Mutation Detection in Machado-Joseph Disease Using Repeat Expansion Detection

Kerstin Lindblad,* Astrid Lunkes,† Patricia Maciel,‡§ Giovanni Stevanin,‖ Cecilia Zander,* Thomas Klockgether,# Tim Ratzlaff,† Alexis Brice,‖ Guy A. Rouleau,‡ Tom Hudson, Georg Auburger,† and Martin Schalling***

*Neurogenetics Unit, Department of Molecular Medicine, Karolinska Hospital, Stockholm, Sweden
†Department of Neurology, University of Düsseldorf, Düsseldorf, Germany
‡Department of Neuroscience, Montreal General Hospital Research Institute, Montreal, Canada
§UnIGENe, IBMC, and Laboratório de Genética Médica, ICBAS, University of Porto, Porto, Portugal
‖INSERM U289, Hopital de Salpêtrière, Paris, France
#Department of Neurology, University Hospital of Tübingen, Germany
**Whitehead Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.

ABSTRACT

Background: Several neurological disorders have recently been explained through the discovery of expanded DNA repeat sequences. Among these is Machado-Joseph disease, one of the most common spinocerebellar ataxias (MJD/SCA3), caused by a CAG repeat expansion on chromosome 14. A useful way of detecting repeat sequence mutations is offered by the repeat expansion detection method (RED), in which a thermostable ligase is used to detect repeat expansions directly from genomic DNA. We have used RED to detect CAG expansions in families with either MJD/SCA3 or with previously uncharacterized spinocerebellar ataxia (SCA).

Materials and Methods: Five MJD/SCA3 families and one SCA family where linkage to SCA1–5 had been excluded were analyzed by RED and polymerase chain reaction (PCR).

Results: An expansion represented by RED products of 180–270 bp segregated with MJD/SCA3 ($p < 0.00001$) in five families ($n = 60$) and PCR products corresponding to 66–80 repeat copies were observed in all affected individuals. We also detected a 210-bp RED product segregating with disease ($p < 0.01$) in a non–SCA1–5 family ($n = 16$), suggesting involvement of a CAG expansion in the pathophysiology. PCR analysis subsequently revealed an elongated MJD/SCA3 allele in all affected family members.

Conclusions: RED products detected in Machado-Joseph disease families correlated with elongated PCR products at the MJD/SCA3 locus. We demonstrate the added usefulness of RED in detecting repeat expansions in disorders where linkage is complicated by phenotyping problems in gradually developing adult-onset disorders, as in the non–SCA1–5 family examined. The RED method is informative without any knowledge of flanking sequences. This is particularly useful when studying diseases where the mutated gene is unknown. We conclude that RED is a reliable method for analyzing expanded repeat sequences in the genome.

INTRODUCTION

Machado-Joseph disease (MJD) is an autosomal dominant neurodegenerative disorder involving predominantly cerebellar, pyramidal, extrapyra-

Address correspondence and reprint requests to: Kerstin Lindblad, Neurogenetics Unit, Department of Molecular Medicine, Karolinska Hospital, S-171 76 Stockholm, Sweden. E-mail:KELI@GEN.KS.SE

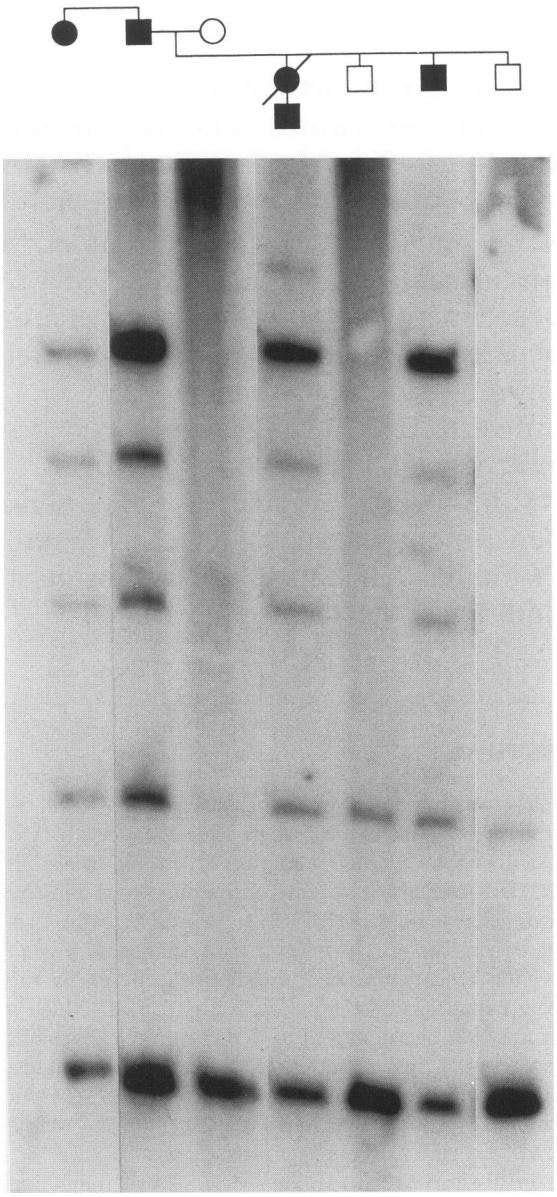

FIG. 1. Autoradiograph of RED products from part of a MJD/SCA3 disease family using a $(CTG)_{10}$ oligonucleotide

In RED, a thermostable ligase is used to ligate repeat oligonucleotides following annealing at adjacent bases to repeat sequences in genomic DNA. Products are size separated on a denaturing polyacrylamide gel, electroblotted, and hybridized to a complementary ^{32}P-labeled probe. The longest reaction product corresponds to the longest repeat size present in the genome. Reactions were run as previously (6), with cycling conditions being 94°C for 5 min followed by 396 cycles of either 80°C for 30 sec and 94°C for 10 sec. The lowest band in each lane represents 90 bp and each additional band represents an additional 30 bp. Note RED products with largest band size at 210 bp in affected individuals in generation 1 and 2 whereas the affected offspring in the third generation (fourth lane from left) has a 240-bp product correlating with a larger repeat size. Unaffecteds only show products of 120 bp.

midal, motor neuron, and oculomotor systems (1). It is one of the most common hereditary spinocerebellar degenerative disorders (2) and has been mapped to chromosome 14q (3), in the same region as the spinocerebellar ataxia 3 (SCA3) locus (4). Kawaguchi and coworkers recently found a candidate disease gene for Machado-Joseph disease using a trinucleotide probe (5). The gene they isolated mapped to chromosome 14q32.1. It contained a 1776-bp long open reading frame, was expressed in human brain, and contained an inframe CAG repeat that was expanded in Machado-Joseph patients but not in normal individuals and unaffected family members. In normal individuals the repeat length in the Machado-Joseph gene was 13–36 repeat copies and in 12 tested patients 68–79 copies (5).

To facilitate detection of long repeat sequences we developed the repeat expansion detection method (RED) (6). RED is an automatable method for detecting repeat expansions directly from genomic DNA. In this method, a repeat oligonucleotide that is complementary to one of the strands of genomic DNA is allowed to hybridize at a temperature close to the melting point. Multimers of the oligonucleotide are subsequently formed using a thermostable ligase in a cycling procedure. Products are detected as a ladder of bands on a gel (Fig. 1). The band with the highest molecular weight represents the largest repeat expansion in that particular genome. The RED method has previously been used to chromosomally map a novel expanded repeat locus (RED-1)(6), to identify repeat expansions in myotonic dystrophy (6), and to detect four novel expanded trinucleotide repeat motifs (7). Furthermore, RED has been used to associate bipolar affective disorder with elongated CAG repeat sequence alleles (8).

Here we analyzed six spinocerebellar ataxia families using both RED and conventional polymerase chain reaction (PCR). One of these families had previously been excluded from the MJD/SCA3 locus on chromosome 14q by linkage analysis. Screening families using the RED method proved useful in identifying ex-

panded repeats in MJD/SCA3 as well as in the previously undiagnosed family.

MATERIALS AND METHODS

Families

Affected and unaffected individuals from six families with spinocerebellar ataxia were analyzed. Three Portuguese-American (9), one German (K. Bürk et al., submitted), and one French (4) family (in total 31 affected and 29 unaffected individuals) had previously been characterized as Machado-Joseph disease and linked to chromosome 14q. For a sixth family (eight affected, eight unaffected individuals), linkage had been excluded to the MJD/SCA3 locus as well as to the loci of SCA1, SCA2, SCA4, and SCA5 (unpublished data).

Sources of Genomic DNA

Human genomic DNA was prepared from blood samples using a standard phenol/chloroform extraction procedure (10).

RED Analysis

All reactions were performed on a GeneAmp PCR System 9600 (Perkin Elmer Cetus, Norwalk, CT, U.S.A.), using the following conditions: Reactions (10 μl) containing 1 μg of genomic DNA, 50 ng of 5'-end phosphorylated $(CTG)_{10}$ oligonucleotide (Pharmacia Biotech, Uppsala, Sweden) and 10 U of Ampligase (Epicentre Technologies, Madison, WI, U.S.A.) with the supplied Ampligase buffer were incubated at 94°C for 5 min. Thereafter, samples were taken through 396 cycles of 80°C for 30 sec and 94°C for 10 sec (4).

Samples were heat denatured in 50% formamide for 5 min, before electrophoresis on a 6% denaturing polyacrylamide/6 M urea gel. The gel was subsequently transferred to 3MM filter paper and DNA electrotransferred onto Hybond N^+ membrane using 2 A for 40 min in 1 × TBE. Following UV immobilization, membranes were hybridized for 16 hr at 60°C to a $(CAG)_{10}$ oligonucleotide, that had been 3'-end labeled (11) using terminal deoxynucleotidyl transferase (Amersham, Little Chalfont, United Kingdom) and ^{32}P dATP (NEG 012Z; NEN Du Medical, Wilmington, DE, U.S.A.) to a specific activity of 2–9 × 10^9 cpm/μg. Membranes were washed in 1 × SSC, 0.1% SDS for 2 hr at 60°C, and autoradiographed 1–3 days on NEN X-ray film using an intensifying screen.

A ligation of three $(CTG)_{10}$ oligonucleotides resulting in a band at 90 bp, serving as a positive control for the RED assay, was present in all samples used for the calculations presented in this paper. A 90-bp product most likely represents ligations at multiple short repeat loci in the genome (6). Families segregating two CAG expansions of different size were analyzed both with respect to highest molecular weight band and to subsets of bands with increased intensity. (see Fig. 4)

PCR Analysis

PCR conditions were as reported by Kawaguchi and coworkers (5), using either the MJD70 (3'-AACCCTCACTAGATCCATTC-5') + MJD52 (5'-CCAGTGACTACTTTGATTCG-3') or MJD25 (3'-AAGTGTAGGTACACTTTCCGGT-5') + MJD52 primer combinations. PCR products were separated either on an agarose gel or heat denatured in 50% formamide for 5 min, before separation on a 6% denaturing polyacrylamide/6 M urea gel. DNA from the polyacrylamide gel was transferred by capillary blotting onto Hybond N^+ membranes for 2 hr. Following UV immobilization, the membranes were hybridized to a $(CAG)_{10}$ oligonucleotide as described above, and autoradiographed for 0.5–12 hr.

Statistical Analysis

Initially, RED products of 210 or 240 bp were defined as the disease allele. We subsequently analyzed RED products of 180–270 bp defined as the disease allele. The probability of cosegregation between phenotype and disease allele was calculated using chi-square analysis or Fisher's exact test. Spearman's correlation coefficient was used to correlate PCR and RED product sizes.

RESULTS

Analysis by PCR across the CAG repeat in the MJD/SCA3 gene from affected individuals (n = 31) showed expansions of 66–80 repeat copies in all except two cases (Table 1). In addition, PCR products in the disease allele size range was seen in six asymptomatic individuals (A in Fig. 2). In two individuals (C and D in Fig. 2) initially scored as being affected, no expanded alleles

TABLE 1. Distribution of elongated CAG allele sizes in 44 MJD/SCA3 family members

Repeat Copy Number	Number of Individuals
64	1[a]
65	—
66	2
67	1
68	8
69	3
70	1
71	5
72	3
73	1
74	5
75	5
76	4
77	—
78	—
79	1
80	4

[a]Not clinically affected at age 45.

were detected. One of these individuals (C in Fig. 2) had two distinct alleles in the normal size range. When reevaluated, this individual was found to be unaffected. The other individual showed only one PCR product in the normal range, but no expansion could be detected.

Analysis by RED showed ligation products of 180–270 bp that segregated with the disease phenotype in all affected cases ($n = 31$) except one (Figs. 1 and 2). This individual (C in Fig. 2) was found to be unaffected when reevaluated, as mentioned above. Using the definition that any RED product of 210/240 bp correspond to a MJD/SCA3 allele, the probability of linkage between disease allele and phenotype as initially determined, was $p < 0.00001$ ($n = 60$). With the RED interval extended to 180–270 bp, an equal significance of $p < 0.00001$ was observed.

Most RED analysis resulted in relatively even intensity of the RED products. In two families (Fig. 2, Family V, and Fig. 3) a longer as well as a shorter RED product was observed in certain individuals (Fig. 4). In these cases, different alleles were identified by analyzing differences in band intensity of the RED products within one given lane on a gel (Fig. 4), as well as by observation of independent segregation of the different band intensities (Fig. 2, Family V, and Fig. 3). In such samples bands at 210 bp and below were distinctly more intense than bands above 210 bp, most likely representing superimposition of products from two alleles of 210 and 420 bp respectively. In both families the 210-bp allele segregated with disease whereas the 420-bp allele was completely unrelated to disease.

Nine asymptomatic individuals (A and B in Fig. 2) showed RED products of 180–270 bp. Six of these nine displayed an expanded repeat also when analyzed with PCR and are thus individuals at risk of developing MJD/SCA3 (A in Fig. 2). The remaining three (B in Fig. 2) had normal size alleles at the MJD/SCA3 locus. The RED products seen in these individuals would thus represent expansions at other loci, possibly related to a different phenotype. RED product sizes correlated well with the actual repeat sizes ($p < 0.001$), preferentially giving a representation somewhat shorter than the actual repeat size as determined by PCR. For example the 210-bp RED products were caused by repeat sizes ranging between 66 and 76 repeat copies (198–228 bp). However, one individual (E in Fig. 2) that had a PCR expansion of 80 repeat copies showed a 270-bp RED product. One asymptomatic male of 45 years (F in Fig. 2) had a RED product of 180 bp and a PCR product of 64 repeat copies, well above the normal range of 13–36 copies (5,9).

Analysis of an SCA family where linkage had previously been excluded for the loci of MJD/SCA3 and SCA1, 2, 4, and 5, identified a 210-bp RED product that segregated with the phenotype ($p < 0.01$, $n = 16$) (Fig. 3). All affected individuals showed 210-bp RED products. The size, sequence, and banding pattern of these novel RED products were identical to those of most MJD/SCA3 patients analyzed. The MJD/SCA3 locus was therefore investigated further despite the previous exclusion by linkage. We subsequently identified a repeat expansion in the MJD/SCA3 gene by PCR confirming a diagnosis of Machado-Joseph disease. In this family, three unaffected individuals showed 210-bp RED products. Of these, one showed an expansion also by PCR and is at risk of acquiring the disease later. Two affected individuals showed a RED product above 270 bp, together with the 210-bp RED product (Fig. 4) analogous to Family V (Fig. 2). This longer RED allele segregated independent of the disease phenotype.

DISCUSSION

Several neurological disorders have been explained at the molecular level by expanding CAG repeat sequences (12), including Machado-Joseph disease (5). In six MJD/SCA3 families analyzed here, the size range of expansions was 66–80 repeat copies (198–240 bp) as determined by PCR (Table 1), extending the initially reported pathological size range (68–79) (5). Accordingly, a pathological size range of 62–84 copies has been reported (9). A case of particular note is Individual F in Fig. 2. This individual is clinically unaffected at age 45 and carries an allele of 64 copies, which is well beyond any normal size allele we have observed. In all likelihood, the expanded allele of Individual F was transmitted from his affected father where it underwent a five-copy reduction. It is quite possible that Individual F will develop symptoms later in life, thereby extending the lower end of the pathological expansion size range in our MJD/SCA3 study. If not, it would be the longest polyglutamine tract reported to date without concomitant neurological disease. As in other trinucleotide associated disorders, MJD/SCA3 displays anticipation coupled to an increase in repeat expansion size within families. We observed several such changes (Fig. 2) represented by increases in size of both RED and PCR products in later generations.

Despite the fact that the expanded repeat sequence is relatively short in MJD/SCA3, we observed a PCR amplification bias towards the shorter allele in many samples. Initially, some expansions were not clearly seen, but deduced by the presence of only one short allele, where the parental genotype was incompatible with homozygozity. This was resolved by blotting the PCR products to a membrane and subsequent hybridization with an internal repeat oligonucleotide that had been 3'-end-labeled to enhance the specific activity. In this context, we observed an affected individual (D in Fig. 2) who showed only one normal size PCR product and had a 180-bp RED product. We have not been able to determine if this individual is a homozygote or if a longer PCR allele fails to amplify. Since this individual is affected and has a RED product in the pathological range, it may be that RED analysis in this case is more robust than PCR. Alternatively, this individual is in fact homozygous for the normal size allele, with the expanded RED product originating from a different locus and consequently a misdiagnosis.

The RED method produced expansion sizes of 180–270 bp that segregated with the disease and were close to the respective repeat sizes as determined by PCR, demonstrating that RED is a reliable method for detection of repeat expansions. Approximately 29% of the normal population show short CAG expansions with no apparent phenotype (8), and consequently some individuals will show RED products unrelated to disease, but within the disease expansion size range (B in Figs. 2 and 3). Nevertheless, RED products relating to a phenotype can often be identified by examining the segregation of a RED product of a given size within a family (exemplified in Fig. 2, Family V, and Fig. 3).

In some cases we have assigned individuals two RED product sizes in the present paper. This was based on the observation of two differently sized, independently segregating RED products within a family. To identify individuals where both alleles are present requires careful analysis. Yet, if RED product size and band intensity are measured, the two expanded alleles can be separated. It is thus frequently possible to detect a shorter expansion in the presence of a long expansion, by analyzing the increase in intensity of a subset of bands produced by superimposition of two repeat alleles within one lane. For future diagnostic purposes, this type of analysis may be substituted by a modified RED protocol, which we are currently developing. In this modification, a different size oligonucleotide complementary to a sequence adjacent to a given repeat locus is used in combination with the repeat oligonucleotide, producing a second RED product that appears in close proximity to each of the repeat-based RED products. Any ladder not containing these double bands would thus represent another locus.

The fact that RED is not limited to one locus per analysis is a great advantage. As shown in this paper, it can be used to identify and link expansions to disease in families where the gene or its locus are unknown. This may be particularly useful in the search for genes involved in adult-onset neurodegenerative disorders. Due to the adult onset, individuals in these families may be carrying the mutant allele without expressing the disease, as exemplified by individuals labeled A in Figs. 2 and 3. Alternatively phenocopies may be seen, as exemplified by Individual C (Fig. 2) who was first considered affected, but then reevaluated as unaffected after genetic analysis. In both instances linkage analysis may fail. RED offers an additional advantage as it detects the

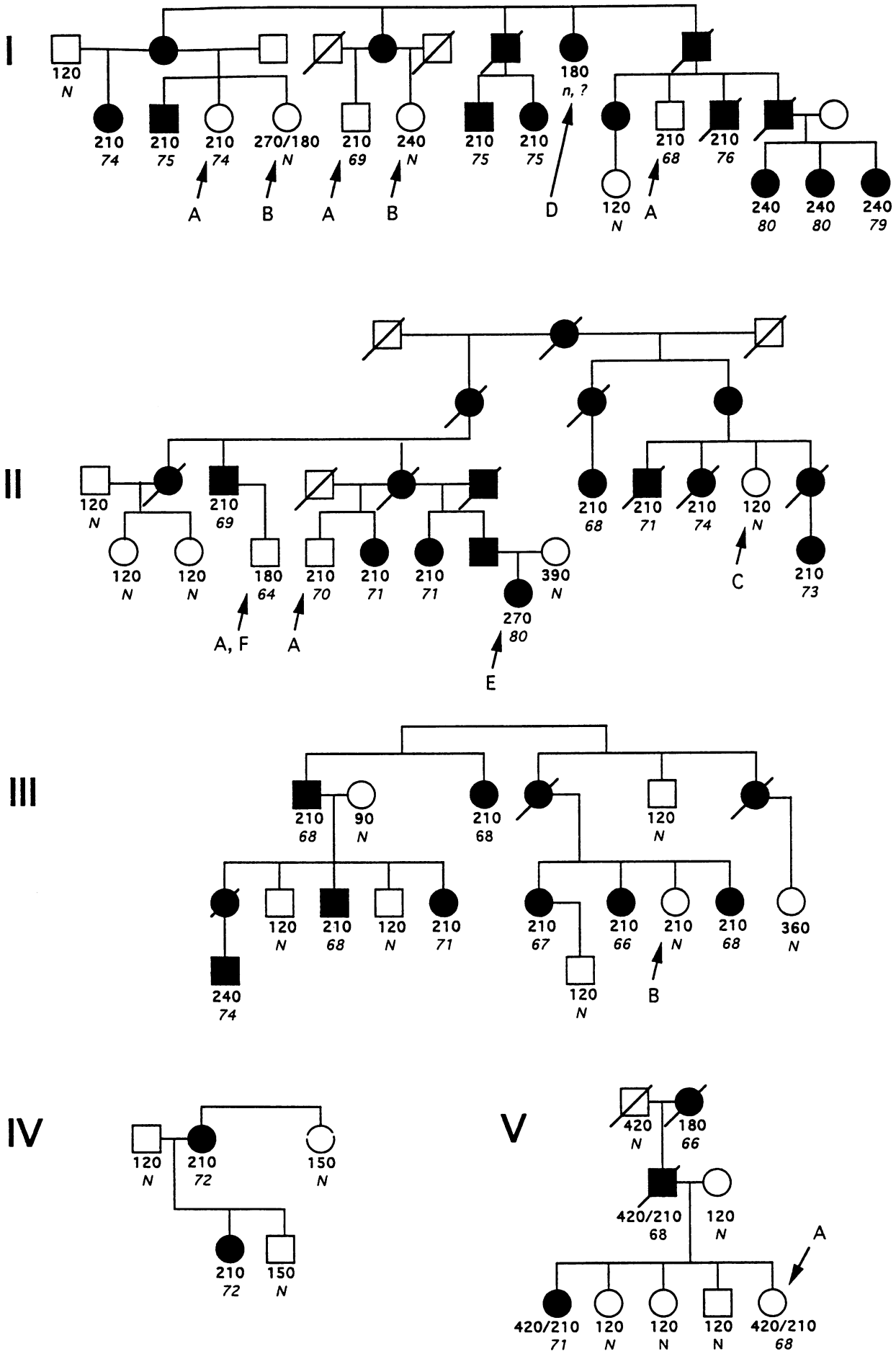
I
120 N
210 74
210 75
210 74
A
270/180 N
B
210 69
A
240 N
B
210 75
210 75
180 n, ?
D
120 N
210 68
A
210 76
240 80
240 80
240 79
II
120 N
120 N
120 N
210 69
180 64
A, F
210 70
A
210 71
210 71
270 80
E
390 N
210 68
210 71
210 74
120 N
C
210 73
III
210 68
90 N
210 68
120 N
120 N
210 68
120 N
210 71
240 74
210 67
120 N
210 66
210 N
B
210 68
360 N
IV
120 N
210 72
150 N
210 72
150 N
V
420 N
180 66
420/210 68
120 N
A
420/210 71
120 N
120 N
120 N
420/210 68

FIG. 2. Pedigrees from five MJD/SCA3 disease families showing repeat sizes determined by RED and PCR

RED product sizes (bp) are indicated in bold characters (upper numbers), whereas repeat copy number of expansions detected by PCR are indicated below in italics. Two PCR alleles in the normal range are designated *N*, whereas cases with only one detectable size are designated *n*, ? Individuals marked A are clinically asymptomatic but show RED and PCR expansions. Individuals marked B show RED products in the MJD/SCA3 range, but are asymptomatic and lack detectable PCR expansions. Individual C was first judged to be affected, but did not show either PCR or RED expansions. She was therefore reevaluated and found to be unaffected. Individual D is affected and has a RED expansion but shows only one short PCR product and no expansion. Individual E has the longest repeat size of 80 copies and also the longest RED product of 270 bp. Individual F is clinically unaffected at the age of 45, with a repeat size of 64 copies.

actual disease-causing mutation (the repeat expansion), thereby excluding recombination between the genetic marker and the mutation. Although this does not eliminate the need for cosegregation analysis and the influence of incorrect phenotyping, it reduces the statistical model to a simple chi-square analysis or Fisher's test, thereby increasing power of detection. This is illustrated by the detection and linkage of MJD/SCA3 expansions in the present paper, despite the presence of individuals at risk (A in Figs. 2 and 3) as well as one individual (C in Fig. 2) where the phenotyping was incorrect. The most interesting example is present in the sixth family, which had previously been excluded from the MJD/SCA3 locus by linkage analysis using simple sequence length polymorphisms. This exclusion was mainly due to phenotyping errors in two individuals. Using RED the correlation between the expansions and phenotype was strongly significant leading to a correct diagnosis of this family. A strong hint that this family was affected with MJD/SCA3 was the similarity in RED product size and band pattern to those seen in individuals diagnosed with MJD/SCA3.

Because RED is a genome-wide screening method, it represents a shortcut in identifying a subtype of mutations compared to linkage and positional cloning. It thus permits identification of a pathological repeat expansion in a genetically uncharacterized disease, without any knowledge of the chromosomal localization of the mutation or candidate disease genes. The RED method should be particularly useful in the

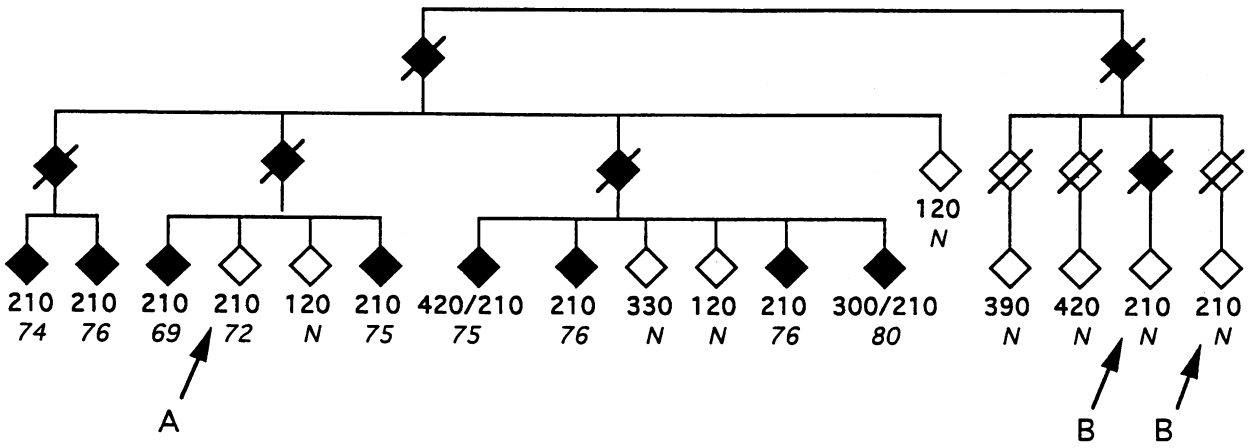

FIG. 3. Pedigree from an uncharacterized spinocerebellar ataxia family where cosegregation of a RED product and disease lead to the diagnosis of Machado-Joseph disease

RED product sizes (bp) are indicated in bold characters above and repeat copy number of expansions detected by PCR are indicated below in italics. Two PCR alleles in the normal range are designated *N*. Individual A is clinically unaffected but show both RED and PCR expansions. Individuals marked B show RED products in the MJD/SCA3 range, but are unaffected and lack PCR expansions. Individuals have been scrambled and marked as unisex to prevent identification.

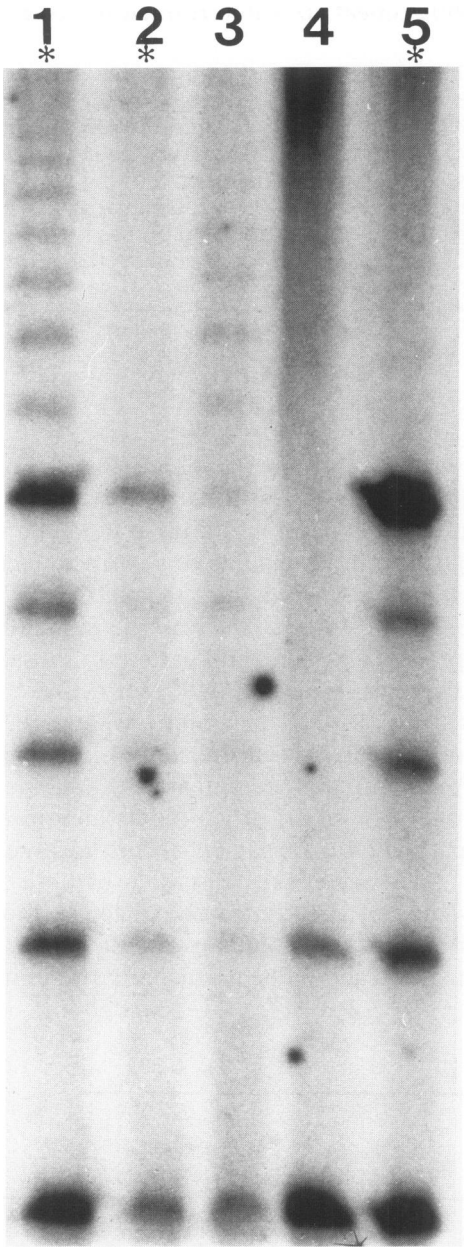

FIG. 4. Autoradiograph showing the detection of two different RED products in siblings of a MJD/SCA3 family

In the unaffected individual in Lane 3 a ladder of RED products of even intensity can be seen, indicating that this represents a single expansion. In the affected individual in Lane 1, an increased intensity of the 210-bp band compared with the longer products in the ladder is seen. This individual is therefore scored as having both a 210-bp RED product (segregating with MJD/SCA3) and a 420-bp product (with no apparent phenotype). The two alleles most likely represent independent loci in the genome. The lowest band in each lane represents 90 bp and each additional band represents an additional 30 bp. Stars denote affected individuals. Pedigree information was excluded to prevent identification.

study of dominant inherited disorders displaying anticipation, particularly those with a neurological phenotype.

ACKNOWLEDGMENTS

We thank the families for their collaboration. This study was funded by the EC Biomed 1 (Grant Gene CT93-0060), Swedish Medical Research Council (Grants B94-13XA-10995-02B and B95-13XB-10909-02B), Svenska Läkarsällskapet, Thuring stiftelse, Söderström-Königska stiftelse, funds from the Karolinska Institute and Karolinska Hospital, and a NARSAD Established Investigator Award to Martin Schalling. Patricia Maciel is a recipient of a scholarship from Junta Nacional de Investigacao Cientifica e Tchnologica, Portugal. Guy A. Rouleau is supported by the Medical Research Council of Canada and the Fonds de Recherches en Santé du Québec and the National Institutes of Health.

REFERENCES

1. Lima L, Coutinho P. (1980) Clinical criteria for diagnosis of Machado-Joseph disease: Report of a non-Azorean Portuguese family. *Neurology* **30:** 319–322.
2. Silveira I, et al. (1996) Frequency of spinocerebellar ataxia type 1, dentatorubralpallidoluysian atrophy and Machado-Joseph disease mutations in a large group of spinocerebellar ataxia patients. *Neurology* **4** (1).
3. Takiyama Y, et al. (1993) The gene for Machado-Joseph disease maps to human chromosome 14q. *Nat. Genet.* **4:** 300–304.
4. Stevanin G, et al. (1994) A third locus autosomal dominant cerebellaataxia type I maps to chromosome 14q24.3–qter: Evidence for the existence of a forth locus. *Am. J. Hum. Genet.* **54:** 11–20.
5. Kawaguchi Y, et al. (1994) CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nat. Genet.* **8:** 221–228.
6. Schalling M, Hudson TJ, Buetow KH, Housman DE. (1993) *Nat. Genet.* **4:** 135–139.
7. Lindblad K, Zander C, Schalling M, Hudson T. (1994) Detection of ATG, CCT, CTT and TGG trinucleotide repeat expansions in the human genome. *Nat. Genet.* **7:** 124.
8. Lindblad K, et al. (1995) Association be-

tween genomic CAG repeat sequence size and bipolar affective disorder. *Neurobiol. Dis.* **1:** 55–62.
9. Maciel P, et al. (1995) Correlation between CAG repeat length and clinical features in Machado-Joseph disease. *Am. J. Hum. Genet.* **57:** 54–61.
10. Sambrook J, Fritsch E, Maniatis T. (1989) *Molecular Cloning: A Laboratory Manual.* 2nd Ed. Cold Spring Harbor Press, Cold Spring Harbor, NY.
11. Schalling M. (1992) In: Conn PM (ed). *Gene Expression in Neural Tissues.* Academic Press, New York, pp. 231–255.
12. Willems PJ. (1994) Dynamic mutations hit double figures. *Nat. Genet.* **8:** 213–215.

Contributed by H. Wigzell on October 24, 1995.