Am. J. Hum. Genet. 56:202–209, 1995

Refinement of the Spinal Muscular Atrophy Locus by Genetic and Physical Mapping

C. H. Wang,[1,2] P. W. Kleyn,[1] E. Vitale,[1] B. M. Ross,[1] L. Lien,[3] J. Xu,[1] T. A. Carter,[1] L. M. Brzustowicz,[1] S. Obici,[1] S. Selig,[3] L. Pavone,[5] E. Parano,[5] G. K. Penchaszadeh,[1] T. Munsat,[4] L. M. Kunkel,[3] and T. C. Gilliam[1]

[1]Departments of Psychiatry and Genetics and Development, and [2]Departments of Pediatrics and Neurology, College of Physicians and Surgeons at Columbia University, and the New York State Psychiatric Institute, New York; [3]Department of Genetics, Harvard Medical School, and Division of Genetics, Department of Pediatrics, and the Howard Hughes Medical Institution, Children's Hospital, and [4]Department of Neurology, Tufts University—New England Medical Center, Boston; and [5]Pediatric Clinic, University of Catania, Catania, Italy

Summary

We report the mapping and characterization of 12 microsatellite markers including 11 novel markers. All markers were generated from overlapping YAC clones that span the spinal muscular atrophy (SMA) locus. PCR amplification of 32 overlapping YAC clones shows that 9 of the new markers (those set in italics) map to the interval between the two previous closest flanking markers (D5S629 and D5S557): cen - D5S6 - D5S125 - D5S435 - *D5S1407* - D5S629-*D5S1410-D5S1411/D5S1412-D5S1413-D5S1414-D5Z8-D5Z9*-CATT1-*D5Z10/D5Z6*-D5S557-*D5S1408-D5S1409*-D5S637-D5S351-MAP1B-tel. Four of these new markers detect multiple loci in and out of the SMA gene region. Genetic analysis of recombinant SMA families indicates that D5S1413 is a new proximal flanking locus for the SMA gene. Interestingly, among the 40 physically mapped loci, the 14 multilocus markers map contiguously to a genomic region that overlaps, and perhaps helps define, the minimum genetic region encompassing the SMA gene(s).

Introduction

Spinal muscular atrophy (SMA) is a motor neuron disease characterized by degeneration of spinal cord anterior horn cells and muscular atrophy. Clinical presentations include muscle fasciculation and symmetric weakness involving the proximal limbs and lower cranial nerves. Electromyography shows a neurogenic pattern, and muscle biopsy reveals group atrophy with fiber size variation. The childhood-onset SMAs are often divided into three categories based on degree of disability and age at onset: SMA type I, or Werdnig-Hoffmann disease, is the most severe manifestation, leading to death in the first years of life; SMA type II, or intermediate SMA, and SMA type III (Kugelberg-Welander disease) are milder, with later age at onset and considerable clinical heterogeneity. The childhood-onset SMAs are recessively inherited and affect ~1:10,000 live births, with a carrier frequency of 1/40–1/60.

All three forms of childhood-onset SMA have been mapped to a single locus at chromosome 5q13 (Brzustowicz et al. 1990; Gilliam et al. 1990; Melki et al. 1990*a*, 1990*b*). Pedigree analyses have further refined the SMA disease locus to a 1–2-cM region defined by flanking microsatellite markers (Brzustowicz et al. 1992; Francis et al. 1993; Soares et al. 1993; Clermont et al. 1994). YAC contigs have been constructed that span the disease gene region (Francis et al. 1993; Kleyn et al. 1993; Melki et al. 1994), and a radiation hybrid map of the markers spanning this region has been assembled (Thompson et al. 1993). Recently, Melki et al. (1994) have generated numerous physically mapped genetic markers across the SMA region, several of which detect two or more loci in the region. Genotyping of 201 SMA families with these markers indicated the presence of de novo and inherited deletions in 9 unrelated SMA patients. To further characterize the genetic region defining the SMA locus, we have generated new microsatellite markers directly from a set of five overlapping YAC clones and from lambda phage subcloned from these YACs that encompass this region. A modified vectorette PCR method (Riley et al. 1990; Pandolfo 1992) was used to isolate 11 new polymorphic microsatellite markers that map between D5S435 and MAP1B. The markers were genotyped against six SMA families containing chromosomal recombination in the disease gene region (Brzustowicz et al. 1992; Soares et al. 1993). The new genetic markers were positioned relative to previously reported genetic markers, by PCR amplification of the set of overlapping YAC clones. Marker D5S1413 is the closest proximal flanking marker to the SMA locus. One new marker located in the minimal genetic region (D5S1414) detects significant linkage disequilibrium to the disease locus in a subset of 32 families from Poland (Brzustowicz et al. 1995 [in this issue]).

Received July 12, 1994; accepted for publication October 3, 1994.

Address for correspondence and reprints: Dr. T. Conrad Gilliam, College of Physicians and Surgeons at Columbia University, 722 West 168th Street, Unit 23, New York, NY 10032.

© 1995 by The American Society of Human Genetics. All rights reserved.

0002-9297/95/5601-0026$02.00

Methods

Description of YAC, Cosmid, and Bacteriophage Resources

A contig of overlapping YAC clones spanning the SMA region has been described elsewhere (Kleyn et al. 1993). All YAC clones described in this study are from the CEPH1 midi-YAC library (Albertsen et al. 1990). Total DNA was prepared from YACs 55 (284 D1), 88 (428 C5), and 98 (81 B11), by a modification of the protocol of Green and Olson (1990). Set amounts of DNA were mixed with decreasing amounts of *Mbo*I, and the degree of partial digestion was monitored by gel electrophoresis. The four partial digests with the most ethidium-positive DNA in the 10–20-kb size range were pooled and separated on low-melt agarose. DNA in the 10–20-kb range was excised and electroeluted from the agarose. After ethanol precipitation, the DNA was quantitated, and 0.5 μg was mixed with 1 μg of EMBL-3 phage arms previously cleaved with *Bam*HI (Stratagene). After ligation at 15°C overnight, the DNA was packaged into phage particles (Gigapack II Gold; Stratagene). Phage were titered and then plated onto LE392 bacterial lawns where nitrocellulose filter lifts were preformed. Phage containing human inserts were identified by hybridization with P^{32}-labeled total human DNA. Positively hybridizing phage were picked and replated as grids of 100. These libraries contained an average insert size of 15 kb. Homologous cosmid clones were selected from a chromosome 5-enriched fluorescence-amplified flow-sorted cosmid library (SCos/1[AO5NCO19]) developed by the Los Alamos National Laboratories.

Isolation of CA-Repeat Microsatellite Markers

A modified vectorette PCR method (Pandolfo 1992) was used to isolate CA-repeat microsatellite markers. Purified YAC inserts, lambda phage DNA, or cosmid DNA were restriction digested (with either *Pst*I, *Bgl*II, *Dde*I, or *Bfa*I) and were ligated with vectorette linkers as described elsewhere (Kleyn et al. 1993). PCR amplification was initiated using either $(CA)_{15}$ or $(GT)_{15}$ oligonucleotides together with linker primer which is identical to one strand of the vectorette sequence. This strategy depends on the annealing of the CA/GT primer to initiate synthesis of the complementary strand which can be recognized by the linker primer. The resultant PCR amplification product was purified, and DNA was sequenced as described elsewhere (Kleyn et al. 1993). Next, the complementary oligonucleotide primer, or new primers selected from the flanking DNA sequence, was used to amplify the DNA sequence on the opposite side of the CA repeat. A schematic presentation of this method is presented in figure 1.

Physical Mapping of Microsatellite Markers

Each microsatellite marker was used to PCR-amplify a panel of genomic DNA and DNA prepared from somatic hybrid cell lines including total human DNA; total hamster

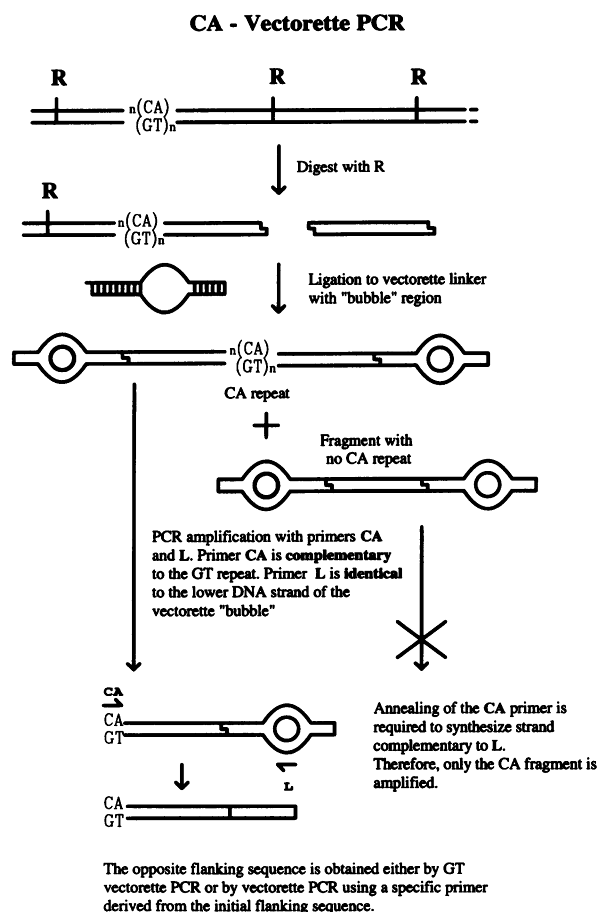

Figure 1 Vectorette PCR method to identify DNA sequence flanking CA-repeat sequences. Double-stranded DNA from YAC, cosmid, or bacteriophage clones is limit-digested with one of several restriction endonucleases (*Pst*I, *Bgl*II, *Dde*I, or *Bfa*I), then ligated to vectorette linker. Complementary strand for the linker primer (L) is primed by either $(CA)_{15}$ or $(GT)_{15}$ oligonucleotide. Thus, PCR product is only obtained for restriction fragment–vectorette linker products containing CA/GT-repeat sequences. The PCR product is purified and sequenced. The second round is primed with the alternate oligonucleotide or with primer complementary to the flanking DNA.

(Chinese hamster ovarian [CHO]) DNA; human whole chromosome 5 only/CHO hybrid cell line (HHW105); and human deleted chromosome 5 (deleted for 5q11.2-q13.3) only/CHO hybrid cell line (HHW1064) (Gilliam et al. 1989). This strategy places markers relative to the ~20-cM deletion region that includes the SMA locus. Each marker was used to PCR-amplify DNA from a panel of overlapping YAC clones (Kleyn et al. 1993), in order to determine their relative positions along the contig. Individual YAC clones were picked from glycerol stocks, colonies were amplified, and YAC DNA was either isolated as aqueous genomic DNA, by a modification of the method of Green and Olson (1990), or purified in agarose blocks (Schwartz and Cantor 1984) and separated from yeast genomic DNA, by pulsed-field gel electrophoresis by using a

Pharmacia LKB apparatus with a hexagonal electrode array. PCR amplification conditions were similar to those described previously by Soares et al. (1993). In brief, reaction mix was denatured at 94°C for 3 min, followed by 20–25 cycles of 94°C/1 min, 55°C/1 min, and 72°C/ 1 min. A final step of 72°C/5 min completed the elongation. PCR products were analyzed on 6% acrylamide nondenaturing gel after separation for 2–4 h at 50 W. The gels were dried onto 3-mM Whatman paper and were visualized by autoradiography after a 1–16-h room-temperature exposure to Kodak XAR film. In our experience, PCR amplification of YAC DNA can give variable results, so that each map location is based on several repeat amplifications and, in some cases, on Southern blot hybridization.

Marker Characterization and Genetic Analysis

Oligonucleotide primers for amplification of microsatellite markers were selected by analysis of flanking DNA sequence using the program PRIMER (version 0.5; Whitehead Institute for Biomedical Research). Oligonucleotides were synthesized with an Applied Biosystems model 392 DNA/RNA synthesizer. For single-locus markers, allele sizes and frequencies and marker heterozygosities were determined using a panel of 80 unrelated Caucasians, primarily CEPH parents. For multilocus markers, the extensive overlap of alleles from different loci prevented accurate calculation of allele frequencies and heterozygosities. Each polymorphic marker was typed against six SMA families known to display meiotic recombination between markers JK348 and MAP1B (Brzustowicz et al. 1992; Soares et al. 1993). One new recombinant family was identified in this analysis. Family 4677 is from Sicily and contains two affected children. Both patients conform with internationally established criteria for SMA (Munsat and Davies 1992) based on clinical criteria, muscle biopsy, and nerve conduction studies. The firstborn patient was able to sit but not stand or walk without support and was diagnosed with SMA type II. The second-born child showed reduced fetal movements and was diagnosed at age 2 mo, showing an inability to support her head or to sit unattended. She was diagnosed with SMA type I.

Results

We have previously reported five SMA families that showed genetic recombination between the disease locus and either D5S435 or MAP1B (Brzustowicz et al. 1992; Soares et al. 1993). The minimum genetic region defined by this analysis was mapped and cloned as a collection of overlapping YAC clones (Kleyn et al. 1993). Several of the key YACs have been subcloned into lambda bacteriophage vector and used to generate a total of 12 microsatellite markers with the CA-vectorette PCR method outlined in figure 1. Six markers were isolated from lambda phage DNA subcloned from YAC55 (D5S1407, D5S1411, and D5S1413), from YAC88 (D5Z8), and from YAC98 (D5Z10 and D5S1408). Six markers were isolated directly from YAC55 (D5S1412 and D5S1410), from YAC88 (D5S1414), from YAC97 (D5Z6), and from YAC107 (D5S1409). From YAC98 we further characterized marker 9847 (D5S557), which we and others have previously reported (Kleyn et al. 1993; Francis et al. 1993). The marker D5Z9 was isolated from a cosmid clone obtained by screening a chromosome 5–specific cosmid library (Lawrence Livermore National Laboratories) with the YAC-end clone Y88T (Kleyn et al. 1993). DNA sequence comparison of the end of YAC88 and the cosmid clone verified that the two clones contained identical segments of genomic DNA. Our analysis indicates that marker 9847 (Kleyn et al. 1993) amplifies the identical locus as D5S557 (Francis et al. 1993). By comparison, the 9847 forward primer differs by one base pair from D5S557, while the reverse primers are overlapping and identical. The 9847 primers, however, detect a second, equally intense, but nonpolymorphic locus on chromosome 5 (outside the SMA region), which is barely detectable with the published D5S557 primer set (data not shown).

Isolation and Characterization of Microsatellite Markers

Table 1 lists the 12 microsatellite markers identified by the CA-vectorette PCR method. Characteristic of this region of the genome, 4 of the 12 markers (D5Z6, 8, 9, and 10) amplify two or more loci when analyzed on diploid or haploid DNA samples. These markers suggest the presence of duplicated DNA sequence and are useful for characterizing regions of repeat-element DNA across the disease gene region. Marker D5Z9 (fig. 2) illustrates the complexity of these multilocus systems. Two PCR amplification bands are detected in several YAC clones from the SMA region (YACs 88, 97, and 98), in cosmid 155 and in phage clone 9822 (fig. 2*A*). These data indicate two closely spaced, homologous loci mapping to the SMA region. Amplification of total genomic DNA (fig. 2*A* and *B*) suggests the presence of two or more homologous loci, all of which appear to map on chromosome 5 (HHW105) and within the 5q11.2-13.3 deletion region (HHW1064). D5Z8 likewise maps solely to the chromosome 5 deletion region, while D5Z6 and D5Z10 map on chromosome 5, both inside and outside the deletion region (data not shown). Heterozygosity values for these four markers are not presented in table 1, owing to the inability to reliably distinguish alleles from each distinct locus. For these markers, the listed allele size represents the length of the original sequence from which the flanking primers were chosen. The remaining markers detect a single locus and were genotyped against a collection of ≥80 unrelated individuals to generate the allele numbers, size range, and heterozygosity values shown in table 1. The base-pair constitutions of the microsatellite repeat sequences are also listed.

Table 1

Characterization of Microsatellite Markers

Locus	Marker Name	Source	Repeat Sequence	Primers	No. of Alleles	Size of Alleles (bp)	Heterozygosity
D5S1407	B4	Y55 phage	$(CA)_{16}(TA)_7$	CATAAGTTAAGTGATCATGCATATGTGATC CCTGGGCGACAGAGTGAGAC	11	152–172	.88
D5S1410	B18	Y55	$(TG)_{18}(AG)_4$	CCTGTGATCAGGCACTTAGATAAGC TCGTCATTGTACTCCATCCTGGGCA	13	147–171	.84
D5S1411	D9	Y55 phage	$(CA)_{21}$	AAGATGATGGCTGGATCC AGTGATTTTCAACCCTGAGACATTC	12	122–146	.71
D5S1412	B12	Y55	$(CA)_{23}$	ACCATGATGACATCTGCCAG TTATAGGTATGGGCGGTAAAGG	15	120–152	.82
D5S1413	B8	Y55 phage	$(CA)_{15}$TTANTAT$(AC)_5$	AAAATAGGCTTGTGAAACCAACGC GCTACAGGCCAGATGAGGGAAATAG	11	126–146	.56
D5S1414	B15	Y88	$(TG)_{24}$	AGCAGTGTGGACATGACTGATAAGC CGCGCCATTGCACTCCAG	16	152–188	.83
D5Z8	8896	Y88 phage	$(CA)_{20}$	GACCCAAAGGAAATACTTGGGTATAG GATCACGCCCCTGCATGCC	...	262[a]	...
D5Z9	CMS1[b]	cos 155	$(AC)_{20}$	TAATCCCACCCTGGAGAGTTTAT TCATGGAGTATGACTCCTGAGAG	...	134[a]	...
D5Z10	9870	Y98 phage	$(CA)_{23}$	CAGCTTGAATATAAATTGGTTGATACTG GTAAATCCTAGTAGTTTGCTGGTGC	...	197[a]	...
D5Z6	9762	Y97	$(GT)_8(AT)_{11}$	CACTAAAACTGAGAATATAGAGACC GTTGTGTATCACTTGCTTCTTCG	...	137[a]	...
D5S1408	9816	Y98 phage	$(CA)_{23}$	GCGTATAATTACTTCATCTCTTTCATTG CGGCCTCTTTTTGTTTTTATTC	12	103–125	.75
D5S1409	1075	Y107	$(CA)_{16}$	GGCTCATATATTCAGGAAGGGGTC GAGAGAAGTAGCTCACAGAGATACTGAG	5	224–232	.76

NOTE.—Markers with a "D5S" prefix are single-locus markers that detect the listed number and size range of alleles in a sample of 80 chromosomes from unrelated CEPH pedigree parents. Heterozygosities are calculated with the PIC program (Ott 1991). Markers with a D5Z" prefix detect more than one locus, at least one of which cosegregates with SMA and maps to the YAC contig on 5q13.

[a] Allele sizes for these markers represent the length of the original sequence from which the flanking primers were chosen.

[b] This marker was reported, in preliminary form, by Kleyn et al. (1993).

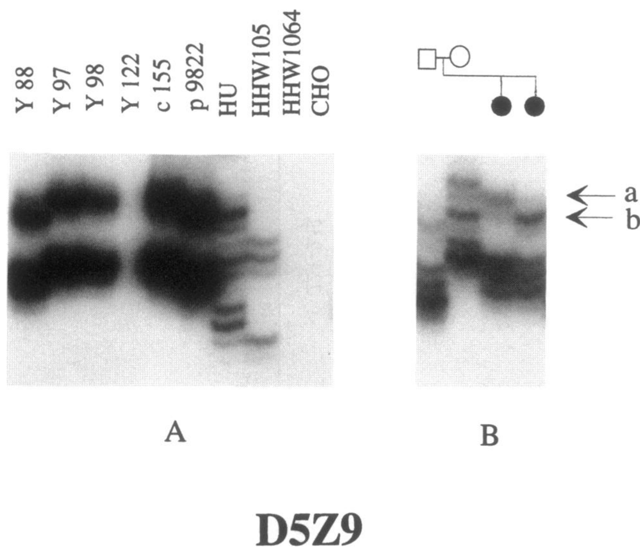

Figure 2 Multilocus marker D5Z9. Radiolabeled oligonucleotide primers for the marker D5Z9 were used to PCR-amplify purified DNA samples (see Methods). Amplification products were separated on acrylamide, as described in Methods. In panel *A*, the first four lanes represent purified aqueous DNA from the YAC colonies indicated. "c155" is a homologous cosmid clone isolated from a flow-sorted, chromosome 5–enriched library. "p9822" is a bacteriophage subclone derived from YAC 98. Total human DNA (HU), total CHO DNA, and the two chromosome 5 hybrid cell lines (HHW105 and HHW1064) are described in Methods. Panel *B* shows family 4677, described in Methods and illustrated in fig. 5. Lane 3 contains DNA from the firstborn child diagnosed with SMA type II. Lane 4 contains DNA from the second-born child diagnosed with the severe form of SMA (type I). The arrows designated "a" and "b" point to allelic variation between the two siblings.

Physical Mapping of Microsatellite Markers

Oligonucleotide primers specific for the 12 microsatellite markers characterized in this study, together with several previously characterized markers, were used to PCR-amplify a set of 32 overlapping YAC clones that span the SMA region (Kleyn et al. 1993). All of the probes we tested did amplify a subset of YAC clones. The physical locations of the microsatellite markers are shown, along with previously characterized STS sequences, in figure 3. Several marker sets identify an identical subset of YAC clones and generate similar genetic information when typed against the six recombinant families and therefore cannot be positioned relative to one another. Markers D5S1411 and D5S1412 could map in the reverse order indicated in figure 3, as could markers D5Z10 and D5Z6. Likewise, the order of the previously published markers (D5F153, 149, 150, and CATT1) (Burghes et al. 1994; Melki et al. 1994) cannot be distinguished by this map. The markers D5S557 and D5S435 were shown to define the genetic limits of the SMA locus (Francis et al. 1993; Soares et al. 1993). In an independent study, markers D5S629 and D5S637 were presented as the closest markers flanking the SMA locus (Clermont et al. 1994). In figure 3, markers and STSs that clearly identify a single locus are indicated by blackened circles and squares, respectively, while multilocus markers are indicated by unblackened circles (or squares for multilocus "STSs"). On the basis of our independent analysis of D5S557, we have listed it as a multilocus marker. A region encompassing most of YAC98 and defined by the centromeric end of YAC98 (Y98T) and the centromeric end of YAC112 (Y112T) is distinguished from the remaining contig region, in that it is composed largely of multilocus markers and YAC end-clones. The marker CATT1 that detects at least 16 alleles (Burghes et al. 1994) is mapped to YAC98 in this region.

Genotyping of SMA Pedigrees

All microsatellite markers were genotyped against six SMA families in whom meiotic recombination was known to occur between loci D5S435 and MAP1B (Brzustowicz et al. 1992; Soares et al. 1993). D5S435 was previously shown to recombine proximal to the disease locus in families 165 and 1938, while MAP1B was shown to recombine distal to the disease locus in families 730, 1250, and 2001. Family 4677 (fig. 2*B*) has not been previously reported. Figure 4 illustrates the most likely position of meiotic breakpoints deduced from genotypic analysis of the 12 newly characterized microsatellite markers, together with 9 previously reported markers. D5S435 is the most centromeric marker, and MAP1B is the most distal marker. In family 165, all markers distal to D5S435 were nonrecombinant or uninformative. In family 1938, recombination is detected by markers extending distally from D5S435, up to and including D5S1413, with nonrecombinants beginning with D5Z8 and all distal markers. Figure 5 shows the result of genotyping of marker D5S1413 against family 1938, illustrating the putative recombinant event occurring between this marker and the SMA locus in this family. The recombination presumably occurs between D5S1413 and D5Z8 and indicates that D5S1413 is the new proximal flanking marker. This interpretation makes the assumption that family 1938 is correctly diagnosed with SMA and that the disease mutation in this family is linked to 5q13. A new family (4677) was identified who is nonrecombinant with marker D5Z8 and with several proximal markers, "indeterminate" for a number of distal multilocus markers, and clearly recombinant for marker D5S637. Those markers indicated as indeterminate in figure 4 reliably detect allelic differences in the affected siblings. For some of these markers, it appears that the first affected sibling fails to receive a full complement of maternal alleles, which is suggestive of a de novo deletion. Other markers, like D5Z9 (fig. 2*B*), reproducibly detect an allele (allele "a") that is not clearly present in the parental genotypes. We are not currently able to interpret the seggregation pattern of these marker alleles in family 4677, but it is presumed that they reflect the genomic instability that characterizes this region.

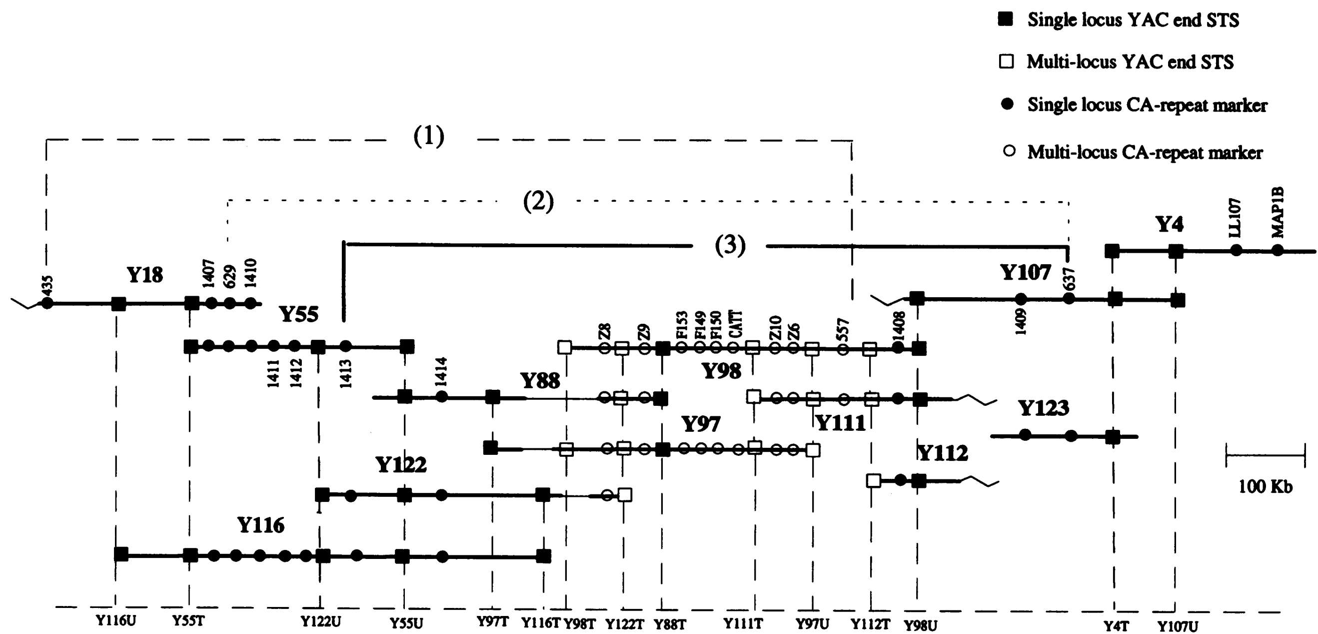

Figure 3 Physical and genetic map of the SMA locus. Horizontal lines represent CEPH1 midi-YAC clones described by Kleyn et al. (1993). Blackened squares and dots represent STS's and single-locus microsatellite markers, respectively. Unblackened circles represent multilocus microsatellite markers, and unblackend squares represent multilocus PCR fragments. YAC end-clones are represented at the bottom of the figure and have been described elsewhere (Kleyn et al. 1993). The minimum genetic region for the SMA locus, defined by Francis et al. (1993), is spanned by flanking markers D5S435 and D5S557 as indicated by (1). Clermont et al. (1994) defined the SMA region between markers D5S629 and D5S637 (2). In this study we define the centromeric flanking marker D5S1413 and the telomeric flanking marker D5S637 as shown in (3). Multilocus markers D5F153, D5F149, and D5F150 were described by Melki et al. (1994) as the sites of inherited and de novo deletions associated particularly with the severe form of SMA.

Discussion

We have used a YAC contig spanning the SMA locus to physically map 12 newly characterized microsatellite markers relative to previously characterized markers. Nine markers were identified that map to the interval defining the minimal genetic region previously identified as harboring the SMA gene (Francis et al. 1993; Soares et al. 1993; Clermont et al. 1994). Analysis of a collection of six SMA families characterized with recombination breakpoints very close to the disease locus (Brzustowicz et al. 1992; Soares et al. 1993) indicates that marker D5S1413 is the closest flanking marker proximal to the SMA locus. Physical mapping places D5S1413 close to locus D5S629, which was shown by Clermont et al. (1994) to be the closest proximal flanking marker to the disease locus. The close convergence between two independent laboratories for the designation of the proximal boundary for the SMA gene is encouraging. It is interesting that marker D5S1414 detects significant linkage disequilibrium with SMA in a collection of SMA families from Poland (Brzustowicz et al. 1995). The fact that D5Z9 was originally isolated using a YAC end probe locates this marker very close to the distal end of YAC 88. The exact physical location of marker D5S1413 on YAC 55 is less certain. Therefore the precise physical distance between markers D5S1413 and D5Z9 is difficult to obtain. Recent physical mapping evidence suggests that markers D5Z9 and D5Z8 are in very close proximity to one another, i.e., within two overlapping phage (data not shown). YAC122 contains both D51413 and D5Z8 and was estimated to be ~450 kb (Kleyn et al. 1993). This analysis shows, however, that YAC122 is deleted for an unknown length of genomic DNA. Thus, the distance from the proximal marker D5S1413 to D5Z9 is expected to be ⩾450 kb.

A number of independent studies have now documented the prevalence of low-copy-repeat DNA segments in the region of the SMA gene (Francis et al. 1993; Kleyn et al. 1993; Thompson et al. 1993; Burghes et al. 1994; Melki et al. 1994). The repeat sequences include both noncoding DNA and gene-coding DNA segments. For example, the pro-melanin-concentrating hormone gene on chromosome 12q23-q24 was shown to exist as two variants that map to 5q12-13 and to 5p14 (Pedeutour et al. 1994), and human elongation factor EF1β5a was mapped to 5q12-q14 as well as to chromosomes 5, 15, and X (Pizzuti et al. 1993). FISH studies have been used to map single cosmid and YAC clones to 5q13 and 5p14, as well as to distal 5q (Francis et al. 1993). In this study, the physical mapping of 12 newly characterized microsatellite markers, together with 9 previously reported markers, clearly localizes a region of repeat-element DNA to a genomic region

largely represented in YAC98 and bounded by the centromeric end of YAC 98 (Y98T) and the centromeric end of YAC112 (Y112T). If the repeat-element DNA is related to SMA disease etiology, then these physical boundaries should be considered together with the genetically defined disease gene boundaries as the most likely site for the SMA gene(s). A recent analysis indicates that these repeat elements are involved in the SMA disease etiology, presumably by a mechanism that increases genomic instability in the region (Melki et al. 1994). By the same reasoning, one would predict that the genomic region encompassing these repeat elements would be prone to deletion or rearrangement on amplification as clonal DNA in bacterial or yeast cells. Indeed, we have reported several regions of genomic instability within the "midi-YAC" contig (Kleyn et al. 1993; C. H. Wang and T. C. Gilliam, unpublished data), and Melki et al. (1994) report instability within their "mega-YAC" contig. Our preliminary analysis of genomic cosmid clones likewise indicates deletions within contiguous genomic DNA (C. H. Wang and T. C. Gilliam, unpublished observations). The physically mapped microsatellite markers reported in this study further limit the SMA minimal genetic region. They will be useful in evaluating deletion or rearrangement events in SMA families that might further pinpoint the disease locus.

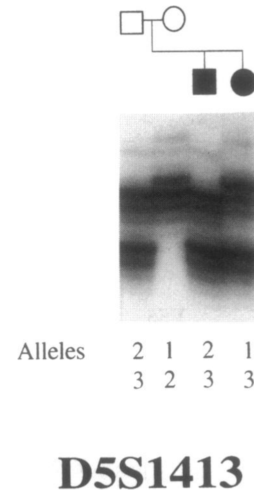

Figure 5 Meiotic recombination between D5S1413 and the SMA locus. "Recombinant" SMA family 1938 genotyped with the new centromeric flanking marker D5S1413. The affected children both inherit the "3" allele from father, but the firstborn child inherits a "2" allele from mother while the second-born child inherits the "1" allele.

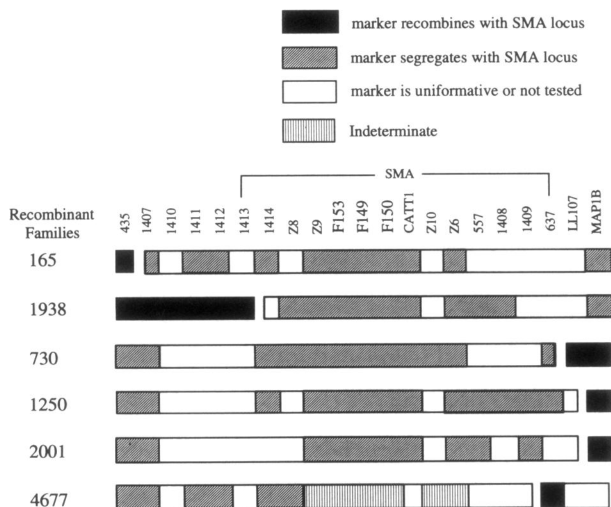

Figure 4 Summary of recombination events in the SMA region. Six "recombinant" families were genotyped with 22 microsatellite markers that span the ~0.7-cM region between the centromeric flanking marker D5S435 and the telomeric flanking marker MAP1B (Soares et al. 1993). Markers D5F153, D5F149, and D5F150 are described by Melki et al. (1994); CATT1 by Burghes et al. (1994); D5S557 by Francis et al. (1993); and D5D637 by Clermont et al. (1994). Families 165 and 1938 show meiotic breakpoints centromeric to the SMA locus. In family 1938, locus D5S1413 and the five centromeric markers recombine with the disease locus as indicated by the blackened rectangle. Markers distal to D5S1413 either segregate with SMA or are uninformative. The allele pattern for D5Z9 and family 4677 is illustrated in fig. 3*B*. In this family, the single-locus marker D5S637 clearly recombines with SMA, while six multilocus markers (D5Z9, D5F153, D5F149, D5F150, D5Z10, and D5Z6) all show "indeterminate" allelic differences in the two affected siblings (see fig. 2).

Acknowledgments

This work was supported by the Families of SMA (Chicago), Andrew's Buddies, Inc., the Muscular Dystrophy Association of America (to L.M.K. and T.C.G.), and NIH grants NS28877 (to T.C.G.) and NS23740 (to L.M.K.). C.H.W. is a recipient of an NIH Clinical Investigator Award, NSO1576, and receives support from the Colleen Giblin Foundation for Pediatric Neurology Research. L.M.K. is an associate investigator of the Howard Hughes Medical Institute. We thank Dr. Larry Deaven for the flow-sorted chromosome 5 cosmid library.

References

Albertsen HM, Abderrahim H, Cann HM, Dausset J, LePaslier D, Cohen D (1990) Construction and characterization of a yeast artificial chromosome library containing seven haploid human genomc equivalents. Proc Natl Acad Sci USA 87:4256–4260

Brzustowicz LM, Kleyn PW, Boyce FM, Lien LL, Monaco AP, Penchaszadeh GK, Das K, et al (1992) Fine-mapping of the spinal muscular atrophy locus to a region flanked by MAP1B and D5S6. Genomics 13:991–998

Brzustowicz LM, Lehner T, Castilla LH, Penchaszadeh GK, Wilhelmsen KC, Daniels R, Davies KE, et al (1990) Genetic mapping of chronic childhood-onset spinal muscular atrophy to chromosome 5q11.2-13.3. Nature 344:540–541

Brzustowicz LM, Wang CH, Matseoane D, Kleyn PW, Vitale E, Das K, Penchaszadeh GK, et al (1995) Linkage disequilibrium

and haploytype analysis among Polish families with spinal muscular atrophy. Am J Hum Genet 56:210–215 (in this issue)

Burghes AHM, Ingraham SE, McLean M, Thompson TG, McPherson JD, Kote-Jarai Z, Carpten JD, et al (1994) A multicopy dinucleotide marker that maps close to the spinal muscular atrophy gene. Genomics 21:394–402

Clermont O, Burlet P, Burglen L, Lefebvre S, Pascal F, McPherson J, Wasmuth J, et al (1994) Use of genetic and physical mapping to locate the spinal muscular atrophy locus between two new highly polymorphic DNA markers. Am J Hum Genet 54: 687–694

Francis MJ, Morrison KE, Campbell L, Grewal PK, Christodoulou Z, Daniels RJ, Monaco AP, et al (1993) A contig of nonchimaeric YACs containing the spinal muscular atrophy gene in 5q13. Hum Mol Genet 2:1161–1167

Gilliam TC, Brzustowicz LM, Castilla LH, Lehner T, Penchaszadeh GK, Daniels RJ, Byth BC, et al (1990) Genetic homogeneity between acute and chronic forms of spinal muscular atrophy. Nature 345: 823–825

Gilliam TC, Freimer NB, Kaufmann CA, Powchik PP, Bassett AS, Bengtsson U, Wasmuth JJ (1989) Deletion mapping of DNA markers to a region of chromosome 5 that cosegregates with schizophrenia. Genomics 5:940–944

Green ED, Olson MV (1990) Systematic screening of yeast artificial chromosome libraries by use of the polymerase chain reaction. Proc Natl Acad Sci USA 87:1213–1217

Kleyn PW, Wang CH, Lien LL, Vitale E, Pan J, Ross BM, Grunn A, et al (1993) Construction of a yeast artificial chromosome contig spanning the spinal muscular atrophy disease gene region. Proc Natl Acad Sci USA 90:6801–6805

Lien LL, Boyce FM, Kleyn P, Brzustowicz LM, Nenninger J, Ward DC, Gilliam TC, et al (1991) Mapping the human microtubule-associated protein 1B in proximity to the spinal muscular atrophy locus at 5q13. Proc Natl Acad Sci USA 88:7873–7876

Melki J, Abdelhak S, Sheth P, Bachelot MF, Burlet P, Marcadet A, Aicardi J, et al (1990*a*) Gene for chronic proximal spinal muscular atrophies maps to chromosome 5q. Nature 344:767–768

Melki J, Lefebvre S, Burglen L, Burlet P, Clermont O, Millasseau P, Reboullet S, et al (1994) De novo and inherited deletions of the 5q13 region in spinal muscular atrophies. Science 264: 1474–1477

Melki J, Sheth P, Abdelhak S, Burlet P, Bachelot M-F, Lathrop MG, Frezal J, et al (1990*b*) Mapping of acute (type 1) spinal muscular atrophy to chromosome 5q12-q14. Lancet 336:271–273

Munsat TL, Davies KE (1992) Workshop report: International SMA Consortium meeting. Neuromuscul Disord 2:423–428

Ott J (1991) Analysis of human genetic linkage. Johns Hopkins University Press, Baltimore

Pandolfo, M (1992) A rapid method to isolate $(GT)_n$ repeats from yeast artificial chromosomes. Nucleic Acids Res 20:1154

Pedeutour F, Szpirer C, Nahon J-L (1994) Assignment of the human pro-melanin-concentrating hormone gene (PMCH) to chromosome 12q23-q24 and two variant genes (PMCHL1 and PMCHL2) to chromosome 5p14 and 5q12-q13. Genomics 19: 31–37

Pizzuti A, Gennarelli M, Novelli G, Colosimo A, Cicero SL, Caskey CT, Dallapiccola B (1993) Human elongation factor EF-1β: cloning and characterization of the EF1β5a gene and assignment of the EF-1β isoforms to chromosomes 2,5,15 and X. Biochem Biophys Res Commun 197:154–162

Riley J, Butler R, Ogilvie D, Finniear R, Jenner D, Powell S, Anand R, et al (1990) A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. Nucleic Acids Res 18:2887–2890

Schwartz DC, Cantor CR (1984) Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell 37:67–75

Soares VM, Brzustowicz LM, Kleyn PW, Knowles JA, Palmer DA, Asokan S, Penchaszadeh GK (1993) Refinement of the spinal muscular atrophy locus to the interval between D5S435 and MAP1B. Genomics 15:365–371

Thompson TG, Morrison KE, Kleyn P, Bengtsson U, Gilliam TC, Davies KE, Wasmuth JJ, et al (1993) High resolution physical map of the region surrounding the spinal muscular atrophy gene. Hum Mol Genet 2:1169–1176