Linkage disequilibrium at the SCA2 locus

Olivier Didierjean, Géraldine Cancel, Giovanni Stevanin, Alexandra Dürr, Katrin Bürk, Ali Benomar, Agnès Lezin, Samir Belal, Myriem Abada-Bendid, Thomas Klockgether, Alexis Brice

Abstract

Spinocerebellar ataxia type 2 (SCA2) is caused by the expansion of an unstable CAG repeat encoding a polyglutamine tract. Repeats with 32 to 200 CAGs are associated with the disease, whereas normal chromosomes contain 13 to 33 repeats. We tested 220 families of different geographical origins for the SCA2 mutation. Thirty three were positive (15%). Twenty three families with at least two affected subjects were tested for linkage disequilibium (LD) between the SCA2 mutation and three microsatellite markers, two of which (D12S1332-D12S1333) closely flanked the mutation; the other (D12S1672) was intragenic. Many different haplotypes were observed, indicating the occurrence of several ancestral mutations. However, the same haplotype, not observed in controls, was detected in the German, the Serbian, and some of the French families, suggesting a founder effect or recurrent mutations on an at risk haplotype.

(*J Med Genet* 1999;**36**:415–417)

Keywords: linkage disequilibrium; SCA2; trinucleotide repeat expansion; founder effect

INSERM U289, Hôpital de la Salpêtrière, 75651 Paris Cedex 13, France
O Didierjean
G Cancel
A Dürr
G Stevanin
A Brice

Departement of Neurology, University of Tübingen, D-72076 Tübingen, Germany
K Bürk
T Klockgether

Service de Neurologie (Pr T Chkili), Hôpital des Spécialités, Rabat, Morocco
A Benomar

Laboratoire de Biologie Moléculaire, CTS-Hôpital Perre Zobda-Quitman, Fort-de-France, Martinique
A Lezin

Institut National de Neurologie, La Rabta, 1001 Tunisia
S Belal

Service de Neurologie, Hôpital de Ben-Aknoum, CHU Alger-ouest, Algiers, Algeria
M Abada-Bendid

Correspondence to:
Professor Brice.

Received 22 July 1998
Revised version accepted for publication 22 October 1998

Autosomal dominant cerebellar ataxias (ADCA) are clinically and genetically heterogeneous neurodegenerative disorders. Patients typically present with progressive cerebellar ataxia and additional neurological signs.

Molecular genetic studies have permitted identification of seven distinct ADCA loci[1–3]: spinocerebellar ataxia 1 (SCA1) on chromosome 6p22-p23, SCA2 on 12q24.1, SCA3 or Machado-Joseph disease (SCA3/MJD) on 14q32.1, SCA4 on chromosome 16q22.1, SCA5 on chromosome 11q, SCA6 on chromosome 19p13, and SCA7 on 3p12-p21.1. Five of the seven genes have been identified, and all contain a trinucleotide $(CAG)_n$ repeat expansion in the coding region. The same type of mutation is also responsible for Huntington's disease, spinobulbar muscular atrophy, and dentatorubral-pallidoluysian atrophy (DRPLA).

The SCA2 mutation was found in 33 out of 220 families of different geographical origins with ADCA type I (15%), 32 of which were previously reported.[4] Twenty six unrelated SCA2 families with at least two affected members were selected. Eight were from France, four from Germany, three from Austria, three from other European countries, five from North Africa, and three from the French West Indies (table 1).

Genotypes were determined for three micosatellite markers that span 550 kilobases (Kb) on chromosome 12p24.1 as follows: cen-D12S1332-350 Kb-D12S1672-200 Kb-D12S1333-tel.[5] D12S1332 and D12S1333 flank the SCA2 gene. D12S1672 is in the first intron, close to the CAG repeat.[6] Microsatellite polymorphisms were analysed by PCR with pairs of primers, one of which was 5' end labelled with either FAM, TET, or HEX (PE Applied Biosystems). PCR was carried out as previously described.[7] One µl of each PCR product was then mixed with 3 µl formamide, 0.5 µl loading buffer (dextran blue 50 mg/ml, containing 25 mmol/l of EDTA), and 0.5 µl of Genescan 500-TAMRA (Perkin Elmer), denatured at 94°C for two minutes, then cooled on ice and electrophoresed through a 5% acrylamide gel containing 6 mol/l of urea in 1 × TBE buffer for two hours. Product sizes were determined using GeneScan™ software version 3.0, and alleles were numbered according to the order of product sizes.

Families were separated into two groups for analyses of linkage disequilibrium (LD): northern Europe (Germany, Austria, France, Belgium, and Serbia, n=17) and the whole population (northern Europe, Morocco, Algeria, Tunisia, and Portugal, n=23). Owing to the lack of appropriate controls, the three families from the French West Indies were not used for LD analysis. For calculation of LD, allele frequencies at each locus were determined, for the first group, in spouses and unrelated controls from northern Europe (n=77). For the second group, populations corresponding to the 23 families tested were used (n=104). The difference in allele distribution between normal and carrier chromosomes was evaluated by χ^2 and two tailed Fisher's exact tests. Yates's correction was used when appropriate. The most frequent allele on disease bearing chromosomes was defined as a single allele, and the others were pooled to form a second allele. The presence of LD was tested by $D = \chi^2/N$. A difference of $p<0.01$ was considered to be statistically significant. The proportion of carrier chromosomes bearing the original associated allele was calculated with the equation $\delta=(P_D-P_N)/(1-P_N)$, where P_D and P_N are the frequencies of carrier and normal chromosomes, respectively.[8] Only the ancestral haplotype from each family was used, if it could be determined.

As shown in table 1, 15 disease linked haplotypes (D12S1332-D12S1672-D12S1333) were found. The three French West Indian

Table 1 D12S1332-D121672-D12S1333 haplotype in SCA2 families according to their origin

Family No	Origin	D12S1332*	D12S1672†	D12S1333‡
SAL 221	Austria	6	7	2
SAL 222	Austria	6	7	2
SAL 223	Austria	6	7	2
SAL 312	France	6	7	2
SAL 350	France	6	7	2
SAL 016	France	4	5	8
SAL 315	France	5	5	8
SAL 383	France	5	5	8
SAL 054	Germany	5	5	8
SAL 060	Germany	5	5	8
SAL 061	Germany	5	5	8
SAL 216	Germany	5	5/7	8
SAL 332	Serbia	5	5	8
SAL 365	France	8	11	2
SAL 086	France	4	5/7	9
SAL 174	France	4	5/7	9
SAL 333	Belgium	8	4	9
AAD 147	Portugal	6	4	8
RBT 010	Morocco	6	7/3	2/9
RBT 014	Morocco	4	6	2
RBT 016	Morocco	6	11	2
AAD 002	Algeria	5	7	8
TUN 2	Tunisia	6	5/7	8
JOS	French West Indies	3	11	5
LAR	French West Indies	3	11	5
MAR	French West Indies	5	11	5

When the allele segregating with the disease could not be determined, both are indicated.
*Allele 3=208 bp, 4=206 bp, 5=204 bp, 6=202 bp, 8=198 bp. †Allele 4=285 bp, 5=283 bp, 6=281 bp, 7=279 bp, 11=271 bp. ‡Allele 2=239 bp, 5=233 bp, 8=227 bp, 9=225 bp.

families had the same haplotype for the two markers that flank the mutation (D12S1672-11, D12S1333-5), suggesting a common origin. In a previous study,[9] using more distant flanking markers, only two French West Indian kindreds of the three shared the same haplotype. The four German and the Serbian families shared a common haplotype for the three markers (5-5-8). Another haplotype (6-7-2) was shared by the three Austrian families suggesting the existence of a common but different founder in both countries. In France, four different haplotypes were found in eight families, two of which were the 5-5-8 haplotype found in the German and Serbian families, and the 6-7-2 haplotypes shared with the Austrians. This might reflect common founders in countries that are geographically close.

No common haplotypes were found in North Africans families. The haplotypes in the Algerian, Moroccan, Belgian, Tunisian, and Portuguese kindreds probably resulted from different ancestral mutations. The D12S1672 alleles, segregating with the disease, were different in all but two of these kindreds. Since D12S1672 is within 20 kb of the SCA2 mutation,[10] recombination events separating both loci are very unlikely, as well as mutations in D12S1672 that would generate at least five different alleles. This suggests an independent origin for the SCA2 mutation among these families. However, a very ancient founder effect cannot be excluded. The existence of multiple ancestral mutations is supported by haplotype analysis in smaller series of SCA2 families.[10 11]

We observed significant LD between the SCA2 mutation and the three microsatellite markers. Alleles 5 of D12S1332 and D12S1672 and allele 8 of D12S1333 were significantly more frequent in patients than in controls in the northern European group ($p<0.0017$), as well as in the group as a whole ($p<0.0017$) (table 2). In both groups, the most significant LD was obtained for D12S1672-5 and D12S1333-8, as shown by the *D* values (table 2). This could be explained by the fact that the SCA2 CAG repeat is located between D12S1672 and D12S1333. The δ value for D12S1672 reached 0.48 in the northern European population, whereas the largest δ value in the whole population was for D12S1333 (δ=0.48). The decrease in the δ value for D12S1672 in the whole population (0.34), compared with northern Europeans, suggests that recombination events occurred independently in northern Europeans, and in other populations, especially in North Africans.

Haplotype 6-7-2 which was found on 31% of the disease bearing chromosomes and 9% of controls was not in significant LD with the SCA2 mutation (table 2). Haplotype 5-5-8, absent in the controls (n=50), represented 35% of disease bearing chromosomes ($p<0.001$), confirming that these French, German, and Serbian alleles derived from a common founder (table 2).

Independent ancestral mutations, with founder effects in limited regions, have also been reported for other ADCA loci. At the

Table 2 Linkage disequilibrium between 12q24.1 markers and the SCA2 mutation

Locus	Allele	Patients (frequency)	Controls (frequency)	χ^2	p	D	δ
Northern European							
D12S1332	5	7/17 (0.41)	7/70 (0.10)	9.85	0.0017	0.34	0.34
	6	5/17 (0.29)	40/70 (0.57)	4.21	0.0401	0.22	−0.65
D12S1672	5	7/14 (0.50)	3/77 (0.04)	21.24	<0.0001	0.48	0.48
	7	5/14 (0.36)	54/77 (0.70)	4.73	0.030	0.23	−1.13
D12S1333	2	6/17 (0.35)	9/66 (0.14)	4.28	0.038	0.23	0.24
	8	8/17 (0.47)	3/66 (0.04)	17.71	<0.0001	0.46	0.45
Whole population							
D12S1332	5	8/23 (0.35)	10/104 (0.10)	9.8	0.0017	0.28	0.28
	6	9/23 (0.39)	48/104 (0.46)	0.14	NS	—	—
D12S1672	5	7/18 (0.39)	8/101 (0.08)	13.30	0.0003	0.33	0.34
	7	6/18 (0.33)	66/101 (0.65)	6.55	0.0105	0.23	−0.91
D12S1333	2	8/22 (0.36)	9/96 (0.09)	10.57	0.0011	0.30	0.30
	8	11/22 (0.50)	4/96 (0.04)	29.88	0.0000	0.50	0.48
D12S1332–D12S1672–D12S1333 haplotype							
	6–7–2	5/16 (0.31)	4/46 (0.09)	3.21	NS		
	5–5–8	6/17 (0.35)	0/50 (0.00)	15.29	<0.0001	0.48	0.35

NS = not significant.

SCA1 locus, haplotypes differed according to ethnic origin,[12] and LD with linked markers was reported in Japanese and Italian families.[13 14] For SCA3/MJD, a founder effect or recurrent mutations were found in the French[15] and in the Japanese populations.[16 17] Probably, several founders accounted for the SCA6 mutations.[18] Finally, LD exists between closely linked markers and the SCA1, SCA3, and SCA6 mutations in several countries, but several distinct ancestral mutations were found in different ethnic groups. SCA2 linked haplotypes might represent at risk chromosomes for CAG expansion, as suggested for SCA3,[15] but the de novo cases, which are needed to confirm this hypothesis, have not yet been described.

This study was supported financially by the Association Française contre les Myopathies (AFM), the Verum Foundation, the Association pour le Développement de la Recherche sur les Maladies Génétiques Neurologiques et Psychiatriques, and the Biomed. GS held a fellowship from RETINA-France. We are grateful to Dr Merle Ruberg for critical reading of the manuscript, Christiane Penet, Yolaine Pothin, Agnès Camuzat, and Jacky Bou for their technical help and Drs Fayçal Hentati, Christophe Vial, João Guimarães, and José Vale Santos for clinical evaluation.

1 Dürr A, Brice A. Genetics of movement disorders. *Curr Opin Neurol* 1996;**9**:290-7.
2 Koshy BT, Zoghbi HY. The CAG/polyglutamine tract diseases: gene products and molecular pathogenesis. *Brain Pathol* 1997;7:927-42.
3 David G, Abbas N, Stevanin G, *et al*. Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nat Genet* 1997;**17**:65-70.
4 Cancel G, Dürr A, Didierjean O, *et al*. Molecular and clinical correlations in spinocerebellar ataxia 2: a study of 32 families. *Hum Mol Genet* 1997;**6**:709-15.
5 Pulst SM, Nechiporuk A, Nechiporuk T, *et al.* Moderate expansion of normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nat Genet* 1996;**14**:269-76.
6 Sahba S, Nechiporuk A, Figueroa KP, *et al.* Genomic structure of the human gene for spinocerebellar ataxia type 2 (SCA2). *Genomics* 1998;**47**:359-64.
7 Stevanin G, Le Guern E, Ravise N, *et al*. A third locus for autosomal dominant cerebellar ataxia type I maps to chromosome 14q24.3-qter: evidence for the existence of a fourth locus. *Am J Hum Genet* 1994;**54**:11-20.
8 Risch N, de Leon D, Ozelius L, *et al*. Genetic analysis of idiopathic torsion dystonia in Ashkenazi Jews and their recent descent from a small founder population. *Nat Genet* 1995;**9**:152-9.
9 Lezin A, Cancel G, Stevanin G, *et al.* Autosomal dominant cerebellar ataxia I in Martinique (French West Indies): genetic analysis of three unrelated SCA2 families. *Hum Genet* 1996;**97**:671-6.
10 Geschwind DH, Perlman S, Figueroa CP, Treiman LJ, Pulst SM. The prevalence and wide clinical spectrum of the spinocerebellar ataxia type 2 trinucleotide repeat in patients with autosomal dominant cerebellar ataxia. *Am J Hum Genet* 1997;**60**:842-50.
11 Giunti P, Sabbadini G, Sweeney MG, *et al*. The role of the SCA2 trinucleotide repeat expansion in 89 autosomal dominant cerebellar ataxia families - frequency, clinical and genetic correlates. *Brain* 1998;**121**:459-67.
12 Kwiatkowski TJ Jr, Orr HT, Banfi S, *et al*. 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* 1993;**53**:391-400.
13 Wakisaka A, Sasaki H, Takada A, *et al*. Spinocerebellar ataxia 1 (SCA1) in the Japanese in Hokkaido may derive from a single common ancestry. *J Med Genet* 1995;**32**:590-2.
14 Frontali M, Sabbadini G, Novelletto A, *et al*. Genetic fitness in Huntington's disease and spinocerebellar ataxia 1: a population genetics model for CAG repeat expansions. *Ann Hum Genet* 1996;**60**:423-35.
15 Stevanin G, Lebre AS, Mathieux C, *et al*. Linkage disequilibrium between the spinocerebellar ataxia 3/Machado-Joseph disease mutation and two intragenic polymorphisms, one of which, X359Y, affects the stop codon. *Am J Hum Genet* 1997;**60**:1548-52.
16 Takiyama Y, Igarashi S, Rogaeva EA, *et al.* Evidence for inter-generational instability in the CAG repeat in the *MJD1* gene and for conserved haplotypes at flanking markers amongst Japanese and Caucasian subjects with Machado-Joseph disease. *Hum Mol Genet* 1995;**4**:1137-46.
17 Endo K, Sasaki H, Wakisaka A, *et al.* Strong linkage disequilibrium and haplotype analysis in Japanese pedigrees with Machado-Joseph disease. *Am J Med Genet* 1996;**67**:437-44.
18 Stevanin G, Dürr A, David G, *et al*. Clinical and molecular features of spinocerebellar ataxia type 6. *Neurology* 1997;**49**:1243-6.