Am. J. Hum. Genet. 59:554–560, 1996

The Relationship between Trinucleotide (GAA) Repeat Length and Clinical Features in Friedreich Ataxia

Alessandro Filla,[1] Giuseppe De Michele,[1] Francesca Cavalcanti,[1,3] Luigi Pianese,[2,3] Antonella Monticelli,[2] Giuseppe Campanella,[1] and Sergio Cocozza[2]

Departments of [1]Neurology and [2]Molecular and Cellular Biology and Pathology and CEOS, Federico II University, Naples; and [3]Mediterranean Institute of Neurosciences, Pozzilli

Summary

Friedreich ataxia (FA) is associated with the expansion of a GAA trinucleotide repeat in the first intron of the X25 gene. We found both alleles expanded in 67 FA patients from 48 Italian families. Five patients from three families were compound heterozygotes with expansion on one allele and an isoleucine→phenylalanine change at position 154 on the other one. We found neither expansions nor point mutations in three patients. The length of FA alleles ranged from 201 to 1,186 repeat units, with no overlap with the normal range, and showed a negatively skewed distribution with a peak between 800 and 1,000 repeats. The FA repeat showed meiotic instability with a median variation of 150 repeats. The lengths of both larger and smaller alleles in each patient inversely correlated with age at onset of the disorder. Smaller alleles showed the best correlation, accounting for ~50% of the variation of age at onset. Mean allele length was significantly higher in patients with diabetes and in those with cardiomyopathy.

Introduction

Friedreich ataxia (FA) is the most frequent early-onset autosomal recessive inherited ataxia. Birth incidence has been estimated to be 4.9×10^{-5}, in a personal study in southern Italy (Filla et al. 1992). Estimated carrier frequency ranges from 1/110 to 1/70 (Harding 1981; Filla et al. 1992). FA is characterized by progressive ataxia with onset within 20 years of age, absence of lower-limb tendon reflexes, dysarthria, Babinski sign, limb weakness, decreased vibration sense, skeletal deformities, cardiomyopathy, and diabetes (Geoffroy et al. 1976). No treatment is available to delay the progressive course of the disease, and patients become wheelchair-bound, on average, after 15.5 years of disease (Harding 1981). Clinical variability is higher than expected in an autosomal recessive disorder, in particular concerning age at onset and presence of cardiomyopathy and diabetes. Variation of onset age has usually been explained by genetic and environmental factors. Data of intrafamily variation of onset age suggested a significant role for genetic factors (Winter et al. 1981; De Michele et al. 1989).

Chamberlain et al. (1988) mapped the FA locus (FRDA) to chromosome 9 by linkage to the anonymous marker D9S15. Further studies showed no evidence of genetic heterogeneity (Chamberlain et al. 1989; Fujita et al. 1989) and restricted the candidate region to 150 kb (Montermini et al. 1995). We previously reported 21 patients with onset between 21 and 36 years of age (late-onset FA [LOFA]) and showed that age at onset should no longer be considered a diagnostic criterion, since LOFA is genetically homogeneous to FA (De Michele et al. 1994). Furthermore, the rare FA families in which patients had preserved tendon reflexes (FA with retained reflexes [FARR]) have also been mapped to the FRDA region (Palau et al. 1995). LOFA and FARR might represent allelic mutations, but the coexistence in the same families of patients with typical and late onset and that of patients without and with tendon reflexes suggested one mutation.

Within a cooperative study, we recently isolated the defective gene (X25) causing FA (Campuzano et al. 1996). X25 encodes a 210–amino acid protein with unknown function called "frataxin." We also demonstrated reduction of frataxin mRNA levels in FA patients. An expanded unstable trinucleotide (GAA) repeat is present in the X25 first intron on FA chromosomes. Normal chromosomes contained 7–22 units, whereas FA chromosomes carried 200 to above 900 GAA repeats. In addition, we described three rare point mutations. In particular, an amino acid change (isoleucine to phenylalanine at position 154, I154F) was found in X25 exon 4 in southern Italian patients.

The discovery of the FA molecular defect may have

Received April 19, 1996; accepted for publication June 10, 1996.

Address for correspondence and reprints: Dr. Giuseppe De Michele, Clinica Neurologica, Università Federico II, via Pansini 5, 80131, Naples, Italy. E-mail: cocozza@ds.cised.unina.it

0002-9297/96/5903-0010$02.00

implications for the diagnosis of the disease, both in symptomatic and asymptomatic persons, and for explanation of the phenotypic variability of the disease. We looked for the expanded GAA sequence and for point mutations in a large series of personally examined FA patients. In addition, we investigated whether the same mutation is responsible for typical FA, LOFA, and FARR phenotype. To examine the possibility that genetic instability accounts for clinical variability, we studied the relationship between molecular findings and clinical features, as ages at onset and at becoming wheelchair bound and presence of cardiomyopathy and diabetes.

Subjects and Methods

The diagnosis of FA was based on the following criteria: autosomal recessive inheritance or sporadic occurrence, onset within 20 years of age, progressive unremitting ataxia of limbs and gait, and absence of knee and ankle jerks. Furthermore, at least one of the following signs was present in the index cases: dysarthria, extensor plantar response, or echocardiographic evidence of cardiomyopathy (Filla et al. 1990).

DNA was extracted from blood leukocytes. The molecular analysis was performed by PCR and separation on agarose gel. The size of alleles was estimated by least-square fit of fragment size to gel mobility, using the computer program DNAFRAG (version 3.03). To analyze the GAA repeat in the first intron of the X25 gene, we used the following primers: GAA-104F: 5′-GGCTTAAACTTCCCACACGTGTT-3′; and GAA-629R: 5′-AGGACCATCATGGCCACACTT-3′. These primers flank the GAA repeat and generate a PCR product of 500 + 3*n* bp (*n* = number of GAA triplets). We demonstrated that enlarged alleles are generated by GAA expansion (Campuzano et al. 1996). On the other hand, we cannot exclude that size differences among normal alleles can be also generated by other polymorphic variations of the region. Amplifications were conducted using the long PCR protocol (Boehringer Mannheim Long Expand), in 10 cycles composed of the following steps: 94°C for 10 s, 60°C for 30 s, 68°C for 3 min, followed by further 20 cycles in which the length of the 68°C step was increased by 20 s/cycle. PCR products were separated on 1% agarose gel. The estimated error in size determination was ±30 triplets for expanded alleles and ±3 triplets for normal alleles.

Results

GAA Expansion and Point Mutations in the X25 Gene

We analyzed 75 patients from 54 Italian families. Forty-five families were from southern Italy, and 40 of them were from Campania, the region whose main town is Naples. Sixty patients had typical FA, 12 had onset after 20 years of age (LOFA), and 2 had preserved tendon reflexes (FARR). One patient had both late onset (35 years) and preserved tendon jerks.

Sixty-seven patients from 48 families were homozygous for the expanded GAA sequence. Five patients were compound heterozygotes with expansion on one allele and I154F mutation on the other one. They were from three families, two from Campania and one from Rome. We found no expanded alleles in three patients (212, 288, and 298) from three families. We excluded the presence of I154F mutation in these patients. In addition, SSCP analysis on the X25 exons showed no alteration. Review of clinical findings of these patients confirmed that they fulfilled the diagnostic criteria for FA, with onset at 3, 29, and 11 years, progressive ataxia, absence of lower limb reflexes, and dysarthria. Extensor plantar response, diabetes, and electrocardiographic or echographic signs of cardiomyopathy were absent in all. Some atypical findings were also present: head tremor in case 212 and marked distal muscle atrophy in case 298. Vitamin E level was normal in patient 212.

Excluding patients 212, 288, and 298, among 102 FRDA chromosomes from 51 unrelated patients, 99 (97%) carried an expanded allele and the remaining 3 (3%) the I154F mutation. No expansion was detected on 124 normal chromosomes from healthy individuals with no relationship to FA patients. I154F mutation was found in 1 of 470 normal chromosomes.

The Size of GAA Repeats on FA Chromosomes

The expanded repeat length of FA chromosomes was estimated in the 67 patients homozygous for the expansion. The median size of the expanded alleles was 839 repeats with range of 201–1,186 (mean ± SD = 816 ± 188). The distribution of the number of repeats on the mutated alleles, shown in figure 1, was significantly different from normal ($\chi^2 = 30.5$; df = 7; $P < .01$) with negative skewness (coefficient of skewness = −.87; $P < .01$). The peak of the GAA expansion was between 800 and 1,000 repeats. We detected two classes of alleles on 124 normal chromosomes. The estimated number of GAA repeats was 10 in 87% of alleles and 20 in 13%. No allele of intermediate size (21–200) was detected.

Meiotic Instability of the GAA Expansion

Fourteen of our patients were from nine consanguineous marriages (in seven marriages, parents were first cousins; in two, first cousins once removed). They were expected to be homozygous by descent at the FRDA locus. The median difference between alleles was 184 repeats (0–389), and only three patients had expansion of the same length on both alleles.

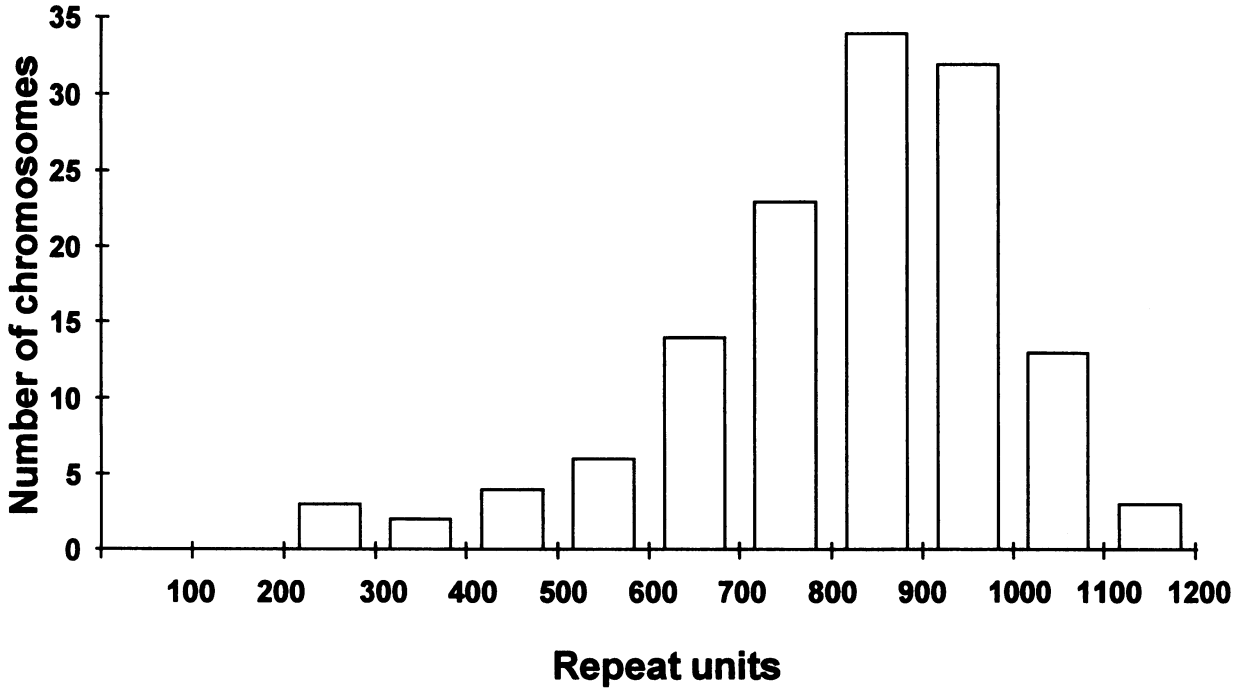

Figure 1 Frequency distribution of (GAA)n repeat unit number on 134 FA chromosomes from 48 independent families

Since we could not identify maternally and paternally derived alleles in the patients, we chose to study variation of transmitted alleles in parent-carrier child pairs. On the basis of previous linkage studies, we identified five pairs from different pedigrees in which parent and heterozygous child carried the same FA chromosome. We found GAA repeat length variation in each pair. The alleles contracted in three meioses and expanded in two. Median variation was of 150 repeats (range 76–208).

Age at Onset and GAA Repeat Size

Genotype-phenotype relationship was studied in the 67 patients homozygous for the expansion. We conventionally named GAA1 the smaller allele and GAA2 the larger one in each patient. Median size was 769 units (range 201–1,093) for GAA1 and 921 (549–1,186) for GAA2.

Age at onset was available in 64 individuals. Mean age at onset was 14.7 ± 7.5 years (range 2–36). Table 1A shows Pearson's correlation coefficients between age at onset and repeat number of GAA1, GAA2, and GAA1-GAA2 mean. We found a significant inverse correlation between age at onset and the number of repeats in each case. The best correlation was obtained for GAA1 (fig. 2). Nearly 50% (r^2 = .47) of the variation in the age at onset can be explained by the number of GAA repeats. To confirm this finding, we divided patients into three groups according to age at onset (group 1, ≤10 years; group 2, 11–20 years; group 3, ≥21 years). Mean ages ± SD were 7.0 ± 2.5 years (n = 17) in group 1, 14.0 ± 2.2 years (n = 35) in group 2, and 27.8 ± 4.9 years (n = 12) in group 3. Trinucleotide repeat length significantly differed between groups at analysis of variance (table 1B). The highest F-value was obtained for GAA1. At least significant difference test, group 3 (LOFA patients) differed from both group 1 and group 2 for GAA1 size ($P < .01$). The statistical difference between groups 1 and 2 was less significant ($P < .05$).

Diabetes Mellitus, Hypertrophic Cardiomyopathy, and GAA Repeat Size

Presence of diabetes mellitus was assessed in 63 patients. Diagnosis of diabetes mellitus was made when fasting venous plasma glucose was found to be ≥140 mg/dl on at least two separate occasions. Mean allele sizes in diabetic FA patients were greater than in normoglycemic patients. Table 1B shows analysis of variance for allele sizes in 8 diabetic and 55 normoglycemic patients. Figure 3 shows the distribution of patients according GAA1 size and presence of diabetes.

Cardiomyopathy was diagnosed by B-mode echocardiography in 25 of 56 patients. Left ventricular hypertrophy, concentric or asymmetric, was the most frequent diagnosis (Pentland and Fox 1983). Patients with cardiomyopathy carried larger alleles, in comparison with

Table 1

Statistical Analysis

Characteristic	GAA1	GAA2	GAA1-GAA2 Mean
	A. Pearson's Correlation Coefficient		
Onset age (df = 62)	−.69***	−.36**	−.64***
Age at becoming wheelchair bound (df = 31)	−.48**	−.22	−.45**
Disease duration to wheelchair (df = 31)	−.16	−.31	−.27
	B. Analysis of Variance		
Age at onset:			
≤10 years (n = 17)	856 ± 133	967 ± 117	911 ± 110
11–20 years (n = 35)	755 ± 115	926 ± 120	841 ± 101
>20 years (n = 12)	462 ± 197	804 ± 145	633 ± 107
F-ratio (df = 2.61)	29.9***	6.4**	26.2***
Diabetes:			
Present (n = 8)	956 ± 82	981 ± 74	968 ± 70
Absent (n = 55)	698 ± 180	911 ± 132	805 ± 132
F-ratio (df = 1.61)	15.7***	2.1	11.8**
Cardiomyopathy:			
Present (n = 25)	801 ± 124	969 ± 108	885 ± 101
Absent (n = 31)	690 ± 221	883 ± 140	787 ± 150
F-ratio (df = 1.54)	5.0*	6.4*	7.8**

NOTE.—GAA1 and GAA2 represent GAA number on the smaller and on the larger allele in each patient, respectively. Values are means ± SD.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

those with normal echocardiographic findings. The results of the statistical analysis are shown in table 1B. LOFA patients had neither cardiomyopathy nor diabetes.

Disease Progression and GAA Repeat Size

Thirty-eight patients were wheelchair bound. Age at becoming wheelchair bound was personally observed or reliably referred in 33 of them. Mean value ± SD was 24.8 ± 7.7 years after a mean disease duration of 13.1 ± 4.8 years. Median GAA1 length was 807 repeats (range 203–1,093). Age at becoming wheelchair bound showed an inverse correlation with sizes of GAA1 and GAA1-GAA2 mean (table 1A). No correlation was found between disease progression (onset of symptoms→dependence on wheelchair) and expansion size.

Phenotype of Patients Compound Heterozygotes for GAA Expansion and I154F Point Mutation

Five patients were compound heterozygotes for GAA expansion and I154F point mutation. Mean size of the expansion was 846 ± 137 triplets (median 921; range 625–959). Mean disease onset was at 9.6 ± 5.2 years (range 2–15 years). This value was lower than that found in patients homozygous for the expansion, although this difference was not statistically significant (t = 1.94; $P < .1$). Dysarthria and Babinski sign were present in four, cardiomyopathy in three, and diabetes in none. The clinical picture of these patients was indistinguishable from that of patients homozygous for the expansion.

Discussion

Trinucleotide repeat expansions are now a well-recognized mutational mechanism in human genetics. FA is the first autosomal recessive disease that is due to an expanded trinucleotide repeat. The repeat is located in a noncoding region of the X25 gene, whose function is unknown (Campuzano et al. 1996).

We performed a molecular analysis on 75 FA patients originating mainly from southern Italy. We found an expanded GAA repeat on both alleles in all patients but eight. Five patients from three independent families were compound heterozygotes for the expansion and for the I154F point mutation (Campuzano et al. 1996). Three

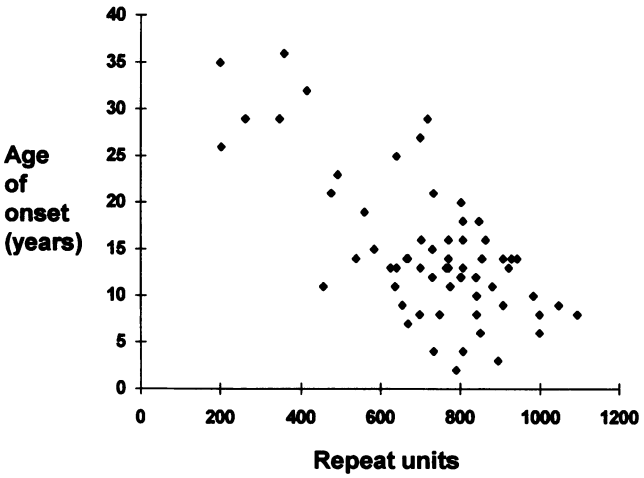

Figure 2 Correlation between age at onset and size of the expansion on the smaller allele (GAA1) in 64 patients homozygous for the expansion.

further patients had both alleles in the normal range and no apparent genetic alteration in the coding sequence. The hypothesis that they carry not-yet-recognized mutations on both alleles cannot be excluded, but it appears unlikely. The presence of slight atypical findings in two of them suggests that these patients might have a different disease mimicking FA.

Our data suggest that expanded GAA repeat analysis has a high diagnostic value in FA. The presence of an expanded repeat on both alleles appears to be highly specific for the disease, since no overlap has been seen between normal and affected individuals up to now. The sensitivity of the analysis of expansion needs to be assessed, since the complete spectrum of the genetic alterations causing FA has not yet been elucidated. Patients with FA phenotype and expansion on one allele only or with no expansion should be investigated for point mutations. In our opinion, clinical diagnosis of FA should be confirmed by molecular studies, at least in atypical cases.

Other applications from this study are carrier and prenatal diagnosis in at-risk families. Prenatal diagnosis has already been performed in FA (Wallis et al. 1989). A direct molecular test allows an accurate assessment of the risk without the need to obtain samples from many family members. Genetic heterogeneity hampered carrier detection in other autosomal recessive disorders. Testing partners may be cumbersome, expensive, and often not conclusive, because of the large number of mutations. The overall genetic homogeneity in FA (97%) could allow carrier diagnosis in individuals who do not belong to at-risk families. However, these issues require more knowledge about the frequency of point mutations of X25, cost-benefit analysis, and careful evaluation of ethical implications.

In this study, we estimated the GAA repeat size of 134 expanded alleles. Their distribution was not normal, and it was negatively skewed, the tail of the smallest alleles corresponding to LOFA patients. We found no allele between 21 and 200, but we cannot exclude that further studies will fill this gap. This study gave also evidence of meiotic variability of GAA repeat size. We found variation (contraction or expansion) of repeat size in all five parent-offspring pairs studied. These data, together with those from consanguineous families, confirmed the presence of meiotic instability in the GAA expansion of X25.

The present study shows a clear relationship between expansion size and phenotype variability. We have shown an inverse correlation between the number of GAA repeats and age at onset, as already demonstrated for other triplet diseases (Bates and Lehrach 1994). However, a broad range of onset ages may be associated with a given repeat length. For instance, in the GAA1 800–900 repeat range, which contains 25% of patients, onset age varies from 3 to 20 years.

LOFA patients represent 17% of cases in our series. LOFA patients have a more benign course and cardiomyopathy is rare. In a previous linkage study, we showed that LOFA mapped to the FRDA region (De Michele et al. 1994). In the present study, we demonstrated the expansion of both alleles in all LOFA patients but one (case 288). The number of GAA1 repeats in LOFA ranged from 201 to 734, and mean repeat length markedly differed from that of patients with onset be-

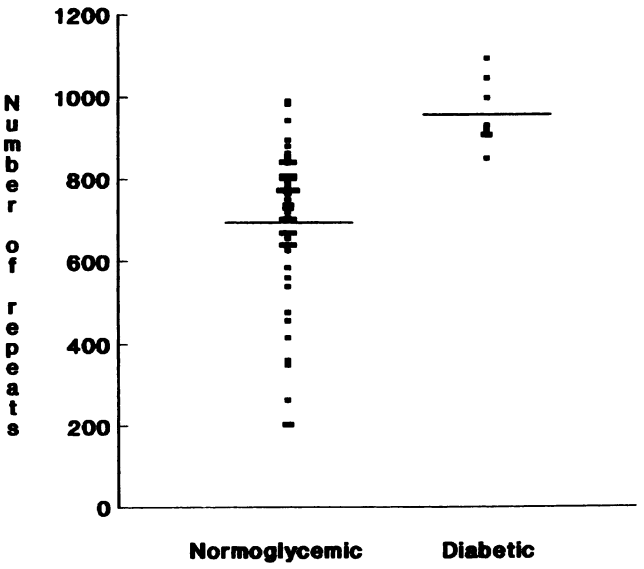

Figure 3 Distribution of 63 FA patients according to expansion size on the smaller allele (GAA1) and diabetes.

fore 20 years of age (table 1B). Our data confirm the genetic homogeneity between FA and LOFA and explain the phenotype differences on the basis of differences in expansion size.

Diabetes mellitus is present in 14% of FA patients (Filla et al. 1990), and it develops in late stages of the disease, after a mean disease duration of 15 years (Harding 1981). We found that mean expansion size is higher in patients with diabetes. No patient with GAA1 size <850 repeats had diabetes. Allele size, together with disease duration, seems to be the major determinant of diabetes.

Occurrence of cardiomyopathy depends on the tool of ascertainment. It appears to be constant at pathological examination, very frequent at ECG, and less common when echocardiography is used (Filla et al. 1990). Assessing cardiomyopathy by echocardiography we found significantly larger expansions in cardiopathic patients, even though there was a large overlap between patients with and without cardiomyopathy.

GAA1 size was higher in patients confined to wheelchairs, and it inversely correlated with age at becoming wheelchair bound. However, this age might not be a reliable measure of progression rate of the disease, since it depends on the age at onset. Time from onset to wheelchair, which is independent from onset age, did not correlate significantly with the expansion length. This finding requires further analysis, since a selection bias might be present. In fact, patients with the fastest progression are the most likely to be already wheelchair bound and to be included in the analysis.

Finally, we found no genetic peculiarity in the patients with the FARR phenotype and no characteristic phenotype in the I154F patients. This amino acid change is located within a highly conserved domain and may affect the protein function severely. We suggest that I154F mutation had the same effect of an expanded allele.

GAA1 size correlated with clinical features better than GAA2 or GAA1-GAA2 mean (table 1). The reduction of mRNA levels in FA patients and the presence of rare disruptive causing disease mutations suggest that FA may be classified as a "loss-of-function" disease (Campuzano et al. 1996) For a recessive disorder, one could hypothesize that severity of the disease would be a function of how much normal transcript could be made. Therefore, it is not unexpected that disease severity strongly correlated with the GAA1 size, which could represent the "remaining activity." This hypothesis should be confirmed by studies about the relationship between allele size and mRNA levels.

Although our data demonstrate a relationship among GAA repeat length and clinical variability, the size of the expanded GAA sequence determined in leukocyte DNA is only a partial determinant of the clinical features. X25 gene showed highest expression in heart, intermediate in liver, skeletal muscles, and pancreas, and minimal in other tissues, including whole brain. Within the CNS, highest expression was in the spinal cord (Campuzano et al. 1996). Somatic mosaicism has been demonstrated in several triplet diseases (Bates and Lehrach 1994). Tissue mosaicism caused by mitotic instability might be a possible further mechanism that could explain clinical variability among FA patients. Studies on X25 transcription and translation in the CNS, the heart, and the pancreas will elucidate this issue.

Acknowledgments

This work was supported by Italian Telethon grant 722 (to S.C.) and by grants from the Italian Ministries of Health and Education and from CNR (91.04180).

References

Bates G, Lehrach H (1994) Trinucleotide repeat expansions and human genetic disease. Bioassays 16:277–284

Campuzano V, Montermini L, Moltò MD, Pianese L, Cossèe M, Cavalcanti F, Monròs E, et al (1996) Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. Science 271:1423–1427

Chamberlain S, Shaw J, Rowland A, Wallis J, South S, Nakamura Y, von Gabain A, et al (1988) Mapping of the mutation causing Friedreich's ataxia to human chromosome 9. Nature 334:248–250

Chamberlain S, Shaw J, Wallis J, Rowland A, Chow L, Farrall M, Keats B, et al (1989) Genetic homogeneity at the Friedreich ataxia locus on chromosome 9. Am J Hum Genet 44: 518–521

De Michele G, Filla A, Barbieri F, Perretti A, Santoro L, Trombetta L, Santorelli F, et al (1989) Late onset recessive ataxia with Friedreich's disease phenotype. J Neurol Neurosurg Psychiatry 52:1398–401

De Michele G, Filla A, Cavalcanti F, Di Maio L, Pianese L, Castaldo I, Calabrese O, et al (1994) Late onset Friedreich's disease: clinical features and mapping of mutation to FRDA locus. J Neurol Neurosurg Psychiatry 57:977–979

Filla A, De Michele G, Caruso G, Marconi R, Campanella G (1990) Genetic data and natural history of Friedreich's disease: a study of 80 Italian patients. J Neurol 237:345–351

Filla A, De Michele G, Marconi R, Bucci L, Carillo C, Castellano AE, Iorio L, et al (1992) Prevalence of hereditary ataxias and spastic paraplegias in Molise, a region of Italy. J Neurol 239:351–353

Fujita R, Agyd Y, Trouillas P, Seck A, Tommasi-Davenas C, Driesel A, Olek K, et al (1989) Confirmation of linkage of Friedreich ataxia to chromosome 9 and identification of a new closely linked marker. Genomics 4:110–111

Geoffroy G, Barbeau A, Breton G, Lemieux B, Aube M, Leger C, Bouchard JB (1976) Clinical description and roentgen-

ologic evaluation of patients with Friedreich's ataxia. Can J Neurol Sci 3:279–286

Harding AE (1981) Friedreich's ataxia: a clinical and genetic study of 90 families with an analysis of early diagnostic criteria and intrafamilial clustering of clinical features. Brain 104:589–620

Montermini L, Rodius F, Pianese L, Moltò MD, Cossèe M, Campuzano V, Cavalcanti F, et al (1995) The Friedreich ataxia critical region spans a 150-kb interval on chromosome 9q13. Am J Hum Genet 57:1061–1067

Palau F, De Michele G, Vilchez J, Pandolfo M, Monròs E, Cocozza S, Smeyers P, et al (1995) Early onset ataxia with cardiomyopathy and retained tendon reflexes maps to the Friedreich's ataxia locus on chromosome 9q. Ann Neurol 37:359–362

Pentland B, Fox K (1983) The heart in Friedreich's ataxia. J Neurol Neurosurg Psychiatry 46:1138–1142

Wallis J, Shaw J, Wilkes D, Farrall M, Williamson R, Chamberlain S, Skare JC, et al (1989) Prenatal diagnosis of Friedreich's ataxia. Am J Med Genet 34:458–461

Winter RM, Harding AE, Baraitser M, Bravery MB (1981) Intrafamilial correlation in Friedreich's ataxia. Clin Genet 20:419–427