Am. J. Hum. Genet. 64:768–775, 1999

Location Score and Haplotype Analyses of the Locus for Autosomal Recessive Spastic Ataxia of Charlevoix-Saguenay, in Chromosome Region 13q11

Andrea Richter,[1] John D. Rioux,[4] Jean-Pierre Bouchard,[5] Jocelyne Mercier,[1] Jean Mathieu,[6] Bing Ge,[2] Josée Poirier,[1] Dominique Julien,[1] Gabor Gyapay,[7] Jean Weissenbach,[7] Thomas J. Hudson,[2,3,4] Serge B. Melançon,[1] Kenneth Morgan[2,3]

[1]Service de Génétique Médicale, Centre de Recherche, Hôpital Sainte-Justine, Université de Montréal, [2]Montreal General Hospital Research Institute, and [3]Departments of Human Genetics and Medicine, McGill University, Montreal; [4]Whitehead Institute for Biomedical Research, MIT, Cambridge, MA; [5]Service de Neurologie, Hôpital de l'Enfant-Jésus, Québec; [6]Service de Neurologie, Hôpital de Chicoutimi, Chicoutimi, Canada; and [7]Généthon, CNRS URA 1922, Évry, France

Summary

Autosomal recessive spastic ataxia of Charlevoix-Saguenay (ARSACS) is a clinically homogeneous form of early-onset familial spastic ataxia with prominent myelinated retinal nerve fibers. More than 300 patients have been identified, and most of their families originated in the Charlevoix-Saguenay region of northeastern Quebec, where the carrier prevalence has been estimated to be 1/22. Consistent with the hypothesis of a founder effect, we observed excess shared homozygosity at 13q11, among patients in a genomewide scan of 12 families. Analysis of 19 pedigrees demonstrated very tight linkage between the ARSACS locus and an intragenic polymorphism of the γ-sarcoglycan (*SGCG*) gene, but genomic DNA sequence analysis of all eight exons of *SGCG* revealed no disease-causing mutation. On the basis of haplotypes composed of seven marker loci that spanned 11.1 cM, the most likely position of the ARSACS locus was 0.42 cM distal to the *SGCG* polymorphism. Two groups of ARSACS-associated haplotypes were identified: a large group that carries a common *SGCG* allele and a small group that carries a rare *SGCG* allele. The haplotype groups do not appear to be closely related. Therefore, although chromosomes within each haplotype group may harbor a single ARSACS mutation identical by descent, the two mutations could have independent origins.

Received July 7, 1997; accepted for publication December 23, 1998; electronically published February 12, 1999.

Addresses for correspondence and reprints: Dr. Andrea Richter, Service de Génétique Médicale, Hôpital Sainte-Justine, 3175 chemin de la Côte Sainte-Catherine, Montréal, Québec H3T 1C5, Canada. E-mail: richtera@ere.umontreal.ca; and Dr. Kenneth Morgan, Division of Medical Genetics, Montreal General Hospital, 1650 Cedar Avenue, Montreal, Quebec H3G 1A4, Canada. E-mail: ken@population.genetics.mcgill.ca

0002-9297/99/6403-0014$02.00

Introduction

Patients with autosomal recessive spastic ataxia of Charlevoix-Saguenay (ARSACS [MIM 270550]) present in early childhood with spastic gait ataxia. The disease progresses rapidly in young adults, and patients are wheelchair bound by their 5th decade. Hypermyelinated retinal nerve fibers are a unique diagnostic feature (Bouchard 1991). Sensory-nerve conduction is abolished, and motor-nerve velocity is reduced (Bouchard et al. 1979). Magnetic-resonance images show cerebellar vermis atrophy, and autopsy studies reveal both the absence of Purkinje cells (Bouchard 1991) and the presence of swollen neurons with dense, lipofuscin-like granules (Richter et al. 1996). The hypermyelinated retinal fibers and the hypoplasia of the superior cerebellar vermis point to developmental defects in myelination of the central and peripheral nervous systems in ARSACS, whereas the progressive axonal degeneration in the corticospinal and spinocerebellar tracts accounts for the peripheral clinical signs (Harding 1993).

Families with ARSACS originate from the population, of French descent, of Charlevoix County and the Saguenay–Lac St. Jean region of northeastern Quebec (hereafter we will refer to Saguenay-Lac St. Jean as "Saguenay"). In the Saguenay population, for the period 1941–85, the incidence of ARSACS has been estimated to be 1/1,932 births, and the carrier prevalence has been estimated to be 1/22; the relatively high incidence has been attributed to founder effect (De Braekeleer et al. 1993). There are relatively high incidences of several Mendelian diseases in the Charlevoix-Saguenay population (see Bouchard and De Braekeleer 1991).

Recently, we reported that the ARSACS locus maps to chromosome region 13q11 and that the candidate region appears to be 5.5 cM, by inspection of haplotypes in recombinant individuals (Bouchard et al. 1998). Subsequently, we constructed a genetic map of the region,

on the basis of combined linkage and physical mapping data from both this study and public databases; more precisely mapped the ARSACS locus by location-score analysis; and constructed haplotypes to significantly narrow the candidate-gene region. We report here the details of the genetic mapping of the ARSACS locus.

Families and Methods

Families

Over a 20-year period, >200 ARSACS families including 320 patients living in the Charlevoix-Saguenay region were identified (Bouchard et al. 1978, 1998; Bouchard 1991; Richter et al. 1993). The 1997 population sizes are 30,659 and 296,537 for Charlevoix County and Saguenay-Lac St. Jean, respectively (Quebec's Bureau of Statistics). The patients were identified and followed in the neuromuscular-disease clinics in Quebec City (J.-P.B.) and Chicoutimi (J.M.). After informed consent had been obtained, to comply with institutional and national standards (Knoppers and Laberge 1991), we collected blood, for DNA extraction and to establish lymphoblastoid cell lines, from 84 affected individuals and 316 other relatives from 25 extended families. Pedigree information for 12 families has been published elsewhere (Richter et al. 1993).

DNA Analysis and Genomewide Scans

DNA was prepared as described elsewhere (Richter et al. 1993). Eight families (31 patients, 108 individuals) were studied by means of 121 highly polymorphic microsatellite markers (distributed approximately every 25–30 cM), by the method of rapid genotyping, at Généthon (Gyapay et al. 1994). No evidence for linkage was obtained. A semiautomated, second genomewide scan was done on 29 patients and 18 parents from 12 families, by a modified version of the Cooperative Human Linkage Center (CHLC) Human Screening Set (version 6.0) of polymorphic markers (Dubovsky et al. 1995). This screening set comprised 298 fluorescently labeled microsatellite markers (Research Genetics) with an average heterozygosity of .75 and with an average spacing of 12 cM between markers. Multiplexed PCR products were analyzed on an ABI 377 DNA sequencer (PE Applied Biosystems). Gels were visually inspected for shared homozygosity among patients, and an increase was noted for the tetranucleotide repeat CHLC.GATA23C03 (*D13S787*) located in chromsome region 13q11 (Genome Database; Cayanis et al. 1998).

Genotyping of Microsatellite Loci in Chromosome Region 13q11

Nine microsatellite loci (*D13S175, D13S1236, D13S1275, D13S232, D13S292, D13S787, D13S1243, D13S283,* and *D13S1285*) were typed in 243 individuals, including 76 carrier parents and 69 affected offspring, from 19 pedigrees. Nine of the pedigrees were nuclear families, and the remainder were extended families. The marker alleles were amplified by PCR, with one of the primers end labeled with [^{32}P]-γATP, under standard conditions. The PCR products were separated by electrophoresis on acrylamide-urea sequencing gels and were visualized by autoradiography. Allele sizes were determined by means of a sequencing ladder.

Exon Sequencing of the γ-Sarcoglycan (SGCG) *Gene*

Eight DNA samples from obligate carriers were used to screen for variation in the exons of the *SGCG* gene, by comparative sequencing. Specifically, each exon was amplified from the individual genomic DNA samples, by means of primers designed by McNally et al. (1996). The PCR products were purified by the solid-phase reversible-immobilization procedure (Hawkins et al. 1994), and both strands were sequenced by means of FS dye primer cycle sequencing kits (PE Applied Biosystems) and then were run on ABI 377 sequencers. The exonic sequences were aligned by the Staden (GAP4) package (Staden 1996). We identified two single-nucleotide substitutions, C→T transitions, in exon 8, at nucleotide positions 38 and 222 of GenBank accession number U63395 (McNally et al. 1996). They define single-nucleotide polymorphisms (SNPs) that have been reported by McNally et al. (1996). The first SNP is a silent substitution in a leucine residue (codon 235), and the second is located in the 3′ UTR, 13 nucleotides beyond the stop codon. At least one carrier parent and one ARSACS offspring from each of 19 nuclear families were genotyped for both SNPs, by the method described above.

Genetic Map of Microsatellite Loci in Chromosome Region 13q11

Genotypes of nine marker loci for eight Centre d'Etude du Polymorphisme Humain (CEPH) Genotype Database families were obtained from two sources. Data for *D13S175, D13S1236, D13S1275, D13S292, D13S1243, D13S283,* and *D13S1285,* were retrieved from the CEPH Genotype Database (version 8.1), and data for *D13S232* and *D13S787* were retrieved from the Center for Medical Genetics, Marshfield Medical Research Foundation genotype database (Broman et al. 1998). CEPH Genotype Database family data for *D13S787* genotypes that were uncertain (J. Weber, personal communication) were omitted from the analysis. In the ARSACS families, *D13S175* and *D13S1243* were not typed in one nuclear family and in one large, extended pedigree.

The order of the nine marker loci was determined on

the basis of various linkage, physical, and integrated maps (Chumakov et al. 1995; Guilford et al. 1995; Hudson et al. 1995; Shaw et al. 1995; Collins et al. 1996; Dib et al. 1996, and supplementary material in an extended reprint; Broman et al. 1998; Cayanis et al. 1998; authors' unpublished data). The physical-map order was given precedence particularly where the local order of markers was not reliably established on the linkage maps. The CEPH Genotype Database and ARSACS pedigrees were analyzed for linkage between adjacent marker loci. The LOD scores for the families were combined to obtain maximum-likelihood estimates of recombination fraction (θ) values.

Two-Point Linkage Analysis

LOD scores for data from genomewide scans were computed by LINKAGE programs (version 5.1) (Lathrop et al. 1984, 1985). LOD scores and θ values both between adjacent pairs of the nine marker loci and between each marker and the ARSACS locus were computed by FASTLINK versions (4.0P) (Cottingham et al. 1993; Schäffer et al. 1994; Schäffer 1996) of the LINKAGE programs. Marker-allele frequencies in the 19 ARSACS pedigrees were estimated by FASTLINK/ILINK, and linkage equilibrium was assumed (Terwilliger and Ott 1994). The ARSACS disease-allele frequency was estimated to be .023, on the basis of a previous genetic epidemiological study (De Braekeleer et al. 1993), and complete penetrance of the autosomal recessive trait was assumed.

Location Scores and Haplotypes of the ARSACS Candidate-Gene Region

Location-score analysis (Lathrop et al. 1984) and haplotyping were performed by Markov-chain Monte Carlo methods (Sobel and Lange 1996; Lange 1997, by the SIMWALK2 program (versions 2.01 and 2.27) (Weeks et al. 1995; E. Sobel, personal communication), with computation of the multipoint LOD scores by the MENDEL package (version 3.35) (Lange et al. 1988; Lange 1997). In location-score analysis, the most likely initial genetic-descent graph (also known as the "inheritance vector") of a pedigree was found by simulated annealing in three parallel runs. Two thousand sampled pedigrees for each original pedigree were obtained by random walk from the initial position. These representative pedigrees are completely typed and proportional to their true likelihood. The MENDEL program was used to compute multipoint LOD scores for the sample of representative pedigrees. In haplotyping, the best marker-haplotype vector for each pedigree was found by simulated annealing in three parallel runs. Two analyses were conducted for location scores and haplotypes; both analyses used the genetic map for the combined CEPH Genotype Database and ARSACS family data, for seven of the marker loci (*D13S1236, D13S232, D13S1275, D13S292, D13S787, D13S283,* and *D13S1285*). For the first analysis, marker-allele frequencies were estimated in the 19 ARSACS pedigrees by maximum likelihood, with linkage equilibrium being assumed, and these frequencies were also used in the two-point linkage analyses. For the second analysis, marker-allele frequencies were estimated by weighting the conditional allele frequencies in ARSACS and normal haplotypes of 68 carrier parents by the ARSACS and normal allele frequencies (.023 and .977, respectively) in the population. The marker haplotypes were constructed by the first SIMWALK2 analysis and were identified as either ARSACS associated or normal, on the basis of cosegregation with the deduced ARSACS allele, in each of the pedigrees. The haplotypes of the 68 obligate carrier parents obtained in the second SIMWALK2 analysis were used for further analysis. Alternative estimates of the conditional frequencies of haplotypes were based on 55 disease-associated and 58 normal chromosomes, obtained by not counting copies that were considered to be identical by descent within a pedigree. The conditional frequencies of ARSACS-associated and normal haplotypes may be representative of the population, although the families were not selected to constitute a population-based sample.

Results

In the first genomewide scan of eight multiplex ARSACS families, no evidence for linkage was obtained. A LOD score greater than -2 at $\theta = .01$–$.2$ was obtained for 121 microsatellite loci (data not shown). The second genomewide scan of parents and patients from 12 nuclear families was undertaken with a different set of microsatellite markers. Shared homozygosity for *D13S787* allele 4 (259 bp) was observed in all the patients except for those in two nuclear families in which the patients were heterozygous 3/4 (255 and 259 bp, respectively) and 4/5 (259 and 262 bp, respectively). There was significant evidence for linkage to *D13S787* (LOD score 3.66 at $\theta = 0$), a result that mapped the ARSACS locus to chromosome region 13q11. In the first genomewide scan, chromosome 13q was sparsely covered by three markers>16 cM distal to *D13S787*.

To better localize the ARSACS locus, nine microsatellite loci were chosen from the Généthon and Marshfield linkage maps, and 243 individuals from 19 ARSACS pedigrees were genotyped. No recombinants were observed, by two-point linkage analysis, between ARSACS and either *D13S232* (maximum LOD score 24.1 at maximum θ [$\hat{\theta}$] 0) or *D13S787* (maximum LOD score 24.7 at $\hat{\theta} = .005$) (table 1).

We attempted to obtain a better estimate of the genetic

Table 1

Two-Point Linkage between ARSACS and Marker Loci in Chromosome Region 13q11

Marker Locus	LOD Score at θ = 0	.01	.03	.05	.07	.09	.11	Maximum LOD Score	$\hat{\theta}$
D13S175	−∞	4.99	8.75	9.94	10.34	10.36	10.14	10.39	.081
D13S1236	−∞	17.97	20.15	20.33	19.89	19.16	18.24	20.36	.043
D13S1275	−∞	28.19	27.47	26.25	24.87	23.41	21.92	28.19	.010
D13S232	24.13	23.54	22.35	21.15	19.94	18.73	17.51	24.13	.000
D13S292	−∞	19.55	18.95	18.08	17.13	16.14	15.13	19.56	.008
D13S787	24.39	24.61	23.75	22.66	21.49	20.27	19.03	24.69	.005
D13S1243	−∞	17.96	19.53	19.47	18.91	18.10	17.15	19.60	.038
D13S283	−∞	12.46	12.94	12.73	12.29	11.74	11.12	12.94	.029
D13S1285	−∞	12.71	15.86	16.54	16.47	16.03	15.37	16.57	.057

map of this region, for multipoint-linkage and haplotype analyses. First, the order of published physical maps was validated by additional physical mapping of YACs and bacterial artificial chromosomes in our laboratory (data not shown). Second, we estimated θ values between adjacent marker loci by combining the results of linkage analysis of CEPH Genotype Database family data from public genotype databases and the results of linkage analysis of the ARSACS families. The estimated genetic length of the region spanned by the nine loci, in the combined family data, was 19.3 cM (table 2). Between *D13S1243* and the flanking loci—*D13S787* and *D13S283*—there was apparently greater recombination in the ARSACS families than in the CEPH Genotype Database families, although the 1-LOD-unit support intervals for the two sets of family data overlapped (data not shown). Since *D13S787, D13S1243,* and *D13S283* are located on a 1.71-Mb YAC (906_c_1) (CEPH-Généthon Integrated Map; Chumakov et al. 1995), the length of our map of the nine marker loci may be overestimated. We found it difficult to score *D13S1243* genotypes, although we could determine them in nuclear families such that there were no Mendelian segregation errors. Therefore, for multipoint mapping and haplotype analysis of the candidate region of the ARSACS gene, we chose a subset of seven loci (*D13S1236, D13S1275, D13S232, D13S292, D13S787, D13S283,* and *D13S1285*) that did not include *D13S175* (the most proximal locus) and *D13S1243*. We estimated $\hat{\theta}$ between *D13S787* and *D13S283* to be .033 (maximum LOD score 38.5). In addition, no recombination (maximum LOD score 45.8) was observed between *D13S292* and *D13S787,* but both loci are located on a 600-kb YAC (768_a_4) (CEPH-Généthon Integrated Map). We arbitrarily assigned θ = .001 to this pair of loci, for location-score and haplotype analyses. Finally, the estimated genetic length of the region spanned by these seven marker loci is 11.1 cM, under the assumption that there is no interference (fig. 1).

Location scores were computed for the ARSACS locus by the SIMWALK2 and MENDEL programs, under the assumption of linkage equilibrium for two sets of marker-allele frequencies (fig. 1). The best haplotype vectors for the pedigrees were estimated by SIMWALK2

Table 2

Genetic Maps of Marker Loci in Chromosome Region 13q11 in CEPH Genotype Database and ARSACS Families and Combined Family Data

Marker Locus	Genetic Location Database Composite Location (Mb)	$\hat{\theta}$ (Maximum LOD Score) in[a] CEPH Genotype Database Families	ARSACS Families	All Families	Map Distance, for All Families (cM)
D13S175	16.201	.037 (12.24)	.066 (26.49)	.060 (38.57)	0
D13S1236	16.380	.051 (23.23)	.028 (44.72)	.036 (67.67)	6.4
D13S1275	18.570	.000 (30.67)	.005 (45.83)	.003 (76.33)	10.1
D13S232	19.141	.022 (23.12)	.007 (27.62)	.013 (50.55)	10.4
D13S292	19.250	.000 (16.30)	.000 (29.50)	.000 (45.80)	11.7
D13S787	19.370	.017 (14.96)	.039 (32.36)	.034 (47.17)	11.8
D13S1243	19.490	.000 (26.19)	.029 (25.72)	.017 (51.01)	15.4
D13S283	20.490	.018 (29.64)	.026 (26.99)	.022 (56.59)	17.1
D13S1285	21.080				19.3

Note.—For location-score and haplotype analyses, θ = .001 was assigned between *D13S292* and *D13S787,* and θ = .033 was estimated for *D13S787* and *D13S283* (see text).

[a] The estimates are for adjacent loci and pertain to this locus and that on the line below.

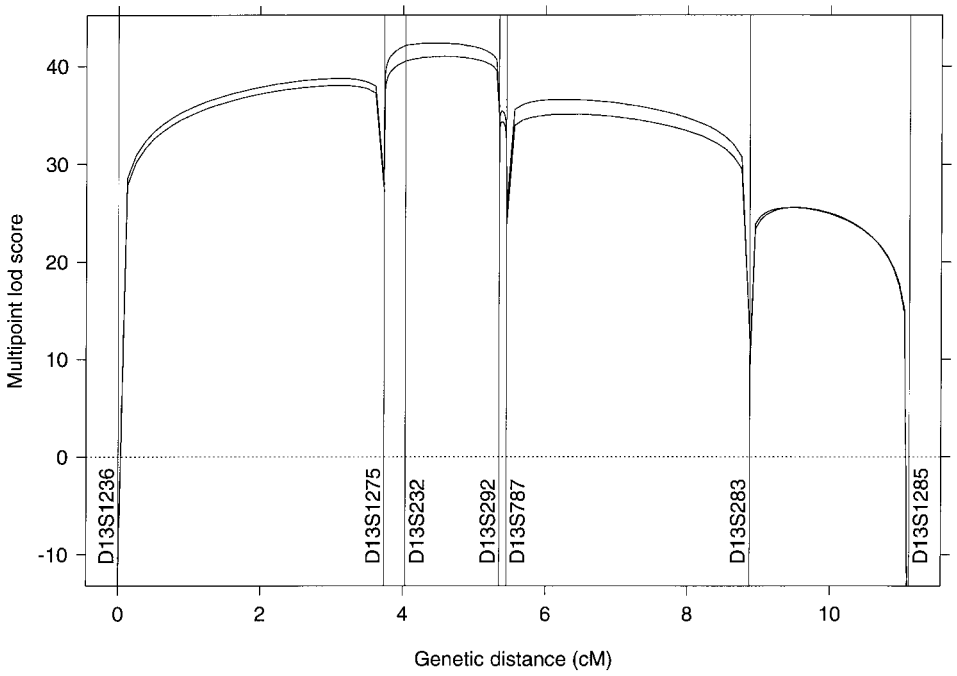

Figure 1 Location-score analysis of ARSACS locus. The SIMWALK2 program was used to generate 2,000 sampled pedigrees for each original pedigree, by Markov-chain Monte Carlo methods. These representative pedigrees are completely typed and proportional to their true likelihood. The MENDEL program was used in conjunction with SIMWALK2, to compute the multipoint LOD scores (E. Sobel, personal communication). The run with the larger LOD scores used marker-allele frequencies that were weighted estimates from ARSACS-associated and normal haplotypes (see Methods); the run with the smaller LOD scores used sample estimates of marker-allele frequencies.

(see below). The linkage analyses will tend to underestimate the LOD scores, because linkage disequilibrium was not taken into account (Clerget-Darpoux 1982). As in the two-point linkage analysis (table 1), there was no evidence of recombination between the ARSACS locus and *D13S232*. Maximum multipoint LOD scores were 41.0 at positions 4.56 and 4.61 cM distal to *D13S1236*, in the first analysis, and 42.3 at positions 4.43 and 4.48 cM, in the second analysis (fig. 1). On the results of the latter location-score analysis, the most likely position of the ARSACS locus would be 0.42 cM distal to *D13S232*. The 2-LOD-unit support interval for the map position of the ARSACS locus extends distal to *D13S1275* and proximal to *D13S292*. Thus, the likely candidate region for the ARSACS gene has been narrowed 1.58 cM.

We used SIMWALK2 to estimate the seven-marker haplotypes in the 19 pedigrees. There are two distinct groups of ARSACS-associated haplotypes: a large group consisting of 22 haplotypes and a small group consisting of 2 haplotypes (table 3). In the large group, 14 of 53 chromosomes were inferred to have the same seven-marker haplotype as a major ancestral haplotype (but they are not necessarily identical by descent for the entire 11.1-cM region). A core haplotype of the 1.6-cM central map region (spanned by *D13S1275, D13S232, D13S292,* and *D13S787*) that harbors the ARSACS locus was shared by 41 of 53 chromosomes in the large haplotype group. Only 3 of 58 normal chromosomes carried a core haplotype (2-5-1-3, 2-5-1-4, and 3-5-1-4) that was present in the ARSACS chromosomes.

On the basis of the genotyping of two SNPs in exon 8 of the *SGCG* gene in a subset of individuals from 19 nuclear families, all ARSACS-associated haplotypes carrying *D13S232*5* carried *SGCG* exon 8 haplotype C-C (table 4). Furthermore, the rare ARSACS-associated haplotypes that carried *D13S232*2* (table 3) had *SGCG* exon 8 haplotype T-C. *SGCG* exon 8 haplotype C-C was the most frequent haplotype in both ARSACS chromosomes and normal chromosomes, whereas haplotype C-T was not observed in the sample of 19 nuclear families, and no particular exon 8 haplotype was absolutely associated with *D13S232* alleles in normal chromosomes (table 4).

Discussion

The Quebec population of French descent has its origins in the immigration of 8,527 settlers during 1608–1759 (Charbonneau and Robert 1987). The settlement of the Charlevoix region began in 1675 (Jetté et al. 1991). During 1838–1911, 13,800 individuals immigrated from Charlevoix to Saguenay (Gauvreau et al.

Table 3

ARSACS-Associated Haplotypes in 55 Parents of ARSACS Patients

Allele at							
D13S1236 (.0 cM)	*D13S1275* (3.74 cM)	*D13S232* (4.04 cM)	*D13S292* (5.35 cM)	*D13S787* (5.45 cM)	*D13S283* (8.87 cM)	*D13S1285* (11.12 cM)	No. of Parents
1	2	5	1	3	7	8	1
1	2	5	1	4	2	4	1
1	2	5	1	4	6	6	1
1	2	5	1	4	7	7	1
1	2	5	1	4	7	4	3
1	2	5	1	4	7	6	2
1	2	5	1	4	7	8	14
4	2	5	1	4	7	8	8
4	2	5	1	4	7	6	2
2	2	5	1	4	7	8	2
2	2	5	1	4	7	6	1
3	2	5	1	4	7	8	5
5	2	5	1	4	7	8	1
4	2	5	2	4	7	6	2
3	2	5	2	4	7	6	1
4	3	5	1	4	7	6	1
5	3	5	1	4	2	4	1
1	4	5	1	4	7	8	1
3	4	5	1	4	7	8	1
9	6	5	1	4	7	8	2
4	7	5	1	4	7	8	1
2	7	5	1	4	7	6	1
2	3	2	1	5	10	3	1
3	3	2	1	5	1	6	1

Haplotypes for 55 ARSACS and 58 normal chromosomes were obtained from 68 obligate carrier parents by not counting copies that were considered to be identical by descent within a pedigree. Haplotypes were inferred for seven parents who were not sampled. Among the 58 normal chromosomes there was one each of the *D13S1275-D13S232-D13S292-D13S787* core haplotypes, 2-5-1-4, 3-5-1-4, and 2-5-1-3; the remainder were different than the ARSACS-associated core haplotypes. The ARSACS locus was estimated by location-score analysis to be at 4.45 cM with a 2-LOD-unit support interval from 3.78 to 5.35 cM.

1991); therefore, Charlevoix made the major contribution to the founding of the Saguenay population (Bouchard et al. 1988). The genetic epidemiology of ARSACS indicates that a founder effect occurred in the Charlevoix population (De Braekeleer et al. 1993).

We computed location scores (fig. 1) and constructed haplotypes (table 3) for an 11.1-cM region spanned by seven microsatellite loci in chromosome region 13q11, using the SIMWALK2 and MENDEL programs. The ARSACS locus was estimated to be 0.42 cM distal to *D13S232* (multipoint LOD score 42.3; 2-LOD-unit support interval 1.58 cM). No recombination was observed between the ARSACS locus and *D13S232,* and, since *D13S232* is an intragenic polymorphism of *SGCG,* we sequenced all eight exons of the latter gene; but we did not find a disease-causing mutation. The overwhelming majority of ARSACS chromosomes (53 of 55 chromosomes from parents of patients) carried allele *D13S232*5, SGCG*-exon 8 C-C (based on a sample of 19 nuclear families), and both of the remaining chromosomes were observed to carry the rare allele *D13S232*2, SGCG*-exon 8 T-C. Thus, two groups of ARSACS-associated haplotypes are minimally defined by different *SGCG* alleles: a large group consisting of 22 haplotypes and a small group consisting of 2 haplotypes (table 3). In the large haplotype group, an inferred, major ancestral haplotype, 1-2-5-1-4-7-8, was present on 14 (26%) of 53 ARSACS chromosomes.

The haplotypes in the large group that contain alleles *D13S292*2* and *D13S787*3* are not obviously recombinant haplotypes. These alleles may have originated by mutation on the major ancestral chromosome. Both alleles are smaller, by one repeat, than the ARSACS-associated allele at the corresponding loci. Gain or loss of a single repeat unit is the predominant mutational event, as a consequence of replication slippage, of short tandem-repeat sequences (Weber and Wong 1993; Brinkmann et al. 1998). In any event, these haplotypes do not further narrow the ARSACS candidate-gene region. The two chromosomes in the small group of ARSACS-associated haplotypes appear to be related, since they share a core haplotype (*D13S1275, D13S232, D13S292,* and *D13S787*) that is distinctly different than the haplotypes of the large group. However, the most recent common

Table 4

***SGCG* Alleles Defined by *D13S232* Allele and Exon 8 SNP Haplotype**

	Exon 8 SNP Haplotype (No. of Chromosomes)	
D13S232 Allele	ARSACS[a]	Normal
1	…	C-C (5)
		T-C (1)
		T-T (1)
2	T-C (2)	C-C (1)
4	…	C-C (1)
		T-C (1)
5	C-C (34)	C-C (10)
		T-C (2)
6	…	C-C (3)
		T-C (3)
		T-T (2)
7	…	C-C (4)
		T-C (2)

[a] An ellipsis (…) denotes that the allele was not observed.

ancestral haplotype of the small group is probably not very recent, since the two haplotypes differ in both flanking regions, most likely as a consequence of recombination. Our hypothesis is that the overwhelming majority (96%) of ARSACS chromosomes that carry allele *D13S232*5, SGCG*-exon 8 C-C, also harbor the same disease-associated mutation identical by descent. Furthermore, the remaining (4%) ARSACS chromosomes that carry rare allele *D13S232*2, SGCG*-exon 8 T-C, may harbor another mutation of independent origin.

The ARSACS gene region (defined by the 2-LOD-unit support interval in the location-score analysis) lies distal to *D13S1275* and proximal to *D13S292* and spans 1.58 cM (fig. 1). The Genetic Location Database composite map distance of this interval is 680 kb (table 2). We estimated the ARSACS locus to be 0.42 cM distal to *D13S232* and have constructed a BAC contig that includes *SGCG, D13S232, D13S292,* and *D13S787* and is ⩽675 kb (J. Engert, T. J. Hudson, and A. Richter, unpublished data). Excluding the BAC that is positive for *D13S787* reduces the maximum physical size of this genomic segment to 525 kb, for a search of the ARSACS gene.

Acknowledgments

We wish to acknowledge the contributions of France Gosselin and Claude Prevost, for sample collection and pedigree information; Christina Vestegaard and Jessica Zuckerman, for technical assistance at the Whitehead Institute; Mary Fujiwara and Celia Greenwood, for suggestions on the manuscript; and Eric Sobel, both for advice on analysis and assistance in the use of the SIMWALK2 program and for kindly providing version 2.27. Our research was supported by grants from the Muscular Dystrophy Association of Canada (to A.R.); Association Canadienne de l'Ataxie de Friedreich-Fonds de la Recherche en Santé du Québec, programme conjoint (to A.R.); Réseau de Génétique Médicale Appliquee-Fonds de la Recherche en Santé du Québec (to S.B.M.); the Canadian Genetic Diseases Network (Federal Networks of Centers of Excellence Program) (to K.M.); the Canadian Genetic Analysis and Technology Program (to T.J.H. and K.M.); and the Medical Research Council of Canada (to T.J.H. and K.M.). T.J.H. is a recipient of a Clinician Scientist award from the Medical Research Council of Canada.

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

Center for Medical Genetics, Marshfield Medical Research Foundation, http://www.marshmed.org/genetics
CEPH Genotype Database, http://www.cephb.fr/cephdb
CEPH-Généthon Integrated Map, http://www.cephb.fr/ceph-genethon-map.html
Genetic Location Database, http://cedar.genetics.soton.ac.uk/public_html
Genome Database, http://www.gdb.org
Online Mendelian Inheritance in Man, http://www.ncbi.nlm.nih.gov/Omim (for ARSACS [MIM 270550])
Quebec's Bureau of Statistics, http://www.bsq.gouv.qc.ca

References

Bouchard G, De Braekeleer M (eds) (1991) Histoire d'un génome: population et génétique dans l'est du Québec. Presses de l'Université du Québec, Sillery

Bouchard G, Laberge C, Scriver CR (1988) Reproduction démographique et transmission génétique dans le nord-est de la province de Québec (18e–20e siecles). Eur J Popul 4: 39–67

Bouchard J-P (1991) Recessive spastic ataxia of Charlevoix-Saguenay. In: de Jong JMBV (ed) Handbook of clinical neurology. Vol 16: Hereditary neuropathies and spinocerebellar degenerations, Elsevier Science, Amsterdam, pp 451–459

Bouchard J-P, Barbeau A, Bouchard R, Bouchard RW (1978) Autosomal recessive spastic ataxia of Charlevoix-Saguenay. Can J Neurol Sci 5:61–69

Bouchard J-P, Barbeau A, Bouchard R, Bouchard RW (1979) Electomyography and nerve conduction studies in Friedreich's ataxia and autosomal recessive spastic ataxia of Charlevoix-Saguenay. Can J Neurol Sci 6:185–189

Bouchard J-P, Richter A, Mathieu J, Brunet D, Hudson TJ, Morgan K, Melançon SB (1998) Autosomal recessive spastic ataxia of Charlevoix-Saguenay. Neuromuscul Disord 8: 474–479

Brinkmann B, Klintschar M, Neuhuber F, Hühne J, Rolf B (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. Am J Hum Genet 62:1408–1415

Broman KW, Murray JC, Sheffield VC, White RL, Weber JL

(1998) Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am J Hum Genet 63:861–869

Cayanis E, Russo JJ, Kalachikov S, Ye X, Park SH, Sunjevaric I, Bonaldo MF, et al (1998) High-resolution YAC-cosmid-STS map of human chromosome 13. Genomics 47:26–43

Charbonneau H, Robert N (1987) The French origins of the Canadian population 1608–1759. In: Harris RC (ed) Historical atlas of Canada. Vol 1: From the beginning to 1800. University of Toronto Press, Toronto, plate 45

Chumakov IM, Rigault P, Le Gall I, Bellanné-Chantelot C, Billault A, Guillou S, Soularue P, et al (1995) A YAC contig map of the human genome. Nature Suppl 377:175–297

Clerget-Darpoux F (1982) Bias of the estimated recombination fraction and lod score due to an association between a disease gene and a marker gene. Ann Hum Genet 46:363–372

Collins A, Frezal J, Teague J, Morton NE (1996) A metric map of humans: 23,500 loci in 850 bands. Proc Natl Acad Sci USA 93:14771–14775

Cottingham RW Jr, Idury RM, Schäffer AA (1993) Faster sequential genetic linkage computations. Am J Hum Genet 53: 252–263

De Braekeleer M, Giasson F, Mathieu J, Roy M, Bouchard J-P, Morgan K (1993) Genetic epidemiology of autosomal recessive spastic ataxia of Charlevoix-Saguenay in northeastern Quebec. Genet Epidemiol 10:17–25

Dib C, Fauré S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, et al (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 380:152–154

Dubovsky J, Sheffield VC, Duyk GM, Weber J (1995) Sets of short tandem repeat polymorphisms for efficient linkage screening of the human genome. Hum Mol Genet 4:449–452

Gauvreau D, Guérin M, Hamel M (1991) De Charlevoix au Saguenay: mesure et caractéristiques du mouvement migratoire avant 1911. In: Bouchard G, De Braekeleer M (eds) Histoire d'un génome: population et génétique dans l'est du Québec. Presses de l'Université du Québec, Sillery, pp 145–159

Guilford P, Dodé C, Crozet F, Blanchard S, Chaïb H, Levilliers J, Levi-Acobas F, et al (1995) A YAC contig and an EST map in the pericentromeric region of chromosome 13 surrounding the loci for neurosensory nonsyndromic deafness (DFNB1 and DFNA3) and limb-girdle muscular dystrophy type 2C (LGMD2C). Genomics 29:163–169

Gyapay G, Morisette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, et al (1994) The 1993–94 Généthon human genetic linkage map. Nat Genet 7:246–339

Harding A (1993) Clinical features and classification of inherited ataxias. In: Harding A, Deufel T (eds) Advances in neurology. Vol 61: Hereditary ataxias, Raven Press, New York, pp 1–14

Hawkins TL, O'Connor-Morin T, Roy A, Santillan C (1994) DNA purification and isolation using a solid-phase. Nucleic Acids Res 22:4543–4544

Hudson TJ, Stein LD, Gerety SS, Ma J, Castle AB, Silva J, Slonin DK, et al (1995) An STS-based map of the human genome. Science 270:1945–1954

Knoppers BM, Laberge C (1991) DNA sampling and informed consent. CMAJ 144:128–129

Jetté R, Gauvreau D, Guérin M (1991) Aux origines d'une région: le peuplement fondateur de Charlevoix avant 1850. In: Bouchard G, DeBraekeleer M (eds) Histoire d'un génome: population et génétique dans l'est du Québec. Presses de l'Université du Québec, Sillery, pp 75–106

Lange K (1997) Mathematical and statistical methods for genetic analysis. Springer-Verlag, New York

Lange K, Weeks D, Boehnke M (1988) Programs for pedigree analysis: MENDEL, FISHER and dGENE. Genet Epidemiol 5:471–472

Lathrop GM, Lalouel JM, Julier C, Ott J (1984) Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci USA 81:3443–3446

Lathrop GM, Lalouel JM, Julier C, Ott J (1985) Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. Am J Hum Genet 37:482–498

McNally EM, Passos-Bueno MR, Bönnemann CG, Vainzof M, de Sá Moreira E, Lidov HGW, Ben Othmane K, et al (1996) Mild and severe muscular dystrophy caused by a single γ-sarcoglycan mutation. Am J Hum Genet 59:1040–1047

Richter A, Morgan K, Bouchard J-P, Mathieu J, Lamarche J, Rioux J, Hudson T, et al (1996) ARSACS: Possibly a lysosomal storage disease? Am J Hum Genet Suppl 59: A379

Richter A, Morgan K, Bouchard J-P, Poirier J, Mercier J, Gosselin F, Melançon SB (1993) Clinical and molecuar genetic studies on autosomal recessive spastic ataxia of Charlevoix-Saguenay. In: Harding A, Deufel T (eds) Advances in neurology. Vol 61: Hereditary ataxias, Raven Press, New York, pp 97–103

Schäffer AA (1996) Faster linkage analysis computations for pedigrees with loops or unused alleles. Hum Hered 46: 226–235

Schäffer AA, Gupta SK, Shriram K, Cottingham RW Jr (1994) Avoiding recomputation in linkage analysis. Hum Hered 44: 225–237

Shaw SH, Farr JEW, Thiel BA, Matise TC, Weissenbach J, Chakaravarti A, Richard CW III (1995) A radiation hybrid map of 95 STSs spanning human chromosome 13q. Genomics 27:502–510

Sobel E, Lange K (1996) Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 58:1323–1337

Staden R (1996) The Staden sequence analysis package. Mol Biotechnol 5:233–241

Terwilliger JD, Ott J (1994) Handbook of human genetic linkage. Johns Hopkins University Press, Baltimore

Weber JL, Wong C (1993) Mutation of human short tandem repeats. Hum Mol Genet 2:1123–1128

Weeks DE, Sobel E, O'Connell JR, Lange K (1995) Computer programs for multilocus haplotyping of general pedigrees. Am J Hum Genet 56:1506–1507