Am. J. Hum. Genet. 55:244–252, 1994

Molecular and Clinical Correlations in Spinocerebellar Ataxia Type 1: Evidence for Familial Effects on the Age at Onset

Laura P. W. Ranum,[1] Ming-yi Chung,[1] Sandro Banfi,[3] Alan Bryer,[4] Lawrence J. Schut,[2] Raj Ramesar,[5] Lisa A. Duvick,[1] Alanna McCall,[3] S. H. Subramony,[6] Lev Goldfarb,[7] Christopher Gomez,[2] Lodewijk A. Sandkuijl,[8] Harry T. Orr,[1] and Huda Y. Zoghbi[3]

[1]Department of Laboratory Medicine and Pathology and Institute of Human Genetics and [2]Department of Neurology, University of Minnesota, Minneapolis; [3]Departments of Pediatrics and Molecular and Human Genetics, Baylor College of Medicine, Houston; [4]Department of Neurology, Groote Schuur Hospital, and [5]MRC Research Unit for Medical Genetics, Department of Human Genetics, University of Cape Town, Observatory, South Africa; [6]University of Mississippi Medical Center, Jackson; [7]National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda; and [8]Institute of Clinical Genetics, Erasmus University, Rotterdam

Summary

The spinocerebellar ataxias are a group of debilitating neurodegenerative diseases for which a clinical classification system has proved unreliable. We have recently isolated the gene for spinocerebellar ataxia type 1 (SCA1) and have shown that the disease is caused by an expanded, unstable, CAG trinucleotide repeat within an expressed gene. Normal alleles have a size range of 19–36 repeats, while SCA1 alleles have 42-81 repeats. In this study, we examined the frequency and variability of the SCA1 repeat expansion in 87 kindreds with diverse ethnic backgrounds and dominantly inherited ataxia. All nine families for which linkage to the SCA1 region of 6p had previously been established showed repeat expansion, while 3 of the remaining 78 showed a similar abnormality. For 113 patients from the families with repeat expansion, inverse correlations between CAG repeat size and both age at onset and disease duration were observed. Repeat size accounted for 66% of the variation in age at onset in these patients. After correction for repeat size, interfamilial differences in age at onset remained significant, suggesting that additional genetic factors affect the expression of the SCA1 gene product.

Introduction

The autosomal dominant spinocerebellar ataxias (SCAs) are a clinically heterogeneous group of neurodegenerative disorders characterized by gait and limb ataxia, dysarthria, and variable degrees of brain-stem dysfunction. The neuropathological findings in SCA include selective loss of neurons in the cerebellum and brain stem, as well as degeneration of spinocerebellar tracts (Greenfield 1954). Inter- and intrafamilial variation in the age at onset (Schut 1950; Zoghbi et al. 1988; Orozco Diaz et al. 1990), severity of the disease, and associated findings such as retinal degeneration, optic atrophy, spasticity, dementia, and extrapyramidal signs have made the clinical classification of this group of disorders difficult and unreliable. Almost 20 years ago Yakura et al. (1974) suggested that SCA maps to the short arm of chromosome 6, on the basis of a study of a small kindred. Jackson et al. (1977) confirmed that one type of SCA maps to 6p, based on linkage to the human leukocyte antigen (HLA) complex in a large kindred. The 6p-linked form was subsequently designated "type 1 spinocerebellar ataxia" (SCA1) (McKusick 1978). Several other families were also found to have SCA1, on the basis of linkage to HLA (Nino et al. 1980; Haines et al. 1984; Zoghbi et al. 1988) or to DNA markers that were subsequently found to be closer to the SCA1 locus than to the HLA loci (Ranum et al. 1991; Zoghbi et al. 1991). Genetic studies on a number of autosomal dominant ataxia and Machado-Joseph disease (MJD) kindreds (Sakai et al. 1983; Kumar et al. 1986; Auburger et al. 1990; Carson et al. 1992; Ranum et al. 1992) excluded the SCA1 region of 6p and demonstrated genetic heterogeneity among the dominant ataxias. Recently, genetic studies allowed the mapping of a second ataxia locus (SCA2) to chromosome 12q (Gispert et al. 1993) and of the MJD locus to chromosome 14q (Takiyama et al. 1993). These linkage studies provide a framework for the partial classification of the dominantly inherited ataxias and for the identification of the genes responsible for these disorders.

The gene for SCA1 is the first ataxia gene to be isolated by positional cloning (Orr et al. 1993). Genetic and physical mapping studies allowed the precise localization of this gene on 6p (Ranum et al. 1991; Zoghbi et al. 1991; Banfi et al. 1993; Kwiatkowski et al. 1993). Anticipation, an increase in the severity of disease in later generations, observed in SCA1 (Schut 1950; Zoghbi et al. 1988) led to the hypothesis that an expanded and unstable trinucleotide re-

Received February 2, 1994; accepted for publication March 30, 1994.

Address for correspondence and reprints: Dr. Laura P. W. Ranum, Institute of Human Genetics, University of Minnesota Medical School, 420 Delaware Street S.E., Box 206, UMHC, Minneapolis, MN 55455-0392.

0002-9297/94/5502-0004$02.00

peat could be the mutational mechanism for SCA1, as it was for other neurologic disorders displaying anticipation—e.g., fragile X syndrome, myotonic dystrophy, and Huntington disease (Kremer et al. 1991; Verkerk et al. 1991; Brook et al. 1992; Buxton et al. 1992; Fu et al. 1992; Harley et al. 1992; Mahadevan et al. 1992; The Huntington's Disease Collaborative Research Group 1993). Based on this hypothesis, a systematic search for trinucleotide repeats in the 1.2-Mb SCA1 candidate region was carried out. This search resulted in the identification of a highly polymorphic CAG repeat, which was found to be unstable and expanded in individuals with SCA1 (Orr et al. 1993). In a study of 27 SCA1 individuals representing three SCA1 kindreds, a direct correlation between the size of the CAG repeat and the age at onset of disease was found, with larger repeats occurring in juvenile cases (Orr et al. 1993).

The identification of an expanded trinucleotide repeat in SCA1 provides a means for rapid genotypic analysis of families with dominantly inherited ataxia, to identify families with SCA1. In this study we report the genotypic data for the CAG repeat at the SCA1 locus for 87 kindreds with diverse ethnic backgrounds and dominantly inherited ataxia. Furthermore, we characterize the correlation between the size of the CAG repeat and both the age at onset and the severity of disease in a cohort of nine SCA1 kindreds.

Families and Methods

Families

Over the past 8 years we have had the opportunity to evaluate members representing 87 kindreds with dominantly inherited ataxia. Nine kindreds of diverse ethnic background were already known to have SCA1, on the basis of linkage to the HLA loci or to D6S89. Genotypic analysis of the SCA1 CAG repeat was carried out on all nine kindreds to determine whether all known SCA1 families had the same mutational mechanism involving repeat expansion. Most of the study participants were personally examined by one of us. The affected status was always confirmed by a neurologist, but the age at onset was based on historical information from the patient and/or other family members. Severity of disease was measured by the age at death minus the age at onset. Detailed characterization of the repeat variability and its correlation with both age at onset and severity of disease was carried out for all nine kindreds. To identify additional kindreds with a CAG expansion at the SCA1 locus, affected individuals from 78 newly identified families with dominantly inherited ataxia were clinically examined. Blood was collected from at least one affected individual from each of these kindreds and was screened by DNA analysis for the presence of a CAG repeat size within the expanded range (≥42 repeats). Although we have no evidence that these 78 individuals are related, there is a chance that some of the affected patients come from the same families. To assess further the distribution of CAG repeat sizes on normal chromosomes, the number of CAG repeats was determined for 304 normal chromosomes from unrelated individuals of various ethnic backgrounds.

Molecular Studies

Blood samples were used to establish lymphoblastoid cell lines by Epstein-Barr virus transformation. Genomic DNA was isolated either directly from venous blood or from lymphoblastoid cell lines. Blood samples were collected from these patients over the past 8 years, during which time 29 patients died. PCR reactions were performed using the Rep1 (TTGACCTTTACACCTGCAT) and Rep2 (CAACATGGGCAGTCTGAG) primers (Orr et al. 1993). Fifty nanograms of genomic DNA was mixed with 5 pmol of each primer in a total volume of 20 μl containing 1.25 mM $MgCl_2$, 250 μM dNTPs, 50 mM KCl, 2% formamide, 10 mM Tris-HCl pH 8.3 and 1 unit Amplitaq (Perkin Elmer Cetus). The Rep1 primer was labeled at the 5′ end with [gamma-^{32}P] ATP. Samples were denatured at 94°C for 4 min, followed by 30 cycles of denaturation (94°C, 1 min), annealing (55°C, 1 min), and extension (72°C, 2 min). Six microliters of each PCR reaction was mixed with 4 μl formamide loading buffer and was denatured at 90°C for 2 min and electrophoresed through a 6% polyacrylamide/7.65 M urea DNA-sequencing gel. Allele sizes were determined by comparing migration relative to an M13 sequencing ladder.

Statistical Analyses

The relationship between age at onset and CAG repeat number on both the affected and the normal chromosomes of patients was evaluated through linear-regression analyses. Similarly, the relationship between repeat length and duration of disease was quantified. Ages at onset were used directly in these analyses, as well as after logarithmic and square-root transformation. Although the latter transformation provided the best approximation to a normal distribution, results obtained were consistent between analyses before and after transformation. Analysis of variance was performed to detect differences among the families in the mean age at onset, after correction for the effect of the CAG repeat number on age at onset. In addition, the sex of the transmitting parent was included as a possible explanatory variable for variations in age at onset. All regression and variance analyses were carried out with the SPSS package of computer programs, versions 4.0.1.

Results

Family Studies

We have gathered clinical data and DNA from individuals representing 87 different kindreds with dominantly inherited ataxia. Nine of these kindreds were previously

known to have SCA1, on the basis of genetic linkage to HLA and/or D6S89 on chromosome 6p (Jackson et al. 1977; Nino et al. 1980; Haines et al. 1984; Zoghbi et al. 1988; Goldfarb et al. 1989; Bryer et al. 1992). These nine kindreds have diverse ethnic backgrounds: three are Caucasian American (Dutch, German, and Anglo-Saxon) (Currier et al. 1972; Nino et al. 1980; Haines et al. 1984); two are African American (Zoghbi et al. 1988; Keats et al. 1991); two are from South Africa and are of mixed ethnic origin (Bryer et al. 1992); and two are from Siberia and are of the native Iakut people (Goldfarb et al. 1989). For the remaining 78 dominant-ataxia kindreds, at least one affected individual was genotyped for the CAG repeat, to determine whether any contained a CAG repeat expansion in the SCA1 gene.

All affected individuals from the nine known SCA1 kindreds had an expanded trinucleotide repeat on one of their alleles. No repeat expansions were observed among eight kindreds previously shown by linkage analyses not to be SCA1. These eight kindreds were examined for the SCA1 gene expansion to confirm the linkage results.

Among the 70 other dominant-ataxia families analyzed, three (4%) were found to have an expanded CAG repeat on one of their alleles. Although we cannot exclude the possibility that some of these 70 families may be related, of all of the dominant kindreds that we have studied, 12 (14%) of 87 have an expanded CAG repeat at the SCA1 locus. While the sample size reported here is relatively small, and although both estimates are arguably biased to exclude or select for SCA1 kindreds, expanded CAG repeat tracts within the SCA1 gene clearly account for only a small fraction of this complex group of diseases. The distribution of the CAG repeat number from normal controls and from ataxic individuals who did not have an expansion were similar (data not shown). These data argue against the involvement of the CAG repeat at the SCA1 locus in these families. However, it is still possible that some of these small families have other mutations at the SCA1 locus.

The typical clinical findings in the genetically proved SCA1 kindreds were gait and limb ataxia, dysarthria, pyramidal tract signs (spasticity, hyperreflexia, and extensor plantar responses), and variable degrees of oculomotor findings, which include one or more of the following: nystagmus, slow saccades, and ophthalmoparesis. In the later stages of the disease course, bulbar findings consistent with dysfunction of cranial nerves IX, X, and XII became evident. Also, dystonic posturing and involuntary movements including choreoathetosis became apparent in the later stages of the disease. Motor weakness, amyotrophy, and mild sensory deficits manifested as proprioceptive loss were also detected. The Appendix lists the clinical features for the nine large SCA1 kindreds studied. Although ataxia, dysarthria, and cranial nerve dysfunction were consistently present in every SCA1-affected individual, considerable intrafamilial variability was noted with regard to all of the other clinical features. Juvenile onset (at age ≤18 years) was observed in four kindreds (TX, MN, LA, and S.Af.-a). Of interest is the finding that juvenile-onset cases typically inherited the disease gene from an affected father (Chung et al. 1993). Several of the kindreds that did not have an expanded SCA1 CAG repeat displayed the same clinical findings as were observed in SCA1 kindreds, confirming the inherent difficulty in clinically classifying this group of disorders. While it is possible that some of these kindreds have other mutations at the SCA1 locus, the disease locus (loci) for eight of these families has also been excluded from the SCA1 region, by linkage analyses.

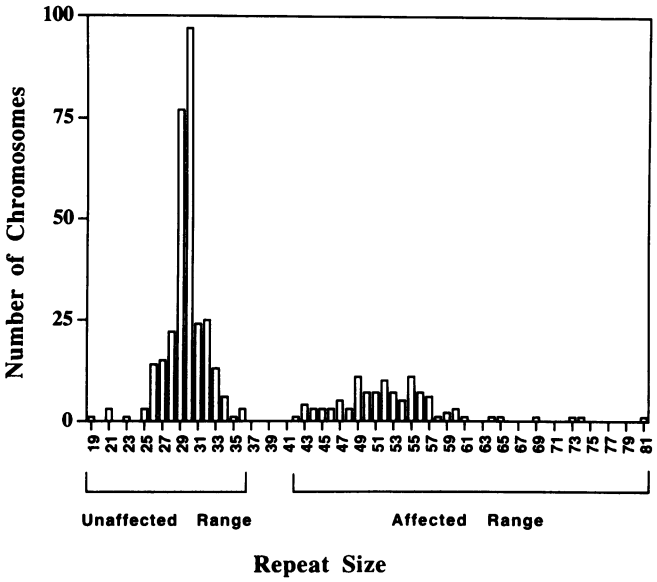

Figure 1 Distributions of CAG repeat lengths in unaffected control individuals and in SCA1 alleles. The size range of normal is 19–36 repeat units, while disease alleles contain 42–81 repeats.

Although many of the SCA1 and non-SCA1 kindreds have identical clinical features, several of the non-SCA1 kindreds displayed clinical features not observed in SCA1 families, including optic atrophy, dementia, and Parkinsonian signs (rigidity and bradykinesia). As additional ataxia loci are identified, it will be easier to establish which clinical features are characteristic of the various genetic forms of ataxia.

Repeat Analysis on Normal and SCA1 Chromosomes

Figure 1 shows the size distribution of the CAG repeats on (*a*) 304 chromosomes from unaffected control individuals who are not at risk for ataxia and (*b*) 113 expanded alleles from individuals affected with the disease. The size range of the normal alleles is 19–36 CAG repeat units. More than 95% of the normal alleles contain 25–33 CAG repeat units; and the majority (65%) of them contain 28–30 repeats. The mean repeat sizes on normal chromosomes

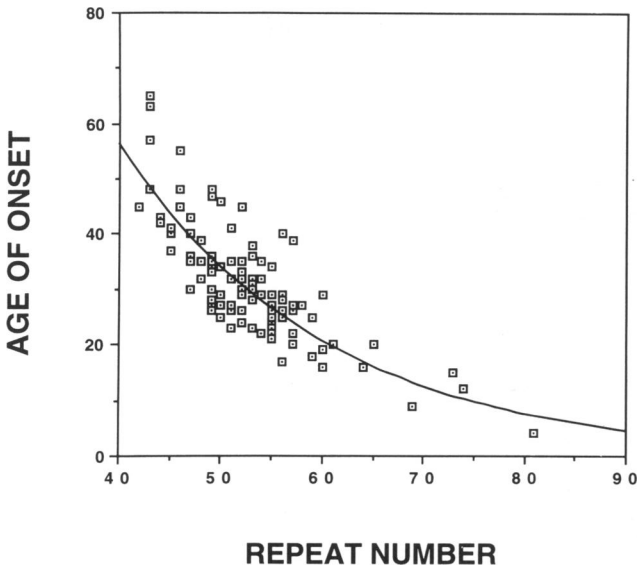

Figure 2 Relationship between the age at onset and the repeat length of the expanded allele, in 113 persons affected with SCA1. A linear correlation coefficient r of $-.774$ ($P \leq .0001$) was obtained. When ages at onset were transformed to square-root values, $r = -.8145$ ($P \leq .0001$) was calculated, indicating that 66% ($r^2 = .66$) of the variation in age at onset can be accounted for by the size of the CAG repeat length on the disease chromosome.

in the African American, Caucasian, and South African populations are very similar—29.1, 29.8, and 29.4 CAG repeat units, respectively. Combined heterozygosity for the CAG repeat at the SCA1 locus was .809 for the populations examined, giving an overall PIC value of .787. No change in CAG repeat length was observed for 135 meioses of SCA1 alleles containing CAG repeat tracts within the normal range; that is, all were inherited in a Mendelian fashion. In contrast, 41 of the 62 meioses involving expanded SCA1 alleles changed in repeat size. The rate of repeat instability for female meioses is 60%, while the instability observed for males was 82%. The number of CAG repeats found on SCA1 chromosomes from 113 affected individuals was always greater than the number of repeats on normal chromosomes, with a range of 42–81 and a mean of 52.6 (fig. 1).

Molecular and Clinical Correlations

Figure 2 shows the correlation of the age at onset of disease with the number of CAG repeats on SCA1 chromosomes. The age-at-onset estimates were based on information provided by the patient and/or other family members, for 113 individuals affected with SCA1. A linear correlation coefficient r of $-.774$ ($P \geqslant .0001$) was obtained, indicating that nearly 60% ($r^2 = .597$) of the variation in the age at onset can be accounted for by the number of CAG repeat units on the disease chromosome. When the ages at onset were transformed to square-root values, $r = -.8145$ ($P \leq .0001$) was calculated, indicating that 66% ($r^2 = .66$) of the variation in age at onset can be accounted for by the size of the CAG repeat length on the disease chromosome.

Multivariate regression analyses were performed to determine whether the age at onset (square root) also depends on the size of the repeat on the normal chromosome as well as on the disease chromosome. While the multivariate analysis again confirms a clear effect of the repeat size of the disease chromosome on age at onset, no significant effect of repeat size of the normal chromosome was observed.

The correlation between the severity of disease, as measured by duration (age at death minus age at onset), and the number of CAG repeats is shown in figure 3. A linear correlation of $-.58$ ($P = .0008$) was obtained where $n = 29$, indicating that $\sim$34% ($r^2 = .34$) of the variation in the duration of the disease is due to the number of CAG repeats.

Familial Variation in Repeat Size and Age-at-Onset Distributions

The repeat sizes and ages at onset for affected individuals were examined separately for each of the seven largest kindreds. A clustering of repeat size and age at onset was observed for six of the seven kindreds. Figure 4 shows that, for most of the SCA1 kindreds examined, the CAG repeat size and age-at-onset data cluster within a limited range. To facilitate comparisons between the different kindreds, each of the age-at-onset/repeat-number graphs (fig. 4*a–g*) was divided in two by a line generated from the mean repeat number for all 113 affected individuals studied. In six

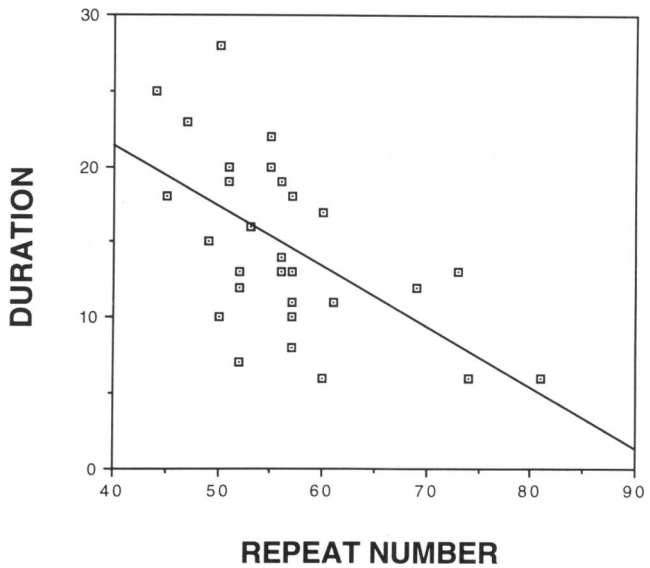

Figure 3 Relationship between CAG repeat lengths of the expanded allele and the severity of disease as measured by duration (age at death minus age at onset). A linear correlation coefficient r $-.58$ ($P = .0008$) was obtained, where $n = 29$.

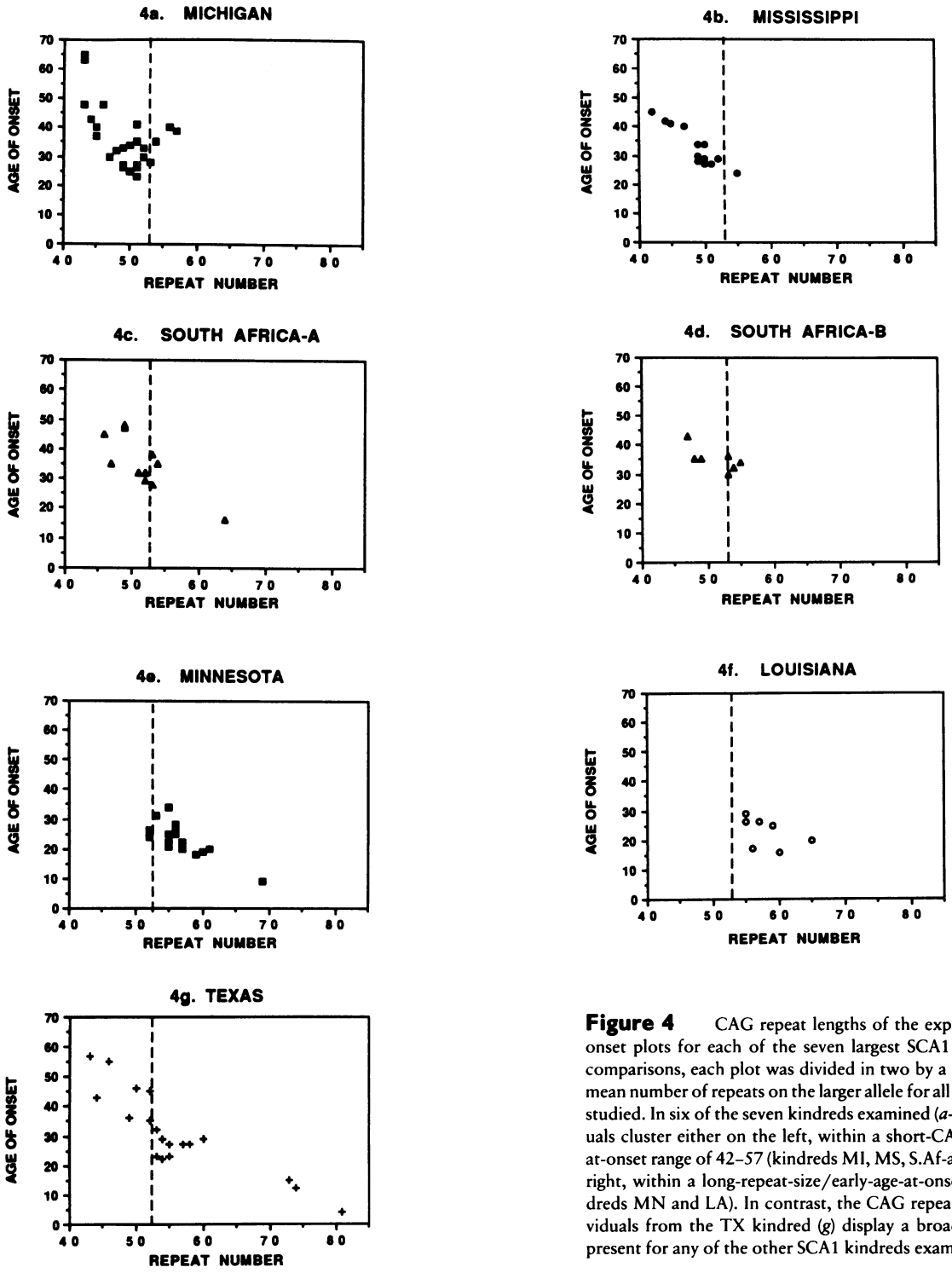

Figure 4 CAG repeat lengths of the expanded allele, vs. age-at-onset plots for each of the seven largest SCA1 kindreds. To facilitate comparisons, each plot was divided in two by a line generated from the mean number of repeats on the larger allele for all 113 affected individuals studied. In six of the seven kindreds examined (*a–f*), the affected individuals cluster either on the left, within a short-CAG-repeat-size/late-age-at-onset range of 42–57 (kindreds MI, MS, S.Af-a, and S.Af.-b), or on the right, within a long-repeat-size/early-age-at-onset range of 51–69 (kindreds MN and LA). In contrast, the CAG repeat sizes for affected individuals from the TX kindred (*g*) display a broader distribution than is present for any of the other SCA1 kindreds examined.

of the seven kindreds examined, the affected individuals cluster either on the left, within a short-CAG-repeat-size/late-age-at-onset range of 42–57 repeats (kindreds MI, MS, S.Af-a, and S.Af.-b), or on the right, within a long-CAG-repeat-size/early-age-at-onset range of 51–69 repeats (kindreds MN and LA). In contrast, the CAG repeat sizes for affected individuals from the TX kindred (fig. 4*g*) display a broader distribution than is present for any of the other SCA1 kindreds examined. In the TX kindred, 7 (37%) of 19 of the affected individuals have smaller repeat sizes and later ages at onset, while 12 (63%) of 19 have larger repeats and earlier ages at onset. Furthermore, three

of the affected individuals from the TX kindred have repeat numbers that are higher than any observed in the other six kindreds.

Analysis of variance showed significant evidence for interfamilial differences in age at onset after correction for differences in the repeat number. Differences between the MN and MS, MN and MI, and LA and MI kindreds were significant ($P<.05$ after correction for multiple tests). For this analysis, the CAG repeat number on the disease chromosome accounts for 66% ($P\leq.001$) of the variation in age at onset, while overall interfamilial differences explain ~5% ($P=.005$) of the age-at-onset variation.

In addition to interfamilial differences, sex of the transmitting parent was included in the analysis of variance. Jointly, the two factors could explain 7.1% of the variation in age at onset ($P=.001$), after elimination of the effect of the repeat number. In this analysis, age at onset in offspring of affected mothers was later than that in patients who inherited the gene from their fathers. At this point, ~27% of the variation in the age at onset is due to unknown factors.

Discussion

The cloning of the SCA1 gene allows the definitive diagnosis of one type of the dominantly inherited ataxias by using a simple blood test. This represents the first step toward an unequivocal molecular classification of the dominant ataxias. A simple and reliable classification system for the ataxias is important because the clinical symptoms overlap extensively between the SCA1 and the non-SCA1 forms of the disease. Furthermore, a molecular test for the only known SCA1 mutation permits presymptomatic diagnosis of disease in known SCA1 families and allows for the identification of sporadic or isolated CAG repeat expansions where there is no family history of the disease.

Previous observations using a limited number of affected individuals demonstrated an inverse correlation between the number of CAG repeats at the SCA1 locus and the age at onset (Orr et al. 1993). In the present study we have expanded this analysis to include 113 SCA1 chromosomes from nine kindreds representing various ethnic backgrounds. These results support the earlier observation of an inverse correlation between repeat size and age at onset. Furthermore, the data reveal a direct correlation between disease severity and repeat size. Individuals with more repeat units (or longer repeat tracts) tend to have both an earlier age at onset and a more severe disease course. To define further the relationship between the length of the CAG repeat tract and the progression of disease, it will be important to undertake prospective studies in which genotypic and detailed clinical data are collected from individuals at risk for SCA1.

While an expanded SCA1 CAG repeat is clearly predictive of disease, its value in predicting age at onset is less certain. Although the repeat size accounts for 66% of what determines the age at onset, a precise age at onset cannot be predicted by repeat size. For example, in figure 2, among the six individuals with a CAG repeat length of 50 repeats, the age-at-onset range is 25–46 years. We have found that interfamilial differences independent of repeat size explain ~5% of the age-at-onset variation. Because the MN, MI, MS, and LA families do not share a common marker haplotype for the SCA1 region, the mutations for these kindreds are likely to have arisen independently. One of several possible explanations for the family-dependent differences in ages at onset is that there may be other differences, in the SCA1 coding and/or regulatory regions among the SCA1 kindreds, that may affect SCA1 gene expression or protein stability. In addition, after correction for repeat size, the sex of the transmitting parent accounts for 2% of the age-at-onset variation. We have also examined the effect of the CAG repeat length on the normal allele in combination with that of the disease allele and found no evidence that the length of the normal allele affects age at onset. Analysis of a larger sample size with additional families may help elucidate the relative contributions of other factors. Identification of these additional factors may provide important insights into the biology of the SCA1 gene.

Dramatic differences in the number of CAG repeats are observed among the seven largest SCA1 kindreds examined (fig. 4). For six kindreds the ranges for the expanded alleles are clustered in one of two groups containing either longer or shorter CAG repeats. These differences represent kindreds with generally early or late ages at onset. The TX kindred, on the other hand, displays a wide distribution of affected individuals with repeat sizes that span a wide range of expanded allele sizes. The distributions in figure 4 are snapshots of the repeat size and age at onset observed for a given kindred at a specific point in time. The differences in the distributions may reflect the age of the mutation in a given family. Alternatively, although intergenerational changes in repeat size are observed in all families, the extreme variation in the sizes of expanded alleles in the TX kindred suggests the possibility that the repeat may be less stable for this kindred than for the other kindreds. Chung et al. (1993) have shown that the size of the expanded repeat at the SCA1 locus is not a significant factor in predisposing an expanded CAG repeat to further intergenerational instability. This result, in combination with the family-specific repeat-distribution data, suggest that the presence of another genetic factor(s) that differs among families may influence the stability of the CAG repeat at the SCA1 locus.

There are striking parallels between SCA1, spinobulbar muscular atrophy (SBMA), Huntington disease (HD), and dentatorubral pallidoluysian atrophy (DRPLA). All of these diseases are caused by the expansion of an unstable CAG repeat that codes (or likely codes) for polyglutamine

tracts within the protein (Biancalana et al. 1992; Huntington's Disease Collaborative Research Group 1993; Li et al. 1993; Orr et al. 1993; Koide et al. 1994; Nagafuchi et al. 1994; Banfi et al., in press). The size ranges for normal and expanded repeat tracts are almost identical (Andrew et al. 1993; Duyao et al. 1993; Goldberg et al. 1993; Huntington's Disease Collaborative Research Group 1993; Li et al. 1993; Norremolle 1993; Orr et al. 1993; Snell et al. 1993; Zuhlke et al. 1993; Koide et al. 1994; Nagafuchi et al. 1994). For SCA1, HD, and DRPLA, paternal transmission of the mutant alleles shows higher rates of increase in size and a larger average size of expansion (Chung et al. 1993; Duyao et al. 1993; Telenius et al. 1993; Koide et al. 1994; Nagafuchi et al. 1994). In addition, for both SCA1 and HD there is selective loss of specific populations of neurons, yet the gene is widely expressed (Orr et al. 1993; Strong et al. 1993). Telenius et al. (1993) have data suggesting the presence of a familial factor that makes some families more likely to have children with juvenile HD. Further study of family-specific effects in SCA1, HD, SBMA, and DRPLA are needed for an understanding of family-specific contributions to CAG repeat instability. In addition, it will be important to determine whether family-specific differences in the age at onset, independent of repeat length, are also observed for other diseases, to further define additional factors that contribute to disease onset.

Although, to date, all families that have been shown to have SCA1 by linkage analysis show repeat expansions, the classification of SCA1 is complicated by the genetic heterogeneity of the ataxias. Many families are too small for one to establish a priori, via linkage analysis, whether they are 6p families. Although the distribution of alleles among the expansion-negative affected families is similar to that among the normals (data not shown), one cannot construct from all the affecteds a simple comparison between SCA1 and "normal" alleles. Alleles in affected patients may be negative for expansion either because they represent a mutation at another locus or because they have a different abnormality in the SCA1 gene. Chung et al. (1993) have shown that 98% of normal alleles have CAT interruptions in their CAG repeat tracts. To date, all CAG repeat tracts >21 units in length on normal chromosomes are interrupted by one or more CATs. Thus, there is a possibility that either differences in the repeat configuration or other mutations within the SCA1 gene may also lead to ataxia.

The cloning of the SCA1 gene creates many opportunities for additional studies aimed at understanding the molecular mechanism underlying this progressive neurologic disorder. Furthermore, studies to estimate the prevalence of this subtype of ataxia can now be carried out. Because variable ages at onset and anticipation have been reported for other dominant ataxias, including Machado-Joseph disease (Coutinho and Andrade 1978) and spinocerebellar ataxia type 2 (Orozco Diaz et al. 1990), the expansion of trinucleotide repeats is likely to be a predominant mutation for other forms of autosomal dominant ataxia. The genetic heterogeneity of the ataxias provides a unique opportunity for studying the various molecular pathways that can lead to loss of specific neuronal populations within the central nervous system.

Acknowledgments

We thank the ataxia families for their participation. In addition, we gratefully acknowledge research support from the National Institutes of Health, the Muscular Dystrophy Association, the National Ataxia Foundation, the MRC of South Africa, the Mauerberger Foundation, and the University of Cape Town Staff Research Fund.

Appendix

Clinical Features of SCA1 Kindreds

Observed in every molecularly proved SCA1 patient:
- Gait and limb ataxia
- Dysarthria
- Dysfunction of cranial nerves IX, X, and XII

Observed in every molecularly proved SCA1 family but not in every SCA1 patient:
- Oculomotor deficits
- Motor weakness and amyotrophy
- Proprioceptive sensory deficits
- Pyramidal tract deficits
- Dystonic posturing and adventicious movements

References

Andrew SE, Goldberg YP, Kremer B, Telenius H, Theilmann J, Adam S, Starr E, et al (1993) The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. Nature Genet 4:398–403

Auburger G, Diaz GO, Capote RF, Sanchez SG, Perez MP, del Cueto ME, Meneses MG, et al (1990) Autosomal dominant ataxia: genetic evidence for locus heterogeneity from a Cuban founder-effect population. Am J Hum Genet 46:1163–1177

Banfi S, Chung M-y, Kwiatkowski TJ Jr, Ranum LPW, McCall AE, Chinault AC, Orr HT, et al (1993) Mapping and cloning of the critical region for the spinocerebellar ataxia type I gene in a yeast artificial chromosome contig spanning 1.2 Mb. Genomics 18:627–635

Banfi S, Servadio A, Chung M-y, Kwiatkowski TJ Jr, McCall AE, Duvick LA (1994) Identification and characterization of the gene causing type 1 spinocerebellar ataxia. Nature Genet (in press)

Biancalana V, Serville F, Pommier J, Julien J, Hanauer A, Mandel JA (1992) Moderate instability of the trinucleotide repeat in spinobulbar muscular atrophy. Hum Mol Genet 1:255–258

Brook JD, McCurrah ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, et al (1992) 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

Bryer A, Martell RW, duToit ED, Beighton P (1992) Adult onset spinocerebellar ataxia linked to HLA in a large South African kindred of mixed ancestry. Tissue Antigens 40:111–115

Buxton J, Shelbourne P, Davies J, Jones C, Van Tongeren T, Aslanidis C, de Jong P, et al (1992) Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. Nature 355:547–548

Carson WJ, Radvany J, Farrer LA, Vincent D, Rosenberg RN, MacLeod PM, Rouleau GA, et al (1992) The Machado-Joseph disease locus is different from the spinocerebellar ataxia locus (SCA1). Genomics 13:852–855

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. Nature Genet 5:254–258

Coutinho P, Andrade C (1978) Autosomal dominant system degeneration in Portuguese families of the Azores. Neurology 28: 703–709

Currier RD, Glover G, Jackson JF, Tipton AC (1972) Spinocerebellar ataxia: study of a large kindred. I. General information and genetics. Neurology 22:1040–1043

Duyao M, Ambrose C, Myers R, Novelletto A, Persichetti F, Frontalli M, Folstein S, et al (1993) Trinucleotide repeat length instability and age of onset in Huntington's disease. Nature Genet 4:387–392

Fu Y-H, Kuhl DPA, Pizzuti A, Pieretti M, Sutcliffe JS, Richards S, Verkerk AJMH, et al (1992) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell 67:1047–1058

Gispert S, Twells R, Orozco G, Brice A, Weber J, Heredero L, Scheufler K, et al (1993) Chromosomal assignment of the second locus for autosomal dominant cerebellar ataxia (SCA2) to chromosome 12q23-24.1. Nature Genet 4:295–299

Goldberg YP, Kremer B, Andrew SE, Theilmann J, Graham RK, Squitieri F, Telenius, H, et al (1993) Molecular analysis of new mutations for Huntington's disease: intermediate alleles and sex of origin effects. Nature Genet 5:174–178

Goldfarb LG, Chumakov MP, Petrov PA, Fedorova NI, Gajdusek DC (1989) Olivopontocerebellar atrophy in a large Iakut kinship in eastern Siberia. Neurology 39:1527–1530

Greenfield JG (1954) The spino-cerebellar degenerations. Charles C Thomas, Springfield, IL

Haines JL, Schut LJ, Weitkamp LR, Thayer M, Anderson VE (1984) Spinocerebellar ataxia in a large kindred: age at onset, reproduction, and genetic linkage studies. Neurology 34: 1542–1548

Harley HG, Brook JD, Rundle SA, Crow S, Reardon W, Buckler AJ, Harper PS, et al (1992) Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. Nature 355:545–546

Huntington's Disease Collaborative Research Group, The (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell 72: 971–983

Jackson JF, Currier RD, Terasaki PI, Morton NE (1977) Spinocerebellar ataxia and HLA linkage—risk prediction by HLA typing. N Engl J Med 296:1138–1141

Keats BJB, Pollack MS, McCall A, Wilensky MA, Ward LJ, Lu M, Zoghbi HY (1991) Tight linkage of the gene for spinocerebellar ataxia to D6S89 on the short arm of chromosome 6 in a kindred for which close linkage to both HLA and F13A1 is excluded. Am J Hum Genet 49:972–977

Koide R, Ikeuchi T, Onodera O, Tanaka H, Igarashi S, Endo K, Takahashi H, et al (1994) Unstable expansion of CAG repeat in hereditary dentatorubral-pallidoluysian atrophy (DRPLA). Nature Genet 6:9–13

Kremer EJ, Pritchard M, Lynch M, Yu S, Holman K, Baker E, Warren ST, et al (1991) Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. Science 252:1711–1714

Kumar D, Blank CE, Gelsthorpe K (1986) Hereditary cerebellar ataxia and genetic linkage with HLA. Hum Genet 72:327–332

Kwiatkowski TJ Jr, Orr HT, Banfi S, McCall AE, Jodice C, Persichetti F, Novelletto A, et al (1993) The gene for autosomal dominant spinocerebellar ataxia (SCA1) maps centromeric to D6S89 and shows no recombination, in nine large kindreds, with a dinucleotide repeat at the AM10 locus. Am J Hum Genet 53:391–400

Li S-H, McInnis MG, Margolis RL, Antonarakis SE, Ross CA (1993) Novel triplet repeat containing genes in human brain: cloning, expression, and length polymorphisms. Genomics 16: 572–579

McKusick VA (1978) Mendelian inheritance in man. John Hopkins University Press, Baltimore

Mahadevan M, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, Neville C, et al (1992) Myotonic dystrophy mutation: an unstable CTG repeat in the 3′ untranslated region of the gene. Science 255:1253–1255

Nagafuchi S, Yanagisawa H, Sato K, Shirayama T, Ohsaki E, Bundo M, Takeda T, et al (1994) Dentatorubral and pallidoluysian atrophy expansion of an unstable CAG trinucleotide on chromosome 12p. Nature Genet 6:14–18

Nino HE, Noreen HJ, Dubey DP, Resch JA, Namboodiri K, Elston RC, Yunis EJ (1980) A family with hereditary ataxia: HLA typing. Neurology 30:12–20

Norremolle A, Riess O, Epplen JT, Fenger K, Hasholt L, Sorensen SA (1993) Trinucleotide repeat elongation in the Huntington gene in Huntington disease patients from 71 Danish families. Hum Mol Genet 2:1475–1476

Orozco Diaz G, Nodarse Fleites A, Cordoves Sagaz R, Auburger G (1990) Autosomal dominant cerebellar ataxia: clinical analysis of 263 patients from a homogeneous population in Holguin, Cuba. Neurology 40:1369–1375

Orr HT, Chung M-y, Banfi S, Kwiatkowski TJ Jr, Servadio A, Beaudet AL, McCall AE, et al (1993) Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. Nature Genet 4:211–226

Ranum LPW, Duvick LA, Rich SS, Schut LJ, Litt M, Orr HT (1991) Localization of the autosomal dominant HLA-linked spinocerebellar ataxia (SCA1) locus, in two kindreds within an 8-cM subregion of chromosome 6p. Am J Hum Genet 49:31–41

Ranum LPW, Rich SS, Nance MA, Duvick LA, Aita JF, Orr HT, Anton-Johnson S, et al (1992) Autosomal dominant spinocerebellar ataxia: locus heterogeneity in a Nebraska kindred. Neurology 43:344–347

Sakai T, Ohta M, Ishino H (1983) Joseph disease in a non-Portuguese family. Neurology 33:74–80

Schut JW (1950) Hereditary ataxia: clinical study through six generations. Arch Neurol Psychiatry 63:535–586

Snell RG, MacMillan JC, Cheadle JP, Fenton I, Lazarou LP, Davies P, MacDonald ME, et al (1993) Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. Nature Genet 4:393–397

Strong TV, Tagle DA, Valdes JM, Elmer LW, Boehm K, Swaroop M, Kaatz KW, et al (1993) Widespread expression of the human and rat Huntington's disease gene in brain and nonneural tissues. Nature Genet 5:259–265

Takiyama Y, Nishizawa M, Tanaka H, Kawashima S, Sakamoto H, Karube Y, Shimazaki H, et al (1993) The gene for Machado-Joseph disease maps to human chromosome 14q. Nature Genet 4:300–304

Telenius H, Kremer HPH, Theilmann J, Andrew SE, Almqvist E, Anvret M, Greenberg C, et al (1993) Molecular analysis of juvenile Huntington disease: the major influence on (CAG)n repeat length is the sex of the affected parent. Hum Mol Genet 2:1535–1540

Verkerk AJMH, Pieretti M, Sutcliffe JS, Fu Y-H, Kuhl DPA, Pizzuti A, Reiner O, et al (1991) 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

Yakura H, Wakisaka A, Fujimoto S, Itakura K (1974) Hereditary ataxia and HLA genotypes. N Engl J Med 291:154–155

Zoghbi HY, Jodice C, Sandkuijl LA, Kwaitkowski TJ Jr, McCall AE, Huntoon SA, Lulli, P, et al (1991) The gene for autosomal dominant spinocerebellar ataxia (SCA1) maps telomeric to the HLA complex and is closely linked to the D6S89 locus in three large kindreds. Am J Hum Genet 49:23–30

Zoghbi HY, Pollack MS, Lyons LA, Ferrell RE, Daiger SP, Beaudet AL, et al (1988) Spinocerebellar ataxia: variable age of onset and linkage to human leukocyte antigen in a large kindred. Ann Neuroal 23:580–584

Zuhlke C, Riess O, Schroder K, Siedlaczck I, Epplen JT, Engel W Thies U (1993) Expansion of the (CAG)n repeat causing Huntington's disease in 352 patients of German origin. Hum Mol Genet 2:1467–1469