Meiotic stability and genotype–phenotype correlation of the trinucleotide repeat in X-linked spinal and bulbar muscular atrophy

Albert R. La Spada[1], Daniel B. Roling[1], Anita E. Harding[2], Carolyn L. Warner[3], Roland Spiegel[4], Irena Hausmanowa-Petrusewicz[5], Woon-Chee Yee[6] & Kenneth H. Fischbeck[1]

Expansion of the trinucleotide repeat $(CAG)_n$ in the first exon of the androgen receptor gene is associated with a rare motor neuron disorder, X-linked spinal and bulbar muscular atrophy. We have found that expanded $(CAG)_n$ alleles undergo alteration in length when transmitted from parent to offspring. Of 45 meioses examined, 12 (27%) demonstrated a change in CAG repeat number. Both expansions and contractions were observed, although their magnitude was small. There was a greater rate of instability in male meiosis than in female meiosis. We also found evidence for a correlation between disease severity and CAG repeat length, but other factors seem to contribute to the phenotypic variability in this disorder.

[1]*Department of Neurology, University of Pennsylvania School of Medicine, CRB 250, 422 Curie Blvd, Philadelphia, Pennsylvania 19104-6146, USA*
[2]*Institute of Neurology, The National Hospital, Queen Square, London WC1N 3BG, UK*
[3]*Dent Neurologic Institute, Millard Fillmore Hospital, Buffalo, New York 14209, USA*
[4]*Institute for Medical Genetics, University of Zurich, 8001 Zurich, Switzerland*
[5]*Neuromuscular Unit, Polish Academy of Sciences, 02-097 Warsaw, Poland*
[6]*Department of Neurology, Washington University School of Medicine, St. Louis, Missouri 63110, USA*

X-linked spinal and bulbar muscular atrophy (SBMA, Kennedy's disease) was first described as a discrete clinical disorder over 20 years ago[1]. About 50 families with SBMA have subsequently been reported. The disorder is characterized by the adult onset of proximal muscle weakness, atrophy and fasciculations. SBMA follows a progressive course complicated by involvement of the bulbar muscles. Affected patients also often have signs of androgen insensitivity such as gynecomastia, reduced fertility and testicular atrophy[2]. Although the presentation of SBMA is characteristic and readily diagnosed in isolated patients without the X-linked family history[3], the clinical severity may nevertheless be quite variable. Some patients show signs of androgen insensitivity in the second decade before the development of weakness. Patients with early involvement of the bulbar musculature may suffer from recurrent aspiration in the fifth decade of life and die from recurrent aspiration pneumonia. Other patients become wheelchair-bound by the sixth decade of life. Some less severely affected patients remain ambulatory late in life, with normal life expectancy and no signs of androgen insensitivity.

In 1986, we mapped the SBMA gene defect to Xq12-21 (ref.4). When the human androgen receptor (AR) gene was cloned and mapped to the same region[5], it became a candidate gene for this disorder[6]. In evaluating the *AR* gene in SBMA patients, we discovered enlargement of a tandem trinucleotide repeat which was absolutely associated with the SBMA phenotype[7]. The repeat encodes a long tract of glutamine residues beginning at amino acid 58 and is highly polymorphic in length in both normal individuals and in SBMA patients. On average, the length of the CAG repeat in SBMA patients was more than twice the normal length, and no overlap between patients and controls was observed. This lack of overlap and the absolute association between the CAG enlargement and the SBMA phenotype suggested that enlargement of the glutamine tract of the AR is the probable cause of this disorder.

Fragile X syndrome and myotonic dystrophy are also associated with enlargement of trinucleotide tandem repeats[8–13]. These domains are highly unstable and often expand when transmitted from parent to offspring[14,15]. In the present study, we examined the stability of the *AR* CAG tandem repeat in SBMA families and in controls. As has been noted, the course and severity of clinical disease in SBMA is quite variable. Because a range of enlarged CAG alleles is observed in SBMA patients, we reasoned that there may be a correlation between the clinical severity of the disease and the size of the CAG enlargement. We compared the clinical severity and CAG repeat number in 35 SBMA patients in order to determine whether such a correlation exists.

Meiotic stability of the CAG repeat

In our initial study of the CAG tandem repeat[7], and in other reports[16,17], a range of CAG repeat numbers was found for both SBMA patients and controls. Additional analyses confirmed that the CAG repeat is highly polymorphic: the mean CAG number obtained for controls was 22 ± 3, while the mean CAG number determined for SBMA patients is 47 ± 4. There was no overlap in CAG repeat number between normal and SBMA populations (Fig. 1). We found no exception to the association between expansion of the CAG tandem repeat and the SBMA phenotype.

To determine the stability of the CAG trinucleotide

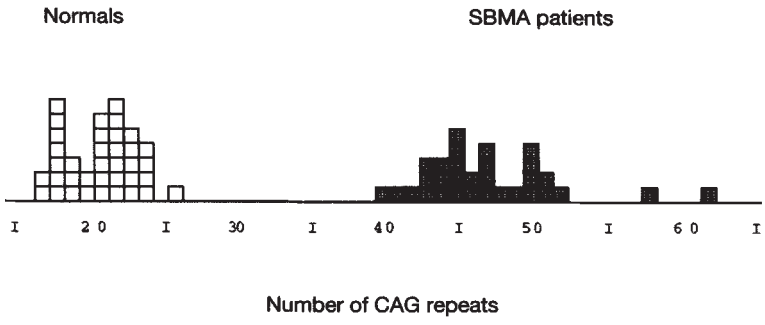

Fig. 1 $(CAG)_n$ allele distribution for 37 normal individuals and 31 unrelated SBMA patients. The CAG repeat is highly polymorphic in both patients and controls. Expansion of the CAG repeat is absolutely associated with the disease, and there is no overlap between control and SBMA samples.

repeat, we tracked the inheritance of the CAG repeat through 12 SBMA families and seven control families. Normal-sized CAG alleles were faithfully transmitted from parent to offspring without any alteration in CAG number in 62 meioses in control and SBMA families. Transmission of expanded CAG alleles, however, was associated with a significant degree of instability (Fig. 2). We found a change in the size of the CAG repeat between parent and offspring in 12 of 45 meioses (Table 1). An alteration in size of the CAG allele was observed in every SBMA family in which we were able to examine at least three meiotic events. CAG allele changes occurred in four of 31 female meioses and four of seven male meioses. Thus, the rate of instability for transmission of expanded CAG alleles was greater in male meiosis. Although our sample size is small, this observation is statistically significant ($p < .0001$).

Both expansion and contraction of the enlarged CAG repeat occurred in the 12 cases of altered transmission which we detected (Table 1). In eight allele changes where the direction of change could be determined, six were expansions and two were contractions. The magnitude of the changes in CAG repeat number ranged from one to seven, with nine of the 12 $(CAG)_n$ allele mutations involving a change in CAG repeat number of two or less. The largest expansion in CAG repeat number which we observed caused an increase of seven CAGs, and the largest contraction in our study involved a loss of four CAGs. We did not see the more dramatic increases in repeat length as found reported in myotonic dystrophy and fragile X syndrome[9,11].

Genotype–phenotype correlation

Analysis of SBMA patients, including SBMA families ascertained since our earlier publication[7], extended the range of the expanded $(CAG)_n$ allele system to include 40–62 CAGs. In order to determine whether a correlation exists between the size of the $(CAG)_n$ expansion and the severity of the clinical phenotype, we compared the ages of specific landmarks in the clinical course of the disease in unrelated SBMA patients as a function of CAG repeat number (Fig. 3). Both the age of onset and the age of stair climbing difficulty correlated inversely with CAG repeat length ($p < .025$ and $p < .05$, respectively). The age of wheelchair dependence and the presence or absence of signs of androgen insensitivity did not correlate significantly with the size of the CAG repeat ($p > 0.1$).

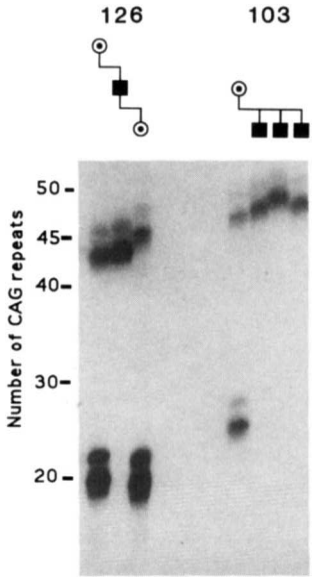

Fig. 2 Meiotic instability of the $(CAG)_n$ repeat in two SBMA pedigrees. PCR amplification of the region containing the $(CAG)_n$ repeat shows changes in the sizes of the expanded alleles when transmitted from parent to offspring. The affected male in pedigree 126 has an expanded CAG repeat number of 43 and his carrier daughter exhibits a $(CAG)_n$ allele of 45, an increase of 2 CAGs. One of the three affected male offspring (the middle son) in pedigree 103 has a $(CAG)_n$ allele of 48, and his brothers have a $(CAG)_n$ allele of 47.

Included in this analysis were SBMA pedigrees with intrafamilial variation in both CAG repeat number and clinical severity. There were two instances of variation in CAG repeat number between affected relatives where clinical information was available. In both of these cases, there were no differences in the ages of onset of weakness, of difficulty climbing stairs and of becoming wheelchair bound. Also, substantial differences in the age of onset (that is, greater than 10 years apart) and rate of progression were noted in affected brothers and affected first cousins with identical CAG repeat numbers. Clinical comparisons of three other related affecteds with identical $(CAG)_n$ alleles in their respective pedigrees revealed no significant

Table 1 $(CAG)_n$ Allele size changes observed during meiosis

	$(CAG)_n$ repeat number		
	Parental	Offspring	Difference
Male meiosis	50	49	–1
	47	52	+5
	43	45	+2
	50 or 51	50 and 51	±1
Female meiosis	62	58	–4
	47	48	+1
	50	51	+1
	43	45	+2
Sex unknown	51 or 52	51 and 52	±1
	43	50	+7
	51 or 52	51 and 52	±1
	44 or 45	44 and 45	±1

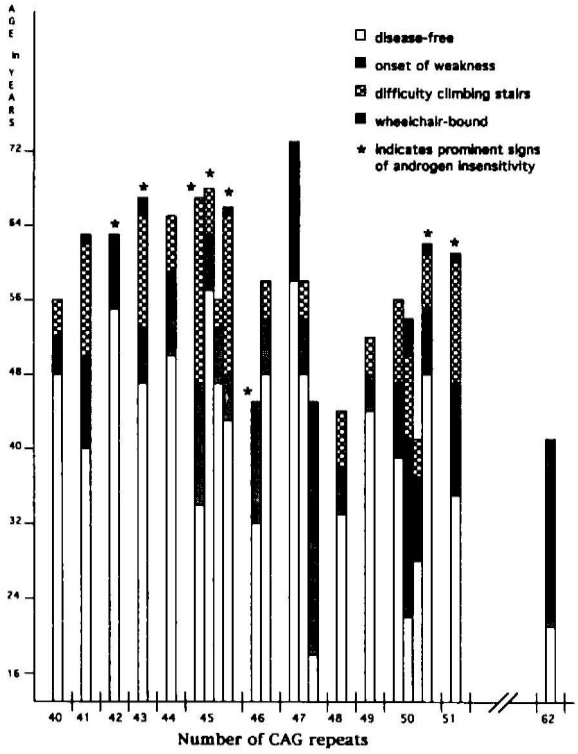

Fig. 3 Progression of clinical disease in SBMA patients plotted as a function of CAG repeat number. Each bar represents an individual SBMA patient. Ages of disease onset and of various clinical landmarks are indicated by changes in the shading of the bars. The patients' clinical courses have been placed along the horizontal axis according to their respective CAG repeat numbers. The age of disease onset and age of difficulty climbing stairs both tend to decrease with increasing CAG repeat number.

differences in the timing and pattern of disease expression. In summary, we found both intrafamilial variation in disease severity without a corresponding variation in CAG repeat length and intrafamilial variation in repeat length without a corresponding variation in disease severity. Thus, in these families, factors other than CAG repeat length are likely to contribute to the severity of the phenotype.

Discussion

The trinucleotide tandem repeat $(CAG)_n$ within the coding region of the *AR* gene is highly polymorphic (heterozygosity rate greater than 90%) both in the normal size range (13–30 CAGs) and in the expanded SBMA size range (40–62 CAGs). No phenotypically normal individuals were found with an expanded $(CAG)_n$ allele, and no normal-sized $(CAG)_n$ alleles were identified in SBMA patients. We found no individuals with an intermediate-sized $(CAG)_n$ allele (31–39 CAGs) nor have we seen any family members with markedly expanded repeats (> 62 CAGs).

We determined the germline stability of the CAG repeat by tracking the inheritance of this repeat through both normal and X-linked SBMA pedigrees. We found that the CAG repeat is stable at the normal size range, as no changes in CAG repeat number were observed in 62 meioses involving normal-sized $(CAG)_n$ alleles. Transmission of expanded $(CAG)_n$ alleles from parent to offspring was associated with significant instability, however, as 27% of the 45 meioses we examined had a change in CAG repeat length. Interestingly, a higher likelihood of alteration was detected in male than in female meiosis (57% versus 13%). Other workers have also observed moderate instability of the expanded CAG repeat and higher likelihood of alteration in male meiosis[17].

Studies of the expanded CGG repeat in the fragile X syndrome and the expanded CTG repeat in myotonic dystrophy have indicated somatic mosaicism in both these disorders[10,18]. PCR-amplified SBMA patient DNA, however, consistently yields a discrete band in the expanded size range. Although various tissue types must be analysed to exclude somatic mosaicism of the expanded CAG repeat, presence of a discrete band after PCR amplification of the expanded CAG repeat indicates that mosaicism is unlikely.

In both fragile X and myotonic dystrophy, increasing expansion of the trinucleotide repeat is correlated with worsening clinical severity of the disease[14,15,19]. We found evidence for such a correlation in the narrower range of expanded SBMA repeats. A similar result has been obtained in another SBMA patient study (Doyu *et al.*, submitted), which supports a relationship between CAG repeat length and disease severity. However, we also found related individuals who have clinical variation without a corresponding variation in repeat length, and vice versa. Therefore, other unidentified factors — genetic, environmental, or both — must be involved in variation in the clinical severity of the SBMA phenotype. Further studies are necessary to sort out the relative roles of expanded repeat length and these other factors in the production of phenotype in SBMA.

The magnitude of the changes in trinucleotide repeat length is much smaller for the expanded CAG repeat in SBMA than for either the CGG repeat in fragile X syndrome or the CTG repeat in myotonic dystrophy[17,20,21]. That we found no dramatically enlarged repeats may be a matter of sample size. Study of additional SBMA families may reveal patients, perhaps with a different phenotype, who have markedly enlarged trinucleotide repeats similar to those found in the more severely affected patients with myotonic dystrophy and fragile X syndrome.

Methodology

SBMA patient evaluation. Samples were collected from 24 SBMA pedigrees and 11 isolated SBMA patients. Referring clinicians were asked to describe presenting symptoms, ages of onset, difficulty climbing stairs, wheelchair dependence (if any), fertility status, degree of gynecomastia, and severity of bulbar and extremity muscle weakness.

DNA analysis. Genomic DNA was isolated from EDTA-anticoagulated blood by phenol/chloroform extraction. PCR amplification of the AR CAG repeat was carried out as described[7]. The sizes of the PCR products and number of CAG repeats were determined by comparison to M13 dideoxy sequencing ladders.

Genotype–phenotype correlation. Without previously reviewing the clinical information obtained for the SBMA patients, we scored each of the patients for CAG repeat number by PCR amplification, as described above. Once $(CAG)_n$ allele sizes were determined, we reviewed the clinical summaries and recorded ages of onset of weakness, of difficulty climbing stairs, and of wheelchair dependence (if any). We also noted individuals with signs of androgen insensitivity, that is, infertility and/or noticeable gynecomastia. We then plotted the clinical course as a function of CAG repeat length.

Acknowledgements

We gratefully acknowledge the cooperation of the SBMA patients who participated in this study and the following clinicians who helped with sample collection: A. Danek, C. Guizion, M. Haguenau, N.G. Laing, J. Mikol, T. Munsat, G. Nicholson, R. Olney, S. Ringel, L. Stern and L. Venier. We also thank Al Edwards. The work was supported by grants from the Muscular Dystrophy Association, the March of Dimes Birth Defects Foundation, and the National Institutes of Health.

Statistical analysis. The tests used for statistical analysis were as follows: χ^2 analysis for comparison of male and female rates of meiotic instability; Spearman's rank correlation test for correlation of clinical severity with CAG repeat length; and Student's t-test for correlation of CAG repeat length with signs of androgen insensitivity.

Received 3 August; accepted 9 September 1992.

1. Kennedy, W.R., Alter, M. & Sung, J.H. Progressive proximal spinal and bulbar muscular atrophy of late onset: A sex-linked recessive trait. Neurology **18**, 671–680 (1968).
2. Arbizu, T., Santamaria, J., Gomez, J.M., Quilez, A. & Serra, J.P. A family with adult spinal and bulbar muscular atrophy, X-linked inheritance and associated testicular failure. *J. neurol. Sci.* **59**, 371–382 (1983).
3. Harding, A.E. *et al.* X-linked recessive bulbospinal neuropathy: a report of ten cases. *J. neurol. neurosurg. Psychiatry* **45**, 1012–1019 (1982).
4. Fischbeck, K.H. *et al.* Localization of the gene for X-linked spinal muscular atrophy. *Neurology* **36**, 1595–1598 (1986).
5. Brown, C.J. *et al.* Androgen receptor locus on the X chromosome: regional localization to Xq 11-12 and description of a DNA polymorphism. *Am. J. hum. Genet.* **44**, 264–269 (1989).
6. Fischbeck, K.H., Souders, D. & La Spada, A.R. A candidate gene for X-linked spinal muscular atrophy. *Advances in Neurology* **56**, 209–214 (1991).
7. La Spada, A.R., Wilson, E.M., Lubahn, D.B., Harding, A.E. & Fischbeck, K.H. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **352**, 77–79 (1991).
8. Kremer, E.J. *et al.* Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CGG)n. *Science* **252**, 1711–1714 (1991).
9. Verkerk, A.J.M.H. *et al.* Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65**, 905–914 (1992).
10. Brook, J.D. *et al.* Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **68**, 799–808 (1992).
11. Harley, H.G. *et al.* Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature* **355**, 545–546 (1992).
12. Buxton, J. *et al.* Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature* **355**, 547–548 (1992).
13. Fu, Y.H. *et al.* An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* **255**, 1256–1258 (1992).
14. Fu, Y.H. *et al.* Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* **67**, 1047–1058 (1991).
15. Mahadevan, M. *et al.* Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* **255**, 1253–1255 (1992).
16. Edwards, A., Hammond, H.A., Jin, L., Caskey, C.T. & Chakraborty, R. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* **12**, 241–253 (1992).
17. Biancalana, V. *et al.* Moderate instability of the trinucleotide repeat in spino bulbar muscular atrophy. *Hum. molec. Genet.* **1**, 255–258, (1992).
18. Pieretti, M. *et al.* Absence of expression of the *FMR-1* gene in fragile X syndrome. *Cell* **66**, 817–822 (1991).
19. Tsilfidis, C., MacKenzie, A.E., Mettler, G., Barcelo, J. & Korneluk, R.G.. Correlation between CTG trinucleotide repeat length and frequency of severe congenital myotonic dystrophy. *Nature Genet.* **1**, 192–195 (1992).
20. Richards, R.I. & Sutherland, G.R. Heritable unstable DNA sequences. *Nature Genet.* **1**, 7–9 (1992).
21. Caskey, C.T., Pizzuti, A., Fu, Y.H., Fenwick, R.G. & Nelson, D.L. Triplet repeat mutations in human disease. *Science* **256**, 784–789 (1992).