194
- Campuzano V, Montermini L, Moltò MD, Pianese L, Cossée M, Cavalcanti F, Monros E, Rodius F, Duclos F, Monticelli A, Zara F, Cañizares J, Koutnikova H, Bidichandani SI, Gellera C, Brice A, Trouillas P, De Michele G, Filla A, De Frutos R, Palau F, Patel FI, Di Donato S, Mandel J-L, Coccoza S, Koenig M, Pandolfo M: Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science* 1996, 271:1423–1427
- Durr A, Cossée M, Agid Y, Campuzano V, Mignard C, Penet C, Mandel J-L, Brice A, Koenig M: Clinical and genetic abnormalities in patients with Friedreich's ataxia. *N Engl J Med* 1996, 335:1169–1175
- 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, Coccoza S: The Friedreich ataxia GAA triplet repeat: premutation and normal alleles. *Hum Mol Genet* 1996, 6:1261–1266
- Warner JP, Barron LH, Goudie D, Kelly K, Dow D, Fitzpatrick DR, Brock DJH: A general method for the detection of large CAG repeat expansions by fluorescent PCR. *J Med Genet* 1996, 33:1022–1026
- Cossée M, Schmitt M, Campuzano V, Reutenauer L, Moutou C, Mandel J-L, Koenig M: Evolution of the Friedreich's ataxia trinucleotide repeat expansion: founder effect and premutations. *Proc Natl Acad Sci USA* 1997, 94:7452–7457
- Epplen C, Epplen JT, Frank G, Mitterski B, Santos EJM, Schöls L: Differential stability of the (GAA)_n tract in the Friedreich ataxia (STM7) gene. *Hum Genet* 1997, 99:834–836
- Filla A, De Michele G, Cavalcanti F, Pianese L, Monticelli A, Campanella G, Coccoza S: The relationship between trinucleotide (GAA) repeat length and clinical features in Friedreich ataxia. *Am J Hum Genet* 1996, 59:554–560
- Fang P, Matsuura T, Teive HAG, Raskin S, Jayakar P, Schmitt E, Ashizawa T, Roa BB: *Am J Hum Genet* 2002, 71:A2243
- Holmes SE, O'Hearn EE, McInnis MG, Gorelick-Feldman DA, Kleiderlein JJ, Callahan C, Kwak NG, Ingersoll-Ashworth RG, Sherr M, Sumner AJ, Sharp AH, Ananth U, Seltzer WK, Boss MA, Viera-Saecker AM, Epplen JT, Riess O, Ross CA, Margolis RL: Expansion of a novel CAG trinucleotide repeat in the 5' region of PPP2R2B is associated with SCA12. *Nat Genet* 1999, 23:391–392
- Fujigasaki H, Verma IC, Camuzat A, Margolis RL, Zander C, Lebre AS, Jamot L, Saxena R, Anand I, Holmes SE, Ross CA, Durr A, Brice A: SCA12 is a rare locus for autosomal dominant cerebellar ataxia: a study of an Indian family. *Ann Neurol* 2001, 49:117–121
- Srivastava AK, Choudhry S, Gopinath MS, Ray S, Tripathi M, Brahmachari SK, Jain S: Molecular and clinical correlation in five Indian families with spinocerebellar ataxia 12. *Ann Neurol* 2001, 50:796–800
- Monrós E, Moltò MD, Martínez F, Cañizares J, Blanca J, Vilchez JJ, Prieto F, De Frutos R, Palau F: Phenotypic correlation and intergenerational dynamics of the Friedreich ataxia GAA trinucleotide repeat. *Am J Hum Genet* 1997, 61:101–110
- Matsuura T, Ashizawa T: Polymerase chain reaction amplification of expanded ATTCT repeat in SCA10. *Ann Neurol* 2002, 51:271–272
- Shöls L, Szymanski S, Peters S, Przuntek H, Epplen JT, Hardt C, Riess O: Genetic background of apparently idiopathic sporadic cerebellar ataxia. *Hum Genet* 2000, 107:132–137