

4 Results

4.1 Clinical Data

The pedigree of the large family with pure spastic paraplegia is studied and shown in Figure 1. This family consists of 66 living members and the blood samples were collected from 42 participating members. HSP was diagnosed in 5 living members. One additional individual (IV:23) had minor symptoms (dorsal reflexes) and was thus diagnosed as presymptomatic. Ages of disease onset in this family vary from 12 years (patient IV:19) to 51 years (patient III:1). Brief case histories are presented below.

Patient III:1

This 60 year old man allegedly had multiple sclerosis since he was 51. Examination showed that reflexes were stronger on the left than on the right. He had mild paraparesis of the legs and a pathological right Babinski reflex, and positive Trömmner signs. There was no sensory disturbance, and no impaired bladder function. Eyes were not involved.

Patient III:12

This 56 year old man had frequently fallen in childhood and acute symptoms began when he was 25. Examination showed paraparesis of the lower limbs and bilateral Babinski signs. He was able to walk without a stick.

Patient III:14

This 57 year old man began to show the symptoms slowly since he was 46. Examination showed paraparesis of the lower limbs and bilateral Babinski signs. Function of his upper limbs was not disturbed and he was able to walk without a stick.

Patient IV:19

This 41 year old woman had gait disturbance since she was 12 years old. Examination showed increased muscular reflexes, spasticity in lower limbs and bilateral Babinski signs. She was able to walk without a stick.

Patient IV:23

This 25 year old woman had Trömmner but no Babinski signs when she was diagnosed as presymptomatic.

The patients of this family showed the common phenotype of spastic paraparesis and various ages of onset. Individuals of both sexes were affected. In addition, male to male transmission was observed (II:3 to III12 and III14). These results are consistent with the autosomal dominant inheritance of SPG in this family. Linkage analysis was performed with markers at the three loci (*SPG3*, *SPG4*, *SPG6*) known at the time 1997 in AD-HSP.

4.2 Genotype and linkage results

4.2.1 Two points lod score of the SPG4 family

Three sets of STRPs markers were chosen for linkage analysis in the large HSP family. The markers had been previously described to be linked to three *SPG* loci (*SPG3*, 4, 6). *D14S75*, *D14S79*, *D14S69*, *D14S266*, *D14S269*, *D14S66* and *D14S63* (*SPG3*). *D2S170*, *D2S146*, *D2S400*, *D2S352*, and *D2S367* (*SPG4*). *D15S128*, *D15S122*, *D15S156* and *D15S165* (*SPG6*). Primer sequences of the polymorphic microsatellite markers are given in Table 6. Two-point lod scores were calculated for markers on chromosome 2p, 14q and 15q. A maximum lod score of 3.63 at $\theta = 0$ was obtained with locus *D2S367* in the large AD-HSP family. Two chromosome 2p markers *D2S352* and *D2S367* show significant evidence of linkage (lod scores >3). Two-point lod scores are summarized in Table 7 and there is no evidence of linkage to chromosome 14q or 15q in ($Z < -2$) in this family. The haplotype analysis and the results of two point lod scores indicate that the disease locus in this AD-HSP family is linked to chromosome 2p (*SPG4*).

Table 7. Two -point lod scores of the SPG family vs chromosome 14q, 15q, 2p markers

SPG locus	Microsatellite	Lod scores Z at recombination fraction (θ)							Z_{\max}	θ_{\max}
		0,00	0,01	0,05	0,10	0,2	0,3	0,4		
SPG3	D14S75	-9,52	-3,92	-2,12	-1,26	-0,47	-0,13	-0,01	0,00	0,50
	D14S79	-12,94	-5,55	-3,04	-1,78	-0,61	-0,13	0,03	0,03	0,40
	D14S69	1,08	1,09	1,05	0,96	0,68	0,36	0,09	1,09	0,01
	D14S266	-7,50	-1,91	-0,96	-0,51	-0,15	-0,04	-0,01	0,00	0,50
	D14S269	-7,50	-1,47	-0,72	-0,51	-0,15	-0,04	-0,01	0,00	0,50
	D14S66	-1,01	-0,87	-0,56	-0,34	-0,13	-0,04	-0,01	0,00	0,50
	D14S63	-16,69	-6,61	-3,79	-2,43	-1,09	-0,42	-0,09	0,00	0,50
SPG6	D15S128	-0,53	-0,49	-0,36	-0,28	-0,18	-0,09	-0,02	0,00	0,50
	D15S122	-4,82	-3,40	-1,83	-0,98	-0,25	-0,02	0,00	0,00	0,50
	D15S156	-9,80	-2,90	-1,34	-0,65	-0,08	0,10	0,10	0,10	0,30
	D15S165	-9,45	-4,49	-2,36	-1,45	-0,62	-0,22	-0,04	0,00	0,50
SPG4	D2S170	$-\infty$	-4,55	-1,88	-0,88	-0,14	0,02	0,00	0,03	0,32
	D2S146	0,11	0,10	0,08	0,07	0,04	0,01	0,00	0,11	0,00
	D2S400	1,73	1,69	1,54	1,34	0,93	0,52	0,15	1,73	0,00
	D2S352	3,35	3,39	3,34	3,09	2,36	1,47	0,55	3,39	0,01
	D2S367	3,63	3,61	3,47	3,19	2,46	1,56	0,61	3,63	0,00

The maximum Lod scores at loci D2S352 and D2S367 were indicated in boldface.

4.2.2 Haplotype of the SPG4 family

The core haplotype includes the region flanked by *D2S146* and *D2S367*. The haplotype was constructed at 5 loci surrounding *SPG4* and is shown in Figure 1. The order of markers of haplotypes is consistent with the previously reported physical map of the region. A common haplotype was found in all 5 affected family members. Individual IV:23, who had very mild symptoms, has the same core haplotype and was considered presymptomatic by clinical criteria. A recombination event was found in IV:22. This reduces the critical interval of 3 cM between *D2S352* and *D2S367*, since the recombination occurred at locus *D2S352*. This recombination was also found by Dubé *et al.*⁴⁷ and Scott *et al.*³⁵. Individuals IV:1 and IV:3 may carry the disease alleles at loci *D2S352* and *D2S367* from their affected father III:1. If

they were carriers, they can not be confidently identified as non-penetrance, though the neurological examination showed asymptomatic at their ages of 30 and 26, respectively.

In section 4.2, two points lod scores and haplotype reveal that the disease locus of the large AD-HSP family is linked to chromosome 2p (SPG4). The lod score > 3 corresponds to the conventional $p < 0.05$ threshold of statistical significance for accepting linkage. Linkage can be rejected if the lod score < -2 . The lod scores between -2 and $+3$ are inconclusive.

4.3 YAC Contig Mapping

To facilitate positional cloning of the *SPG4* gene, a contiguous YAC map which covered the *SPG4* critical region was established.

4.3.1 YAC Contig Spanning the *SPG4* Critical Region between D2S400 and D2S367

The human DNA insert of 26 YACs that had been assigned to 2p21-24 were fine mapped in order to determine whether the YACs are located within the *SPG4* candidate region. The YACs were examined by PCR for presence of the sequences tagged sites (STSs) D2S400, D2S2255, stSG4150, SHGC12567, stSG1707, D2S352, D2S2283, D2S2203, D2S3008, D2S2351, WI9534, SGC32499, SGC32192, WI9668, WI11827, WI16901, SHGC13568, A004H42, stSG3789, WI16788, GAAT-P33068, D2S1794, D2S2347, D2S2005, WI9249, D2S1998, SGC33436, D2S367 which were previously located in the candidate region (Hudson *et al.*⁴⁸). The sequences of STSs primers are listed in Table 6. A total of 24 YACs were positive for one or more STSs, While 15 YACs contain multiple STRs. Compared to the previously described contig (De Jonghe *et al.*⁴⁹), ten additional YACs (715F10, 737D8, 774D8, 785G12, 823C6, 851F4, 884D5, 894E4, 923E10, 947E2) and six STRs (*D2S2255*, *D2S2283*, *D2S2203*, *D2S3008*, *D2S2351*, *D2S2347*) were newly assigned to the YAC contig (see Figure 2).

4.3.2 13 ESTs were newly assigned to the *SPG4* critical region

A total of 13 ESTs were mapped to the YAC contig out of 27 ESTs which had been assigned to 2p21-p24 (<http://www.ncbi.nlm.nih.gov/SCIENCE96/>). Only one of those ESTs, SHGC12567, was found to be part of a known gene, i.e. human xanthine dehydrogenase (100% homology). The remaining ESTs did not have striking homologies with known genes. The YACs spanned the complete candidate region between *D2S400* and *D2S367* (Lau *et al.*⁵⁰). The order of the STRs are consistent with the physical map of Hudson *et al.*⁴⁸. The density of clone coverage is variable along the YAC contig, and ranges from 2 clones at *WI9249* to a maximum of 10 clones at *WI9534* and *D2S2351* (see Figure 2).

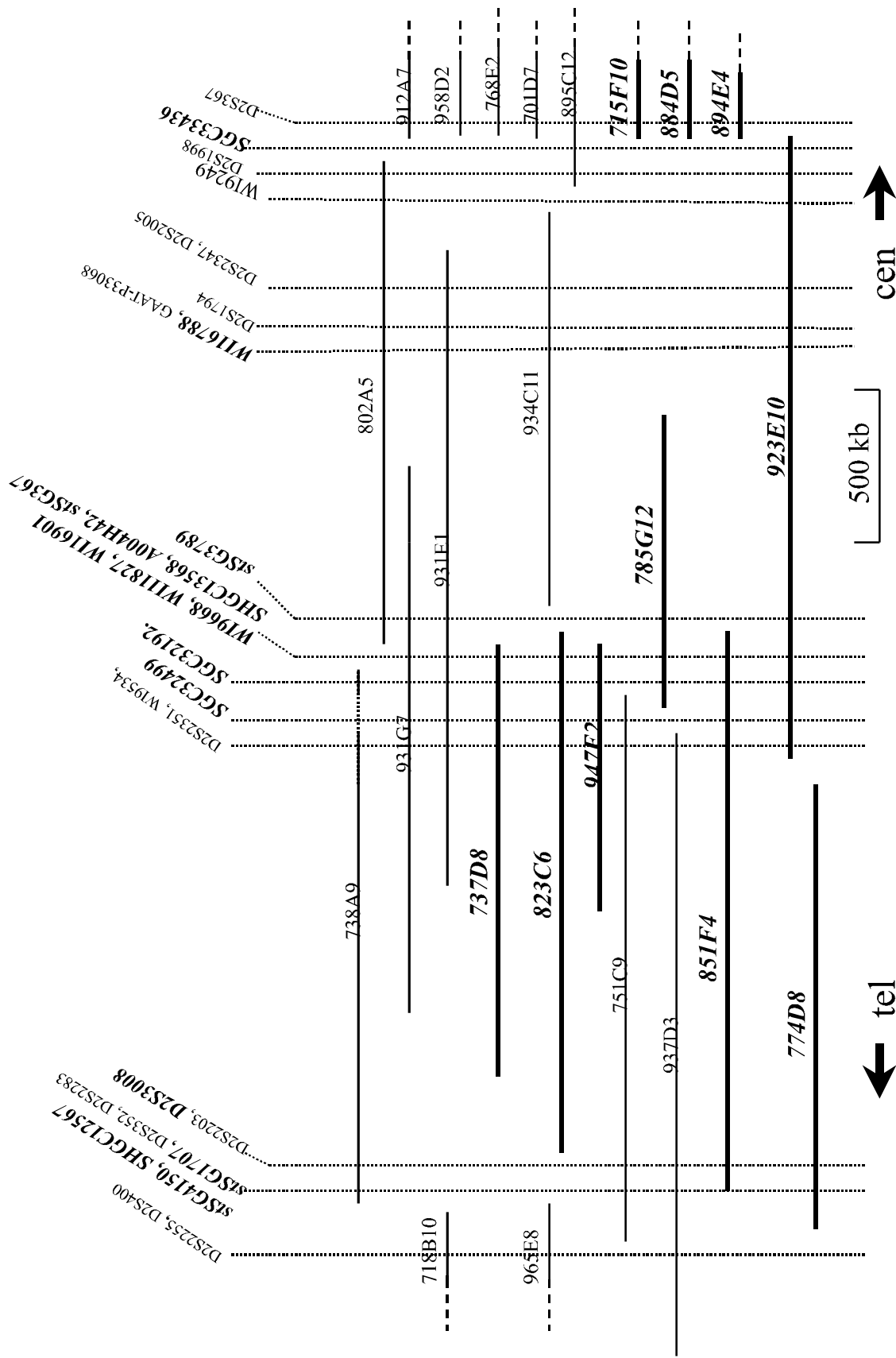


Figure 2. YAC contig of the *SPG4* critical region. Bars indicating newly assigned YACs are printed in bold and newly assigned ESTs and the STRP *D2S3008* are printed in *italics* and bold.

4.3.3 The size of the *SPG4* critical region

The sizes of YACs of the contig were determined by PFGE and are listed in Table 8. The physical size of ~3.6 Mb was estimated for the *SPG4* critical region on the basis of 2 YACs: 937D3 and 923E10. This result is very similar to another estimation reported by Hazan *et al.*³⁷, who constructed a 3.5 Mb YAC contig between *D2S367* and *D2S400*. The overestimation of the physical size of *SPG4* locus was due to the chimerism of YACs in the previous study⁴⁹. Supposed that *SPG4* critical region is 3.6 Mb, the average physical distance of every two STSs in the YAC contig is 124 kb

Table 8. YAC sizes

YAC	Size (kb)
701D7	750
715F10	1200
718B10	1250
737D8	1200
738A9	1050
751C9	1800
774D8	1200
785G12	800
802A5	400;1800
823C6	1600
851F4	1600
884D5	1050
894E4	1000
895C12	1500
912A7	1000
923E10	1800
931E1	1800
931G7	1550
934C11	1600
937D3	1800
947E2	650
958D2	1200
965E8	1600

Note: Different sizes of the same YACs are probably due to *in vivo* truncation of an ancestral YAC

Of 19 STSs tested on the 17 YAC clones, 40 YAC/STS hits reported in contig WC2.3 established by the Whitehead Institute were confirmed and 54 new hits were assigned (see Table 9). One false-positive hit, YAC 774D8 at locus D2S400 was found in CEPH database and one false-negative, YAC 923E10 at locus *D2S2351*, was found in Hazan's publication³⁷. This result is consistent with the estimation of Hudson *et al.*⁴⁸ in which the false positive rate is at the maximum 5% in definite addresses and the 20% average false negative rate is suggested. The false positive and negative signals may result from the recombinations and deletions during the YAC regrowth or technical failures.

Table 9. Newly assigned STSs content of YACs compared with the Whitehead Institute/MIT contig WC 2.3 database

	718B10	937D3	965E10	738A9	751C9	774D8	851F4	931G7	931E1	737D8	823C6	947E2	785G12	923E10	802A5	934C11	895C12
D2S400	v	v	v														
D2S2255	+	+	+														
D2S352		+		v	v	+											
D2S2283		v		v	v	v											
D2S2203		+		+	+	+	+										
D2S3008		+		+	+	+	+										
D2S2351		v		v	+		+	+	v	v	+	v		+			
WI9534		v		v			v	v	v	v	v	+		v			
WI9668							v	v	v	v	+	v	v	+			
WI1827							+	+	+	+	+	+	+	+			
WI16901							+	+	+	+	+	+	+	+			
WI16788									+	+				+	+	+	
GAAT-P33068									v					+	+	v	
D2S1794									v					v	v	v	
D2S2347									+					+	+	+	
D2S2005									+					+	v	v	
WI9249														v	+		
D2S1998														v	v		v
D2S367																	+

Note: YACs are on the top, and STSs are on the left. "+", newly assigned STSs; "v", verified hits.

4.4 Search for CAG /CTG Repeats in the *SPG4* Critical Region 2p21-p24

Several SPG4-linked families have exhibited anticipation which is a hallmark of trinucleotide repeat expansion diseases. Nielsen *et al.*⁵¹ reported (CAG)_n trinucleotide repeats expanded in their SPG4-linked patients by the RED (repeats expansion detection) method⁵². Therefore, to clone the CAG/CTG repeat fragments in *SPG4* critical region was once the strategy to explore the disease mechanism.

4.4.1 Construction of Cosmid Libraries within *SPG4* Critical Region

Cosmid libraries were constructed from 4 YACs (738A9, 937D3, 923E10, and 895C12) spanning the *SPG4* candidate region. These libraries are composed of between 2000 and 2500 clones and were screened with human genomic DNA. Subsequently, positive clones were hybridized with a (CAG)₁₀ oligonucleotide to find those fragments which contain CAG/CTG repeats. The outline of the cloning procedures is shown in Figure 3.

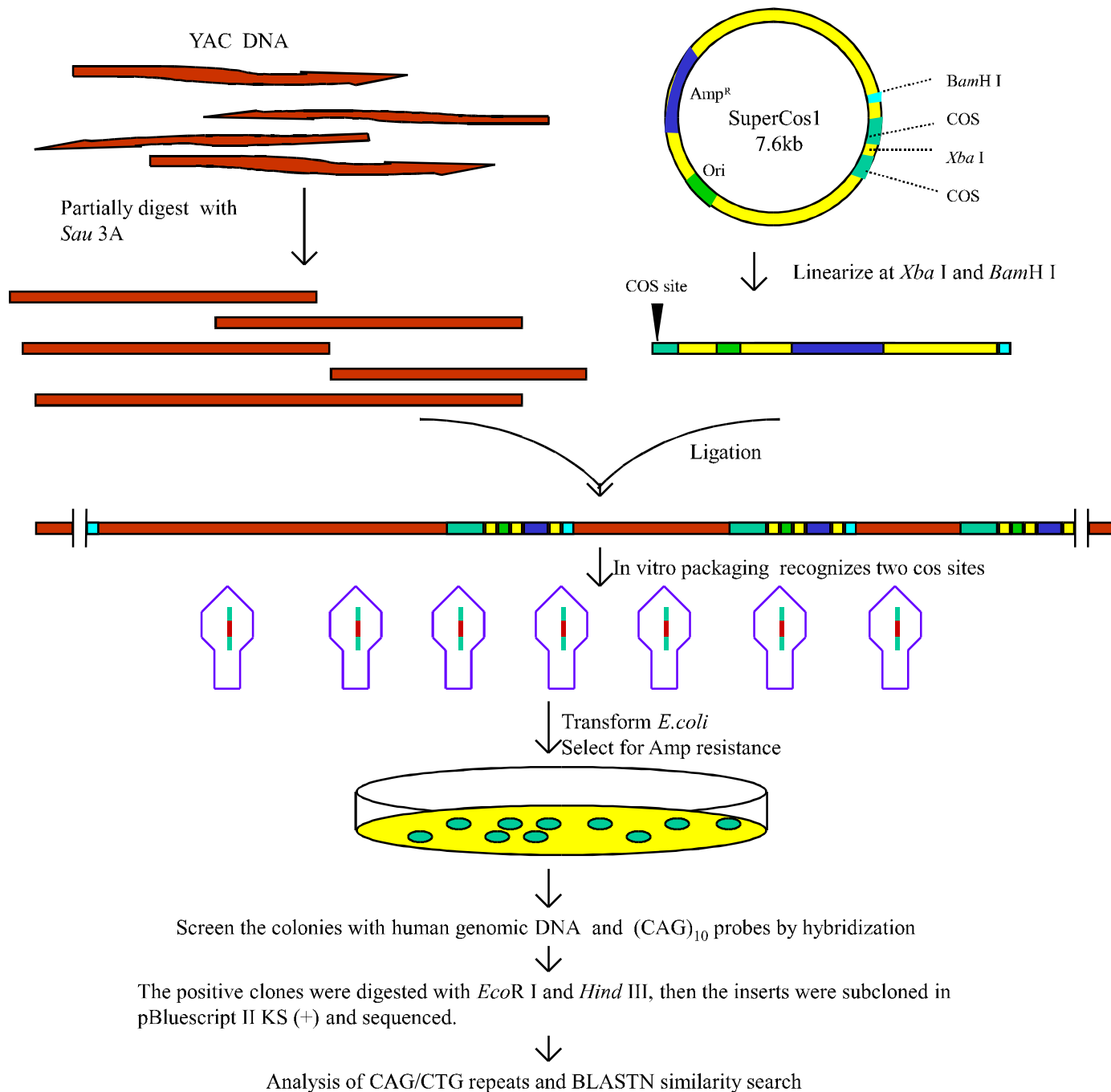


Figure 3. General procedures of isolating CAG/CTG triplet repeat from YAC DNA. SuperCos 1 carries cos sites from the λ phage, plasmid origin of replication and ampicillin resistant gene (Amp^R). To clone YAC DNA into vector SuperCos 1, the vector is linearized with *Xba* I and *Bam*H I, and YAC DNA is partially digested with *Sau* 3A which leaves *Bam*H I compatible ends. Digested YAC DNA fragments are ligated to linearized vector. A λ packaging extract recognizes and packages any ligated DNA flanked by two cos sites. The segments are transformed to *E. coli* and screened with human DNA and (CAG)₁₀ probe. The positive clones are digested with *Eco*R I and *Hind* III and subcloned to pBluescript II KS (+) for sequencing. The sequences are compared with Genbank database by BLASTN similarity search.

4.4.2 Isolation of Six CAG/CTG Repeat Containing Fragments

A total of 6 sets human CAG/CTG repeats containing cosmids were isolated from 3 YACs (738A9, 937D3 and 923E10). YAC 895C12 does not contain CAG/CTG repeats. The clones positive for human DNA and CAG repeats were cleaved with *EcoR* I and *Hind* III (Figure 4) and subcloned into plasmid pBluescript II KS (+). The subclones were sequenced by MWG CL4200 autosequencer.

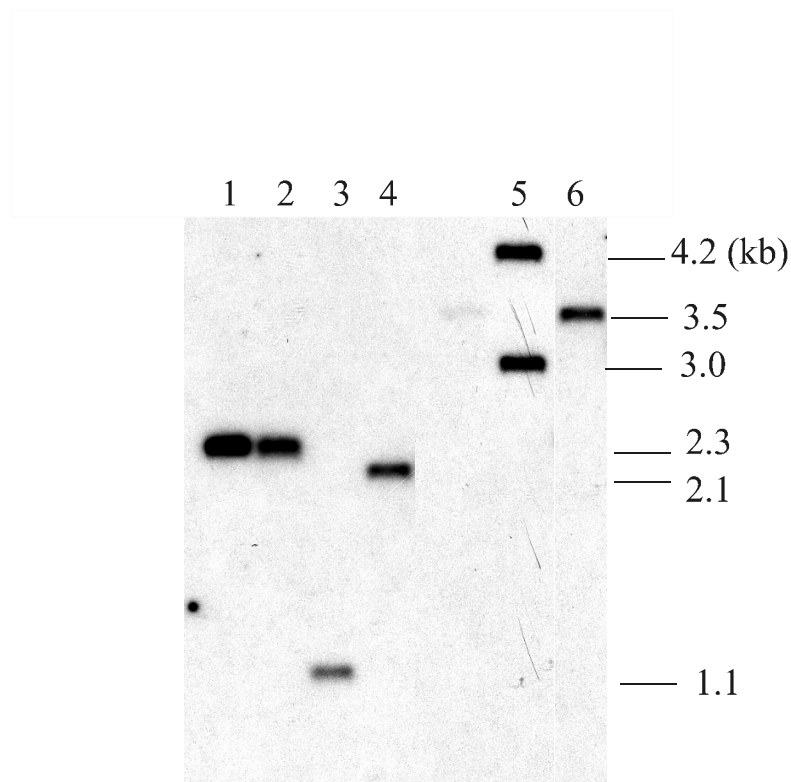


Figure 4. Southern hybridization analysis of cosmid clones. Lanes 1-6 contain clones D 1 , D 4 , D 5 , D 6 , E 1 1 , A 8 . These clones were selected from cosmid libraries by hybridization with human DNA and digested with *EcoR* I/*Hind* III. The digested DNA was electrophoresed on a 1% agarose gel, transferred to a membrane and hybridized with (CAG)₁₀ probe. Sizes of the fragments are indicated on the right. These (CAG)_n positive fragments were subcloned into pBluescript II KS (+) and sequenced. Note: The 3.0 kb fragment does not hybridize with human DNA.

4.4.3 Identification of One New STRP for Indirect Diagnosis

One (CAG)₃ containing fragment, denoted clone A8, was derived from YAC 738A9. Sequence analysis shows that it also includes a (GAAA)_n(GGAA)_n tetranucleotide repeat. The tetranucleotide repeats were amplified using the flanking primers 5'-CCAGCAGGCAGAGTGAGAATC-3' (forward) and 5'-TGTTTAAAGCACAAGGGCAGT-3' (reverse). PCR conditions were 30 cycles at 94°C for 30 sec, 63°C for 60 sec, and at 70°C for 15 sec. The reaction was preceded by 5 min denaturation at 94°C and followed by a final extension at 70°C for 5 min. The sequence was found to be highly polymorphic and a total of 28 alleles ranging in size from 224 to 292 bp were identified in 50 Caucasian DNA samples (see Table 10).

The locus, designated *D2S3008*, was flanked by loci *D2S352* and *D2S2283* distally and *D2S2351* proximally (see Figure 2. Lau *et al.*⁵⁰). The marker has a heterozygosity of 0.81 and a PIC (polymorphism information content) value of 0.78. The heterozygosity (*H*) which refers the degree of polymorphism is defined by $H = 1 - \sum p_i^2$, where p_i is the population frequency of the *i*th allele, and *H* is the probability that a random individual is heterozygous for any two alleles at a gene locus with allele frequencies, p_i . The PIC value is defined as the probability that the marker genotype of a given offspring will allow deduction of which of the two marker alleles of the affected parent it had received. PIC is calculated as $PIC = 1 - \sum p_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2p_i^2 p_j^2$, where p_i is the population frequency of the *i*th allele.

Table 10. Allele sizes at locus D2S3008 in the Caucasian population; 100 chromosomes 2 were studied. PIC: 0.78, heterozygosity: 0.81

Allele	Allele sizes (bp)	Allele frequency
1	224	1
2	229	1
3	230	1
4	232	3
5	234	3
6	236	6
7	238	3
8	240	4
9	243	7
10	245	5
11	247	5
12	249	6
13	250	5
14	252	3
15	254	5
16	256	4
17	258	5
18	260	4
19	262	3
20	264	4
21	268	4
22	270	6
23	273	3
24	275	1
25	278	2
26	280	4
27	283	1
28	292	1

In addition, a 254 bp allele at *D2S3008* segregated in the patients of the SPG4 family (see Figure 5). This highly polymorphic STRP can be used as indirect diagnosis of SPG4.

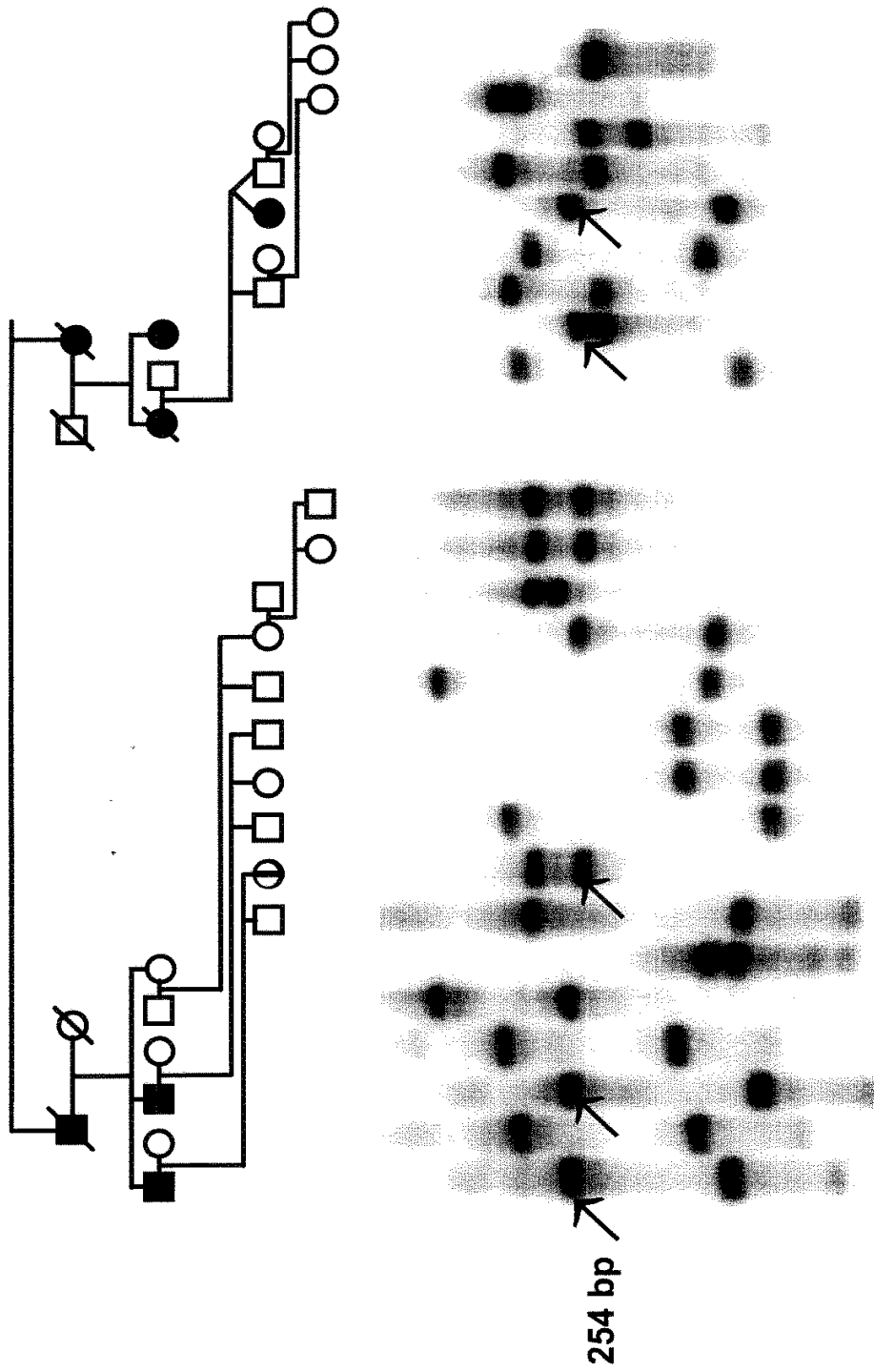


Figure 5. Segregation of alleles at *D2S3008* in a large SPG4 family. The 254 bp allele is found in a patient's daughter (IV:23) who has very mild symptoms at the age of 20 years.

4.4.4 Development of Five New STS and Anchoring in the *SPG4* Critical Region.

One CAG triplet rich fragment, clone D1, was isolated from YAC 937D3 and this clone contains a sporadic single CAG unit. Two (CAG)₂ repeats fragments denoted clones D5, D6 were also isolated from YAC 937D3. Two (CTG)₆-containing fragments were isolated (one denoted clone D4 from YAC 937D3, and one denoted clone E11 from YAC 923E10).

A BLASTN similarity search (Altschul *et al.*⁵³) with the sequence of each CAG/CTG containing fragments against GenBank division's dbEST and dbSTS, showed identifies for sequences of clones A8, D1, D4, D5, D6 and E11 (see Table 11). Clone A8 matched to BAC 41M14 of human CITB_978_SKB library with 96% sequence identity. Clone D1 showed 98% identity with human BAC RP11-62F14 (accession number AC009301). Clone D4 showed 97% identity with BAC RP11-62F14. Clone D5 showed 98% identity to the human chromosome 2 working draft sequence located in the cluster Hs2_5351. Clone D6 was 99% identical to human BAC RP11-62F14. Clone E11 was 99% identical to the human chromosome 2 working draft sequence located in the cluster Hs2_5351. These six clones which contain CAG repeats did not have striking homologies with known genes.

Table 11. Summary of the both human DNA and (CAG)_n positive fragments from YACs 738A9, 923E10, and 937D3 within *SPG4* locus

YAC	Clone	Fragment length (kb)	CAG/CTG repeat units	BLASTN analysis
738A9	A8	3.5	3	human BAC clone 41M14 of CITB_978_SKB library
923E10	E11	4.2	6	human chromosome 2 working draft cluster Hs2_5351
937D3	D1	2.3	1	human BAC clone RP11-62F14
	D4	2.3	6	human BAC clone RP11-62F14
	D5	1.1	2	human chromosome 2 working draft cluster Hs2_5351
	D6	2.1	2	human BAC clone RP11-62F14

Furthermore, five new STSs were developed on the basis of clones D1, D5, D6, D4, E11. They are denoted by *cag1c1p*, *cag2c5p*, *cag2c6p*, *ctg6c4*, *ctg6c11* and were anchored in the *SPG4* candidate region (Figure 8). The new STSs and *D2S3008* are listed in Table 12.

Table 12. Newly developed STSs in this study

Name	GenBank Accession number	Primer	Sequence	Product size (bp)
<i>cag1c1p</i>	G67940	cag1c1p.a cag1c1p.b	GCCCCAGAAACAGCCAAATAA CGCCAGCAACGCCAGTC	385
<i>cag2c5p</i>	G67941	cag2c5p.a cag2c5p.b	TGGGGAGTGGTTTAGTAGCAGAAG AGCGACATACATACAAGGGGAAAAA	250
<i>cag2c6p</i>	G67942	cag2c6p.a cag2c6p.b	AACAGCCCCCTATCACATTCCAGA GGCACATCCCAGCATTTAGCAGTA	332
<i>ctg6c4</i>	G67794	ctg6c4.a ctg6c4.b	TCTACACATCCCATTGCTTTCT TAGCCTCTCCGAGCCTTGGTT	256
<i>ctg6c11</i>	G67795	ctg6c11.a ctg6c11.b	GGGCGCGAGACACAGGATTC TGACAAGTGGTGCGATTTATT	422
<i>D2S3008</i>	AF095704	D2S3008.a D2S3008.b	CCAGCAGGCAGAGTGAGAATC TGTTTAAAGCACAAGGGCAGT	224-292

4.4.5 Analysis of the CAG/CTG Repeats in SPG4 Patients

Since the CAG/CTG repeats involved in neuron degenerative diseases are usually more than 3 repeat units in coding region, we used two cloned sequences which included (CTG)₆ to generate PCR primers for repeats expansion analysis. The alleles were amplified with PCR in which [α^{32} P]-dCTP was incorporated into the synthesized PCR fragments. The radioactive labeled products were electrophoresed on denaturing polyacrylamide gels. The controls and patients have 256bp and 422bp homozygous alleles at loci *ctg6c4* and *ctg6c11*, respectively. None of (CTG)₆ repeats are expanded in affected members of the SPG4 linked family. In addition, no differences were observed in the number of repeat units between human and YAC DNA (see Figure 6 and 7).

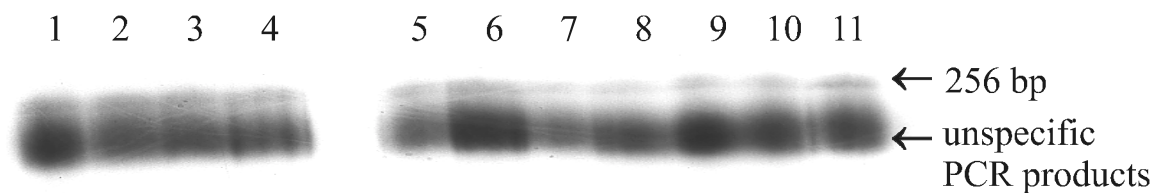


Figure 6. Allele expansion analysis of SPG4 patients at *ctg6c4*. Lanes 1-10 represent DNAs from the SPG4 family members III17, III9, IV16, III7, III4, III12, III14, III11, IV19, III1, respectively. Lane 11: YAC937D3. Lanes 1-5 contain DNAs from unaffecteds and Lanes 6-10 DNAs from patients. A 256 bp homozygous allele is detected in both controls and patients. There is no evidence of allele expansion.

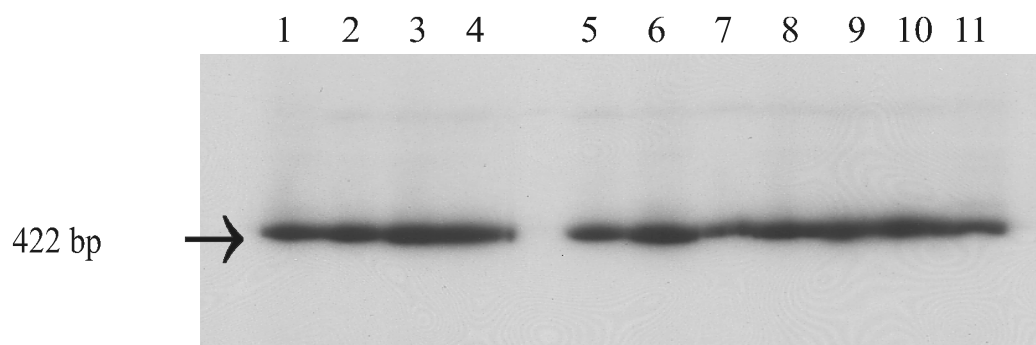


Figure 7. Allele expansion analysis of SPG4 patients at *ctg6c11*. Lanes 1-10 contain DNAs from SPG4 family members III17, III9, IV16, III7, III4, III12, III14, III11, IV19, III1, respectively. Lane11: YAC923E10. Lanes 1-5 contain DNAs from unaffecteds and lanes 6-10 contain DNAs patients. A 422bp homozygous allele is detected in both controls and patients. There is no evidence of allele expansion.

4.5 BAC/PAC Contig Mapping

The failure to identify an expanded CAG/CTG repeat within the *SPG4* candidate region may be caused by a deletion in one of the YACs or high rates of chimerism of YACs used to construct the cosmid libraries. Therefore, we started to construct a BAC/PAC contig.

4.5.1 Updating the STSs to the Contig

To update ESTs and also increase the density of ESTs in the *SPG4* critical region, we searched the databases Gene Map 98 and 99 during the experiment (see Table 13). These selected ESTs from databases were first tested with the YACs in the contig. Twenty ESTs were newly assigned to the *SPG4* critical region. Together with the previously mapped ESTs which were selected from Gene Map 96 in the YAC contig, a total of 36 ESTs and 12 STRs were integrated in the *SPG4* candidate region between *D2S367* and *D2S400*. Primer sequences of ESTs and STRs are shown in Table 6. In addition to SHGC12567 which had 100% homology with human xanthine dehydrogenase⁵⁴, three ESTs were found to be parts of known genes. N52847 was identical with latent transforming growth factor beta binding protein 1 (LTBP1)⁵⁵ and stSG 4150 and stSG1707 were a part of human EH-domain containing protein 3 (EHD3)⁵⁶.

Table 13. YAC contig integrated with ESTs from Gene Map 98 and 99

YAC	EST	stSG30783	stSG2515	stSG44150	stSG15464	stSG47880	stSG51853	stSG62506	stSG32067	stSG22421	stSG52124	stSG63103	stSG48207	stSG29624	stSG31920	Cda0sd05	Cda15g09	A008N33	N52847
718B10	+																		
737D8		+	+	+	+	+	+	+	+	+	+					+	+	+	
738A9		+	+													+		+	
751C9		+	+	+	+		+	+											
785G12							+	+	+	+	+						+		
802A5													+	+	+				+
823C6		+	+	+	+	+	+	+								+	+	+	
851F4		+	+	+	+	+	+	+	+	+	+					+		+	
923E10		+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
931E1		+	+	+	+	+	+	+	+	+	+	+		+	+	+	+	+	
931G7		+	+	+	+	+	+	+			+	+	+			+	+	+	
934C11													+	+					
937D3	+	+	+																
947E2			+	+	+	+	+	+	+	+						+	+	+	
965E8	+																		

Note: YACs are on the left and Ests are on the top. "+", PCR positive.

4.5.2 Assembly of Several BACs and PACs in *SPG4* Critical Region

The BAC library RPCI-11, the two PAC libraries RPCI-1 and RPCI-6, and the chromosome 2-specific PAC library LL02NP04'AI' were screened by hybridization with PCR products for filter libraries or directly screened by PCR using DNA pool libraries against the STSs localized in the *SPG4* critical region between D2S400 and D2S367. A total of 52 BAC/ PAC clones which comprised 19 BACs and 33 PACs were assigned to the *SPG4* candidate region (see Figure 8).

4.5.3 *SPAST* is located in the BAC bG11100

While this contig construction was going on, Hazan *et al.*^{36,37} published 1.5 Mb sequences in the *SPG4* critical region and identified five mutations of *SPG4* patients in a novel gene named *SPAST*. Since Hazan found that *SPAST* is very close to the locus *D2S2351*, the adjacent BACs and PACs in the contig were screened for *SPAST*. Several exons of *SPAST* (exon 1, 2, 11, 12, 16, 17) were tested against those adjacent BACs and PACs (59b24, 69d15, 122b3, N15143, bA16130, bK08201, bP12129, bP13129, bB07176, bC11100, bG11100) by PCR. The PCR primers are shown in Table 4. The result reveals that *SPAST* is located in BAC clone bG11100 (see Figure 8)

4.5.4 Sizes of BACs and PACs

Sizes of several PACs and BACs were determined by PFGE (see Table 14). The contig built in this dissertation is about 2.2 Mb, of which 1.2 Mb is located in SPG4 minimum critical region. This estimation takes into account the sizes of these clones P17234, A0827,105d11, bK13144, bP13129, 59b24, bG11100, B043, G1183, C1011, P2046, H19197, J17106, bP09265, 018191, 126m18, bA2313. The contig covers 80% of *SPG4* minimum region, however there is still 9 gaps in the whole contig (see Figure 8).

Table 14. PACs and BACs sizes

Clone	Size(kb)	Clone	Size(kb)
PACs:		BACs:	
RPCII-59b24	90	RPCIB753A2313	150
RPCII-61d7	20;90	RPCIB753A16130	100
RPCII-69d15	90	RPCIB753A21163	130
RPCII-105d11	150	RPCIB753B07176	120
RPCII-113e2	120	RPCIB753C11100	120
RPCII-113f2	110	RPCIB753D16113	150
RPCII-118c4	150	RPCIB753E16113	140
RPCII-122b3	120	RPCIB753G0743	100
RPCII-126m18	170	RPCIB753G11100	120
RPCII-151k12	150	RPCIB753K13144	145
LLNLP708B043	150	RPCIB753K08201	100
LLNLP708C1011	90	RPCIB753L0314	160
LLNLP708E157	40;50	RPCIB753P09265	170
LLNLP708I213	30;70	RPCIB753P12129	100
LLNLP708J213	30;70	RPCIB753P13129	100
LLNLP708K165	40;50		
LLNLP709A0827	50;70		
LLNLP709A14127	150		
LLNLP709A23175	90		
LLNLP709B0240	30		
LLNLP709E13129	100		
LLNLP709F0733	120		
LLNLP709G1138	50;60		
LLNLP709G04178	120		
LLNLP709H19197	150		
LLNLP709J2428	120		
LLNLP709J17106	170		
LLNLP709M1711	110		
LLNLP709N15143	100		
LLNLP709O0557	120		
LLNLP709O18191	170		
LLNLP709P2046	90		
LLNLP709P17234	120		

Different sizes of the same PACs are probably due to an internal *Not* I restriction site of insert DNA

4.6 Mutation detection of SPG4 patients

The 17 exons of *SPAST* were amplified by PCR from genomic DNA of patients and controls. The primers used for PCR amplification and sequencing are shown in Table 4. Sequence analysis was done with terminators sequencing method using P³³-dd NTP.

4.6.1 Identification of A Novel Mutation 1206-1209del CCTT

One novel mutation 1206-1209delCCTT (Pro361Leu and Ser362Stop) was identified in exon 7 of *SPAST* gene (see Figure 9). The same mutation was not observed in ten independent controls or unaffected members of this family. The mutation resulted in a truncated spastin which lacks three functional domains, the Walker motif A, Walker motif B, and the AAA minimal consensus sequence. The Walker motif A located at amino acid positions 382-389, also called p-loop, corresponds to the ATPase binding domain. The Walker motif B and the AAA minimal consensus are located in amino acids 437-442 and 480-498, respectively.

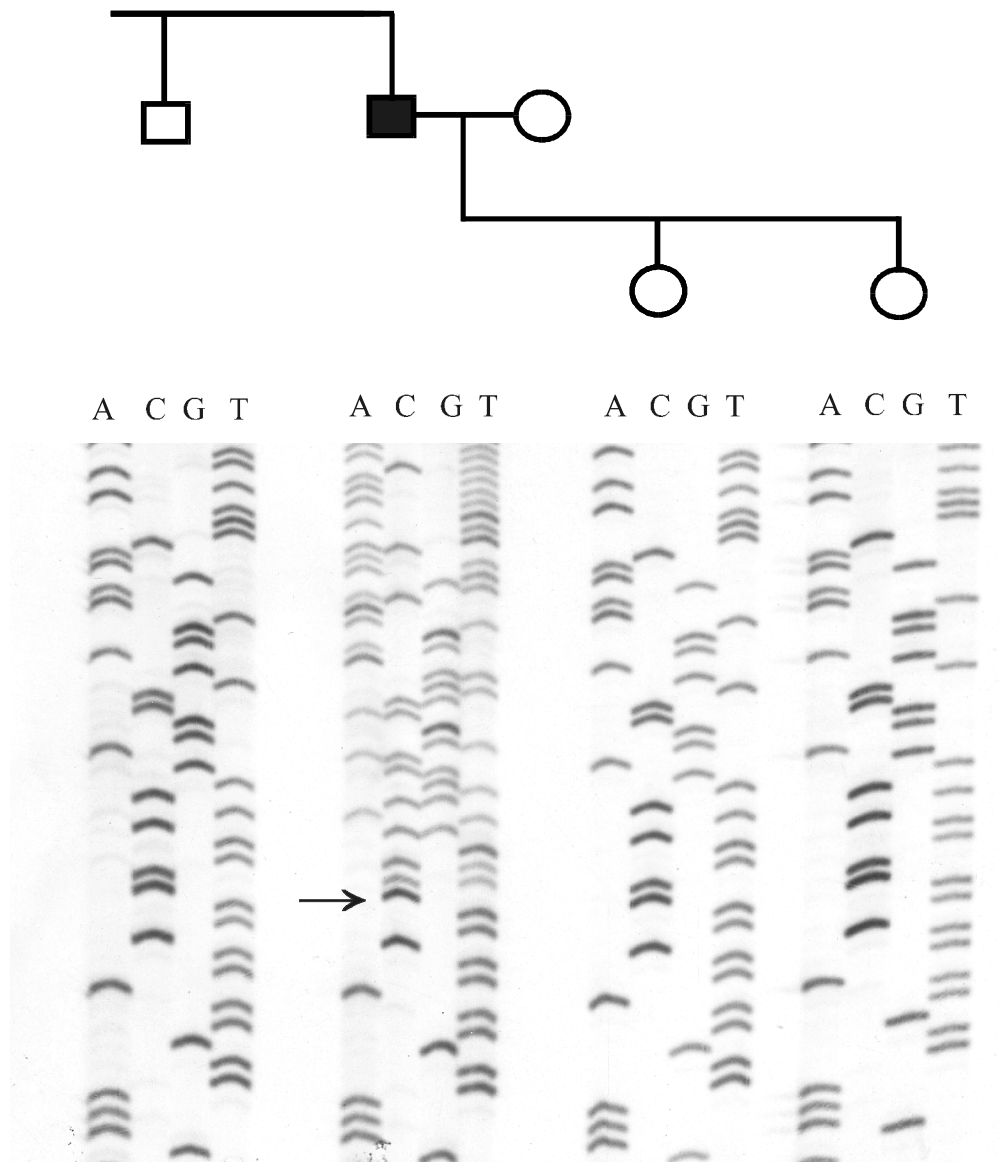


Figure 9. A novel mutation of *SPAST*, 1206-1209del. The SPG patient has a CCTT deletion in exon 7. His brother and two daughters do not carry this mutation. The arrow indicates the start of the deletion.

4.6.2 Identification of An Intronic polymorphism 1298+17A>C

A polymorphism 1298+17A>C was found in intron 8 of *SPAST* in one SPG affected patient (see Figure 10). A total of 100 unrelated individuals were tested against this polymorphism. The same polymorphism was coincidentally found in one unrelated control (Figure 10,11). This polymorphism was amplified using the flanking primers 5'-ACCTGGGAATGGGAAGACA-3'(forward) and 5'-AATAGACTCAAGGACAAGAT-AAAG-3' (reverse). The PCR conditions were 30 cycles at 94°C for 30 sec, 60°C for 60 sec, and at 70°C for 15 sec. The reaction was preceded by 5 min denaturation at 94°C and followed by a final extension at 70°C for 5 min.

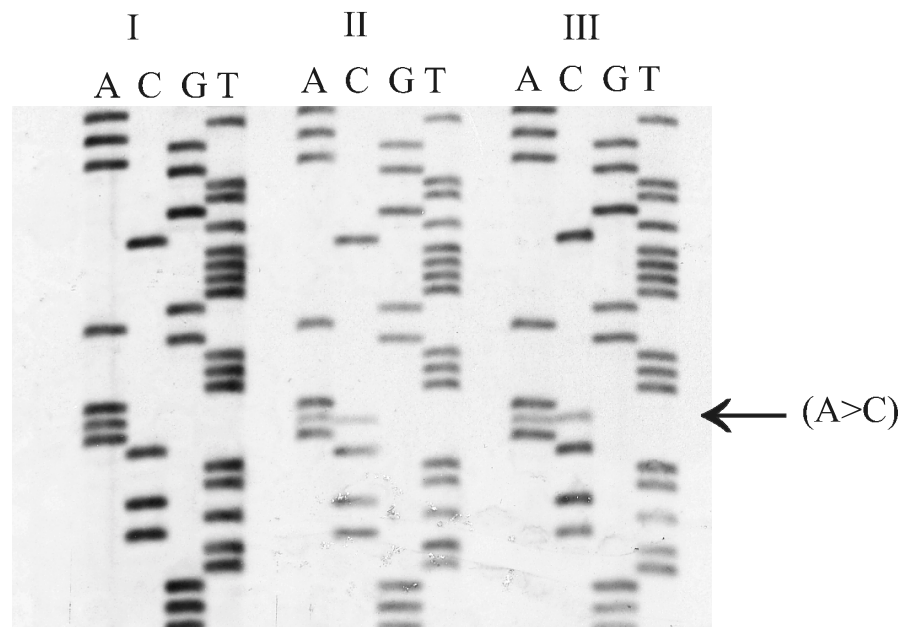


Figure 10. Sequence analysis of the polymorphism in intron 8 of *SPAST*. I: unaffected control. II: a spastic paraplegia patient. III: a control. The arrow indicates the base substitution 1298+17 (A>C) in II and III. This polymorphism was amplified by PCR and the PCR products were directly sequenced.

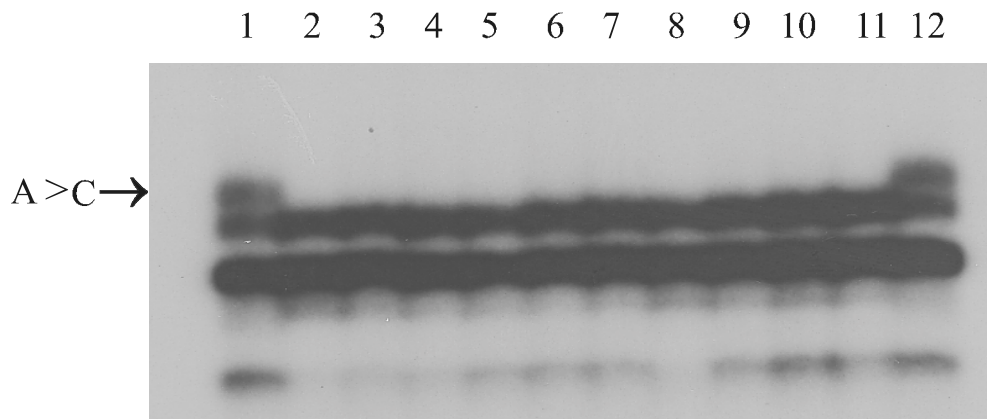


Figure 11. SSCP analysis of polymorphisms in intron 8 of *SPAST*. Lane 1: spastic paraplegia patient. Lane 2-12: unrelated controls. The arrow indicates the band shift due to A>C. However, this polymorphism was not detected in an additional 100 unrelated controls and nine spastic paraplegia patients. The polymorphic fragments were amplified by PCR and electrophoresed on 0.5 x MDE gels at 4°C overnight.

The SPG4 family was screened in all 17 exons, intron 6, 10, 11, the 5' and 3' untranslated region. The sequences of PCR primers for *SPAST* amplification and sequencing are listed in Table 4. However, there was no mutation found in these regions. Furthermore, RT-PCR was performed for this SPG4-linked family. One patient's blood sample was prepared for RNA isolation and subsequent cDNA synthesis. Five pairs of PCR primers were adapted from Genoscope database to amplify four overlapping PCR products in *SPAST* cDNA. These primers are shown in Table 3. The first PCR fragment flanked by SPA-Db/Dm and SPA-Dc/Dn can not be amplified in our samples. This may be due to degradation of RNA or a high GC content in exon 1. The other three PCR overlapping fragments were amplified. However, an aberrant PCR product amplified by SPA-Ba/Bm was detected in this patient and two normal lab-control samples. Exon 8 is skipped in this aberrant PCR product (see Figure 12 and 13). This aberrant product may result from the degradation of RNA which was isolated from the post delivered blood sample. In the control samples, the same splicing artifact was observed after blood samples were placed for 10 hr at room temperature.

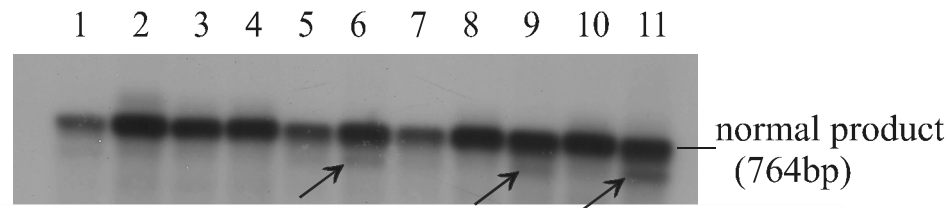


Figure 12. RT-PCR analysis of *SPAST*. Lanes 1-10: normal control samples. Lane 11: a patient from the large SPG4 family. The arrows indicate aberrant PCR products in this patient and two controls. The normal and the aberrant PCR products were sequenced. The results reveals that exon 8 is deleted in this aberrant PCR product. The PCR products were electrophoresed on 1.5% agarose gel, transferred to a membrane and hybridized with [γ - P^{32}]ATP labeled exon 7.

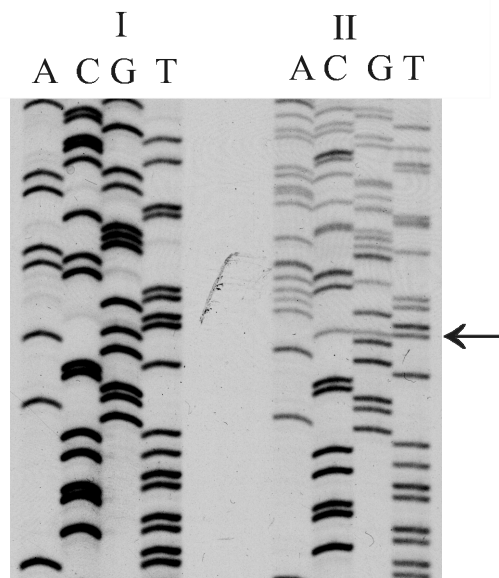


Figure 13. Sequence analysis of RT-PCR products amplified by SPA-Ba/m. I: normal PCR product, II: aberrant PCR product. Sequences reveal that exon 8 is truncated in II. The arrow indicates the first nucleotide of exon 8 in *SPAST* cDNA.