

*Eladio Velasco*<sup>a</sup>  
*Carmen Valero*<sup>a</sup>  
*Elena García*<sup>a</sup>  
*Arancha de la Puente*<sup>b</sup>  
*Jesús Cruces*<sup>b</sup>  
*José Luis San Millán*<sup>a</sup>  
*Ignacio del Castillo*<sup>a</sup>  
*Antonio Coloma*<sup>b</sup>  
*Felipe Moreno*<sup>a</sup>  
*Concepción*  
*Hernández-Chico*<sup>a</sup>

<sup>a</sup> Unidad de Genética Molecular,  
Hospital Ramón y Cajal, Madrid,

<sup>b</sup> Departamento de Bioquímica,  
Facultad de Medicina,  
Universidad Autónoma de  
Madrid, Spain

## Isolation of Microsatellites from the Spinal Muscular Atrophy (SMA) Candidate Region on Chromosome 5q and Linkage Analysis in Spanish SMA Families

.....  
**Abstract**

A locus responsible for autosomal recessive spinal muscular atrophy (SMA) on chromosome 5q11.2-q13.3 has been mapped to a critical interval delimited by markers D5S435 and D5S557. By a modification of the Vectorette-(GT)<sub>n</sub> method, we have isolated three polymorphic CA repeats from two YACs of the SMA region. Two of them (D5S1417 and D5S1416) map within the SMA critical region, and the other (D5S1415) is centromeric to D5S435. Linkage analysis in Spanish SMA families with eleven markers showed that in our families the disease is linked to this region and confirmed that the novel markers are tightly linked to the SMA locus. The most likely order of markers was 5cen-(D5S63/D5S1356)-(D5S125/D5S465)-(D5S435/D5S1417/D5S1416/D5S557)-D5S610-D5S112-D5S127-5qter, with odds against alternative orders >1,000:1. Genetic distances are in agreement with those previously published. However, the recombination fraction between D5S610 and D5S112 is remarkably greater than expected from the physical distance, suggesting a hot spot for recombination in this region. Our results from haplotype and multipoint analyses show that the SMA locus must lie between D5S465 and D5S112, and lend further support to the current location of the SMA locus.

.....  
**Key Words**

Spinal muscular atrophy  
Microsatellites  
Linkage analysis

## Introduction

Spinal muscular atrophy (SMA) is characterized by degeneration of anterior horn cells of the spinal cord leading to progressive symmetrical weakness and wasting in the proximal muscles. Three forms of SMA have been distinguished on the basis of clinical severity and age of onset [1]. Autosomal recessive inheritance has been reported in more than 90% of cases, with a carrier frequency of about 1/50.

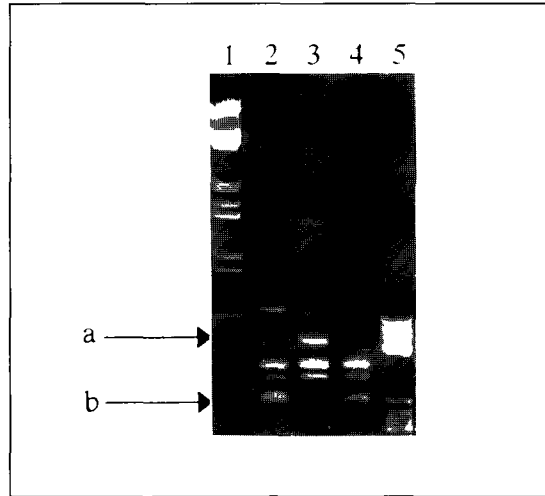
A locus for the three clinical types of SMA has been mapped to chromosome 5q11.2-q13.3 [2–5]. In the last 3 years, the development of new markers has contributed to narrowing the genetic interval for the SMA locus from 11 [2] to 0.7 cM [6]. The SMA locus has recently been mapped to a region flanked by loci D5S435 and D5S557 [7].

Several YAC contigs spanning the chromosomal region between loci D5S125 and MAP1B have been generated [7–9]. To narrow the SMA candidate region, we have characterized highly polymorphic markers from two YACs of the contig reported by Kleyn et al. [8]. Linkage analysis was then carried out in Spanish SMA families with the new and other previously described markers.

## Materials and Methods

### SMA Families

SMA was diagnosed by clinical examination, electromyography and muscle biopsy according to the criteria laid out by the International SMA Consortium [10]. A total of 55 families were collected. Thirty-four of them, consisting of a total of 160 probands including 41 affected individuals, were suitable for linkage analysis. Of these families, 10 were of type I SMA, 14 of type II and 10 of type III. DNA was isolated from peripheral blood lymphocytes by the salting-out method [11].



**Fig. 1.** Vectorette-(GT)<sub>n</sub>-PCR products from *Bgl*II Vectorette libraries separated on 1.4% agarose gels. Lane 1 =  $\lambda$  DNA digested with *Hind*III and *Eco*RI; 2 = YAC clone Y38; 3 = Y116; 4 = 236B1 as control; 5 = pBR322 DNA digested with *Hae*III. Arrows a and b indicate specific PCR products of human origin from YACs Y116 (lane 3) and Y38 (lane 2), respectively.

### Isolation of Microsatellites

Microsatellites were obtained from clones Y38 and Y116 of the YAC contig of Kleyn et al. [8] by a modification of the Vectorette-(GT)<sub>n</sub> method [12]. *Bam*HI, *Bgl*II and *Pst*I Vectorette libraries were prepared from total genomic DNA of clones Y38, Y116 and others carrying non-overlapping YACs as controls, as described by Riley et al. [13]. PCR reactions were then performed using the Vectorette universal primer 224 and the GGGG(GT)<sub>11</sub> primer. The rationale of our experiment was that, for Vectorette libraries made with the same restriction endonuclease, those PCR products common to all YAC clones should correspond to yeast sequences, whereas those specifically amplified from each clone should be of human origin (fig. 1). To prove the human origin of these PCR products, they were hybridized to PFGE blots containing undigested genomic DNA of YAC clones Y38 and Y116. In every case, only the YAC band was revealed, confirming their human origin (data not shown). This protocol does not require the purification of YAC DNA from PFGE gels, which is time-consuming and usually offers poor yields.

**Table 1.** PCR primers and conditions of the novel microsatellites

Locus <sup>a</sup>	Structure	PCR primers	Annealing temperature °C	Number of alleles	Heterozygosity
D5S1415 (38CA2)	(CA) <sub>n</sub>	F: 5'-AATGACTGGATAAAGATAATGCA-3' R: 5'-CCAGGCTGCTACAAATACCAT-3'	58	7	0.74
D5S1416 (116CA1)	(CA) <sub>n</sub>	F: 5'-GCGTTTCGCTACAGGCCAGATGA-3' R: 5'-CAGGGGAAAATAGGCTTGTGAA-3'	60	11	0.53
D5S1417 (116CA3)	(CA) <sub>n</sub>	F: 5'-TTTTCAACCCTGAGACATTCAAC-3' R: 5'-TCCCCATCAGATCAATTAATA-3'	58	9	0.67

Primer pairs were used in a 25- $\mu$ l PCR reaction containing 75 ng of genomic DNA, 15–25 pmol of each primer, 200  $\mu$ M each dNTP, standard PCR buffer, 1.5 mM MgCl<sub>2</sub> and 1 U Taq DNA Polymerase (Perkin Elmer). After an initial denaturation step at 94 °C for 3 min, 30 cycles were performed as follows: denaturation at 94 °C for 1 min, annealing, at the temperature indicated in the table, for 1 min, and extension at 72 °C for 1 min; finally, an additional extension step at 72 °C for 7 min.

<sup>a</sup> Probe in parentheses.

PCR products were cloned into the pGEM-T vector (Promega) and sequenced using Sequenase v2.0 (US Biochemicals) according to the instructions of the manufacturer. An oligonucleotide primer (F primer of 38CA2, 116CA1 and 116CA3 in table 1) was then designed to amplify the other side of the CA repeat by the Vectorette method [13]. Once this PCR product had been sequenced, another primer (R primers of table 1) compatible with the F primer was designed to amplify the microsatellite. Thus, this approach is quicker than subcloning the YAC and screening the resultant library with a poly(CA)<sub>n</sub>(GT) probe.

#### PCR Analysis

Oligonucleotide primers were designed with the program OLIGO™ v4.0 (Research Genetics) to amplify the novel CA repeats. Primers and annealing temperatures are indicated in table 1.

Other microsatellite markers used were: cVS19 (D5S435) [6], 2ae9.1 (D5S557) [7], EF(TG/AG)<sub>n</sub>, JK53CA1/2, YN(CT)<sub>n</sub> (loci D5S125, D5S112 and D5S127, respectively) [14], Afm114ye7 (D5S465) [15], 38.3 (D5S610) [16] and 236B1CA1 (D5S1356) [17]. All these markers were amplified using primers previously reported, and PCR conditions were modified in order to avoid the use of radioactivity. Alleles were visualized on 6–8% non-denaturing polyacrylamide gels after ethidium bromide staining.

The *TaqI* RFLP at the D5S63 locus, which is detected by the probe L407 (Collaborative Research) [18], was typed by a PCR-based assay using primers 407-Taq1 (5'-ATGGTCTGAGCTTTAGTTAGGA-3') and 407-Taq2 (5'-TTAGAAAGCAAAGCAAGATGAT-3') following the conditions detailed in the footnote to table 1 with an annealing temperature of 56 °C.

#### Linkage Analysis

Linkage analysis was performed with the LINKAGE package of programs v5.1 [19]. SMA was coded as an autosomal recessive disorder and full penetrance was assumed in all cases. The gene frequency for SMA was 0.006. Alleles at each marker locus were down-coded according to Ott [20]. Two-point lod scores and multipoint analysis were done with the MLINK and LINKMAP options, respectively. The recombination fractions were sex-averaged. The Kosambi mapping function was used to convert between recombination fractions and map distances. Testing for locus heterogeneity was carried out with the HOMOG program [21].

**Table 2.** Two-point lod scores between SMA and 5q loci

SMA vs.	Recombination fractions								$Z_{\max}$	$\theta_{\max}$
	0.0	0.001	0.01	0.05	0.1	0.2	0.3	0.4		
D5S63	1.56	1.56	1.56	1.49	1.34	0.92	0.51	0.18	1.56	0.000
D5S1356	—∞	0.63	2.48	3.23	3.04	2.13	1.17	0.38	3.24	0.053
D5S125	2.53	2.52	2.45	2.15	1.79	1.14	0.60	0.18	2.53	0.000
D5S465	—∞	2.30	3.19	3.39	3.06	2.10	1.12	0.33	3.42	0.035
D5S435	6.27	6.25	6.07	5.29	4.38	2.76	1.43	0.44	6.27	0.000
D5S1417	6.15	6.13	5.99	5.37	4.57	3.02	1.61	0.53	6.15	0.000
D5S1416	5.24	5.23	5.09	4.48	3.72	2.32	1.18	0.39	5.24	0.000
D5S557	2.23	2.22	2.16	1.87	1.48	0.78	0.32	0.07	2.23	0.000
D5S610	6.84	6.82	6.66	5.91	4.95	3.13	1.59	0.48	6.84	0.000
D5S112	—∞	3.49	4.35	4.37	3.81	2.49	1.31	0.44	4.50	0.026
D5S127	—∞	-1.96	1.80	3.56	3.53	2.52	1.39	0.48	3.64	0.071

## Results and Discussion

### *Isolation and Location of the Microsatellites*

To isolate microsatellite markers within the interval for the SMA locus, we chose YAC clones Y38 and Y116 of the contig described by Kleyn et al. [8]. CA repeats were isolated from these YACs by a modification of the Vectorette-(GT)<sub>n</sub> method [12], as described in the Materials and Methods. In this way, we isolated one microsatellite from Y38 (38CA2, *Bgl*II Vectorette library), and two from Y116 (116CA1, *Bam*HI library, and 116CA3, *Pst*I library).

Microsatellite markers were typed against a panel of more than 50 unrelated individuals and heterozygosities ranging from 0.53 (D5S1416) to 0.74 (D5S1415) could be observed (table 1).

The new markers were mapped between sequence-tagged sites (STSs) previously described by PCR-typing 31 clones from the YAC contig [8]. Marker 38CA2 (D5S1415) was located proximal to D5S435 between STSs Y38U and Y19T. On the other side,

markers 116CA1 (D5S1416) and 116CA3 (D5S1417) were located distal to D5S435, between STSs Y38T and Y122U (D5S1417) and between Y122U and Y55U (D5S1416), and therefore map into the SMA candidate region.

### *Linkage Analysis of Spanish SMA Families*

Thirty-four Spanish SMA families were typed with the 11 genetic markers listed in table 2. These include D5S435 and D5S557, which delimit the genetic interval for the SMA locus, and D5S1417 and D5S1416, which map within that interval. Two-point lod scores between the loci indicated that the new microsatellite markers are linked to markers of the SMA region, with peak lod scores found between D5S610 and D5S1416 (11.0 at  $\theta = 0.025$ ) and between D5S610 and D5S1417 (15.12 at  $\theta = 0.011$ ). Furthermore, pairwise linkage analysis between the marker loci and the SMA locus showed that D5S1417 and D5S1416 were tightly linked to the SMA locus. No recombination of these markers with the disease was found in our study (table 2).

A genetic map of polymorphic markers including D5S1417 and D5S1416 was constructed from genotypic data in our families. The order of the markers D5S1356 (assumed to be at the same position as D5S6) [17], D5S465, D5S112 and D5S127 was fixed, as previously reported [14, 15, 22], and the position of the other markers was varied sequentially throughout the map. The maximum likelihood order of markers was: cen-(D5S63, D5S1356)-0.035-(D5S125, D5S465)-0.018-(D5S435, D5S1417, D5S1416, D5S557)-0.010-D5S610-0.051-D5S112-0.024-D5S127. The odds against alternative orders were >1,000:1. The order of central markers could not be established, as no recombinational event was observed within the group D5S435, D5S1417, D5S1416 and D5S557. Similarly, no recombinational event was observed between D5S63 and D5S1356, or between D5S125 and D5S465. The distances calculated from the pedigrees reported here were consistent with previous results, except for that between D5S610 and D5S112 which is much greater than expected considering the physical distance between them (650–700 kb). Five recombinational events involving this area were observed in 180 meioses in our sample. A high frequency of recombination in this zone has been reported in SMA families by other authors [22]. Furthermore, repeated sequences are known to be located in this region [7–9, 16]. Altogether, these observations suggest that there might be a hot spot for recombination in this interval.

A multipoint linkage analysis was performed to estimate the most likely position of the SMA gene within the genetic map described above. The peak multipoint lod score was 8.6 placing SMA most likely in the 7.9-cM interval between D5S465 and D5S112. The difference between the maximum lod scores for the placement of SMA to this region is greater than 3.0 when compared with all

other possible locations in the interval analysed, strongly supporting the location of SMA in this interval.

A heterogeneity analysis performed with the multilocus lod scores provided no evidence of genetic heterogeneity in this family set (data not shown).

In conclusion, we report three novel microsatellite markers, mapping to the SMA region on chromosome 5q. Our physical and genetic data support the location of two of them (D5S1417 and D5S1416) to the critical interval that must include the SMA locus. Results of the linkage analysis of Spanish SMA families place the disease between D5S465 and D5S112. Further narrowing of this interval was not possible using our set of families. However, given the location and informativeness of D5S1417 and D5S1416, these markers should become very useful tools to precisely locate the SMA locus and perform prenatal diagnosis.

### Acknowledgements

We thank the SMA families and the clinicians who collaborated in this study, particularly Dr. S.I. Pascual Pascual and Dr. I. Pascual-Castroviejo (Hospital La Paz, Madrid). We are also grateful to Denis Le Paslier (CEPH, Paris) for kindly providing the CEPH YAC clones and Susana Gómez for critical reading of the manuscript. This work was supported by grants 93/0027-01 and 93/0027-02 of the Fondo de Investigaciones de la Seguridad Social (Ministerio de Sanidad y Consumo, Spain).

## References

- 1 Dubowitz V: Muscle Disorders in Childhood. Philadelphia, Saunders, 1978.
- 2 Brzustowicz LM, Lehner T, Castilla LH, Penchaszadeh GK, Wilhelmsen KC, Daniels R, Davies KE, Leppert M, Ziter F, Wood D, Dubowitz V, Zerres K, Hausmanowa-Petrusewicz I, Ott J, Munsat TL, Gilliam TC: Genetic mapping of chronic childhood-onset spinal muscular atrophy to chromosome 5q11.2-q13.3. *Nature* 1990;344:540-541.
- 3 Melki J, Abdelhak S, Sheth P, Bachelot MF, Burlet P, Marcadet A, Aicardi J, Barois A, Carrière JP, Fardeau M, Fontan D, Ponsot G, Billette T, Angelini C, Barbosa C, Ferrière G, Lanzi G, Ottolini A, Babton MC, Cohen D, Hanauer A, Clerget-Darpoux F, Lathrop M, Munnich A, Frezal J: Gene for chronic proximal spinal muscular atrophies maps to chromosome 5q. *Nature* 1990;344:767-768.
- 4 Melki J, Sheth P, Abdelhak S, Burlet P, Bachelot MF, Lathrop MG, Frezal J, Munnich A: Mapping of acute (type I) spinal muscular atrophy to chromosome 5q12-q14. *Lancet* 1990;336:271-273.
- 5 Gillian TC, Brzustowicz LM, Castilla LH, Lehner T, Penchaszadeh GK, Daniels RJ, Byth BC, Knowles J, Hislop JE, Saphira Y, Dubowitz V, Munsat TL, Ott J, Davies KE: Genetic homogeneity between acute and chronic forms of spinal muscular atrophy. *Nature* 1990;345:823-825.
- 6 Soares VM, Brzustowicz LM, Kleyn PW, Knowles JA, Palmer DA, Asokan S, Penchaszadeh GK, Munsat TL, Gilliam TC: Refinement of the spinal muscular atrophy locus to the interval between D5S435 and MAP1B. *Genomics* 1993;15:365-371.
- 7 Francis MJ, Morrison KE, Campbell L, Grewal PK, Christodoulou Z, Daniels RJ, Monaco AP, Frischauf AM, McPheerson J, Wasmuth J, Davies KE: A contig of non-chimeric YACs containing the spinal muscular atrophy gene in 5q13. *Hum Mol Genet* 1993;8:1161-1167.
- 8 Kleyn PW, Wang CH, Lien LL, Vitale E, Pan J, Ross BM, Grunn A, Palmer DA, Warburton D, Brzustowicz LM, Kunkel LM, Gilliam TC: Construction of a yeast artificial chromosome contig spanning the spinal muscular atrophy disease gene region. *Proc Natl Acad Sci USA* 1993;90:6801-6805.
- 9 Melki J, Lefebvre S, Burglen L, Burlet P, Clermont O, Millasseau P, Reboullet S, Bénichou B, Zeviani M, Le Paslier D, Cohen D, Weissenbach J, Munnich A: De novo and inherited deletions of the 5q13 region in spinal muscular atrophies. *Science* 1994;264:1474-1477.
- 10 Munsat T, Davies KE: Meeting report: International SMA Consortium Meeting. *Neuromusc Disord* 1992;2:423-428.
- 11 Miller SA, Dykes DD, Polesky HF: A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;16:1215.
- 12 Pandolfo M: A rapid method to isolate (GT)<sub>n</sub> repeats from yeast artificial chromosomes. *Nucleic Acids Res* 1992;20:1154.
- 13 Riley J, Butler R, Ogilvie D, Finniear R, Jenner D, Powell S, Anand R, Smith JC, Markham AF: A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res* 1990;18:2887-2890.
- 14 Daniels RJ, Suthers GK, Morrison KE, Thomas NH, Francis MJ, Matthew CG, Loughlin S, Heiberg A, Wood D, Dubowitz V, Davies KE: Prenatal prediction of spinal muscular atrophy. *J Med Genet* 1992;29:165-170.
- 15 Melki J, Burlet P, Clermont O, Pascal F, Paul B, Abdelhak S, Sherrington R, Gurling H, Nakamura Y, Weissenbach J, Lathrop M, Munnich A: Refined linkage map of the chromosome 5 in the region of the spinal muscular atrophy gene. *Genomics* 1993;15:521-524.
- 16 Thompson TG, Morrison KE, Kleyn PW, Bengtsson U, Gilliam TC, Davies KE, Wasmuth JJ, McPheerson JD: High resolution physical map of the region surrounding the spinal muscular atrophy gene. *Hum Mol Genet* 1993;2:1169-1176.
- 17 Velasco E, de la Puente A, Cruces J, Valero MC, García E, del Castillo I, Coloma A, Moreno F, Hernández-Chico C: Dinucleotide repeats polymorphisms at the D5S1356, D5S1357 and D7S1480 loci. *Hum Mol Genet* 1994;3:1441.
- 18 Weiffenbach B, Falls K, Green P, Shute N, Keith T, Mao J, Donis-Keller H: Genetic map of chromosome 5 with 53 RFLP loci. *Cytogenet Cell Genet* 1989;51:1104.
- 19 Lathrop GM, Lalouel JM, Julier C, Ott J: Multilocus linkage analysis in humans: Detection of linkage and estimation of recombination. *Am J Hum Genet* 1985;37:482-498.
- 20 Ott J: A simple scheme for the analysis of HLA linkages in pedigrees. *Ann Hum Genet* 1978;42:255-257.
- 21 Ott J: Analysis of Human Genetic Linkage, rev ed. Baltimore Johns Hopkins University Press, 1991.
- 22 Wirth B, Pick E, Leutner A, Dadze A, Voosen B, Knapp M, Piechaczek-Wappenschmidt B, Rudnik-Schöneborn S, Schönling J, Cox S, Spurr NK, Zerres K: Large linkage analysis in 100 families with autosomal recessive spinal muscular atrophy (SMA) and 11 CEPH families using 15 polymorphic loci in the region 5q11.2-q13.3. *Genomics* 1994;20:84-93.

### Note added in proof

Two genes for SMA have recently been identified within the candidate region described in this work. *Cell* 1995;80:155-165, 167-178.