Aus dem Institut für Humangenetik der Justus-Liebig-Universität Gießen

# Molecular Analysis of the Autosomal Dominant Spastic Paraplegia Type IV (SPG4)

Inaugural-Dissertation zur Erlangung des Doktorgrades
der Naturwissenschaftlichen Fachbereiche (Fachbereich Biologie)
der Justus-Liebig-Universität Gießen

vorgelegt von
En-Lieng Lau
aus Taiwan

Gießen 2001

Dekan: Prof. Dr. J. Janek

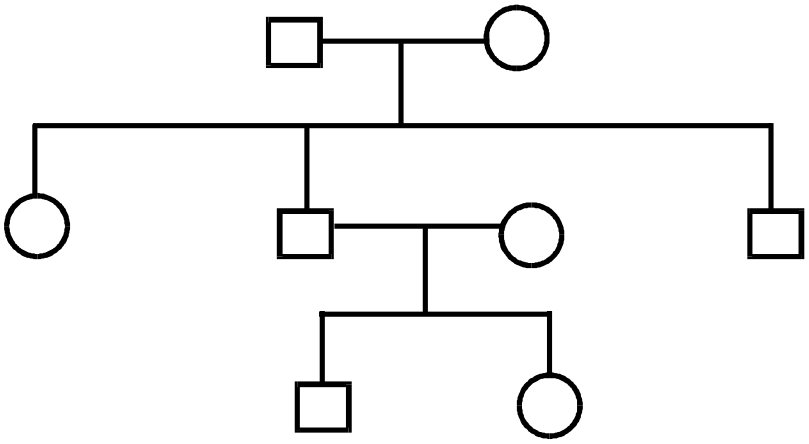
1.Gutachter: Prof. Dr. R. Renkawitz

2.Gutachter: Prof. Dr. U. Müller

Tag der mündlichen Prüfung: 24.09.2001

*Dedication*

*For Family Lau*



## Abbreviation

| | |
|---|---|
| aa | amino acids |
| AD | autosomal dominant |
| Ala | alanine |
| Amp | ampicillin |
| APS | ammoniumpersulfate |
| AR | autosomal recessive |
| Arg | arginine |
| Asn | asparagine |
| Asp | aspartic acid |
| ATP | adenosinetriphosphate |
| BAC | bacterial artificial chromosome |
| bp | base pair |
| BSA | bovine serum albumin |
| cDNA | complementary DNA |
| Ci | Curie |
| cM | centimorgan |
| CNS | central nervous system |
| Cys | cysteine |
| °C | degree Celsius |
| Da | dalton |
| dATP | deoxyadenosintriphosphate |
| dCTP | deoxycytidintriphosphate |
| dGTP | deoxyguanosintriphosphate |
| dTTP | deoxythymidintriphosphate |
| ddATP | dideoxyadeosintriphosphate |
| ddCTP | dideoxycytidintriphosphate |
| ddGTP | dideoxythymidintriphosphate |
| ddTTP | dideoxythymidintriphosphate |
| $dH_2O$ | distilled water |
| DMSO | dimethylsulfoxide |
| DNA | deoxyribonucleic acids |
| DTT | dithiothreitol |
| *E.coli* | *Escherichia coli* |
| EDTA | ethylenediaminetetraacetic acid |
| EST | expressed sequence tag |
| EtBr | ethidium bromide |
| EtOH | ethanol |
| g | gram |
| Gln | glutamine |
| Glu | glutamic acid |
| Gly | glycine |
| *et al* | et altera |
| His | histidine |
| hr | hour |
| HSP | hereditary spastic paraplegia |
| Ile | isoleucine |
| IPTG | isopropyl-β-D-thiogalactoside |

| | |
|---|---|
| kb | kilobase |
| kDa | kilodalton |
| KOAc | potassium acetate |
| *l* | liter |
| *lac* Z | the symbol of β-galactosidase gene |
| LB | Luria-Bertani |
| Leu | leucine |
| LMP | low melting point |
| Lys | lysine |
| M | molar |
| Mb | megabase |
| Met | methionine |
| min | minute |
| ml | milliliter |
| mM | millimolar |
| μ | micro |
| nm | nanometer |
| NaOAc | sodium acetate |
| ng | nanogram |
| OD | optical density |
| ORF | open reading frame |
| ori | origin of replication |
| p | pico |
| PAC | P1-derived artificial chromosome |
| PAGE | polyacrylamide gel electrophoresis |
| PBS | phosphate buffered saline |
| PCR | polymerase chain reaction |
| pH | pH value |
| PIC | polymorphism information content |
| Phe | phenylalanine |
| PMSF | phenylmethylsulfonylfluoride |
| Pro | proline |
| %(v/v) | percent of volume |
| %(w/v) | percent of weight |
| RNA | ribonucleic acids |
| RNase | ribonuclease |
| rpm | rotations per minute |
| RT-PCR | reverse transcriptase PCR |
| sec | second |
| SDS | sodium dodecylsulfate |
| Ser | serine |
| SPG | spastic paraplegia |
| STR | short tandem repeat |
| STS | sequence tagged site |
| TAE | Tris-Acetate-EDTA |
| TBE | Tris-Borate-EDTA |
| TE | Tris-EDTA |
| TEMED | N,N,N',N',-tetramethylenediamine |
| Thr | threonine |

| | |
|---|---|
| Tris | tris(hydroxymethyl)aminomethane |
| Trp | tryptophan |
| Tyr | tyrosine |
| U | units |
| µl | microliter |
| UV | ultraviolet |
| V | voltage |
| Val | valine |
| vs. | *versus* |
| W | watt |
| YAC | yeast artificial chromosome |

# Contents

## 3 Methods

# 1 Introduction

Hereditary spastic paraplegias (HSP) (Mckusick number [1] 182600) are a group of clinically and genetically diverse disorders characterized by progressive, generally severe, spasticity of lower extremity and in some cases upper extremities as well. Two groups of HSPs can be distinguished clinically by whether progressive spasticity occurs in isolation (pure HSP) or in combination with other neurologic abnormalities (complicated HSP), including optic neuropathy, retinopathy, extrapyramidal disturbance, dementia, ichthyosis, mental retardation, and deafness.[2,3] However, there are few cases presenting intermediate clinical features. Pure HSP is more common and also genetically heterogeneous. On the other hand, Complicated HSP consists of many rare conditions which tend to be inherited as an autosomal recessive trait.[4] Both forms can be inherited in an autosomal dominant, autosomal recessive or X-linked recessive pattern. Prevalence of HSP was estimated at $9 \times 10^{-6}$ in Denmark,[5] $140 \times 10^{-6}$ in western Norway[6] and $96 \times 10^{-6}$ in northern Spain.[7]

## 1.1 Pure Hereditary Spastic Paraplegia (Pure HSP)

Pure HSP is the most common form of HSP. Approximately 70-80% of pure HSPs follow an autosomal dominant inheritance pattern (AD-HSP)[8]; autosomal recessive inheritance (AR-HSP)[9] is found in 20-30% of cases and in very rare cases X-linked recessive inheritance (X-HSP)[10] occurs.

### 1.1.1 Clinical Features of Pure HSP

The first cases of pure HSP were described in 1880 by Strümpell[11] who reported two brothers with an onset in middle life of progressive weakness and spasticity in the lower limbs. The clinical feature of pure HSP is the presence of a progressive spastic paraplegia mainly of the lower extremities. The disease severity ranges from completely asymptomatic (10-20%) to

wheelchairbound (10-15%). Patients develop leg stiffness and gait disturbance (stumbling and tripping) due to difficulty in dorsiflexing the foot and weakness of hip flexion. Although weakness and hypertonicity are usually restricted to the lower limbs, upper limb hyperreflexia is also common. Patients show weakness in addition to spasticity. The presence of urinary symptoms is likely. About 50% of patients are affected by urinary frequency, urgency or incontinence. Diminished sense of vibration is present in 20-65% of patients and 30-50% of patients have a Pes cavus. In addition, erectile impotence, mild upper limb incoordination, and absent ankle jerks were found in a small proportion of patients.[2,4,8,12] The symptoms and signs of pure HSP are shown in Table 1.

Harding[9] subclassified autosomal dominant pure HSP in Type I and Type II. Type I has an early onset (before age 35) and slow progression of symptoms. Patients show greater spasticity than weakness. On the other hand, Type II has a later onset (after age 35) and demonstrates faster progression. Dürr *et al.*[12] studied 23 families with pure autosomal dominant spastic paraplegia and found that the mean age of onset was 29, with a range between 1 and 68. Patients showed hyperreflexia in the upper limbs, sphincter disturbances, a decreased sense of vibration, and increased frequency of lower limb muscle weakness with disease duration. However, the clinical manifestations of early-onset (< age 29) and late-onset (> age 29) patients were not significantly different. Age at onset varied both between and within families. Anticipation and imprinting did not occur. These results suggested that only genetic analyses could provide accurate classification of the pure HSP. Fink *et al.*[2] observed that the ages of onset overlap in autosomal dominant HSP kindred and also concluded that there did not appear to be a genetic basis for HSP classification based entirely on age of symptom onset.

Table 1. Symptoms and signs in pure HSP[2,4,8,12]

| |
|---|
| A progressive spastic paraplegia |
| Weakness and hypertonicity of legs |
| Upper limb and jaw jerk hyperreflexia |
| Urinary symptoms (urinary frequency, urgency, or hesitancy) |
| Sphincter disturbances |
| Decreased vibration sense |
| Decreased joint position sense |
| Absence of ankle jerks |
| Babinski signs |
| Hoffman's signs |
| Trömner'signs |
| Pes cavus |
| Clonus at ankles and knees |
| Extensor plantar responses |
| Mild upper limb incoordination |
| Scoliosis (rare) |

### 1.1.2 Pathology

The major neuropathologic feature of pure HSP is an axonal degeneration in the terminal portions of the longest corticospinal fibers and dorsal columns.[8,13] Loss of anterior horn cells is observed in some cases.[4] Neuronal cell bodies of degenerating fibers are normally preserved. There is no evidence of primary demyelination reported. The dorsal root ganglia, posterior roots, and peripheral nerves are normal[13].

## 1.2 Complicated Hereditary Spastic Paraplegia (Complicated HSP)

In complicated HSP, spastic paraplegia appears with other disorders which include amyotrophy, cerebellar signs, optic atrophy, choreoathetosis/dystonia, sensory neuropahty, disordered skin pigmentation, macular degeneration/mental retardation (Kjelllin syndrome), Sjögren-Larsson syndrome, Mast syndrome, MASA syndrome (mental retardation, aphasia, shuffling gait, and adducted thumbs). Although these disorders are rare, spastic paraplegia is the common feature of them. The most common type of complicated HSP is spastic paraplegia associated with amyotrophy. This type frequently results in peroneal muscular atrophy syndrome or hereditary motor and sensory neuropathy. In addition, brisk knee jerks

and extensor plantar responses are specific signs. Distal sensory loss is observed in 50% of patients. The age of onset ranges from early childhood to 50´s and some affected individuals show asymptomatic[3,4].

### 1.3 Molecular Genetics of HSP

Distinguishing between pure and complicated HSP is still a useful criteria. However, molecular genetics allow the definitive distinction of HSP due to the expanding variety of phenotypes. Now at least 15 HSP loci can be differentiated by genetic analysis, and the gene defect has been identified beyond doubt in 4 loci. These findings are a first step towards an understanding o f the pathological mechanism of HSP at the molecular level.

#### 1.3.1 X-linked HSP

Three X-linked HSP loci have been mapped to Xq28 (*SPG1*;312900), Xq22 (*SPG2*;312920),.Xq11.2 (*SPG16*;300266).

***SPG1***

Kenwrick *et al.*[14] reported a complicated X-HSP and mapped the disease locus to *DXS15* and *DXS52* on the long arm of the X chromosome (Xq28). All patients examined had mental retardation. Winter *et al.*[15] presented evidence showing that the family described by Kenwrick *et al.* was an instance of MASA syndrome. Later, Schrander-Stumpel *et al.*[16] proposed that MASA syndrome and X-linked hydrocephalus are allelic. Jouet *et al.*[17] reported that the mutations in L1CAM gene resulted in X-HSP, MASA syndrome, and X-linked hydrocephalus and thus all three syndromes are allelic disorders. L1CAM is a cell surface glycoprotein which is expressed in the axon of postmitotic neurons and is involved in neuronal migration and neurite extension.

***SPG2***

Pure X-HSP is rare. The first genetic linkage of a pure form of spastic paraplegia was presented by Keppen *et al.*[10] The mutation is linked to markers DXS17 and YNH3 located in the Xq21-q22 region. Bonneau *et al.*[18] proposed that this pure X-HSP mapped to Xq21. However, this form could not be separated from the complicated variant because some members had mental retardation. This pure X-HSP was designated *SPG2*. Saugier-Veber *et al.*[19] demonstrated that the gene for proteolipid protein was the closest marker mapped to Xq21 reported by Bonneau *et al*. They also found a His139Tyr mutation in exon 3B of the *PLP* gene in an affected male. *PLP* encodes two myelin proteins, PLP and DM20, and is involved in oligodendrocyte maturation and myelin sheath compaction. The His139Tyr mutation observed in this family results in production of a mutant PLP but wild-type DM20. Kobayshi et al[20] reported the linkage results of the family with X-linked HSP whose symptoms began as pure spastic paraplegia, but later developed nystagmus, dysarthria, sensory disturbance or mental retardation. The half of patients had optic atrophy. Muscle wasting and joint contractures were observed in later symptoms. The patients needed crutches or wheelchair in their early adulthood The disease locus was mapped to the region Xq21.3-q24 which includes the PLP gene.

***SPG16***

Another complicated X-HSP was reported by Steinmüller *et al.* [22] The patients were quadriplegic and had motor aphasia, reduced vision, mild mental retardation, and dysfunction of the bowel and bladder. Onset of the symptoms occurred during the first 3 months of life. Nystagmus and dorsal flexion of the great toes were observed in the beginning development of the symptoms. Spasticity occurred first in the lower and then upper extremities, and motor

development was delayed. The patients never learned to walk and all females of the family were normal. The authors assigned the disease locus to Xq11.2-q23 by linkage and haplotype analysis. They performed mutation analysis of the *PLP* gene which was mapped to this region but no mutation was detected.

### 1.3.2 Autosomal Recessive HSP

Loci for autosomal recessive forms of SPG have been mapped to chromosomes 8p (*SPG5A*;270800), 16q (*SPG7*; 602783), 15q (*SPG11*; 604360), and 3q (*SPG14*; 605229).

#### *SPG5A*

Hentati *et al.*[23] first reported a pure AR-HSP locus and the patients examined showed a mild loss of vibratory and positiona sense of the toes and dysfunction of the bladder sphincter. The age of onset ranged from one to 20 years. This locus was mapped to a 32.2 cM (centimorgen) interval in the pericentromeric region of chromosome 8 between 8p12 and 8q13.

#### *SPG7*

De Michele *et al.*[24] described a large consanguineous family with complicated AR-HSP. All cases had abnormal gait and vibration sense was frequently decreased in legs. Hypernasal, slowed speech, dysphagia, urinary urgency, scoliosis, pes cavus and pale optic disk appeared in some cases. There were no cerebellar or extrapyramidal signs in any of 6 affected individuals. The mean age of onset was 30±8 years in this family. This locus was linked to markers *D16S413* and *D16S303* on 16q24.3. Soon after the mapping, the same team[25] identified a gene within this region that encodes a novel protein, paraplegin that is highly homologous to yeast mitochondrial ATPases, AFG3, RCA1, and YME1, which have both proteolytic and chaperone-like activities at the inner mitochondrial membrane.

Immunofluorescence analysis and import experiments showed that paraplegin localizes to mitochondria. Analysis of muscle biopsies showed typical signs of OXPHOS defects that include cytochrome c oxidase deficient fibers and the presence of subsarcolemmal accumulation of mitochondria of bizarre shape. These data indicated that a mitochondria protein defect in HSP. *Paraplegin* is composed of 17 exons and spans approximately 52 kb. The cDNA length of this gene is approximately 3.2kb and it encodes a 795 amino acid protein. A 9.5kb deletion and two frameshift mutations which were found in one pure and in one complicated HSP family have been reported.

Paraplegin and yeast mitochondrial ATPases (AFG3, RCA1, and YME1) are the members of AAA protein family ( ATPase associated with a variety of cellular activities). AAA proteins play essential roles in peroxisome biogenesis, the assembly of mitochondrial membrane proteins, cell cycle control, mitotic spindle formation, cytoskeletal interactions, vesicle secretion, signal transduction, and transcription. These proteins are characterized by a highly conserved AAA motif. This motif forms a 230-250 amino acids domain that includes the Walker homology sequences of P-loop ATPases and other regions of similarity unique to AAA family such as Walker B and the minimal AAA consensus sequence. The motif confers ATPase activity. Evolutionary conservation of the AAA domain ranges from eubacteria and archaebacteria to modern eukaryotes and tends to be homologous in regions outside the AAA domain. All members of the AAA family are $Mg^{2+}$-dependent ATPases and include metalloproteases, proteins involved in vesicle and organelle biogenesis, cell-cycle regulators , components of the 26S proteasome, and chaperone and chaperone-like protein families ($AAA^{+}$ class).[26,27,28]

***SPG11***

The third AR-HSP locus (*SPG11*) was described by Martínez Murillo *et al.*[29] The patients in eight examined AR-HSP families had the common features of spasticity and weakness in the lower limbs, hyperreflexia, extensor plantar reflexes. Mental retardation was present in two families. The age of onset ranged from one to 50 years. In seven of eight AR-HSP families, the disease locus was mapped to chromosome 15q13-q15, a 6.41 cM region between markers *D15S971* and *D15S118.* The other family showed significant linkage to the previously characterized 8q locus.

***SPG14***

Recently, Vazza *et al.*[30] described a consanguineous Italian family with HSP. The clinical features were spastic gait, hyperreflexia, mild lower limb hypertonicity. Bilateral pes cavus , extensor plantar responses, and mild mental retardation were present in all cases. The average age of onset was 30. All three patients were able to walk without aid. The authors performed a whole genome scan in another AR-HSP family and reported a new AR-HSP locus. The disease locus assigned to 3q27-q28.

### 1.3.3 Autosomal Dominant HSP

To date, 8 AD-HSP loci have been mapped to chromosome 14q (*SPG3*;182600), 2p (*SPG4*;182601), 15q (*SPG6*;600363), 8q (*SPG8*;603563), 10q (*SPG9*;601162), 12p (*SPG10*;604187), 19q (*SPG12*;604805), 2q (*SPG13*;605280).

***SPG3***

The first AD-HSP locus was reported by Hazan *et al.*[31] In one of three AD-HSP families the disease was linked to locus AFM267zd5 on chromosome 14q. The clinical feature was very similar in three families. Spastic gait was the prominent sign in all patients, with a variable

degree of severity. Some affected individuals (18.75%) showed asymptomatic. The age of onset ranged from 2 to 50 years. Another two studies demonstrated that anticipation and early age of onset in two SPG3 linked families.[32,33]

***SPG4***

Hazan *et al*.[34] reported the second AD-HSP locus in one Dutch and five French families. The average age of onset ranged from 20 to 39 years in these families. Wide variations of age of onset were observed within families. The symptoms occurs relatively nonprogressive. The authors mapped this disease locus to a 4 cM region flanked by loci *D2S400* and *D2S367*. Fink *et al*.[2] reviewed seven publications which included 67 pure AD-HSP families. Of the families, 34 were indeterminate and the disease loci of 33 kindreds were mapped to chromosomes 2p, 14q, and 15q. Of these 33 kindreds, the disease loci of 15 kindreds (45%) were linked to chromosome 2p, two (6%) were associated with chromosome 14q, and only one (3%) was associated with chromosome 15q. Scott *et al*.[35] studied 11 Caucasian pedigrees with pure AD-HSP to determine the linkage to the previously identified loci on 2p, 14q, and 15q. Chromosome 15q was excluded in all families. There was evidence of linkage to 2p in five families and linkage to 14q in one family. The other five families remained indeterminate. In pure AD-HSP families, a high proportion have linkage to *SPG4*.

In 1999, Hazan *et al*.[36,37] identified the *SPG4* gene (*SPAST*) by using positional cloning. *SPAST* is composed of 17 exons spanning a region of approximately 90 kb. The cDNA of *SPAST* is 3263 bp containing a 1848 bp ORF preceded by a 125 bp 5' UTR and followed by a 1290 bp 3' UTR. The promoter is predicted to be located at 43 bp upstream from the first base of *SPAST* exon 1. The ORF encodes a 616 amino acid protein which is named spastin. The spastin carboxy terminus has striking homology with several members of AAA family

and contains three conserved ATPase domains, including Walker motifs A and B and the AAA minimal consensus, found at amino acid positions 382-389, 437-442 and 480-498 of spastin, respectively. Walker motif A 'GPPFNFKT` corresponds to the ATP-binding domain. Two leucine-zipper domains are detected at amino acid positions 50-78 and 508-529 and a helix-loop-helix dimerization domain is located between amino acid positions 478 and 486. RT-PCR analysis reveals that human *SPAST* and its mouse homologue are expressed early and ubiquitously in fetal and adult tissues.

***SPG6***

Fink *et al*.[38] demonstrated the third AD-HSP locus in a large kindred extensively affected with pure AD-HSP. Neurologic examination of patients showed hyperreflexia, spasticity in the lower limbs, weakness of hip flexion and ankle dorsiflexion, extensor plantar responses, diminished vibratory sense in the feet, and pes cavus. Bladder disturbance was present in some cases. The patients developed progressive gait disturbance at age 12 to 35 years and the average age of onset was 22±5.3 years. The authors mapped this disease locus to chromosome 15q11.1.

***SPG8***

Hedera *et al*.[39] examined a Caucasian kindred which developed insidiously progressive gait disturbance at age 22 to 60 years. Neurological examination demonstrated spastic diplegic gait disturbance, frank corticospinal -tract deficits in legs, weakness of hip flexion and ankle dorsiflexion, diminished vibratory sensation in the feet, and often, pes cavus. The patients in wheelchairs had mild muscle atrophy and bladder disturbance was present in some cases. The disease locus mapped to chromosome 8q23-q24, a 6.2 cM region between *D8S1804* and *D8S1774*.

***SPG9***

Seri *et al*.[40] studied a large pedigree with complicated AD-HSP. The patients presented with bilateral cataracts, gastroesophageal reflux with persistent vomiting, and spastic paraparesis with amyotrophy. Severity of spastic paraparesis varied. Pes cavus and the Babinski signs, along with different degrees of muscle wasting localized in the hands and "forelegs" (*sic*.) were present in some cases. The authors performed a genome wide scan and assigned this disease locus to a 12-cM region in 10q23.3-q24.2.

***SPG10***

Reid *et al*.[41] reported a large family with pure AD-HSP. All but one affected members had a spastic gait abnormality, in addition to lower-limb reflexia. The asymptomatic affected individual was a carrier who had bilateral lower-limb hyperrefleixa, bilateral extensor-plantar responses, and unilateral sustained ankle clonus. The development of asymptomic to symptomatic history was observed in one patient. The age of onset ranged from 8 to 40 years and the average age of onset was 10.8±9.6 years. The authors performed a genome wide scan data and mapped this pure ADHSP locus to chromosome 12q13, a 9.2 cM interval between markers *D12S368* and *D12S8334.* Later, this interval was narrowed to a 6.95 cM region flanked by *D12S270* proximally and *D12S355* distally by further genetic analysis.

***SPG12***

Recently, Reid *et al*.[42] performed a genome wide linkage screen of a large Welsh family with AD-HSP which had previously reported[43] but no linkage with known pure AD-HSP loci. The authors assigned this novel AD-HSP locus to 19q13, a 16.1 cM interval between markers

*D19S868* and *D19S902*. The candidate region was narrowed to a 5 cM region between markers *D19S868* and *D19S220* by additional linkage analysis.

***SPG13***

Fontaine *et al.*[44] excluded all known AD-HSP loci in a large family of French descent with pure AD-HSP. A genomewide scan demonstrated location of the disease locus in the 5 cM region flanked by *D2S294* and *D2S2195* on 2q24-q34. The authors compared the SPG13 family with 12 SPG4 families. They found SPG13 family had significantly more patients without Babinski signs, with increased reflexes in arms and more severe handicaps than the patients of those 12 SPG4 families they examined. The severe handicaps score was obtained by dividing the disability score by disease duration (in years) and multiplying by 100. The authors defined a three point scale: 1 = normal gait or very light stiffness of the legs, 2 = inability to run, and 3 = either inability to walk without help or confinement to a wheelchair. The currently known HSP loci are summarized in Table 2.

## 1.4 The purpose of this dissertation

This dissertation began with study of a pure AD-HSP German family of three generations (see Figure 1). The family consists of 66 living members and 5 were diagnosed as symptomatic. The purpose of this dissertation is to determine the disease locus of this family and identify the mutation of this locus. The AD-HSP family was first tested for linkage with the 3 known AD-HSP loci at 2p, 14q, and 15q. The linkage results revealed that the disease locus of this family was linked to chromosome 2p (*SPG4*). Therefore, this dissertation focuses on *SPG4* which is also the most common form of pure AD-HSP. Since there was no AD-HSP gene identified during the proceeding of this study, a physical map including YACs, BACs, and PACs was established to facilitate the isolation of gene. In addition, the proposed CAG

trinucleotide repeat expansion was investigated in this dissertation. Finally, using the sequence data of the recently cloned *SPAST* gene, the SPG4 family and members from nine pure HSP families were screened by direct sequencing.

Table 2. Classification of hereditary spastic paraplegias

| HSP loci | Gene location | Gene product | inheritance mode | HSP forms | References |
|---|---|---|---|---|---|
| *SPG1* | Xq28 | L1CAM | X-linked | complicated | Saugier-Veber[14] *et al.* (1994) |
| *SPG2* | Xq21 | PLP | X-linked | complicated or pure (rare) | Jouet[17] *et al.*(1994) |
| *SPG3* | 14q11.2-q24.3 | unknown | AD | pure | Hazan[31] *et al.* (1993) |
| *SPG4* | 2p21-p24 | Spastin | AD | pure or complicated | Hazan[34] *et al* (1994) |
| *SPG5A* | 8p12-q13 | unknown | AR | pure | Hentati[23] *et al.*(1994) |
| *SPG6* | 15q11.1 | unknown | AD | pure | Fink[38] *et al.*(1995) |
| *SPG7* | 16q24.3 | Paraplegin | AR | complicated or pure | Casari[24] *et al.*(1998) |
| *SPG8* | 8q23-q24 | unknown | AD | pure | Hedera[39] *et al.*(1999) |
| *SPG9* | 10q23.1-q24.1 | unknown | AD | complicated | Seri [40]*et al.*(1999) |
| *SPG10* | 12q13 | unknown | AD | pure | Reid[41] *et al.*(1999) |
| *SPG11* | 15q13-q15 | unknown | AR | complicated or pure | Martinez[29] *et al.*(1999) |
| *SPG12* | 19q13 | unknown | AD | pure | Reid[42] *et al.*(2000) |
| *SPG13* | 2q24-q34 | unknown | AD | pure | Fontaine[44] *et al.*(2000) |
| *SPG14* | 3q27-q28 | unknown | AR | complicated | Vazza[30] *et al.*(2000) |
| *SPG16* | Xq11.2 | unknown | X-linked | complicated | Steinmüller[22]*et al.*(1997) |

AD: autosomal dominant inheritance

AR: autosomal recessive inheritance

# 2 Materials

## 2.1 Chemicals

| | |
|---|---|
| Acetic acid | Merck (Darmstadt) |
| Acrylease | Stratagene(La Jolla, USA) |
| Agar Agar | Serva( Heidelberg) |
| Agarose (LMP, ultrapure) | Gibco BRL (Gaithersburg, USA) |
| Agarose NA | Pharmacia (Uppsala, Schweden) |
| Agarose-Pulse Field | Amersham (Braunschweig) |
| Ammoniumacetate | Sigma (München) |
| Ammoniumpersulfate(APS) | Serva (Heidelberg) |
| Ampicilline | Bayer (Leverkusen) |
| Bacto Peptone | Difco (Detroit, USA) |
| Bacto Yeast Extract | Difco (Detroit,USA) |
| Boric acid | Merck ( Darmstadt) |
| Bovine Serum Albumin(BSA) | Sigma (München) |
| 5-Brom-4-chloro-3-indolyl-β-D-Galactopyranoside (X-gal) | Roth ( Karlsruhe) |
| Bromphenolblue | Merck (Darmstadt) |
| Calciumchloride | Merck (Darmstadt) |
| Chloramphenicol | Boehringer (Mannheim) |
| Chloroform | Merck (Darmstadt) |
| Dextransulfate | Sigma (München) |
| D(+)-Glucose | Sigma (München) |
| Dimethylformamide | Merck (Darmstadt) |
| 1,4-Dithiothret (DTT ) | Merck (Darmstadt) |
| D(+)-Saccharose | Roth (Karlsruhe) |
| Ethanol | Merck (Darmstadt) |
| Ethidiumbromide (EtBr) | Serva (Heidelberg) |
| Ethylendiaminteraacetate (EDTA-2Na) | Merck (Darmstadt) |
| Ficoll type 400 | Amersham (Braunschweig) |
| Formaldehyde (37%) | Merck (Darmstadt) |
| Formamide | Merck (Darmstadt) |
| Glutamine | Merck (Darmstadt) |
| Glycerin (87%) | Merck (Darmstadt) |
| Glycogen | Boehringer (Mannheim) |
| Hydrogenchloride (HCl) | Merck (Darmstadt) |
| HEPES | Sigma (München) |
| Isopropanol | Merck (Darmstadt) |
| Isopropyl-β-D-Thio-Galactopyranoside (IPTG) | Roth (Karlsruhe) |
| Kanamycin | Merck (Darmstadt) |
| Polyvinylpyrrolidone | Sigma (München) |
| Potassiumchloride(KCl) | Merck (Darmstadt) |
| Potassiumacetate | Sigma ( München) |
| L-Arginin-HCl | Roth (Karlsruhe) |
| L-Histidine | Roth (Karlsruhe) |
| L-Isoleucine | Roth (Karlsruhe) |
| L-Leucine | Roth (Karlsruhe) |
| L-Methionine | Roth (Karlsruhe) |

| | |
|---|---|
| L-Phenylalanine | Roth (Karlsruhe) |
| L-Threonine | Roth (Karlsruhe) |
| L-Tryptophan | Roth (Karlsruhe) |
| L-Tyrosine | Roth (Karlsruhe) |
| L-Valine | Roth (Karlsruhe) |
| Lysine | Merck (Darmstadt) |
| Long Ranger gel solution | Amersham (Braunschweig) |
| MDE gel solution | FMC (Maine, USA) |
| Magnesiumchloride | Merck ( Darmstadt) |
| Manganchloride | Merck (Darmstadt) |
| 3-[N-Morpholino] propansulfonate (MOPS) | Sigma (München) |
| Mineral oil | Sigma (München) |
| N,N,N´,N´-Tetramethylethylendiamin (TEMED) | Serva (Heidelberg) |
| Sodiumchloride | Merck (Darmstadt) |
| Sodiumdodecylsulfate (SDS) | Serva (Heidelberg) |
| Sodiumhydroxide (NaOH) | Merck (Darmstadt) |
| Sodiumpyrophosphate | Sigma (München) |
| Phenol | Merck (Darmstadt) |
| Rotiphorese Gel 40 | Roth (Karlsruhe) |
| Rubidiumchloride | Sigma (München) |
| Sigmacote | Sigma (München) |
| Spermidin | Sigma (München) |
| Tris,[Tris(hydroxymethyl)-aminoethan] | USB (Braunschweig) |
| Tri-sodiumcitrate | Merck (Darmstadt) |
| Triton X-100 | Merck (Darmstadt) |
| Urea | Merck (Darmstadt) |
| Whatman 3 MM-Paper | Schleicher u. Schuell (Dassel) |
| Xylene cyanol | Merck ( Darmstadt) |

## 2.2 Media

**I** dYT medium : For 1 *l* medium
- 16 g Tryptone
- 10 g Yeast Extract
- 10 g NaCl

**II** LB medium : For 1 *l* medium
- 10 g Tryptone
- 5 g Yeast Extract
- 10 g NaCl

**III** LB plate : For 1 *l* preparation
- 1 *l* LB medium
- 15 g Agar-Agar

**IV** NZYC-Broth: For 1 *l* :
- 5 g NaCl
- 2 g $MgCl_2.6H_2O$
- 10 g NZ amino acids (casein hydrolysate N-Z Amine A)
- 5 g Bacto yeast extract
- 1 g Bacto casamino acids (adjust pH to 7.4 )

**V** 10 x (-) Ura, (-) Trp stock solution : For 1 *l*:

| | |
|---|---|
| 600 ml | 0.5 mg/ml L-Tyrosin |
| 30 ml | 7.5 mg/ml Adenine |
| 30 ml | 7.5 mg/ml L-Arginine |
| 30 ml | 7.5 mg/ml L-Histidine |
| 30 ml | 7.5 mg/ml L- Methionine |
| 40 ml | 7.5 mg/ml L-Lysine |
| 70 ml | 7.5 mg/ml L-Phenylalanine |
| 80 ml | 7.5 mg/ml L-Leucine |
| 80 ml | 7.5 mg/ml L-Valine |
| 40 ml | 7.5 mg/mlL-Isoleucine |

Each amino acid solution is separately prepared , autoclaved and stored at 4°C.

**VI** Yeast nitrogen base stock solution : 6.7 %(w/v) Yeast Nitrogen base solution
The stock solution is sterilized with filter and stored at room temperature.

**VII** (-) Ura, (-) Trp yeast medium:

| | |
|---|---|
| For 100 ml | 2 g Glucose |
| | 10 ml 10 x (-) Ura, (-) Trp stock solution |
| | 10 ml 6.7 % yeast nitrogen base stock solution |

**VIII** (-) Ura ,(-) Trp yeast agar plates

| | |
|---|---|
| For 100 ml | 2 g Agar-Agar |
| | 100 ml (-)Ura,(-)Trp Yeast medium |

**2.3 Solutions and buffers**

**I Antibiotic solutions**

Ampicillin stock solution : 50 mg/ml in dd $H_2O$ (store at -20°C)

Kanamycin stock solution: 10 mg/ml in dd $H_2O$ (store at -20°C)

Chloramphenicol stock solution: 34 mg/ml in ethanol (store at -20°C)

**II Buffers for polymerase chain reaction (PCR)**

10 x reaction buffer:
100 mM Tris-HCl
500 mM KCl
15 mM $MgCl_2$
0.1% Gelatin, pH8.3

Taq polymerase dilution buffer:
20 mM HEPES
100 mM KCl
0.1 mM EDTA, 1mM DTT
10 % Glycerol

PCR stop buffer: 10 mM NaOH
1 mM EDTA
80 % Formamide
0.1 % Bromophenolblue
0.1 % Xylencyanol

**III Buffers for preparation of competent cells**

TFB I-Buffer: 100 mM $RbCl_2$
50 mM $MnCl_2$
30 mM KOAc
10 mM $CaCl_2$ pH 5.8

TFB II Buffer: 10 mM MOPS
10 mMRbCl
75 mM $CaCl_2$
15 % Glycerol pH 7.0

**IV Buffers for DNA preparations**

**(1) Buffers for genomic DNA isolation of human blood**

Nuclei Extraction Buffer: For 100 ml 10.95 g Sucrose
0.5 ml 1 M $MgCl_2$
1 ml Triton X-100
1 ml 1 M Tris, pH 8.0
Filter sterile and store at 4°C

DNA Extraction Buffer: For 100 ml 4 ml 1 M Tris, pH 8.0
4 ml 0.5 M EDTA, pH 8.0
2 ml 5 M NaCl

The nuclei extraction buffer and DNA extraction buffer are sterilized with filters and stored at 4°C.

**(2) Buffers for plasmid DNA minipreparation**

Lysis Buffer I: 50 mM Glucose
10 mM EDTA
25 mM Tris-HCl
2 mg/ml Lysozyme pH 8.0

Lysis Buffer II: 0.2 N NaOH
1% (w/v) SDS

Lysis Buffer III: 3 M Na-Acetate, pH 4.8

RNase A : 10 mg/ml

Ribonuclease A is dissolved in 10 mM Tris/HCl, pH7.5 / 15 mM NaCl, and heated for 15 min at 100°C. After slowly cooing down to room temperature, the solution is aliquoted and store at -20°C.

**(3) Buffers for PAC and BAC DNA isolations**

P1(filter sterilized. store at 4°C): 15 mM Tris,pH8.0
10 mM EDTA ,pH 8.0
100 µg/ml RNase A

P2 (filter sterilized, store room temperature) 0.2 N NaOH
1 % SDS

P3( autoclaved,store at 4°C) 3 M KOAc, pH5.5

**V Solutions for hybridization**

100 x Denhardt´s solution: For 1 *l*: 20 g Ficoll
20 g Polyvinylpyrrolidone
20 g BSA

Salmon sperm DNA: 5 mg/ml $dH_2O$
The solution is autoclaved for 10 min to shear DNA and stored at -20°C.

Hybridization solution: For 1 *l*: 50 g Dextran Sulfate
500 ml Formamide
300 ml 20 x SSC
50 ml 20% SDS
50 ml 100 x Denhardt´s solution
4 ml 5mg/ml Salmon sperm DNA

Prehybridization and hybridization solution for oligo-hybridization:
For 500 ml: 150 ml 20 x SSC
5ml 100 x Denhardt´s solution
2.5 ml 10% Sodiumpyrophosphate
12.5 ml 20% SDS
2.5 ml Salmon sperm DNA(5mg/ml)

**VI Solutions for Southern blot**

Denaturing solution: 0.5 M NaOH ,
0.5 M NaCl

Neutralization solution: 0.5 M Tris-HCl, pH7.0
1 M NaCl

**VII 50 x TAE buffer**: 2 M Tris-base, pH8.0 (adjust pH with acetic acid)
50 mM EDTA-diNa

**VIII 10 x TBE buffer:** 0.5 M Tris-HCl , pH8.0

0.5 M Boric acid
0.5 M EDTA

**IX 20 x SSC buffer**: 0.3 M Tri-sodium citrate, pH7.0
3.0 M NaCl

**X TE buffer** : 10 mM Tris/HCl, pH7.6
1 mM EDTA-diNa

**XI X-gal stock solution**: 20 mg/ml in Dimethylformamide

**XII Proteinase K solution** : 1 mg/ml Proteinase K
0.5 M EDTA
1% Sarcosyl

**XIII 1O x TAE DNA loading buffer**: 50 % (v/v) Glycerol
0.25 % (w/v) Bromophenolblue
0.25 % (w/v) Xylencyanolblue
10 x TAE

**XIV Ethidiumbromide stock solution**: 10 mg/ml $dH_2O$

### 2.4 Dideoxyribonucleic acids (DNA)

| | |
|---|---|
| Fish sperm DNA | Boehringer (Mannheim) |
| Lambda DNA | New England BioLabs (Boston) |

### 2.5 Nucleotides

| | |
|---|---|
| Adenosine-5'-triphosphate-disodium (ATP) | Boehringer (Mannheim) |
| dNTPs | Roth (Karlsruhe) |

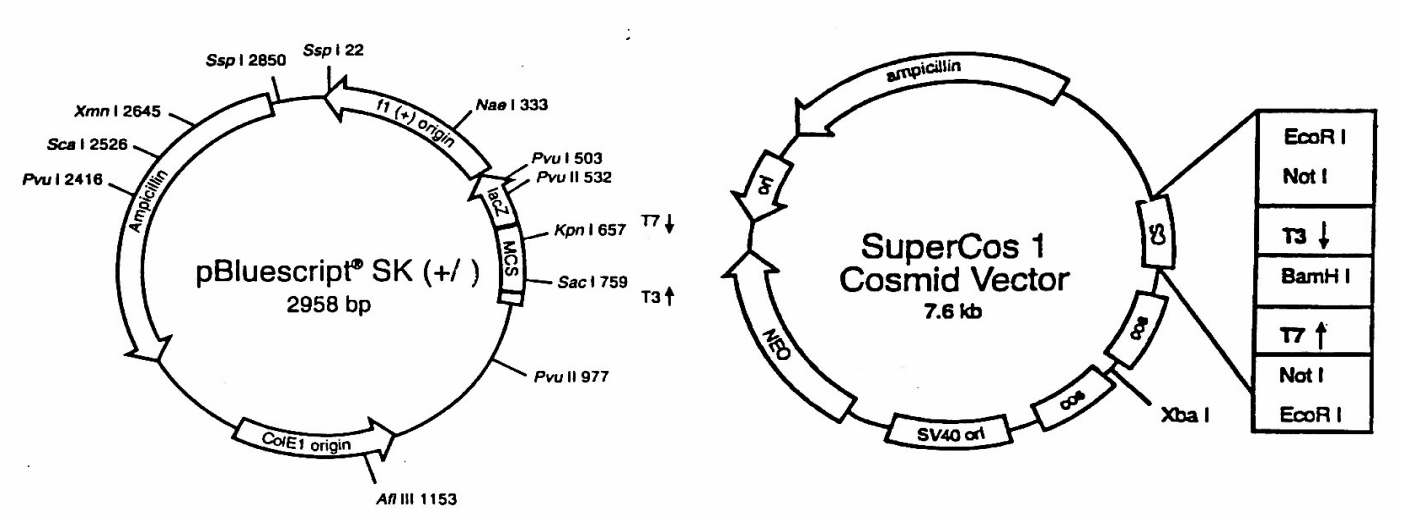
### 2.6 Radioactive Labeled Substances

| | |
|---|---|
| $\alpha$-$^{32}$P-dCTP (~3000 Ci/mmol) | Amersham (Braunschweig) |
| $\gamma$-$^{32}$P-ATP (~3000 Ci/mmol) | Amersham (Braunschweig) |
| $\alpha$-$^{33}$P-ddNTP (~1000 Ci/mmol) | Amersham (Braunschweig) |
| $\alpha$-$^{35}$S-dATP (~1000 Ci/mmol) | Amersham (Braunschweig) |

### 2.7 Cloning Vectors

| | |
|---|---|
| pBluescript II KS (+) | Stratagene ( La Jolla, USA.) |
| SuperCos 1 | Stratagene ( La Jolla, USA.) |

The structures of vectors are given below.




## 2.8 Bacterial strains

All bacteria used in this study were derived from *Escherichia coli* K12 strains

| | |
|---|---|
| XL1-Blue MR | *rec*Al, *end*A1, *gyr*A96, *thi*-1, *hsd*R17,*sup*E44,*rel*A1, *lac*[F´*pro*AB,*lac* $I^q$ ZΔM15, Tn10, $(Tet^r)$]$^c$ |
| JM 83 | $F^-$, *ara*, Δ(*lac-pro*AB), *rps*L ($Str^R$),( Φ80dΔ(*lac*Z)M15) |
| DH5α-MCR | $F^-$ *mcr*AΔ(*mrr-hsd*RMS-*mcr*BC)ϕ80d*lac*ZΔM15Δ(*lac*ZYA-*arg*F) U169*end*A1 *rec*A1deoR *thi*-1 *pho*A *sup*E44$\lambda^-$*gyr*A96 *rel*A1 |
| DH10B | $F^-$*mcr*AΔ(*mrr-hsd*RMS-*mcr*BC)ϕ80d*lac*ZΔM15Δ*lac*X74*end* A1 *rec*A1 deoRΔ(*ara, leu*)7697 *ara*D139*gal*U *gal*K *nup*G *rps*L$\lambda^-$ |

## 2.9 DNA length standards

| | |
|---|---|
| 100 bp ladder(100-1000bp) | Invitrogen |
| 500 bp ladder (500- 8000bp) | Appligene-Oncor |
| 1 Kb ladder( 506-12216bp) | GibcoBRL (Eggenstein) |
| 50 Kb ladder (48.5-1000Kb) | Bio-Rad (München) and New England BioLabs (Boston) |
| λ/Hind III (0.5-23Kb) | New England BioLabs (Boston) |

## 2.10 Enzymes

| | |
|---|---|
| Restrictionendonucleases | New England Biolab (Boston)<br>Boehringer (Mannheim)<br>Gibcol/BRL (Eggenstein)<br>AGS (Heidelberg) |
| Taq DNA polymerase | Institute of Human Genetics, Giessen |
| T4 ligase | Boehringer (Mannheim) |
| T4 polynucleotide kinase | Amersham (Braunschweig) |
| Calf alkaline phosphatase (CIP) | Boehringer (Mannheim) |
| Superscript II RNase H - reverse transcriptase | Gibco/BRL (Eggenstein) |

| | |
|---|---|
| Proteinase K | Merck (Darmstadt) |
| Lysozyme | Sigma (München) |

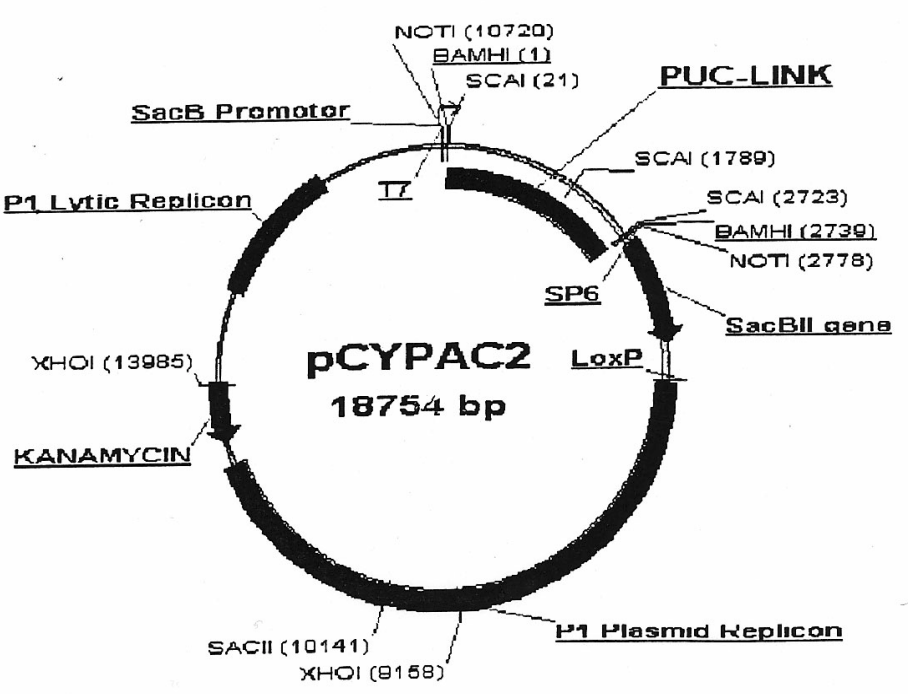
## 2.11 Kits

| | |
|---|---|
| Nonaprimer labelling kit | Appigene-Oncor (Heidelberg) |
| T7 sequenase | Amersham (Braunschweig) |
| Thermal sequenase for fluorescent labeled primers | Amersham (Braunschweig) |
| Thermal sequenase for radioterminator cycle sequencing | Amersham (Braunschweig) |
| Nucleospin extraction kit | Macherey-Nagel (Düren) |
| Gigapack III Gold Packaging kit | Stratagene (La Jolla, USA) |
| QIAamp RNA Blood Mini kit | Qiagen (Düsseldorff) |

## 2.12 Human genomic DNA library

### I. Human PAC library : RPCI1 high density gridded filters

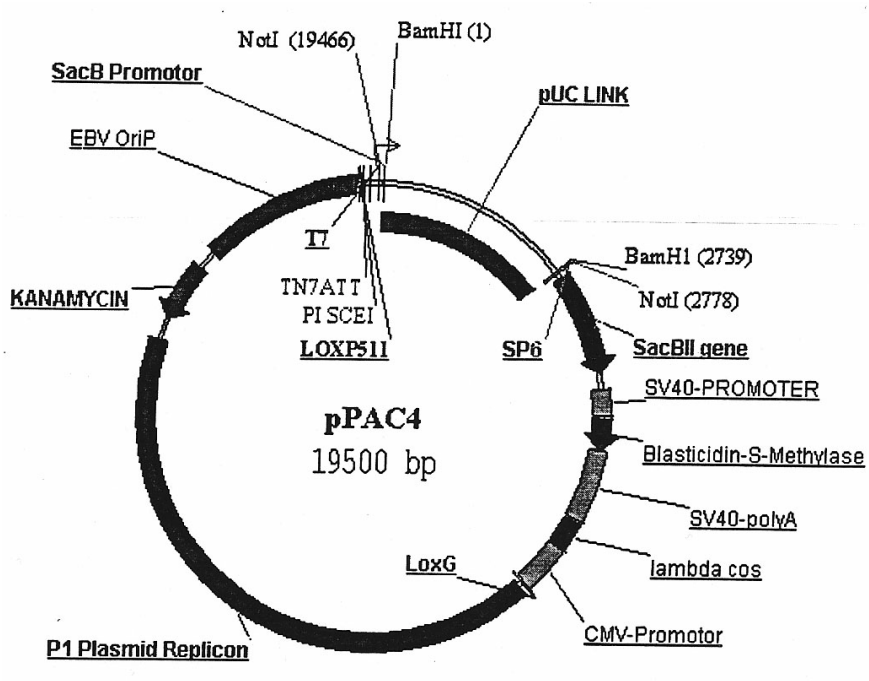
The human PAC library RPCI1 provided by the UK HGMP Resource Center was constructed by Pieter de Jong and his group at the Roswell Park Cancer Institute, Buffalo. The library is constructed in vector pCYPAC2. The source of insert DNA is a normal male blood donor, and the average insert size is about 110kb. There are 25% clones in this library without inserts. The library consists of approximately 120,000 clones in 315 384-well microtitre plates and has been gridded in a 4 x 4 array on 22.2 x 22.2 cm Hybond N nylon membanes (Amersham). Each clone is spotted twice to give 36,864 spots on each membrane. The library consists of 7 membranes. The structure of pCYPAC2 is shown below.




Origin: http://bacpac.med.buffalo.edu/

## II. Human PAC library: RPCI6 high density gridded filters and DNA pools

The human PAC library RPCI6 provided by the Resource Center/Primary Database of the German Human Genome Project was constructed by Pieter de Jong and his group at the Roswell Park Cancer Institute, Buffalo. It is constructed in the vector pPAC4 and *E.coli* DH10B is used as the host cell. The source of insert DNAs is a normal female blood donor, and the average insert size is about 140kb. The library consists of approximately 92160 clones in 240 384-well microtitre plates and has been gridded in a 5 x 5 array on 22 x 22 cm Hybond N+ nylon membanes. Each clone is spotted twice to give 27,648 spots on each membrane. The library consists of 4 membranes. The library is also available in DNA pools for PCR screens. The structure of pPAC4 is shown below.




Origin: http://bacpac.med.buffalo.edu/

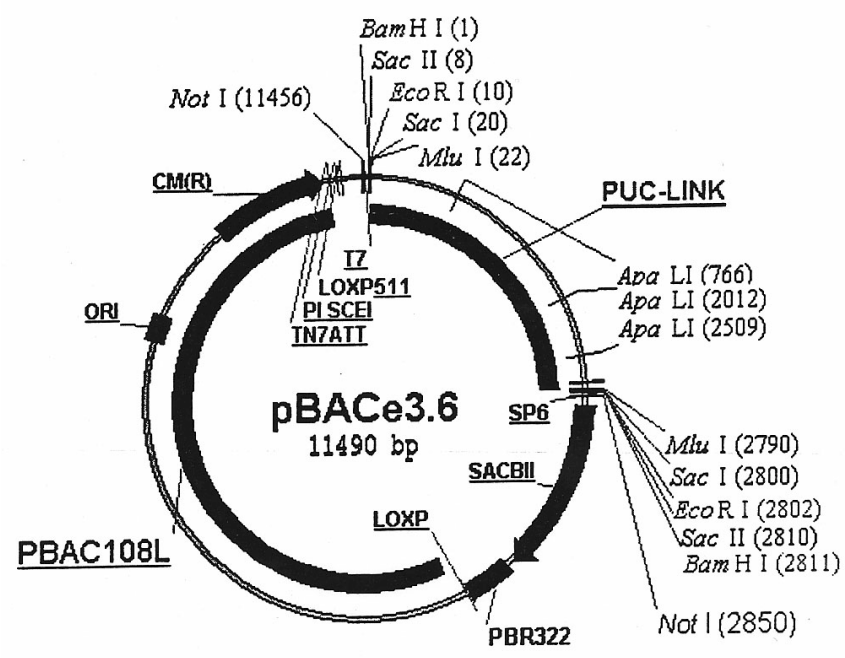
## III. Human chromosome 2 PAC library

The library LL02NP04"AI" provided by the Resource Center/Primary Database of the German Human Genome Project was constructed by Gingrich. Digests of DNA from flow-sorted human chromosome 2 of human-hamster hybrid somatic cell line GM10826 are cloned to vector pCYPAC2 and *E.coli* DH5 alpha MCR is used as the host cell. The library consists of approximately 5800 clones in 15 384-well microtitre plates, and has been gridded in a 3 x 3 array on 22 x 22 cm Hybond N+ nylon membanes. Each clone is spotted twice on one filter.

## IV. Human BAC library : RPCI-11 DNA pools

The human BAC library RPCI-11 provided by Resource Center/Primary Database of the German Human Genome Project was constructed by Osoegawa and Tateno at the Roswell Park Cancer Institute, Buffalo. It is constructed in vector pBACe3.6 and *E.coli* DH10B is used

as the host cell. The source of insert DNA is a normal male blood donor, and the average insert size is about 174kb. The pools consist of 442368 clones in 1152 plates. The library comprises 144 primary pools and represents approximately 25.3 fold genomic coverage. The structure of pBACe3.6 is shown below.




Origin: http:// bacpac.med.buffalo.edu/

## 2.13 Patients

A large German pedigree with AD-HSP was investigated. This family consists of 66 living members (see Figure 1) and the blood samples were collected from 42 participating members of the family. In addition, *SPAST* was sequenced in nine affecteds of small families with pure HSP.

# 3 Methods

## 3.1 Isolation of DNA

### 3.1.1 Isolation of human genomic DNA from peripheral blood

The blood sample was mixed with an equal volume of cold PBS and blood cells were collected by centrifugation at 1500 rpm for 10 min at 4°C. The blood cell pellet was first resuspended in 10-15 ml cold nuclei extraction buffer. The cell suspension was mixed with an additional 20-30 ml nuclei extraction buffer and placed on ice for 30 min. Cells were harvested by centrifugation at 2500 rpm for 10 min at 4°C and resuspended in 500µl DNA extraction buffer. An equal volume of DNA extraction buffer containing SDS and proteinase K was added slowly and the solution was incubated at 50°C overnight. The DNA solution was stepwise extracted with an equal volume of phenol, phenol/chloroform (1:1), and chloroform. The aqueous phase was separated from organic phase by centrifugation at 2500 rpm for 10 min. DNA was precipitated from the water phase with ethanol (EtOH) and washed with 70% ethanol. DNA was dissolved in 500 µl TE containing RNase. The DNA solution was incubated at 37°C for 30 min and stored at 4°C.

### 3.1.2 Isolation of plasmid or cosmid DNA

Bacteria were harvested from 1.5 ml of overnight culture by centrifugation at 9000 rpm for 1 min and resuspended in 100 µl Lysis buffer I. After addition of 200 µl Lysis buffer II, the cell suspension was incubated on ice for 5 min and mixed with 150 µl Lysis buffer III. Cell lysates were separated by centrifugation at 13000 rpm for 10 min. The aqueous phase was extracted with 500 µl phenol/chloroform (1:1) and centrifuged at 13000 rpm for 5 min. DNA in the aqueous phase was precipitated with 1 ml EtOH at -20 °C for 30 min and then collected by centrifugation at 13000 rpm for 10 min. The DNA pellet was washed with 200 µl 70% EtOH and air dried for 10 min. DNA was dissolved with 100 µl TE buffer with RNase and stored at -20°C.

### 3.1.3 Isolation PAC or BAC DNA

Bacteria containing PAC or BAC DNA were cultured in media containing kanamycin or chloramphenicol, respectively. Bacteria were grown as 1.5 ml culture for up to 16 hr at 37°C with vigorous shaking. Bacteria were harvested by centrifugation at 3000 rpm for 10 min. The pellet was resuspended in 0.3 ml P1 solution, mixed with 0.3 ml P2 solution, and placed at room temperature for 5 min. Cell lysates were mixed with 0.3 ml P3, incubated on ice for 5 min, and cleared by centrifugation. The DNA from the aqueous phase was precipitated by addition of 0.8 ml ice-cold isopropanol and collected by centrifugation at 13000rpm for 15 min at 4°C. The DNA pellet was washed with 0.5 ml of 70% EtOH and air dried at room temperature. When the color of DNA pellet turned from white to translucent, DNA was resuspended in 40 µl TE and stored at -20°C.

### 3.1.4 Isolation of YAC DNA

Yeast were cultured in 30 ml (-)Ura (-)Trp Yeast medium at 30°C for 2-3 days. Yeast cells were harvested by centrifugation at 3000 rpm for 10 min and the pellet was washed with 20 ml TE. The yeast cell pellet was resuspended in 500 µ1 of DNA extraction solution and

incubated at 50°C for 5 hr. The solution was stepwise extracted with 500 µ1 phenol, 500 µ1 phenol/chloroform (1:1) and 500 µ1 chloroform. The aqueous phase was separated from the organic phase by centrifugation at 13000 rpm for 15 min. DNA was precipitated from the aqueous phase with 1 ml EtOH containing 50 µ1 3M NaOAc at -20°C for 30 min and collected by centrifugation. The DNA pellet was dissolved in 500 µ1 TE.

### 3.1.5 DNA extraction from agarose gel slices

Macherey-Nagel Nucleospin Extract kit was used to extract DNA from agarose gel slices and isolate PCR products from PCR reaction mixtures.
The desired DNA fragment was excised from an agarose gel and mixed with 300 µl NT1 for each 100 mg agarose gel slices. The agarose gel slices in NT1 buffer were dissolved at 50°C for 10 min and then applied to a Nucleospin Extract column. The DNA in solution was adsorbed to the column during centrifugation. The column was washed with 600µl NT3 buffer and 200 µl NT3 buffer. The DNA was eluted with 40 µl NE elution buffer by centrifugation.

### 3.1.6 Purification of Polymerase Chain Reaction (PCR) Products

PCR product (20 µl) was mixed with 80 µl NT2 buffer and applied to a Nucleospin Extract column. The spin column was centrifuged, the flowthrough discarded and the column was washed with 600 µl NT3 buffer then twice with 200 µl NT3 buffer. The DNA was eluted with 40 µl NE elution buffer by centrifugation.

## 3.2 DNA Quantitation

Absorption (A) of the sample was measured at several different wavelengths to assess purity and concentration of nucleic acids. $A_{260}$ measurement is quantitative for pure nucleic acid preparations in microgram quantities. The ratio of absorption at 260 and 280 nm was used as an indicator of nucleic acid purity. The spectrophotometer was set to zero at 260 nm by using TE buffer as a reference. The reference was removed and the DNA solution was pipetted into the cuvette and read at 260 nm. The same procedure was performed at 280 nm.
To determine the concentration (C) of DNA, the $A_{260}$ reading was used in combination with the following equation: Double-stranded DNA: $C\ (\mu g/ml) = A_{260} / 0.020$
The value of $A_{260} / A_{280} > 1.8$ indicated highly purified preparation of DNA. Proteins that absorbed at 280 nm reduced this ratio.

## 3.3 DNA gel electrophoresis

DNA molecules carry an overall negative electric charge due to phosphate groups. Consequently, when DNA molecules are placed in an electric field they migrate towards the positive pole. Different sizes of DNA fragments can be separated in a gel which is usually made of agarose or polyacrylamide. The gel matrix is composed of a complex network of pores through which DNA molecules must pass to reach the positive pole. The larger the DNA molecule, the slower it migrates through the gel. Thus, different sizes of DNA can be separated through gel electrophoresis. Usually, agarose gels were used for separating DNA fragments which were longer than 500 bp and polyacrylamide gels are used for those fragments shorter than 500 bp.

### 3.3.1 Agarose gel electrophoresis

For 500 to 20 000 bp DNA fragments, a range of agarose gel concentrations from 0.5 to 2% (w/v) was usually used. The higher the agarose gel concentration, the better separation obtained for smaller DNA fragments. The suitable gel concentrations for separating different sizes of DNA fragments are listed in *Current Protocol in Molecular Biology* (Ausbel *et al.*[46]) and shown below.

| Agarose gel (%) | Separation of DNA fragments(kb) |
|---|---|
| 0.5 | 1~30 |
| 0.7 | 0.8 ~12 |
| 1.0 | 0.5 ~10 |
| 1.2 | 0.4~7 |
| 1.5 | 0.2~3 |

Agarose gels were prepared as follows:
Agarose was mixed with TAE buffer and boiled for 3 min. The melted gel solution was cooled down and poured into a gel chamber. The comb was set into the gel before the gel solution became solid. Electrophoresis was carried out in TAE buffer. The DNA sample was mixed with loading dye before loading to the gel. The dye ran as size indicator during electrophoresis. Usually, the voltage used for agarose gel electrophoresis was not more than 100 V. The gel was stained in an ethidium bromide (EtBr) solution (~2.4 μg/ml) for 15 min. The EtBr molecules intercalate with the DNA helix structure, and the EtBr-DNA complex is fluorescent under UV(λ=254 or 300 nm) irradiation allowing DNA fragments on the gel to be visualized.

### 3.3.2 Pulsed field gel electrophoresis (PFGE)

Pulsed field electrophoresis is a technique for resolving large DNA molecules up to chromosomes which are not able to be separated by a standard agarose gel electrophoresis. By changing the electric field between spatially distinct pairs of electrodes, megabase sized DNAs are able to reorient and migrate at different speeds through the pores in an agarose gel. According to the instruction manual of Bio-Rad CHEF-DR III system, there are several parameters influencing PFGE.

a. Agarose concentration: Gel concentration affects the size range of DNA molecules to be separated, and the sharpness or tightness of the bands. A 1.0% (w/v) agarose gel is useful for separation of DNA molecules up to 3 Mb in size. Gel concentrations between 0.5 and 0.9% are useful for extremely high molecular weight DNA of greater than 3 Mb. For improving band tightness, 1.2- 1.5% agarose concentrations are used. However, running time increases proportionately.
b. Buffer concentration and temperature: In PFGE, the mobility of the DNA is sensitive to changes in buffer temperature. It is recommended that the buffer is chilled to 14°C to maintain band sharpness and to dissipate heat generated during prolonged electrophoresis. Tris-Borate-EDTA (TBE) buffer at a concentration of 0.5x is commonly used.
c. Switch times: In PFGE, DNA molecules change direction or reorient themselves in the gel matrix according to the switches of electric field. Larger molecules take longer to reorient and therefore have less time to move during each pulse, so they migrate slower than smaller molecules. As DNA size increases, an increased switch time is needed to resolve DNA of different sizes.

d. Voltage: DNA migration increases with increasing voltage. In general, lower voltage should be applied with increasing sizes. At high field strengths (6 V/cm), some very large DNAs (>3 Mb) can not be resolved on the gel and the voltage needs to be reduced. Some large DNA molecules do not enter the gels at high electric field strengths. Therefore, in selecting the field strength, a compromise between running time and resolution has to be made.
e. Field angle**:** Movement of DNA molecules up to 1 Mb is independent of the angle between 90° and 120°. Decreasing the included angle will decrease the resolution of smaller DNAs by causing them to pile up on each other. The field angle should be decreased to less than 120° when separating DNA fragments greater than 2 Mb.
f. Electrophoresis running time: The running time is determined by the migration rates of the DNA molecules. As the migration rate of the DNA decreases, the running time needs to be increased to adequately resolve the DNA fragments.

### 3.3.2.1 Sample Preparation

Large DNA fragments, such as YACs, are so fragile that they are sheared by mechanical forces during preparation. To avoid breakage of large DNA molecules, intact cells embedded in agarose were lysed and deproteinized in agarose blocks.

#### 3.3.2.1.1 YAC Plug Preparation

The method of yeast culture in medium was the same as described previously[45]. The cells were harvested by centrifugation. The pellet was washed with 20 ml TE followed by 20 ml SCE. The cells were collected by centrifugation at 3000 rpm for 10 min and resuspended in 800 µl Yeast lysis buffer (8 mg lysozyme /ml SCE). The cell solution was mixed with 800 µl 1.5% LMP agarose in SCE at 50°C. The mixtures were aliquoted to each block on ice. The agarose blocks were transferred into 10 ml SCE containing 10 mM DTT and incubated at 37°C for 60 min. The blocks were transferred into 10 ml proteinase K solution and incubated at 50°C for 48 hr, following two washes of 15 ml TE for 15 min. The blocks were incubated in 10 ml TE containing 0.04 mg/ml PMSF at 50°C for 30 min and washed two times with TE as described previously. The blocks were stored in 20 ml 0.5 M EDTA pH 8.0 at 4°C.

#### 3.3.2.1.2 Liquid Samples

PAC and BAC DNA were digested with *Not* I before carrying out PFGE, the DNA solution can be directly loaded into the wells of the gel. In addition, using a thin well comb (0.75 mm) improved the resolution and sharpness of the bands.

### 3.3.2.2 Parameters for PFGE

Equipment: Bio-Rad CHEF-DR III system
Running buffer: 0.5 x TBE
Gel: 1% Agarose-Pulse Field in 0.5 x TBE buffer

| | YAC | PAC/BAC |
|---|---|---|
| DNA | 1/3 YAC plug | 15 µl (~100ng) |
| Switch time | 50→100 sec | 1→20 sec |
| voltage | 6 V | 5V |
| Angle | 120° | 120° |
| Time | 24hr | 22hr |

The PFGE was performed in 0.5 x TBE buffer at 14°C. The gel was stained for 30 min in 500 ml 0.5 x TBE containing 100 µl EtBr (10 mg/ml) and destained for 10min in 0.5 x TBE. DNAs were visualized by UV irradiation. YAC sizes were determined by using the chromosomes of *Saccharomyces cerevisiae* (strain YPH49) as size markers which range from 260 kb to 2000 kb. BAC and PAC sizes were determined by using 50 kb size ladders (New England BioLabs) which range from 48.5 kb -1000 kb.

### 3.3.3 Denaturing polyacrylamide gel electrophoresis

Denaturing polyacrylamide gels can distinguish molecules differing in length by just a single nucleotide. In this study, denaturing polyacrylamide gels are used to separate the sequencing products and the PCR products which were amplified by different polymorphic markers.
Denaturing gels were prepared as follows:
After addition of 40 µl TEMED and 150 µl 10% APS, 60 ml of acrylamide gel solution was poured. The composition of acrylamide gel solution is as follows.

| Components | 6% Polyacrylamide gel | 8% Polyacrylamide gel |
|---|---|---|
| 10 M Urea | 375 ml | 350 ml |
| 10 X TBE | 50 ml | 50 ml |
| 40% Acrylamide/bis-acrylamide | 75 ml | 100 ml |

total volume: 500 ml

Electrophoresis:
The gel was prerun at 60 W for a minimum of 30 min. Samples were denatured with the loading buffer at 94 °C for 5 min before loading on the denaturing gel. Running time depended on the size of the desired DNA fragment using the loading dye as a size indicator. The relationship between gel concentration and dye mobility was adapted from *Current Protocol in Molecular Biology* and is shown below.

| Polyacrylamide gel (%) | Migration of bromphenol blue (bases) | Migration of xylene cyanol (bases) |
|---|---|---|
| 6 | 26 | 105 |
| 8 | 19 | 75 |

After electrophoresis, the gel was dried on Whatman 3MM paper for 45 min at 75 °C on a gel dryer and subjected to autoradiography. Autoradiography required a few hours to several days depending on the intensity of the signal.

### 3.3.4 Single-stranded conformation polymorphism gel electrophoresis

Single-stranded conformation polymorphism (SSCP) analysis was used to detect mutations based on single-nucleotide changes in DNA sequences that alter the mobility of single-

stranded DNA in nondenaturing gels. The DNA fragments containing mutations were amplified using polymerase chain reaction (PCR) and rendered single-stranded by heating in a denaturing buffer. The denatured strands were separated on polyacrylamide gels under conditions that resolve two molecules by conformational changes caused by differences in as little as one base.

SSCP gels were prepared as follows:

| | | |
|---|---|---|
| For 0.5 x MDE gel | 15 ml | 2 x MDE gel solution |
| | 3.6 ml | 10 x TBE buffer |
| | 41.4 ml | $dH_2O$ |

The gel solution was mixed with 500 µl APS and 50 µl TEMED and allowed to polymerize. The electrophoresis was carried out in 0.6 x TBE .

Electrophoresis:
The desired DNA fragments were amplified by a standard PCR procedure incorporating [$\alpha$-$P^{32}$] dCTP in the amplified fragments. PCR stop buffer was added to 20 µl of PCR product and samples were denatured at 90°C for 3 min before loading on a 0.5 x MDE gel. Electrophoresis was carried out at 400 V at 4°C overnight. The gel was dried and subjected to autoradiography.

### 3.4 Enzymatic manipulation of DNA

#### 3.4.1 Digestion of DNA with restriction endonucleases

Restriction endonucleases are isolated from bacteria and can recognize short DNA sequences and cleave double-stranded DNA at specific sites within or adjacent to recognition sequences. The amount of enzyme, DNA, buffer and ionic concentration, and the temperature and duration of the reaction were adjusted according to different applications.

##### 3.4.1.1 Single digestion

Single digestion refers to digesting a single DNA sample with a single restriction endonuclease. A total of 20 µl reaction including 0.1~ 4 µg DNA, 1 to 5 U restriction endonuclease /µg DNA and specific buffer concentration was convenient for analysis by gel electrophoresis. The reaction was incubated at 37°C for 1.5-2 hr and stopped at 70°C for 10 min. The volume of restriction endonucleases was not more than 1/10 the reaction volume because glycerol in enzyme buffer can interfere with the reaction.

##### 3.4.1.2 Double digestion

Double digestion refers to digesting a single DNA sample with two different restriction enzymes. Many restriction endonucleases are active in a wide variety of buffers, so it is usually possible to choose a standard buffer in which two enzymes can retain their activities. Alternatively, DNA can be first digested with the restriction enzyme in low salt buffer followed by the adjusting the reaction mixtures to a high salt condition and digested with the other enzyme.

### 3.4.1.3. Restriction endonuclease digestion of YAC plugs

To cleave the desired DNA fragments from YAC plugs, it was necessary to treat the YAC plugs before restriction enzyme digestion. One YAC DNA plug was cut into three slices and the slices were washed twice with TE for 30 min. The slice was mixed with 60 µl $dH_2O$ and 10 µl 10 x restriction enzyme buffer and dissolved at 70°C for 20 min. The sample was cooled down to 37°C and 15-30 U of restriction enzyme were added. The reaction was incubated at 37°C for 1.5-2 hr.

### 3.4.2 Dephosphorylation of DNA with calf intestine phosphatase (CIP)

Calf intestine phosphatase catalyzes the hydrolysis of 5'-phosphate residues from DNA, and therefore can be used to prevent self-ligation of cloning vector termini. The CIP reaction conditions for dephosphorylation of DNA were as follows:

| | |
|---|---|
| 10 x CIP buffer | 5 µl |
| $dH_2O$ | 14.5 µl |
| Vector-DNA(0.5-1.0 pmol /µl) | 30 µl |
| CIP (20 U/µl) | 0.5 µl |
| total volume: | 50 µl |

The reaction was incubated at 37°C for 1 hr and stopped at 75 °C for 10 min. The dephosphorylated DNA was precipitated with ethanol as described previously.

### 3.4.3 Preparation of blunt ends by repairing of 5' Overhangs with Klenow Fragment

If the 5' termini of vector and insert DNAs were incompatible, both could be modified with Klenow fragment to generate blunt ends for ligation. The reaction conditions are as follows:

| | |
|---|---|
| DNA (1-2 µg) | 15 µl |
| 10 x reaction buffer | 2.5 µl |
| $dH_2O$ | 2.5 µl |
| dNTP mix (125µM) | 4 µl |
| Klenow-polymerase (1-5U) | 1 µl |
| total volume: | 25 µl |

The reaction was incubated at room temperature for 20 min and stopped at 75°C for 10 min. The DNA was precipitated with ethanol.

### 3.4.4 Ligation of DNA with T4 DNA ligase

T4 ligase, originally isolated from phage T4, catalyzes the formation of phosphodiester bonds between 5' phosphate and 3' hydroxyl termini in duplex DNAs. Using ATP as a cofactor, T4 DNA ligase catalyzes the repair of single-stranded nicks in duplex DNA and join duplex DNA restriction fragments with either blunt or sticky ends. The standard reaction conditions are as follows:

| | |
|---|---|
| Vector DNA (100 ng/µl) | 1 µl |
| Insert DNA (100 ng/µl) | 4 µl |
| 10 X ligation buffer | 1 µl |
| $dH_2O$ | 3 µl |
| T4 DNA ligase | 1 µl |
| total volume: | 10µl |

Ligation was performed at 16°C for 16 hr and stopped by incubation at 75°C for 10 min. The ligation product was ready for transformation. The pBluescript II KS (+) was used as a vector in this study to ligate CAG positive fragments subcloned from a cosmid library.

Another method is used when insert DNA is embedded in low melting point (LMP) agarose gel. The procedure is as follows:
An agarose slice containing the desired DNA fragment was excised from a 0.7% LMP gel. The excised agarose volume should not be more than 20 µl. The agaraose slice was dissolved in 60 µl $dH_2O$ at 70°C for 10 min until the agarose slice was dissolved. The agarose concentration was less than 0.2% in the solution. The sample was incubated at 32°C and added with 1 µl vector DNA (100 ng), 8 µl 10 x ligation buffer, 2 µl 10 mM ATP, 1 µl T4 DNA ligase. Ligation was performed at 32°C overnight and stopped by incubation at 75°C for 10 min. The ligation product was ready for transformation.

### 3.5 Transformation of *E.coli* cells with plasmids

#### 3.5.1 Preparation of Competent Cells

A single colony of *E. coli* ( JM83 or XL1-Blue) from a freshly streaked LB agar plate was transferred into 50 ml LB medium followed by incubation at 37°C overnight with vigorous shaking. A 100 µl aliquot of the bacterial culture was transferred to 100 ml dYT medium and incubated at 37°C with shaking until the bacterial concentration reached 5 x $10^7$ cells/ml (O.D≅0.5). Bacteria were collected by centrifugation at 2500 rpm at 4°C for 10 min. The cell pellet was resuspended in 15 ml TBF I and incubated on ice for 60 min. Bacteria were collected by centrifugation at 2500 rpm at 4°C for 10 min and resuspended in 2 ml TBF II. Competent cells were aliquoted and stored at -70°C.

#### 3.5.2 Transformation

Competent cells were thawed on ice for 5 min. Plasmid DNA was mixed with the competent cells, the mixture was incubated on ice for 15 min followed by heat shock at 45°C for 90 sec in a water bath. A 200 µl aliquot of dYT was added to the cells followed by incubation at 37°C for 45 min. Cells were spread on the LB-antibiotics agar plates and incubated at 37°C overnight.

### 3.6 Southern Blotting

Southern blotting is the transfer of DNA fragments from an agarose gel to a membrane via upward capillary transfer of a high-salt buffer and subsequent hybridization with a complementary probe. The transfer results in immobilization of the DNA fragments so that the membrane carries a semipermanent copy of the banding pattern of the gel. After

immobilization, the DNA can be used for hybridization analysis which facilitate identification of bands with a sequence identical to a labeled probe.

DNA fragments were separated by agarose gel electrophoresis, denatured in denaturing solution for 1 hr, and soaked for 45 min in a neutralization solution. The blotting apparatus was set up in a chamber containing 20 x SSC and Whatman 3 MM paper saturated with 20 x SSC. The gel was turned upside down on the saturated Whatman 3 MM paper. The blotting membrane (Hybond N/N+) was laid exactly on the gel. Two sheets of Whatman 3 MM paper were placed on the blotting membrane. Additional paper towels were placed on the Whatman 3 MM paper.A 500 g weight was placed on top of this pile of papers and transfer of the DNA to the membrane was allowed to proceed overnight. The blotting membrane was labeled at the well edges and baked in a vacuum oven at 80°C for 2 hr.

### 3.7 Hybridization

The principle of hybridization is that a single-stranded DNA or RNA molecule (the so-called probe) can bind to a single-stranded DNA or RNA with a complementary sequence. The stability of the hybrid depends on the extent of the base pairing and base composition. In practice, a labeled probe is used to hybridize the target DNA that has been immobilized on a membrane. Hybridization analysis is usually performed with a radiolabeled DNA probe (100-1000 bp) and are sensitive enough to detect single-copy genes in complex genomes.

#### 3.7.1 DNA probe preparation

##### 3.7.1.1 Random primers labeling

Appligene-Oncor Nonaprimer labeling kit was used for random prime labeling and purification. A 1 µl aliquot of human genomic DNA (~300 ng) was mixed with 6 µl $dH_2O$ and denatured at 95-100°C for 10 min. The denatured DNA was chilled on ice for 5 min and mixed with 4 µl Nonaprimer mix, 3 µl dATP/dTTP/dGTP (1:1:1), 5 µl [$\alpha$-$P^{32}$]dCTP and 1 µl Klenow enzyme. The reaction mixture was incubated at 37°C for 30 min. The [$\alpha$-$P^{32}$]dCTP labeled DNA fragments were mixed with 60 µl adsorb solution and 2 µl DNAprep resin. The resin was centrifuged and washed twice with 100 µl wash solution. The labeled DNA fragments were eluted thrice with 100 µl elution solution.

##### 3.7.1.2 5´-end labeling of oligonucleotides

The reaction mixture composed of 1 µl (15 pmol/µl) oligonucleotide, 1 µl 10x reaction buffer, 4 µl [$\gamma$-$P^{32}$ ]ATP, 2 µl $dH_2O$ and 2 µl (1:10 diluted) T4 polynucleotide kinase incubated at 37°C for 1 hr. The reaction was stopped at 65°C for 10 min. The oligonucleotide was precipitated by a solution containing 30 µl EtOH,1 µl Glycogen and 1.1 µl 3 M NaOAc at -20°C for 30 min. The oligonucleotide was collected by centrifugation, washed with 70% EtOH and dissolved in 100 µl TE.

#### 3.7.2 Prehybridization and hybridization

The membrane was prehybridized in prehybridization solution (≥ 1 ml /10 $cm^2$ membrane) at 42°C for at least 4 hr. The DNA probe was denatured at 94 °C for 10 min and added to hybridization solution. The membrane was transferred to the hybridization solution and

incubated at 42 °C overnight. The membrane was stepwise washed by increasing washing stringency. The procedures were as follows:

| Washing buffers | Temperature | Time |
|---|---|---|
| 4 X SSC | room temperature | 3 x 15 min |
| 2 X SSC/0.5% SDS | 45°C | 20 min |
| 1 X SSC/0.5% SDS | 45°C | 20 min |
| 0.5 X SSC/0.2% SDS | 60°C | 30 min |
| 0.2 X SSC/0.5% SDS | 60°C | 30 min |
| 0.1 X SSC/0.1% SDS | 60°C | 30 min |

The radioactive signals were controlled with Geiger counter to avoid an over stringent wash. The membrane was packed in plastic wrap and subjected to autoradiography.

### 3.8 Polymerase chain reaction (PCR)

PCR is a rapid method for *in vitro* amplification of a specific fragments of DNA. The basis of PCR is that primers hybridize to DNA templates and then DNA polymerases synthesize the DNA segments between two primers from 5'→ 3'. These products accumulate with each subsequent cycle of denaturation, annealing to primers, and amplification. If n cycles of PCR are carried out, $2^{n-2}$ fold amplification of product can be synthesized. PCR was performed in 1x PCR buffer containing 0.25 M of each dNTP, 0.5 µM of forward/reverse primers, 2.5 U Taq polymerase and 30 ng genomic DNA in a reaction volume of 20 µl. The standard PCR conditions were 30 cycles at: 94 °C for 30 sec, 45-65°C for 60 sec, and 72°C for 15 sec. The reaction was preceded by 5 min denaturation at 94 °C and followed by a final extension at 72°C for 5 min. To incorporate [$\alpha^{32}$P]-dCTP into the PCR product during strands synthesis, a modified dNTP mix was used. The modified dNTP mix contains less dCTP (0.1 mM) and the concentration of dATP, dGTP and dTTP is unchanged. In this case, 0.1 µl [$\alpha^{32}$P]-dCTP is used in a 20 µl PCR reaction and the other components of PCR are the same as described above.

### 3.9 Reverse transcriptional polymerase chain reaction (RT-PCR)

#### 3.9.1 Total RNA isolation from human blood

QIAamp RNA Mini kit was used for isolation of total RNA. Human blood (2 ml) was mixed with 10ml buffer EL and incubated for 15 min on ice. Blood cells were collected by centrifugation at 400 x g for 10 min at 4°C. The cells pellet was resuspended into 4 ml buffer EL and repelleted. Pelleted leukocytes were resuspended into 600 µl buffer RLT containing β-mercaptoethanol. Lysates were transferred into a QIA shredder spin column and homogenized by centrifugation for 2 min at maximum speed. The homogenized lysates were mixed with 600 µl 70% ethanol and transferred into a new QIAamp spin column. The nucleic acids were adsorbed to the column by centrifugation for 15 sec at ≥8000 x g. The QIAamp spin column was washed with 700 µl buffer RW1 followed by 500 µl buffer RPE containing ethanol. RNA was eluted twice with 30-50 µl RNase-free water and stored at -80°C.

### 3.9.2 First strand cDNA synthesis and RT-PCR

An 11 µl aliquot of total RNA (1-5 µg) was mixed with 1 µl Oligo $(dT)_{15}$ (500µg/ml) and heated at 70°C for 10 min. To the mixtures of RNA and Oligo (dT) 4 µl 5x first strand buffer, 2 µl 0.1 M DTT and 1 µl 10 mM dNTP mix were added. This mixture was first incubated at 42°C for 2 min before the addition of 1 µl (200 U) SUPERSCRIPT II. The cDNA synthesis reaction was incubated at 42°C for 50 min and inactivated at 70°C for 15 min. The cDNA was ready for PCR amplification. The standard PCR procedure was performed by using 1-2 µl cDNA as template. The primers for amplifying *SPAST* cDNA are listed.

Table 3. Primers used for *SPAST* cDNA amplification and sequencing

| Primer | Sequence (5'->3') | Tm (°C) | Product size (bp) |
|---|---|---|---|
| SPA_Db | TAGCAGTGGCTGCCGCCGT | 62 | 655 |
| SPA_Dm | AAGCGGTCCTTGGCCATAAC | | |
| SPA_Dc | GGCGGCAGTGAGAGCTGTG | 60 | 543 |
| SPA_Dn | CTAGCTCTTTCACACTGTTC | | |
| SPA_Ad | AACAGGCCTTCGAGTACATC | 60 | 746 |
| SPA_Am | CTGTGAACAACTCAGGCCTC | | |
| SPA_Ba | CTACAACTGCTACTCGTAAG | 58 | 763 |
| SPA_Bm | CAGTGCTGCATCTTTTGCC | | |
| SPA_Ca | TGGAGATGACAGAGTACTTG | 56 | 766 |
| SPA_Cm | CTGGAATACTTTCATCTGC | | |

The primers are adapted from the database of Genoscope, Centre National de S quen age, France.( http://www.genoscope.cns.fr/)

### 3.10 DNA sequencing

Chain termination (or the Sanger method) is the most commonly used technique for sequencing DNA. All reagents needed for *in vitro* DNA synthesis are combined in this reaction, including a DNA polymerase and a 2', 3'-dideoxynucleotide (ddNTP). ddNTPs can be incorporated by DNA polymerases into growing DNA chains through 5' phosphate groups. However, the 3'-OH group necessary for phosphodiester bond formation and chain elongation is lacking so the chains terminate at the exact point at which the ddNTP is incorporated. Four sets of reactions are performed on each template, differing only by ddNTPs species. The dNTP:ddNTP ratio is selected so that the resulting labeled strands formed a nested set of molecules up to several thousand bases long, each terminating at a specific base. These are able to be separated according to size by denaturing acrylamide gel electrophoresis.

### 3.10.1 Sequencing plasmid DNA with thermo sequenase fluorescent labeled primer cycle sequencing kit

In this study, a thermo sequenase fluorescent labeled primer cycle sequencing kit was used to sequence the subcloned DNA fragments derived from the cosmid library.

To denature the template DNA, 10 µ1 plasmid DNA (~2.0µg) was mixed with 2 µl 2 M NaOH and incubated at room temperature for 10 min. The denatured DNA was precipitated with EtOH and dissolved in 23 µl sterile d $H_2O$.
Two common primers, M13 and reverse primers, were used to sequence the insert DNA cloned into pBluescript II KS (+). The sequences of primers are as follows:

M13(-20) primer: 5'-GTAAAACGACGGCCAGT -3'

Reverse primer : 5'-GGAAACAGCTATGACCATG-3'

The sequencing mixture comprising 1 µl DMSO and 1.5 µl fluorescent labeled primer (1 pmol/µl) was mixed with denatured 23 µl of plasmid DNA solution. Aliquots of 6 µl of this mixtures were mixed with 2 µl of reagents A, C, G, T, respectively. Cycle sequencing conditions were 30 cycles at 94°C for 15 sec, 60°C for 15 sec, and at 70°C for 15 sec. After cycle sequencing, 4 µl stop buffer were added to each reaction and samples were denatured at 90°C for 5 min. The samples were loaded on a 4% denaturing acrylamide gel. Sequences were analyzed by MWG CL4200 autosequencer.

### 3.10.2 Sequencing PCR product with thermo sequenase radiolabeled terminator cycle sequencing kit

A thermal sequenase radiolabeled terminator cycle sequencing kit was used to sequence all 17 exons, 5' UTR, 3'-UTR , intron 6, intron 10, and intron 11 of the *SPAST* gene. Each fragment was amplified by PCR and purified after agarose gel electrophoresis with a Nucleospin extraction kit.

In the cycle sequencing reactions [$\alpha$-$^{33}$P]ddNTPs were used as terminators. The sequencing reaction was performed according to the manufacture instructions. The cycle sequencing conditions were 30 cycles at 94°C for 30 sec, Tm for 30 sec, and 72°C for 90 sec. Tm was dependent on the sequencing primer used. The cycle sequencing products were loaded on a denaturing polyacrylamide gel and the sequences were read on an autoradiograph. The sequence and Tm of primers for amplification and sequencing of the various portions of *SPAST* gene are listed.

Table 4. The primers for amplification and sequencing of *SPAST*.

| Location | Primer | Sequence (5'->3') | Tm ($^{o}$C) | Product size (bp) |
|---|---|---|---|---|
| 5' UTR | spg4utr5af | GCCTTTGCGGTCTGGGTTCTGTGC | 64 | 661 |
| | spg4utr5ar | GGCTTGGGGTCGCTCTGCTGGTTC | | |
| | spg4utr5bf | GGTCGCCTGGCAGAAAAAGAT | 62 | 748 |
| | spg4utr5br | CAAACAGCGGGTAGGAGAAATAGT | | |
| exon1 | spgex1pza | GCCACCGACTGCAGGAGGAGAAGG | 65 | 721 |
| | spgex1pzb | CCGCAGAAAAGGGACGCAGGTGTT | | |
| | spgex1pbf | GTCGGTCTGCGGGAGGCGGGTTAT | 65 | 705 |
| | spgex1pbr | TGTGGGGGAAGGCTGGTGTCTGAA | | |

| | | | | |
|---|---|---|---|---|
| exon2 | spgex2pf<br>spgex2pr | TTTTTATGTATTACCTCTCAA<br>AAAAATAAATAAATAAATAAATAG | 45 | 266 |
| exon3 | spgex3paf<br>spgex3par | CTTTCTTTTGGGTATACATTTTCT<br>AGCAAGCGTCCATCTCAA | 55 | 340 |
| exon4 | spgex4paf<br>spgex4par | ATTTCTGTTATTTTCGTGACT<br>ATGCAAGCTTTATTATTTTATG | 54 | 335 |
| exon5 | spgex5pf<br>spgex5pr | CCTATGAAGATCCTGGTAC<br>TTTTATAGCAAGTTGCCCTG | 56 | 806 |
| exon6 | spgex6pbf<br>spgex6pbr | TCTCTAGTGAATACAGTTTTACC<br>ACAGGGCCCAGTTATTACAA | 60 | 552 |
| intron6 | spgin6pf<br>spgin6pr | TTAAGGGTACTCCGAAAAC<br>GACCAGCTATATCATCAAAT | 54 | 457 |
| exon7 | spgex7paf<br>spgex7par | ACTGGGCCCTGTTTGTAT<br>TATCCATTTTCCTATTCTAT | 55 | 352 |
| exon8 | spgex8pf<br>spgex8pr | CTGTTTGGGAAGATGCT<br>GTAAATAATAGACTCAAGGACAAG | 53 | 273 |
| exon9 | spgex9pf<br>spgex9pr | TGGCCTCATAGCTTACATTTTTAG<br>CCAGCCAGTTTACGGTATTTTATT | 55 | 315 |
| exon10 | spgex10pf<br>spgex10pr | GTAGTACTCTCCCCTTTCTCA<br>TAATGTTTCCAATCGTATCTT | 55 | 307 |
| intron10 | spgin10pf<br>spgin10pr | GTAGTACTCTCCCCTTTCTC<br>TGCTCCCCTTCTCTTCTTTCAC | 60 | 505 |
| exon11 | spgex11pf<br>spgex11pr | ACTCACATAGCTTGGTCTT<br>CATGAGTAAATATTGTCTGTAA | 55 | 418 |
| intron11 | spgin11pf<br>spgin11pr | AGACGCCTAAAAACTGAAT<br>TTGCCTAAAACCGTAACTA | 52 | 482 |
| exon12 | spgex12pf<br>spgex12pr | ATGGCCAAGGTTAAAAATACAA<br>CTGGAAGAAAATAGTGAAT | 54 | 281 |

| | | | | |
|---|---|---|---|---|
| exon13 | spgex13pf | CTTTTCCTGTCATTTGCTGTTT | 60 | 311 |
| | spgex13pr | CATTTTTGTTTCTCTGGGGTAA | | |
| exon14 | spgex14paf | GGAGGCTGAGATGGGAGGATT | 60 | 406 |
| | spgex14par | AAGGCAAAGGAGGTAGAGGATGAG | | |
| exon15 | spgex15paf | AAAAAGCGGGAGGGGAAATA | 60 | 328 |
| | spgex15par | GAAGGCTGGGTGGGAGAATCA | | |
| exon16 | spgex16paf | TTCAACTGCAAAATGTATGTA | 54 | 250 |
| | spgex16par | ATGGTGAAATGCCCTCTC | | |
| exon17 | spgex17p1f | ATAACATTAAGAAACAGCAGCATC | 60 | 845 |
| | spgex17p1r | AAAAGGTAAAAAGAACACAACAAT | | |
| | spgex17p2f | AATTGTTGTGTTCTTTTTAC | 50 | 745 |
| | spgex17p2r | TGACATTTTACATAGCATC | | |
| | spgex17pe | GGAATGCCAAACACTCTT | 53 | 312 |
| | spgex17pf | AAAAATACTGCAGGTCACAT | | |
| 3' UTR | spg4utr3af | ATAACATTAAGAAACAGCAGCATC | 62 | 850 |
| | spg4utr3ar | AAAAGGTAAAAAGAACACAACAAT | | |
| | spg4utr3bf | AATTGTTGTGTTCTTTTTAC | 50 | 747 |
| | spg4utr3br | TGACATTTTACATAGCATC | | |
| | spg4utr3cf | GGAATGCCAAACACTCTT | 53 | 313 |
| | spg4utr3cr | AAAAATACTGCAGGTCACAT | | |

*Note*: The sequences are adapted from the database of Genoscope, Centre National de Séquençage, France ( http://www.genoscope.cns.fr/). The primers are designed according to the sequence of the exons including at least 30 nucleotides of the preceding and followed introns.

### 3.10.3 Sequencing DNA with a T7 sequenase kit

A T7 sequenase kit was used to prepare a basewise size ladder by using phage DNA M13 as a template. The preparation was as follows:
Master mix A was composed of 20 µl DNA (1~2µg), 4 µl primer (4 ng/µl) and 4 µl annealing buffer. Master mix B was composed of 3.6 µl dilute buffer, 6.5 µl label-mix, 1.6 µl $^{35}$S dATP and 0.4 µl T7 polymerase. Master mix A was first incubated at 60°C for 10 min then at room temperature for 10 min. Master mix A and 12 µl Master mix B were mixed and placed at room temperature for 5 min. Aliquots of 9 µl of this mixture were added

to 5 µl ddATP, ddCTP, ddGTP, and ddTTP. This reaction was incubated at 37°C for 5 min and stopped with 10 µl stop buffer at 90°C for 3 min .

### 3.11 Cosmid library construction

Cosmid vectors are valuable tools for isolation and physical mapping of large fragments of DNA ranging in size from 10 to 42 kb. SuperCos is a cosmid vector containing bacteriophage promoter sequences flanking a unique cloning site. It has been engineered to contain genes for amplification and expression of cosmid clones in eukaryotic cells. In addition, most genomic inserts can be excised as a single large restriction fragment using the *Not* I restriction site that flanks the SuperCos 1 polylinker.

#### 3.11.1 Preparation of host cells

XL1-Blue MR was used as the host cell and its genotype was previously described in Material section. A single colony of XL1-Blue MR was inoculated in 50 ml NZYC medium and incubated at 37°C overnight. The host cells from 10 ml culture were harvested by centrifugation at 2500 rpm for 15 min and resuspended into 25 ml 10 mM $MgSO_4$ . The host cells were stored for up to 2 days at 4°C.

#### 3.11.2 Preparation of a -80°C glycerol stock

A 10 ml aliquot of LB medium was inoculated with a single colony of host cells and incubated at 37°C for 6-8 hr. The cells culture was mixed with 4.5 ml solution of a sterile glycerol-LB (1:1) medium. The mixtures were aliquoted and stored at -80°C.

#### 3.11.3 Preparation of vector DNA

SuperCos 1 vector (~85 µg) was restricted with 5 µl *Xba* I (20 U/µl) in a total of 100 µl buffer H at 37°C for 2 hr. The linear SuperCos 1 vector is about 7.6 kb. The volume of the digestion reaction mixtures was adjusted to 200 µl with TE. The solution was extracted with 200 µl phenol/chloroform (1:1) and chloroform. The DNA was precipitated with 500 µl EtOH containing 25 µl 3M NaOAc pH5.5 at -20°C for 30 min and collected by centrifugation. The pellet was washed with 70% EtOH and air dried. The DNA was dissolved in 20 µl 1x dephosphorylation buffer. The *Xba* I cut vector DNA was dephosphorylated with 2µl CIP (10 U/µl) at 37°C for 90 min and extracted with phenol/chloroform as described above. The DNA was dissolved with 74 µl 1X buffer B. The vector was restricted with 6 µl *Bam*H I (20 U/µl) at 37°C for 2 hr. The two DNA fragments are 1.1 and 6.5 kb. The DNA was extracted with phenol/chloroform and dissolved in 20 µl TE. The restricted vector was stored at -20°C.

#### 3.11.4 Partial digestion of YAC DNA

In order to clone DNA fragments into the *Bam*H I site of the SuperCos 1 cosmid vector, the YAC DNA was partially digested with *Sau* 3A. An optimal concentration of *Sau* 3A for obtaining appropriate partial digestion products was previously determined.

The YAC DNA agarose plug was washed twice with TE and dissolved in 120 µl 1x buffer A at 70°C. *Sau* 3A (0.03 U/µl) was added to the solution and incubated at 37°C for 15 min. The reaction was stopped at 70°C for 10 min. The partially digested DNA was dephosphorylated

with 5µl CIP at 37°C for 60 min. The DNA was extracted with phenol/chloroform, precipitated with EtOH /0.3 M NaOAc pH 5.5, and dissolved in 12 µl TE.

### 3.11.5 Ligation

The reaction mix was composed of 12 µl partially digested YAC DNA [Sau 3A-CIP], 3 µl (1 µg/µl) SuperCos 1 DNA [*Xba* I-CIP/*Bam*H I], 2µl 10x ligase buffer, and 2 µl 10 mM ATP and 1 µl T4 DNA ligase. The ligation mixture was incubated in a water bath at 32°C overnight.

### 3.11.6 Packaging

The packaging extracts (Gigapack III Gold Packaging Extract; Stratagene) were quickly thawed at room temperature until the contents of the tube began to thaw. The ligation product was added immediately to the packaging extract and stirred gently with a pipette tip. The packaging mixture was spun quickly for 5 sec and incubated at room temperature (~20°C) for 2 hr. To the mixture 470 µl SM buffer and 10 µl Chloroform were added. The cosmid library was ready for titering and could be stored at 4°C for 1 month.

### 3.11.7 Titration of the cosmid packaging reaction

The cosmid library was diluted 1:10 and 1:50 in SM buffer. The diluted cosmid library (25µl) was mixed with the equal volume of host cells in 10 mM $MgSO_4$ and incubated at room temperature for 30 min. Host cells were regrown in 200 µl dYT at 37 °C for 60 min and collected by centrifugation at 9000 rpm for 1 min. The cell pellet was resuspended in 50 µl LB medium. The cell suspension was spread on LB-Amp agar plates and incubated at 37 °C overnight. The colony forming units per milliliter (cfu/ml) were estimated.

### 3.11.8 Screening the cosmid library

The titer of the cosmid library was determined by serial dilution as described above. The appropriate amount of host cell suspension for plating was calculated and the suspension was diluted and ~$10^4$ cfu was spread on the LB-Amp plates (10 x 15 cm in diameter petri dish). The plates were incubated at 37°C overnight. Hybond N/N+ Nylon membranes (15 cm in diameter) were used to perform colony lifts. Host cells were regrown on LB-Amp plates at 37 °C for 4-6 hr until the bacterial colonies began to reform and then stored at 4°C. Colony lift and hybridization on Nylon membranes were prepared as follows.

The membranes were placed colony-side up for 10 minutes on the surface of Whatman 3 MM paper prewetted with 0.5 M NaOH and transferred to a sheet of Whatman 3MM paper prewetted with 1 M Tris-HCl (pH 7.6) for 10 min. Membranes were transferred to a sheet of Whatman 3 MM paper prewetted with 1 M Tris-HCl (pH 7.6) and 1.5 M NaCl for 15 min. Cells debris were removed by carefully brushing the membranes in the solution of 1 M Tris-HCl (pH 7.6) and 1.5 M NaCl. The membranes were air dried and baked at 80°C for 2 hr.

Prehybridization and hybridization were performed at 42°C. The membranes were washed with 4 x SSC, 0.5% SDS buffer followed by washes of increasing stringency until the background diminished. The membranes were packed with plastic wrap and subjected to autoradiograhy.

### 3.11.9 Purification of the positive clones

The autoradiograph and original host plate were aligned. The positive candidate colonies were transferred in array to a new LB-Amp plate and incubated at 37°C overnight. The selective colonies were rescreened by hybridization as described previously[46]. A 4 ml aliquot of LB-Amp medium was inoculated with each positive clone and incubated at 37°C overnight. The bacterial culture was used for isolating the cosmid DNA as described previously and the remaining 1 ml culture was used for preparation of glycerol stock.

## 3.12 YAC contig construction

### 3.12.1 Yeast artificial chromosome (YAC)

A total of 26 YACs were chosen by searching databases, the Whitehead Institute for Genome Research (http://www-genome.wi.mit.edu/), Genethon (http://www.genethon.fr/genethon-en.html), and the previously reported YAC contig (De Jonghe *et al.*[84]) and obtained from CEPH (Centre d'Etude Polymorphisme Humain, Paris, France). The YAC clones were plated on (-)Ura, (-)Trp agar plates and then single colonies were transferred into selective media for YAC DNA isolation using standard procedures.

### 3.12.2 Sequence tagged site (STS)

STS markers which had been assigned to 2p21-p24 (http://www.ncbi.nlm.nih.gov/SCIENCE96, http://www.ncbi.nlm.nih.gov/genemap98, http://www.ncbi.nlm.nih.gov/genemap99) were used to perform a series of PCR reactions with YAC DNAs. The sequences and annealing temperatures of primers are shown in Table 5.

Table 5. Sequence tagged sites (STSs) for YAC , BAC and PAC mapping

| Name | Primer | Sequence (5'->3') | Tm ( $^{o}$C) | Product size (bp) |
|---|---|---|---|---|
| D2S352 | D2S352f | GCAAAGTCGTTCTCAGGTG | 55 | 205 |
| | D2S352r | CTACAGGGCTTCAGCATCC | | |
| D2S367 | D2S367f | TTCTTTGGTCTAAGGGTCAC | 55 | 148 |
| | D2S367r | AGCTTCTTGTTCACAGGTGT | | |
| D2S400 | D2S400f | AATGTGACAAAGCCCAGTGTTAGC | 55 | 190 |
| | D2S400r | GATAATCTCCCTGAGTATGTGTGCC | | |
| D2S1794 | D2S1794f | TTCTTGCCAGAATGAAGTAAGC | 55 | 124 |
| | D2S1794r | TTGAGGACAAGAGCAGTATAGAGG | | |
| D2S1998 | D2S1998f | TGAAACCCCATTAAGAAGAGATC | 55 | 203 |
| | D2S1998r | GGGAGAACTTTTCAGTTAATATTCC | | |

| | | | | |
|---|---|---|---|---|
| D2S2203 | D2S2203f | AAGTGCTTAGAAGGGTTCCTGAC | 55 | 264 |
| | D2S2203r | ATATGCTCTCTGGTGACTGTAGGTG | | |
| D2S2005 | D2S2005f | ATGAACAGATTTCACTGCTACTCTC | 55 | 226 |
| | D2S2005r | GCCTTTTTCTGAGTTTCTTCTCA | | |
| D2S2255 | D2S2255f | CTCCAGGCTACTTTGAGGATTTT | 55 | 214 |
| | D2S2255r | CCATTTGTTTGCAGATTCTTGTAT | | |
| D2S2283 | D2S2283f | GATCATGGCCCCAATA | 55 | 273 |
| | D2S2283r | GCCCCAGGTAACCACT | | |
| D2S2347 | D2S2347f | TGCCTAAGCCCATGTC | 55 | 272 |
| | D2S2347r | CTGCGTCACTTCACATAGA | | |
| D2S2351 | D2S2351f | TGAACTTTGCAGTGAGAAA | 55 | 241 |
| | D2S2351r | TTTACTGTGTATGTGTGTGTACTCT | | |
| D2S3008 | D2S3008f | CCAGCAGGCAGAGTGAGAATC | 63 | 224-292 |
| | D2S3008r | TGTTTAAAGCACAAGGGCAGT | | |
| A004H42 | A004H42f | AACAAAATTATTACAGGTGG | 55 | 168 |
| | A004H42r | GGTTTATAGTAGTCCATGCTATA | | |
| A008N33 | A008N33f | TAGCAGATACGGTCAAGACTAGT | 56 | 105 |
| | A008N33r | GTGGTTAAATAGTTGACGGATTA | | |
| Cda15g09 | Cda15g09f | GCTTGAAATTCAGGTCCTCTC | 56 | 249 |
| | Cda15g09r | GGGTTATTTGGGGAAGTGTTG | | |
| Cda0sd05 | Cda0sd05f | GCAGTGGCACAATTATAAC | 54 | 155 |
| | Cda0sd05r | AATAAAGTAAGCTCCCGTC | | |
| GAAT-P33068 | GAAT-P33068f | TCCTTCATCATGTGATTCCC | 55 | 152 |
| | GAAT-P33068r | TTACCAGTGGCCATTACTGC | | |
| N52847 | N52847f | TAGCAGCTTGGCTAAGAACG | 56 | 110 |
| | N52847r | AGGCCAGTGATTTTGTTCC | | |
| SGC32192 | SGC32192f | TTTTAAAAAGAGACATGGGTCTCA | 55 | 136 |
| | SGC32192r | CTGTAGTGCCACTACTCTGAAAGC | | |
| SGC32499 | SGC32499f | AAATGTCAGCCCTTCACAGC | 55 | 127 |
| | SGC32499r | TACAGTATAGTCAGTGGCAGTTTCC | | |
| SGC33436 | SGC33436f | TCTGCTTTTAAAGATTCTTCATAGC | 55 | 150 |
| | SGC33436r | AGAGTGGAAATAAGACTGGCTATC | | |

| | | | | |
|---|---|---|---|---|
| SHGC12567 | SHGC12567f | AGAGAGTCCTCAGCAGAGTCTTC | 58 | 224-245 |
| | SHGC12567r | TATTCACCATTTAGGCATAACAGTT | | |
| SHGC13568 | SHGC13568f | TGACCCCCAAGCTTTCTGT | 57 | 139 |
| | SHGC13568r | CCTACAGAAGCAGCAACCGT | | |
| stSG1707 | stSG1707f | TGCACTGCAGTAGTAAGCACG | 59 | 196 |
| | stSG1707r | CCCACTGTACTCAATGCCCT | | |
| stSG2515 | stSG2515f | TGTTGGTGGCAAAGAAAACA | 57 | 186 |
| | stSG2515r | AAGGAAAAGGGTGAAGGGAA | | |
| stSG3789 | stSG3789f | GCAGGTGAAAGCACCTTTTC | 55 | 131 |
| | stSG3789r | CCCAGGGAATTAGCTGATGA | | |
| stSG4150 | stSG4150f | TGCACTGCAGTAGTAAGCACG | 58 | 138 |
| | stSG4150r | GCCCACTGAACCAATCACTT | | |
| stSG47880 | stSG47880f | CTTTGCCATGAATGAGAGCA | 55 | 128 |
| | stSG47880r | TTGAAGAGGGGATCTTGGTG | | |
| stSG15464 | stSG15464f | CAGCTTGATGAATTTTCACAAA | 55 | 204 |
| | stSG15464r | TTTTGTATGGCCTTAAATCTGG | | |
| stSG22421 | stSG22421f | ACGAGTCTCGACGTGCAAG | 56 | 148 |
| | stSG22421r | TCACTGTTATGGTCCTGTCAGG | | |
| stSG29624 | stSG29624f | CAAAGCTCTGCCCTGTAAGG | 56 | 175 |
| | stSG29624r | ACTGTTGTGGCCTTAATGTCG | | |
| stSG30783 | stSG30783f | TCTACCCTCCAGAAGCTTGTG | 55 | 146 |
| | stSG30783r | TGCTCAACTGCACTGTGAAA | | |
| stSG31920 | stSG31920f | TAAATCAATAAAGCCAGCCTCT | 55 | 122 |
| | stSG31920r | AGCAGACATTGCACGGGT | | |
| stSG32067 | stSG32067f | GGTCCGAGTTCAAATTCTGC | 55 | 84 |
| | stSG32067r | CGAGCTCCGTCTCAAAAAAA | | |
| stSG44150 | stSG44150f | TAAAATGTGCTTTCGCTGCA | 55 | 81 |
| | stSG44150r | CGTTATTAAGCTATTGAAGGCC | | |
| stSG48207 | stSG48207f | AGCACGGCAACAAAGAAATT | 56 | 128 |
| | stSG48207r | CGTCTTGAAAGGTGAAAGGTG | | |
| stSG51853 | stSG51853f | AAAAACCCACTTTATGCAACTG | 55 | 135 |
| | stSG51853r | ACCATTTACAGTTCGGTTTTGG | | |

| | | | | |
|---|---|---|---|---|
| stSG52124 | stSG52124f | TAAGCATGGAAGCTACATCAGC | 56 | 189 |
| | stSG52124r | GCCTGGATTTTCCCCTTTAT | | |
| stSG62506 | stSG62506f | ACCTAGTAGGGGCAAGTTTTT | 55 | 140 |
| | stSG62506r | TTCTCTGGTGCTTAAATGATGC | | |
| stSG63103 | stSG63103f | TAAGGGCACCAAAAACCAAC | 55 | 189 |
| | stSG63103r | CCTCATCCCAGCCATGTC | | |
| WI9534 | WI9534f | TTTTATTAACAATTTGTTGGTGGC | 55 | 199 |
| | WI9534r | AAGGAAAAGGGTGAAGGGAA | | |
| WI9668 | WI9668f | AAACAGTTTAATTCAAGCAATCTGG | 55 | 281 |
| | WI9668r | AGTGTTGGAAAAAAGATTAGGGC | | |
| WI11827 | WI11827f | TTTTTTGCTGTAAAACAAAACACA | 55 | 111 |
| | WI11827r | CATAAAAGTTGGGGAAAGAGACC | | |
| WI16788 | WI16788f | TGTCTGGAAATACTTTCCCATT | 55 | 150 |
| | WI16788r | CAGACTCATCAAGTTTCAGCAA | | |
| WI16901 | WI16901f | ACATATGTGCTGTCTACTATGAGCC | 55 | 103 |
| | WI16901r | AAATAGGTACTCCTTGGAATCTGTG | | |

### 3.13 BAC and PAC contig construction

To construct a physical BAC and PAC map, human genomic libraries,RPCI-1, RPCI-6, RPCI-11 and human chromosome 2 specific library LL02NP04"AI" were screened either by hybridization or PCR. The sequences and annealing temperatures of the STSs used to screen several genomic libraries were shown in the section of YAC contig construction.

### 3.13 1 High density filter libraries

The high density filters are ready for prehybridization. A standard hybridization was performed as described previously. Several STSs or ESTs that assigned to chromosome 2p21-24 from the GenBank 98 and 99 Map were chosen as probes for hybridization. The probes were first synthesized by PCR and then labeled with [$\gamma$-$^{32}$P ]ATP at 5'ends. The 5'labeling was described in the hybridization section. To reduce the removal of DNA from filters, the filters were reused for hybridization in subsequent steps without removing the hybridized probes. The new positive hybridizing signals appeared on the autoradiogram after the hybridization with a new probe. If the background was too high, filters were stripped by washing at room temperature in 0.2 M NaOH for 30 min, and soaked in 0.1 x SSC/0.1% SDS/ 0.2 M Tris-HCl pH 7.5 for 15 min.
The interpretation of positive signals was according to the instructions of German Resource Center and UK HGMP Resource center because different spotting strategies of libraries were used. The positive clones were confirmed by PCR and the sizes of the clones were determined by PFGE.

### 3.13.2 DNA pool libraries

Two DNA pool libraries, RPCI-6 and RPCI-11, were used for PCR screening. RPCI-6 was a PAC library consisted of 36 primary pools. RPCI-11 was a BAC library consisted of 144 primary pools. DNA primary pools were first screened and the secondary pools were dettermined by the primary pools screening results. The secondary pools were screened further to obtain the positive candidate clones. The DNA primary pool sample (5 µl) was diluted with 45 µl sterile water and 1 µl diluted DNA was used as template for a standard 20 µl PCR. PCR was performed as described before. The lyophylized DNA secondary pool sample needed to be resuspended in 5 µl sterile water and the subsequent procedures of PCR screening is the same as described above.

### 3.14 Genotyping and linkage analysis

DNA was extracted from peripheral blood as described previously and was amplified with different STRP markers by standard procedures of PCR. [$\alpha^{32}$P]-dCTP was incorporated in the PCR products during amplification. Amplified DNA was electrophoresed on 6% plyacrylamide/7M urea gels, and alleles were scored on the basis of autoradiographs. The microsatellite polymorphic markers are listed in Table 6. Two-point LOD scores were calculated using the LINKAGE version 5.1 program package. The following assumptions were made: a gene frequency of $10^{-5}$ for autosomal dominant SPG; equal allele frequencies for each marker; equal recombination rates in males and females and 90% as the penetrance rate.

Table 6. STRPs markers for linkage analysis at SPG 3,4, 6 loci.

| Name | Primer | Sequence | Tm(°C) | Product size(bp) |
|---|---|---|---|---|
| SPG3 | | | | |
| *D14S63* | D14S63A | GGCCAGGTTTCAATCAGTTT | 55 | 208 |
| | D14S63B | GCCAGAGAGCCACACTGTAT | | |
| *D14S66* | D14S66A | GGCAACAGACTTGACCAATC | 55 | 185 |
| | D14S66B | CGTTCAGTAAGCAGAGAGCA | | |
| *D14S69* | D14S69A | AAAGCCACACTGCTAGTCAC | 55 | 204 |
| | D14S69B | TTCAGATGCCAATTAAGGGA | | |
| *D14S75* | D14S75A | TGTCCCCAGGTGTTA | 55 | 185 |
| | D14S75B | CCAAGTGGCCTTGCC | | |
| *D14S79* | D14S79A | AGGTTGATAGACCATGGAGACA | 55 | 82 |
| | D14S79B | TTTTATTGTTATGTGGCTTTCA | | |
| *D14S266* | D14S266A | ACAAGCCCCATATATTCATG | 55 | 135 |
| | D14S266B | AATAGACTTCCAAATCTTCAGATA | | |
| *D14S269* | D14S269A | CACATGGCATTACCAC | 55 | 221 |
| | D14S269B | GCAACATGCTTGACAGG | | |
| SPG4 | | | | |
| *D2S146* | D2S146A | TCATCCTTACTCTAAGCAAAGATCC | 55 | 189 |
| | D2S146B | CACCACATTCAAATGCCTCC | | |
| *D2S170* | D2S170A | TTGCTCAATAATGTCAGGTG | 55 | 216 |
| | D2S170B | CGCATGAGAGGCGTCT | | |
| *D2S352* | D2S352A | GCAAAGTCGTTCTCAGGTG | 55 | 205 |
| | D2S352B | CTACAGGGCTTCAGCATCC | | |
| *D2S367* | D2S367A | TTCTTTGGTCTAAGGGTCAC | 55 | 148 |
| | D2S367B | AGCTTCTTGTTCACAGGTGT | | |
| *D2S400* | D2S400A | AATGTGACAAAGCCCAGTGTTAGC | 55 | 190 |
| | D2S400B | GATAATCTCCCTGAGTATGTGTGCC | | |
| SPG6 | | | | |
| *D15S122* | D15S122A | GATAATCATGCCCCCCA | 55 | 146 |
| | D15S122B | CCCAGTATCTGGCACGTAG | | |
| *D15S128* | D15S128A | GCTGTGTGTAAGTGTGTTTTATATC | 55 | 206 |
| | D15S128B | GCAAGCCAGTGGAGAG | | |
| *D15S156* | D15S156A | CAGCCACCGCATTCTA | 55 | 229 |
| | D15S156B | CAACACCATTTATTGAAGAGAC | | |
| *D15S165* | D15S165A | GTTTACGCCTCATGGATTTA | 55 | 208 |
| | D15S165B | GGGCACACAGTCCCAA | | |

**3.15 WWW. web sites**

| Organization/Database | Address |
|---|---|
| The National Center for Biotechnology Information | http://www.ncbi.nlm.nih.gov/ |
| Medline | http://www.ncbi.nlm.nih.gov/PubMed/ |
| Online Mendelian Inheritance in Man (Human Genes and Genetic Disorders) | http://www.ncbi.nlm.nih.gov/Omim/ |
| Cooperative Human Linkage Center (Human Genome Map) | http://lpg.nci.nih.gov/CHLC/ |
| Whitehead Institute (Physical Mapping) | http://www-genome.wi.mit.edu/ |
| Online Linkage Analysis Software | http://linkage.rockefeller.edu/ |
| Standford Human Genome Center (Radiation Hybrids Mapping) | http://www-shgc.stanford.edu/ |
| European Bioinformatics Institute (Sequence Database) | http://www.ebi.ac.uk/ |
| Radiation Hybrids database | http://www.ebi.ac.uk/RHdb/ |
| Human Gene Mutation Database | http://www.archive.uwcm.ac.uk/ |
| UK Human Genome Mapping Project Resource Cernter (HGMP) | http://www.hgmp.mrc.ac.uk/ |
| Sanger Institute (Human Genome Sequences) | http://www.sanger.ac.uk |
| Genome database | http://www.gdb.org/ |
| German Resource Center (German Human Genome Project Funded) | http://www.rzpd.de/ |
| CEPH (YAC Mapping) | http://www.cephb.fr/ |
| Genethon (SSLP Maps) | http://www.genethon.fr/ |
| Genoscope (Human Chromosome 2p21-p24 Sequences) | http://www.genoscope.cns.fr/ |

# 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 (*SPG*3, 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 famliy vs chromosome 14q, 15q, 2p markers

| SPG locus | Microsatellite | Lod scores Z at recombination fraction (θ) | | | | | | | $Z_{max}$ | $\theta_{max}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 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 haploptype 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 Dub∈ *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 likage. Linkage can be rejected if the lod score < -2. The lod scores between –2 and +3 are inconclusive.

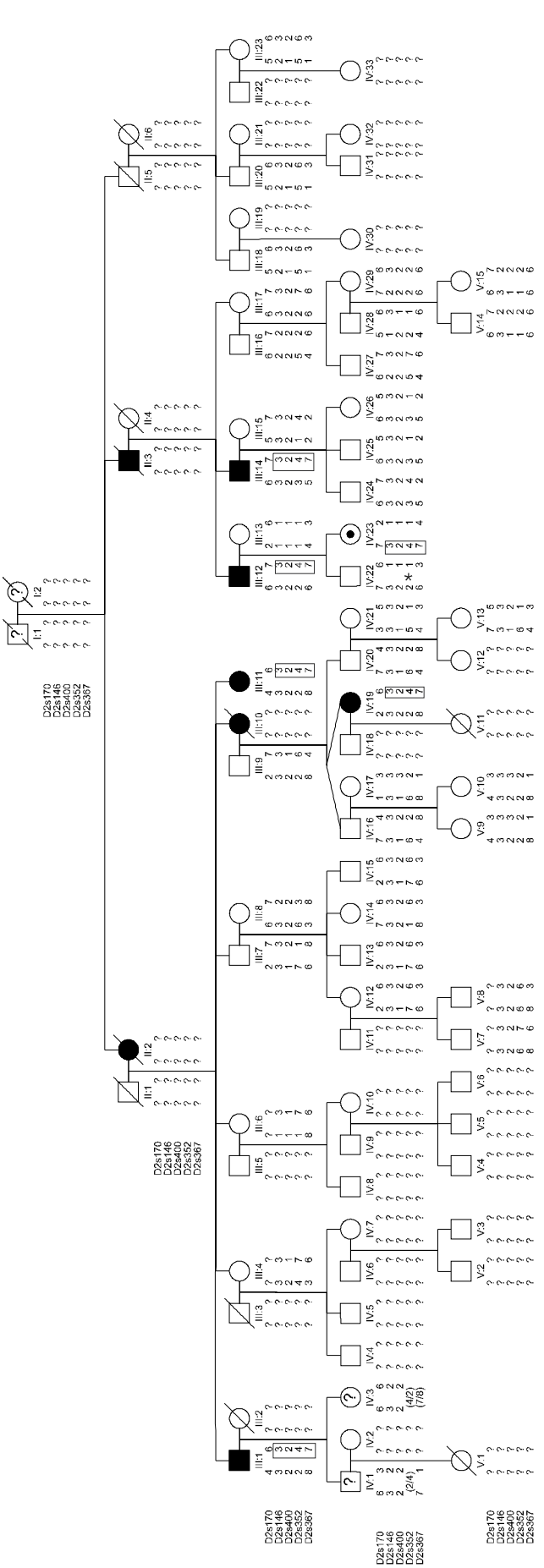


Figure 1. Pedigree of the SPG4 family. Symbols: squares=males;circles=females; slashed symbols =deceased individuals; filled symbols = affected patients; symbols with a dot = presymptomatic individual; question marks = unknown status. The alleles are placed between brackets when phase is unknown. Haplotype of the chromosome 2 loci D2S170, D2S146, D2S400, D2S352 and D2S367 are given below the symbols and one recombinant(IV:22) has been observed. The disease chromosome is boxed.

## 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).

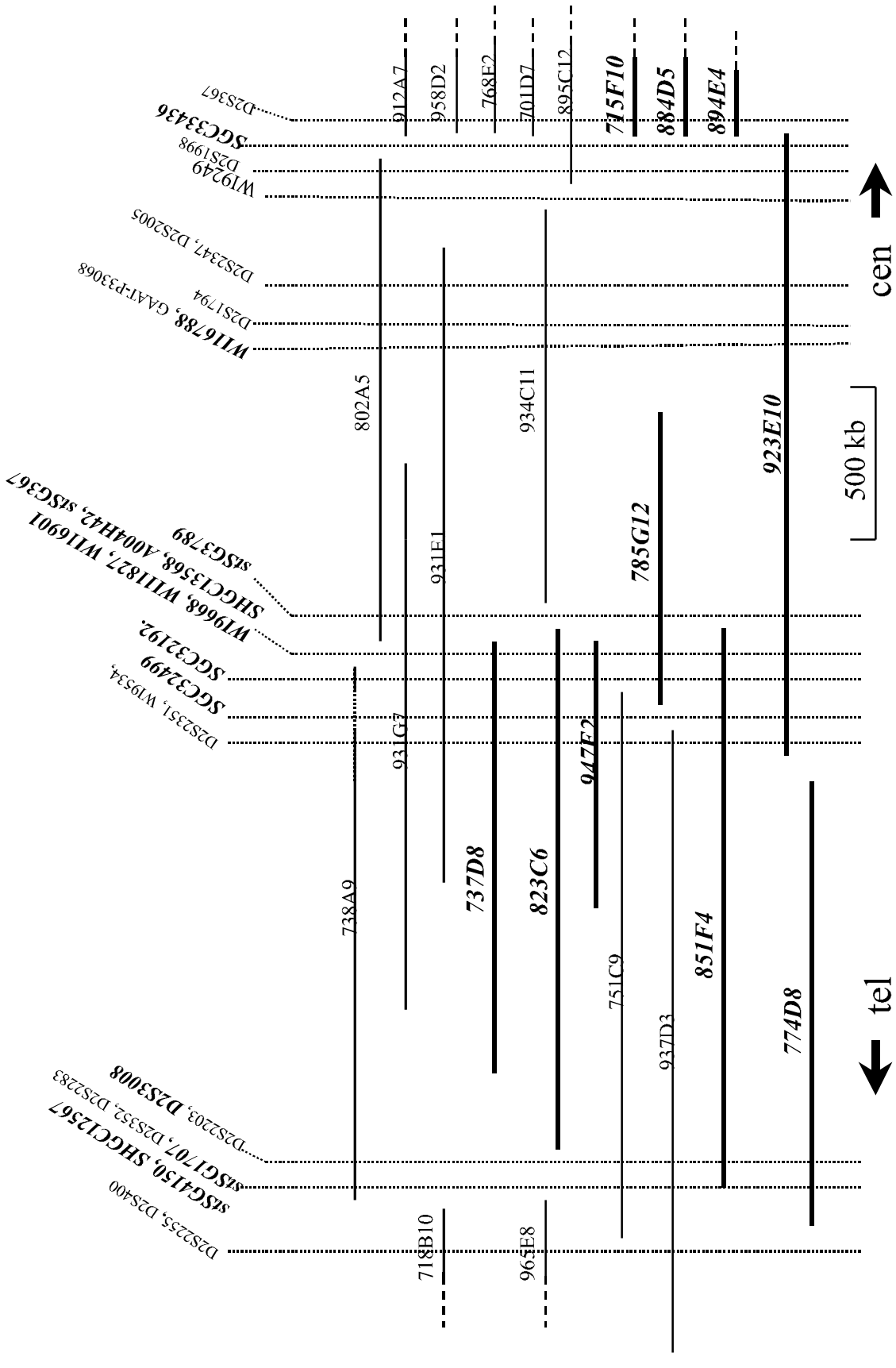



Figure 2. YAC contig of the *SPG4* critical region. Bars indicating newly assigned YACs are printed in bold and newly assigned ESTs and the STRP *D2S3008* are printed in *italics* and bold.

### 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 study[49]. 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 publication[37]. 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 | + | | | |
| WI11827 | | | | | | | + | + | + | + | + | + | + | + | | | |
| 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) method[52]. 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.

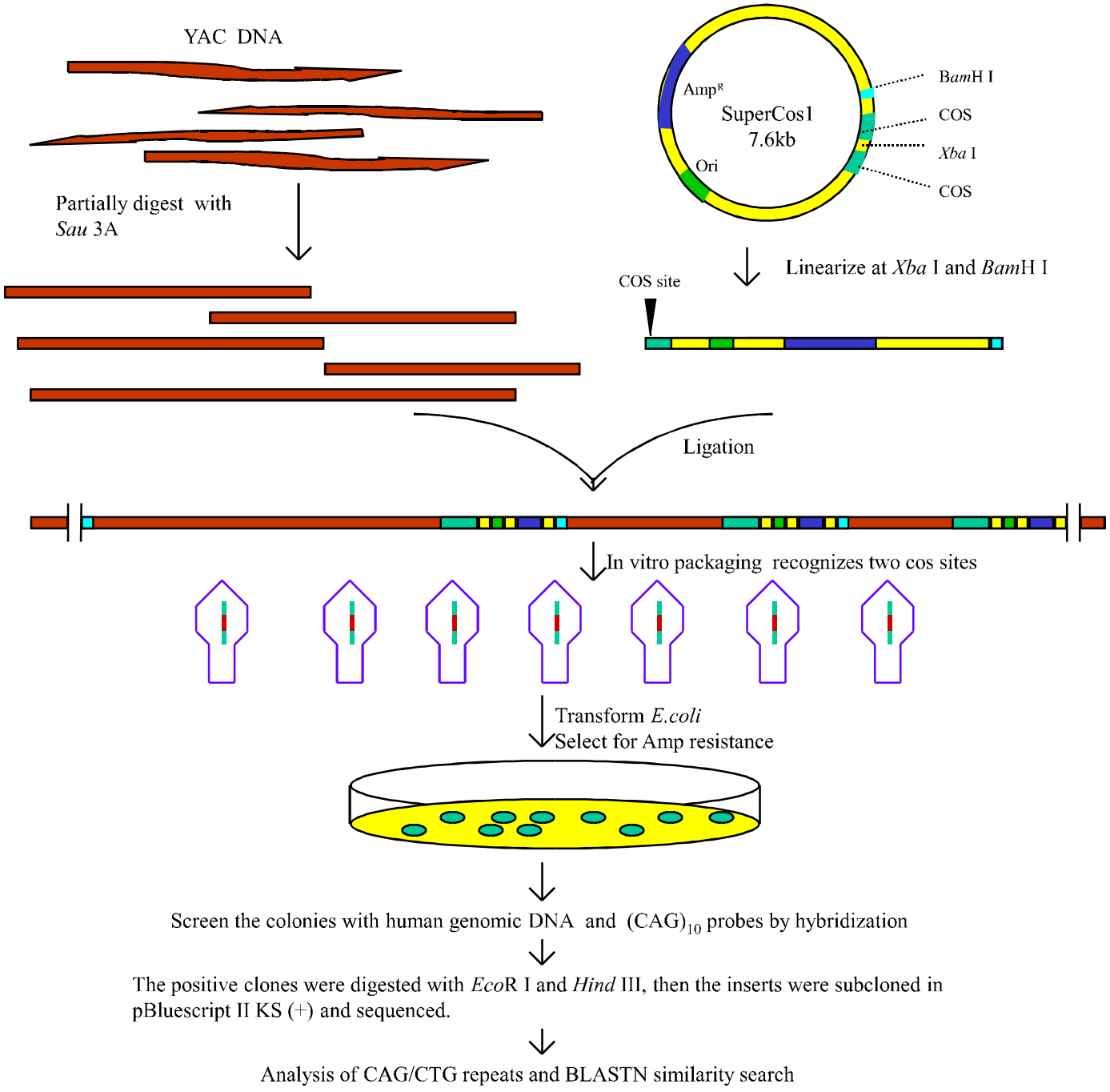



Figure 3. General procedures of isolating CAG/CTG triplet repeat from YAC DNA. SuperCos 1carries cos sites form theλ phage, plasmid origin of replication and ampicillin resistant gene ($Amp^R$ ). To clone YAC DNA into vector SuperCos 1, the vector is linearized with *Xba* I and *Bam*H I,and YAC DNA is 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 ligted 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 *Eco*R I and *Hind* III (Figure 4) and subcloned into plasmid pBluescript II KS (+). The subclones were sequenced by MWG CL4200 autosequencer.

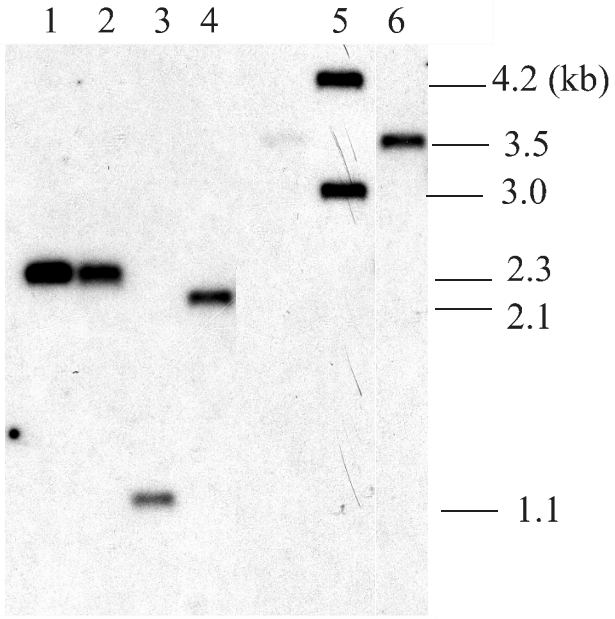



Figure 4. Southern hybridization analysis of cosmid clones. Lanes 1-6 contain clones D1, D4, D5, D6, E11, A8. These clones were selected from cosmid libraries by hybridization with human DNA and digested with *Eco*R I/*Hind* III. The digested DNA was electrophoresed on a 1% agarose gel, transfered to 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 $H = 1 - \Sigma p^2_i$, 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 $PIC = 1 - \Sigma p^2_i - \Sigma^{n-1}_{i=1} \Sigma^n_{j=i+1} 2p_i^2 p_j^2$ , 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.

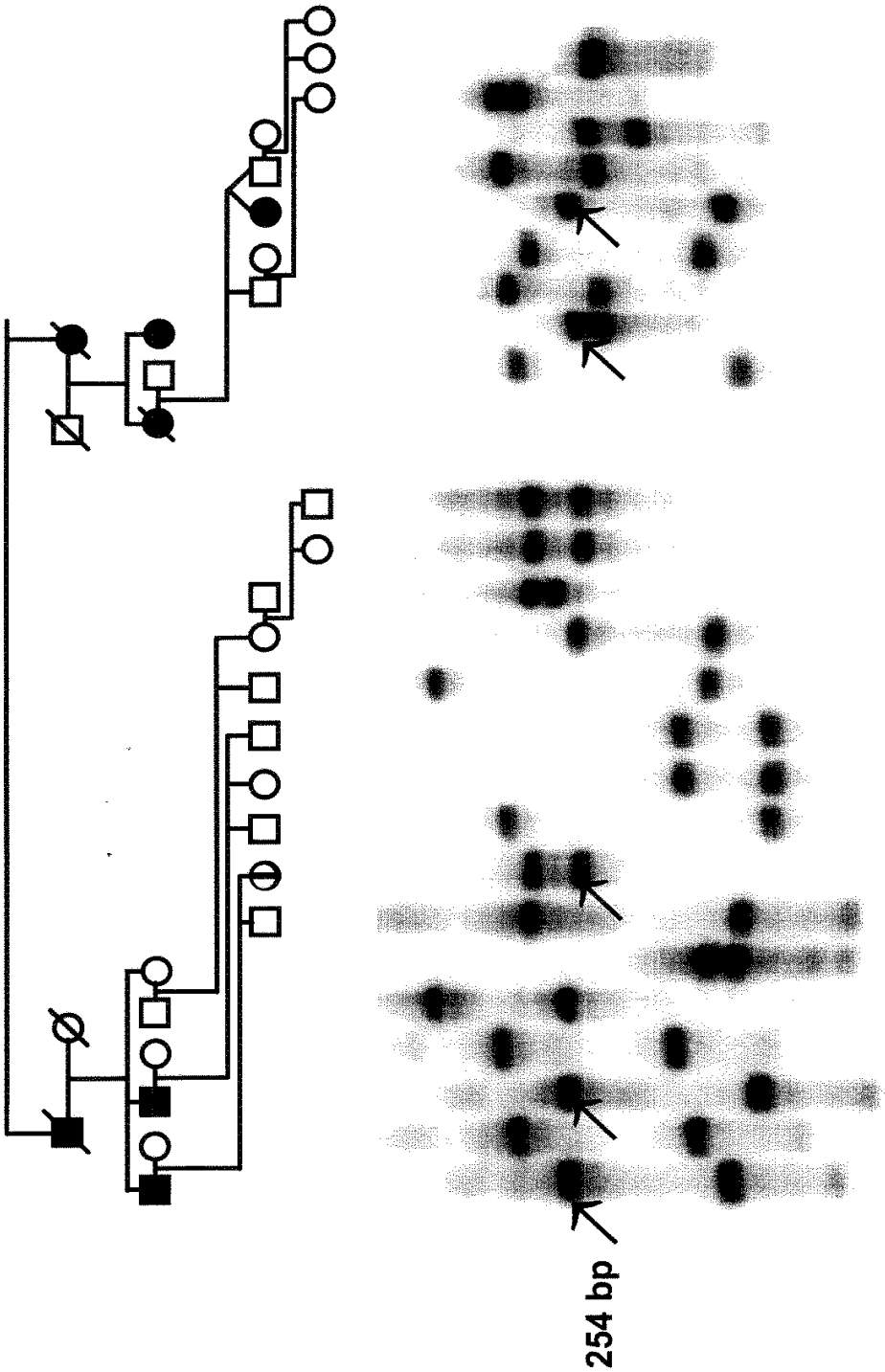



Figure 5. Segregation of alleles at *D2S3008* in a large SPG4 family. The 254 bp allele is found in a patient's daughter (IV:23) who has very mild symptoms at the age of 20 years.

### 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 cloneRP11-62F14 |
| | D4 | 2.3 | 6 | human BAC cloneRP11-62F14 |
| | D5 | 1.1 | 2 | human chromosome 2 working draft cluster Hs2_5351 |
| | D6 | 2.1 | 2 | human BAC cloneRP11-62F14 |

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) |
|---|---|---|---|---|
| *cag1c1p* | G67940 | cag1c1p.a | GCCCCAGAAACAGCCAAATAA | 385 |
| | | cag1c1p.b | CGCCAGCAACGCCAGTC | |
| *cag2c5p* | G67941 | cag2c5p.a | TGGGGAGTGGTTTAGTAGCAGAAG | 250 |
| | | cag2c5p.b | AGCGACATACATACAAGGGAAAAA | |
| *cag2c6p* | G67942 | cag2c6p.a | AACAGCCCCCTATCACATTCCAGA | 332 |
| | | cag2c6p.b | GGCACATCCCAGCATTTAGCAGTA | |
| *ctg6c4* | G67794 | ctg6c4.a | TCTACACATCCCATTTGCCTTTCT | 256 |
| | | ctg6c4.b | TAGCCTCTCCGAGCCTTGGTT | |
| *ctg6c11* | G67795 | ctg6c11.a | GGGCGCGAGACACAGGATTC | 422 |
| | | ctg6c11.b | TGACAAGTGGTGCGATTTATT | |
| *D2S3008* | AF095704 | D2S3008.a | CCAGCAGGCAGAGTGAGAATC | 224-292 |
| | | D2S3008.b | TGTTTAAAGCACAAGGGCAGT | |

### 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 [$\alpha^{32}$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).

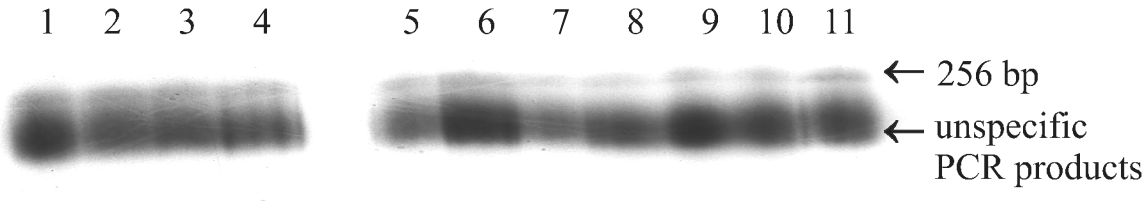



Figure 6. Allele expansion analysis of SPG4 patients at *ctg6c4*. Lanes 1-10 represent DNAs from the SPG4 family members III17, III9, IV16, III7, III4, III12, III14, III11, IV19, III1, respecitively. Lane11: YAC937D3. Lanes 1-5 contain DNAs from unaffecteds and Lanes 6-10 DNAs from patients. A 256 bp homozygous allele is detected in both controls and patients. There is no evidence of allele expansion.

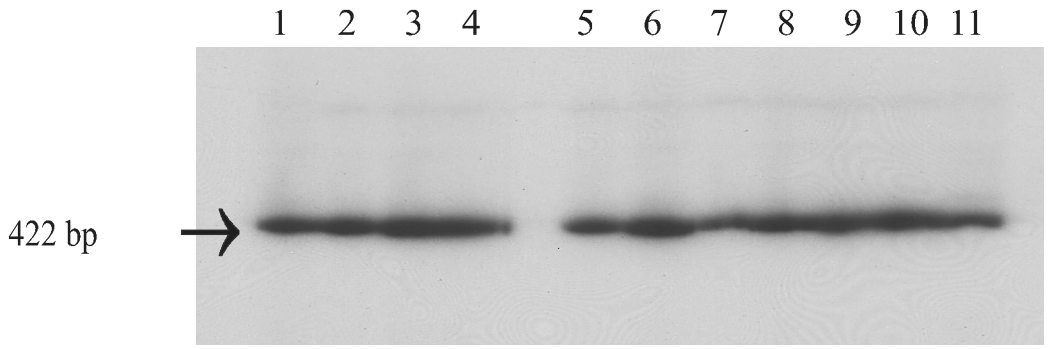



Figure 7. Allele expansion analysis of SPG4 patients at *ctg6c11*. Lanes 1-10 contain DNAs from SPG4 family members III17, III9, IV16, III7, III4, III12, III14, III11, IV19, III1, respectively. Lane11: YAC923E10. Lanes 1-5 contain DNAs from unaffecteds and lanes 6-10 contain DNAs patients. A 422bp homozygous 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 dehydrogenase[54], 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 librairies 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)

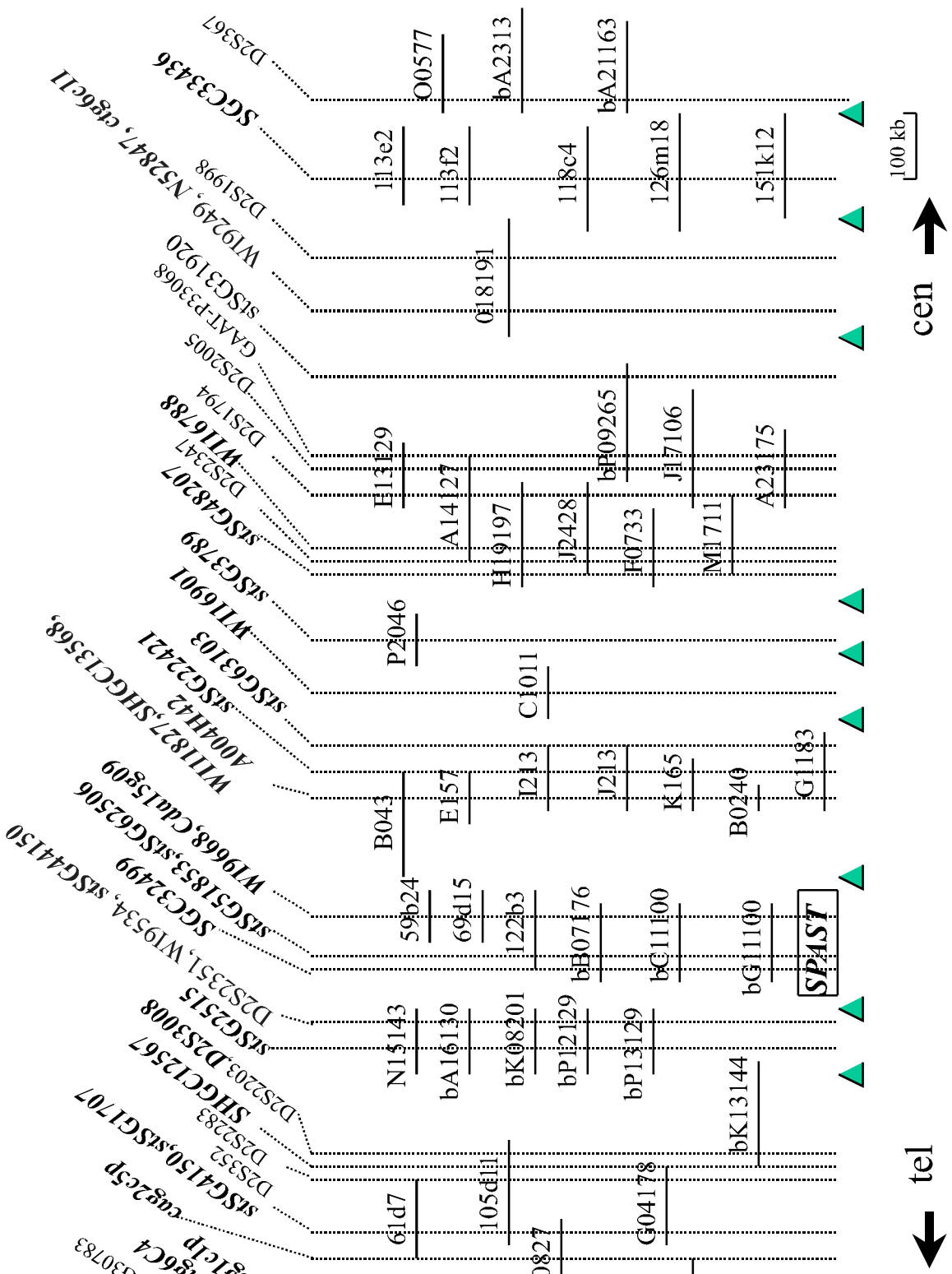



Figure 8. PAC/BAC contig of the SPG4 interval. The STSs are listed on the top and 33 PACs and 15 BACs were assigned to the contig. The BAC name was preceded by 'b'. The narrowed SPG4 interval was defined between D2S352 and D2S2347.The *SPAST* gene of approximately 90 kb is located inclusively in bG11100. The triangles on the bottom indicate the gaps.

### 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: | |
| RPCI1-59b24 | 90 | RPCIB753A2313 | 150 |
| RPCI1-61d7 | 20;90 | RPCIB753A16130 | 100 |
| RPCI1-69d15 | 90 | RPCIB753A21163 | 130 |
| RPCI1-105d11 | 150 | RPCIB753B07176 | 120 |
| RPCI1-113e2 | 120 | RPCIB753C11100 | 120 |
| RPCI1-113f2 | 110 | RPCIB753D16113 | 150 |
| RPCI1-118c4 | 150 | RPCIB753E16113 | 140 |
| RPCI1-122b3 | 120 | RPCIB753G0743 | 100 |
| RPCI1-126m18 | 170 | RPCIB753G11100 | 120 |
| RPCI1-151k12 | 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 $P^{33}$-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.

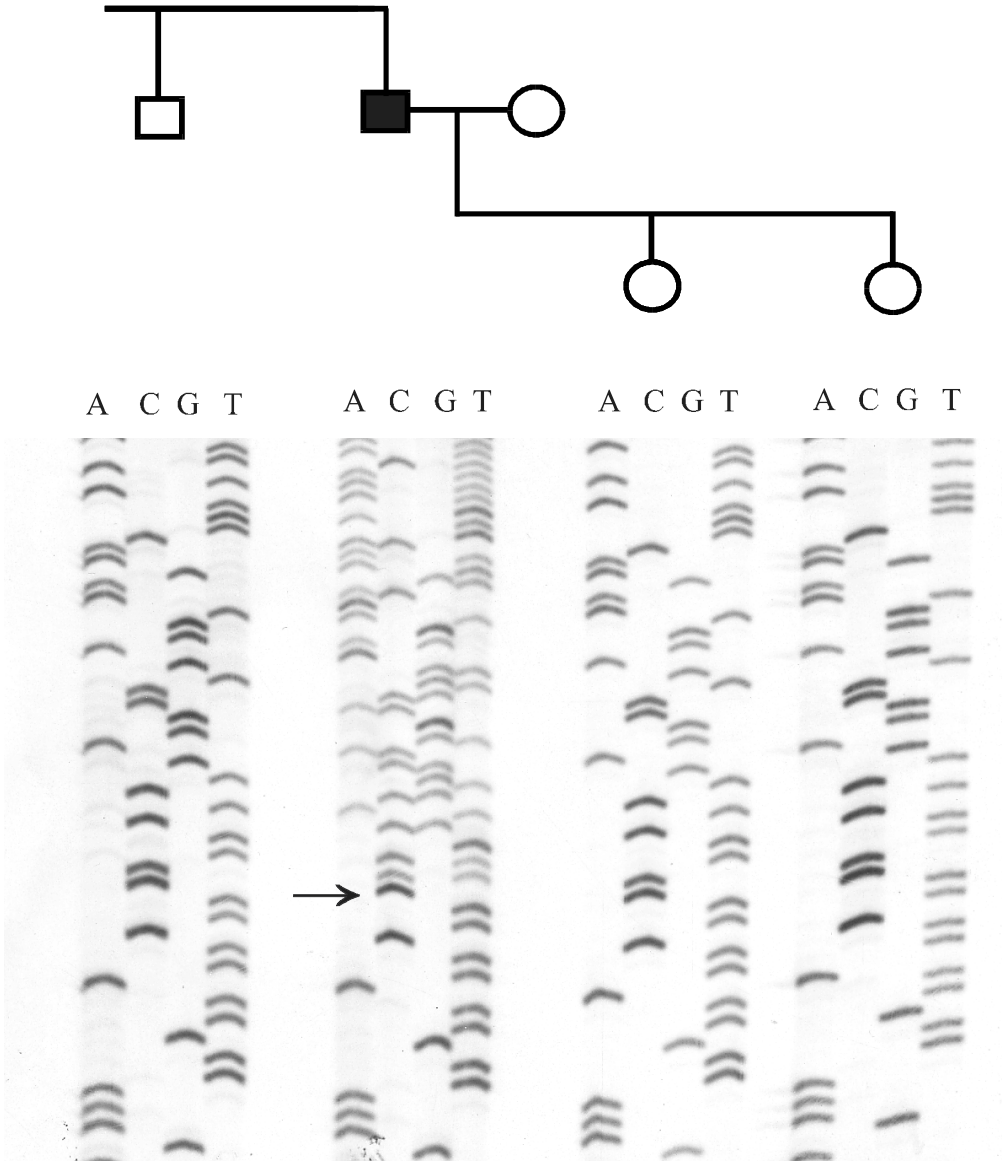



Figure 9. A novel mutation of *SPAST*, 1206-1209del. The SPG patient has a CCTT deletion in exon 7. His brother and two daughters do not carry this mutation. The arrow indicates the start of the deletion.

### 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.

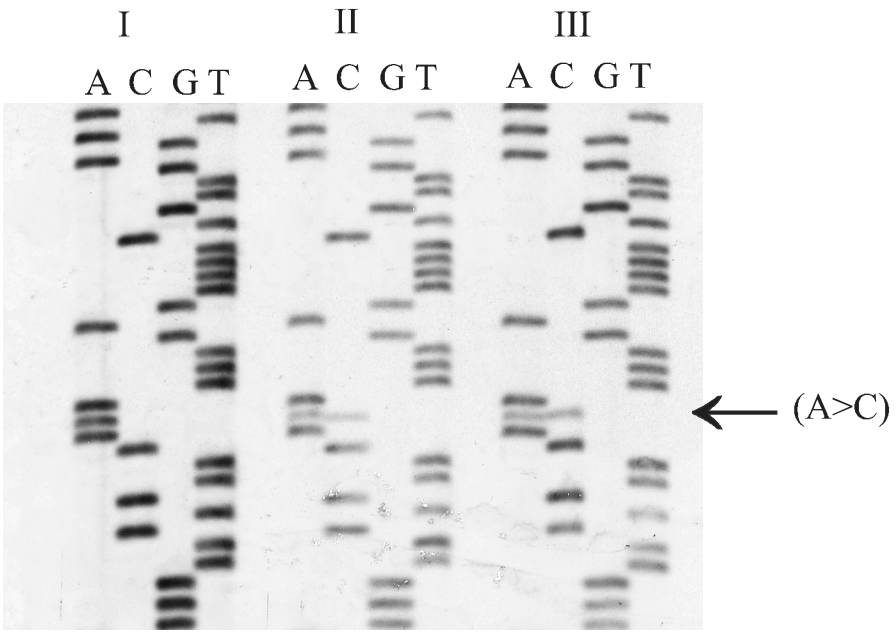



Figure 10. Sequence analysis of the polymorphism in intron 8 of *SPAST.* I: unaffected control. II: a spastic paraplegia patient. III: a control. The arrow indicates the base substitution 1298+17 (A>C) in II and III. This polymorphism was amplified by PCR and the PCR products were directly sequenced.

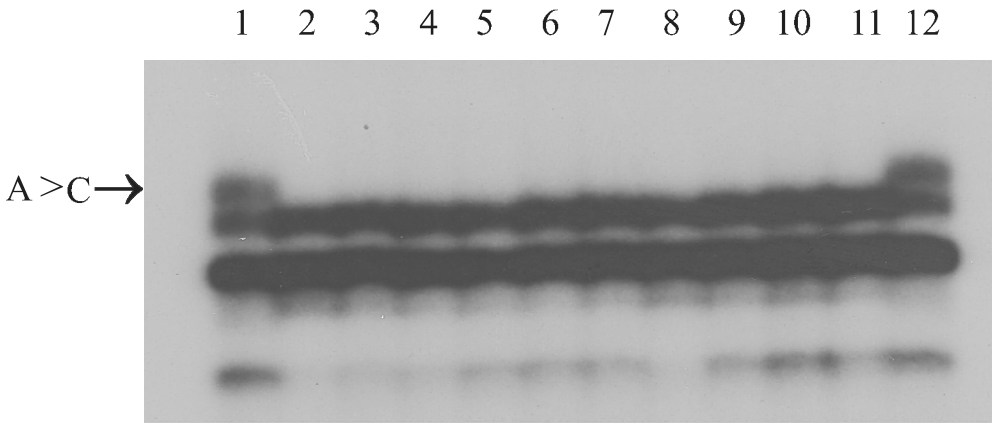



Figure 11. SSCP analysis of polymorphisms in intron 8 of *SPAST*
Lane 1: spastic paraplegia patient. Lane 2-12: unrelated controls. The arrow indicates the band shift due to A>C However, this polymorphism was not detected in an additional 100 unrelated controls and nine spastic paraplegia patients.The 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

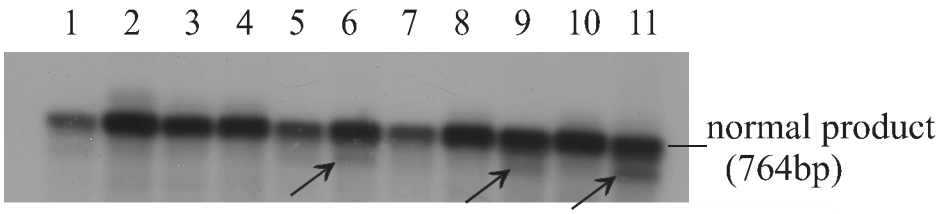



Figure 12. RT-PCR analysis of *SPAST*. Lanes 1-10: normal control samples. Lane 11: a patient from the large SPG4 family. The arrows indicate aberrant PCR products in this patient and two controls. The normal and the aberrant PCR products were sequenced. The results reveals that exon 8 is deleted in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

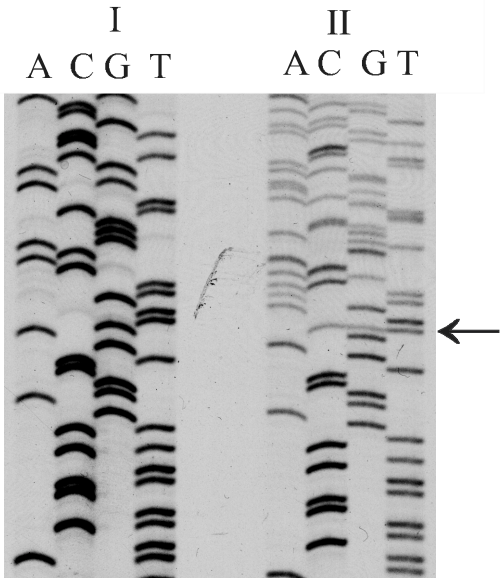



Figure 13. Sequence analysis of RT-PCR products amplified by SPA-Ba/m. I: normal PCR product, II: aberrant PCR product. Sequences reveal that exon 8 is truncated in II. The arrow indicates the first nucleotide of exon 8 in *SPAST* cDNA.

# 5 Discussion

## 5.1 Discovering *SPAST*

In this dissertation, linkage analysis in a large ADHSP family revealed that the disease locus is linked to *SPG4* in this family. In order to explore the *SPG4* locus, a YAC contig covering the *SPG4* candidate region was constructed as the first step of positional cloning of *SPG4*.

Positional cloning[57] allows the detection of disease genes without knowledge of the altered gene products. This method is thus superior to traditional cloning approaches which rely on known gene products and use this information (i.e. the amino acid sequence) to isolate the gene. Positional cloning was introduced to the identification of human disease genes in the early 1980s. Four major steps are involved in positional cloning: (1) chromosomal mapping of the disease gene by either linkage analysis or analysis of chromosomal rearrangements. (2) construction of a physical map of the region containing the disease gene (3) identification of the gene (4) determination of gene function.

Linkage was discovered by Morgan (1911) in drosophila[58] and was also found to be applicable to humans. The first demonstration of linkage in humans was done by Bell and Haldane[59], who showed that in some families, hemophilia and color blindness tended to be inherited together rather than segregating independently. Linkage is defined as the tendency for two genes close together on the same chromosome to be transmitted together through meiosis. The recombinant frequency between loci on the same chromosome is roughly proportional to their distance apart, and that allows the use of recombinant fraction to define chromosomal mapping units. The mathematical relationship between recombination fraction and genetic distance is defined by the mapping function. Assuming that recombination occurs at random along the bivalent and no influence on one another, the appropriate mapping

function is described as Haldane's function[60]: $\varpi = -1/_2\ln(1-2\theta)$, where $\varpi$ is the map distance and $\theta$ the recombinaion fraction. Nowadays, the human genome has been saturated with polymorphic genetic markers that it is often advisable to estimate linkage analysis; the LINKAGE package[61,62] has proven to be very useful. It is usually convenient to work not with the likelihood ratio but with its logarithm, the so-called lod (logarithm of odds) score $Z(\theta)= \log L(\theta)=\log[L(\theta)/L(1/2)]$; where $\theta$ is the recombination fraction, and by construction $Z(1/2)=0$. Two point lod scores are usually estimated at $\theta$ values 0, 0.01, 0.05, 0.1, 0.2 , 0.3,and 0.4. Generally, lod scores are calculated at fixed values of $\theta$ between 0 and 0.5, often in steps of 0.05. A lod score > 3 indicates a significant linkage ($p<0.05$) and a lod score < -2 indicates an exclusion of linkage. Another approach to the mapping of disease genes is cytogenetic studies on chromosomal abnormality (translocations, inversions, deletions, duplication, or other rearrangement) can assist positional cloning of human disease genes[63]. In particular, chromosomal abnormalities have been used to map most of the tumor suppresser genes or other cancer-related genes. For example, the cytogenetic detection of deletion on 13q13.1-q14.5 facilitated the identification of the *RB* gene in Retinoblastoma[64].

There are two types of maps: genetic maps and physical maps. Genetic mapping relies on the genetic methods including linkage analysis of pedigree in case of human or, cross-breeding tests in experimental animals. Genetic maps are defined by mapping functions of recombination fractions. Two loci which show 1% recombination are defined as being 1 cM apart on a genetic map. Physical maps refer to a real distance in units of nucleotides. Different types of physical maps such as radiation hybrid maps[65], restriction maps[66], maps of overlapping, bacteriophage P1[67], cosmid[68], BAC[69], PAC[70], YAC[71] inserts of human genome have been constructed. The clone contig maps are widely used as the sources for DNA sequencing in various Genome projects. The first generation physical contig map[72] of human

genome is so-called YAC contig. Although the large inserts of YAC are very useful in chromosome mapping, many YAC inserts are chimeric or have internal deletions. Therefore, second generation clone contig maps which rely on BACs and PACs are developed. The inserts of BACs and PACs (70-230 kb) are much smaller than YAC inserts (in units of Mb), however, BACs and PACs are without the disadvantages of deletions or insertions. Currently, BACs and PACs contigs are the primary templates for human genome sequencing. To establish a clone contig, for example, YAC, BAC, or PAC contigs, the most common approach is to overlap the cloned DNAs through the gene candidate region. Once a contig has been set up, genes are searched among those cloned DNAs by different methods. One common strategy is based on partial or complete sequence information. Additionally, evolutionarily conserved homology to a relevant gene in a model organism can be used for selecting a candidate gene. Since the first success of positional cloning was for the gene mutated in chronic granulomatous disease, many important disorders such as Duchenne muscular dystrophy, cystic fibrosis, Huntington disease, colorectal cancer, breast cancer, etc have been isolated by positional cloning[73]. Now progress in Human Genome Project is accelerating the identification of disease genes.

The *SPG4* locus was assigned to 2p by linkage analysis in 1994 (Hazan *et al.*[34]). Haplotype and linkage analysis in one Dutch and five French pedigrees delineated *SPG4* within a 4 cM interval flanked by loci *D2S400* and *D2S367*. Since then, many ADHSP families were described in whom the disease gene was assigned to 2p[2,35].

Nielsen *et al.*[51] searched for CAG trinucleotides expansions in *SPG4* due to the observation of posssible anticipation in ADHSP[34,74]. Anticipation refers to earlier disease onset and greater severity of the disorder in consecutive generations of a family. The molecular basis of

anticipation was discovered in trinucleotide repeat expansion disorders. The unstable trinucleotide repeat expansions cause many human hereditary diseases. Such unstable repeat expansions can be divided into two different classes.[73]

The first class contains very large expansions of repeats outside coding regions of the gene. A (CGG)n repeat expansion occurs in the 5'UTR of Fragile-X site A (FRAXA)[75]. A (CCG)n repeat expansion is found in the promoter of Fragile-X site E (FRAXE)[76]. In Friedreich ataxia (FA)[77], a (GAA)n expansion is found in intron 1. In myotonic dystrophy (DM)[78], a (CTG)n expansion is found in the 3'UTR.

The second class contains modest expansions of CAG repeats within coding regions. (CAG)n trinucleotide expansion is the most notable example. (CAG)n repeat expansions are associated with Huntington disease (HD)[79], spinal and bulbar muscular atrophy (SBMA; Kennedy disease)[80], spinocerebellar ataxia type 1 (SCA1)[81], spinocerebellar ataxia type 2 (SCA2)[82] , Machado-Joseph disease (MJD;SCA3)[83], spinocerebellar ataxia type 6 (SCA6)[84], spinocerebellar ataxia type 7 (SCA7)[85], and dentatorubral-pallidoluysian atrophy (DRPLA)[86]. All these CAG repeats are located in the coding region of respective genes and encode polyglutamine tracts within the gene product. When the length of polyglutamine tracts exceeds the threshold, the protein aggregates and forms inclusion bodies that appear to cause apoptosis of neuronal cell. The unstable CAG repeat numbers within coding regions are summarized in Table 15.

Table 15. Expansions of CAG repeats within coding regions of various diseases

| Disease | Wild-type repeat number | Expanded repeat number |
|---|---|---|
| HD | 6-35 | 36- >100 |
| SBMA | 9-35 | 38-62 |
| SCA1 | 6-38 | 39-83 |
| SCA2 | 14-31 | 32-77 |
| SCA3 | 12-39 | 62-86 |
| SCA6 | 4-17 | 21-30 |
| SCA7 | 7-35 | 37-200 |
| DRPLA | 3-35 | 49-88 |

Nielsen *et al.*[51] reported that CAG repeat expansion had been detected in 20 individuals affected with SPG4. They used the repeat expansion detection (RED) method, which does not indicate the location of the CAG repeats. This method, which was based on oligonucleotide amplification and ligation, was developed by Shalling *et al.*[52] in order to assay the unstable repeat expansions in human genome. The strategy for RED is to use a long repeat in genomic DNA as a template for the annealing of multiple complementary oligonucleotide molecules. In this method, oligomers of a given trinucleotide repeat are ligated together by thermostable ligase when they anneal in adjacent positions, forming a mixed population of longer single stranded molecules. Detection of the single stranded molecules is facilitated by a cycling procedure and achieved by separation on a denaturing polyacrylamide gel and subsequent electrotransfer and hybridization.

One year later than Nielsen *et al* detected the CAG repeats in *SPG4*, Benson *et al.*[87] found that in most cases of their ADHSP families the repeat expansion detected by RED is due to non-pathogenic CAG expansions in *SEF2-1* locus[88] on 18q21.1, which encodes a basic helix-loop-

helix DNA binding protein involved in transcriptional regulation or *ERDA1* (expanded repeat domain, CAG/CTG 1) locus[89] on 17q21.3. The CAG expansions were also detected in a subset of affected and at-risk at loci other than *SEF2-1* and *ERDA1* loci.

Due to the possible occurrence of anticipation in ADHSP, the CAG/CTG repeats fragments within the SPG4 critical region were screened in this dissertation. A total of 6 sets of CAG/CTG repeat fragments were isolated from cosmid libraries. Of the CAG/CTG repeat fragments, clone D1 contains a disrupted single CAG unit; clone A8 contains $(CAG)_3$, clones D5 and D6 contain $(CAG)_2$; and clones D4 and E11 contain $(CTG)_6$. Because the repeat number of trinucleotides involved in diseases is usually more than 3, two primer pairs which flanked the two sequences with 6 repeat units were generated in order to test for trinucleotide repeat expansions. The results do not support that CAG repeat expansion is involved in *SPG4* and were later confirmed by Del-Favero *et al*[90]. and Hazan *et al*.[37]. Del-Favero *et al*. searched for CAG/CTG fragments on the basis of homologous recombination in *SPG4* and their results suggested no CAG expansion in *SPG4*. Similar to our studies (Lau *et al*.[50] 1998), Hazan *et al.* later searched for CAG repeat fragments by building a BAC contig of *SPG4* critical region. The authors also did not find any evidence of CAG repeat expansions as the cause of SPG4.

As described previously, BACs/PACs clone contig is prerequisite for a sequence-ready map. While this contig was being constructed as initiated by us, Hazan *et al*.[36] discovered the *SPG4* gene using the same strategy of positional cloning. The authors sequenced 2p21-24 as participants in Human Genome Project. They first constructed a BAC-based sequence ready map of the narrowed *SPG4* interval which was estimated to be approximately 1.5Mb. The contig comprised 37 BAC clones and 32 STSs including 14 ESTs. The BAC clones were isolated from two human genomic libraries CITB_978_SKB and RPCI-11 using various

STSs. Sequencing of 12 of the 37 BAC-inserts and subsequent comparison to nucleic acid and protein databases, revealed a total of 14 putative transcription units in this interval. Among the 14 genes detected by sequence analysis, 6 had been previously identified as ESTs and 3 were known genes; xanthine dehydrogenase[54], steroid 5α-reductase[91] and TGF-β1 binding protein.[55] The remaining 5 genes were only detected DNA sequence applying gene fishing programs to the available. One of these five genes showed homology in the 3' coding region to genes that encode proteins of the AAA family. This gene was called *SPAST* after mutations were detected in SPG4 patients.

### 5.2 Mutation Detection of *SPAST*

Many mutations have been detected in *SPAST* in ADHSP patients[92,93,94,95,96,97]. These mutations are listed in Table 16 and include 72% base substitutions, 24% deletions, and 4% insertions. Of the base substitutions, 60% occurred in the coding region of *SPAST* while the other 40% were found at splicing sites. All insertions occurred in coding regions. Most of the deletions (92%) occurred in the coding region as well, yet 8% were intronic. In total, 70% of the mutations were found in exons and 30% in introns. These mutations were scattered throughout the *SPAST* gene. However, exon 5 was the most frequently mutated exon (7/42).

Table 16. Mutations of *SPAST* found in SPG4 patients

| Mutation class | Location | Mutation | Amino acid change | Consequence |
|---|---|---|---|---|
| Base substitution | Exon 1 | 256C>T | S44L | Missense[94] |
| | Exon 1 | 334G>T | E112Stop | Missense[96] |
| | Exon3 | 702C>T | Q193Stop | Nonsense[92] |
| | Exon 5 | 859C>G | S245Stop | Nonsense[94] |
| | Exon5 | 873A>T | K229Stop | Nonsense[92] |
| | Exon5 | 907C>A | S261Stop | Nonsense[92] |
| | Exon5 | 932C>G | Y269Stop | Nonsense[92] |
| | Exon7 | 1210C>G | S362C | Missense[92] |
| | Exon8 | 1233G>A | G370R | Missense[92] |
| | Exon8 | 1267T>G | F381C | Missense[92] |
| | Exon8 | 1283T>G | N386K | Missense[92] |
| | Exon8 | 1288A>G | K388R | Missense[92] |
| | Exon9 | 1195C>T | R399Stop | Nonsense[96] |
| | Exon9 | 1211C>T | S404F | Missense[96] |
| | Exon10 | 1395A>G | R424G | Missense[94] |
| | Exon10 | 1401C>G | L426V | Missense[92] |
| | Exon10 | 1416C>T | R431Stop | Nonsense[92] |
| | Exon11 | 1447A>G | D441G | Missense[93] |
| | Exon11 | 1468G>A | C448Y | Missense[92] |
| | Exon11 | 1504G>T | R460L | Missense[92] |
| | Exon13 | 1620C>T | R499C | Missense[92] |
| | Exon14 | 1583G>A | G527D | Missense[96] |
| | Exon15 | 1788G>A | D555N | Missense[92] |
| | Exon15 | 1792C>T | A556V | Missense[92] |
| | Exon15 | 1809C>T | R562Stop | Nonsense[92] |
| | Exon17 | 1875G>C | D584H | Missense[94] |
| | Intron4 | 808-2a>g | --- | Missplicing[92] |
| | Intron6 | 1129+2t>g | --- | Missplicing[92] |
| | Intron7 | 1223+1g>t | --- | Missplicing[92] |
| | Intron8 | 1298+1g>a | --- | Missplicing[94] |
| | Intron8 | 1299+1g>a | --- | Missplicing[92] |
| | Intron11 | 1538+3a>c | --- | Missplicing[94] |
| | Intron11 | 1538+5g>a | PTC+6 aa | Exon11 skipping +frameshift[92] |
| | Intron12 | 1618+2t>a | --- | Missplicing[94] |
| | Intron13 | 1661+1g>t | --- | Missplicing[94] |
| | Intron13 | 1661+2t>c | --- | Missplicing[92] |
| | Intron13 | 1662-2a>t | --- | Missplicing[92] |
| | Intron15 | 1812+1g>a | --- | Missplicing[92] |
| | Intron15 | 1812+2t>g | --- | Missplicing[94] |
| | Intron15 | 1813-2a>g | --- | Missplicing[92] |
| | Intron16 | 1853+1g>t | Exon16 skipping | Missplicing[93,95] |
| | Intron16 | 1853+1g>a | --- | Missplicing[92] |
| | Intron16 | 1853+2t>c | --- | Missplicing[94] |

| | | | | |
|---|---|---|---|---|
| Deletion | Exon1 | 411delG | PTC+64aa | Frameshift[94] |
| | Exon5 | 852-862del | PTC+18aa | Frameshift[92] |
| | Exon5 | 906delT | PTC+17aa | Frameshift[92] |
| | Exon7 | 1206-1209del | PTC+1aa | Frameshift[97] |
| | Exon9 | 1299delG | PTC+3aa | Frameshift[92] |
| | Exon9 | 1340-1344del | PTC+35aa | Frameshift[93] |
| | Exon10 | 1406delT | PTC+10aa | Frameshift[94] |
| | Exon11 | 1520delT | PTC+1aa | Frameshift[92] |
| | Exon12 | 1574-1575del | PTC+2aa | Frameshift[92] |
| | Exon12 | 1617-1678del | Exon11-12 and Exon11-13 skipping | Missplicing[93] |
| | Exon13 | 1634-1655del | PTC+18aa | Frameshift[92] |
| | Exon14 | 1685-1688del | PTC+7aa | Frameshift[92] |
| | Intron11 | 1538+3-1538+6del | --- | Missplicing[94] |
| Insertion | Exon2 | 578-579insA | PTC+2aa | Frameshift[92] |
| | Exon4 | 709-710insA | PTC+21aa | Frameshift[96] |
| | Exon5 | 882-883insA | PTC+12aa | Frameshift[92] |
| | Exon14 | 1684-1685insTT | PTC+9aa | Frameshift[92] |

Note: Nucleotide numbers refer to the *SPAST* cDNA sequence. Upper case letters represent bases in exons, lower case letter in introns. Amino acid (aa) numbers refer to the spastin petide sequence. PTC+n,premature termination codon at n amino acids downstream from the location of mutation.

There are several methods for mutation screenings based on various application as follows.

(1) Heteroduplex/SSCP[98]analysis has been the most frequently used method for mutation detection. Many modified methods are based on the formation of heteroduplexes which have abnormal mobility on nondenaturing polyacrylamide gels and abnormal denaturing profiles . SSCP gel electrophoresis is a simple method for distinguishing fragments shorter than 300 bp, which can have insertion, deletion and most but not all single-base substitutions. Denaturing gradient gel electrophoresis[99] (DGGE) and denaturing high performance liquid chromatography (dHPLC) can separate heteroduplexes based on the mobility of the denatured fragments. DGGE requires special primers with a 5' poly(GC) extension (a GC clamp). If optimized, these methods have very high sensitivity. The use of two dimensional DGGE gels can also increase the sensitivity for mutation detection.

(2) Chemical cleavage of mismatch (CCM) is a sensitive method for mutation detection which is used to analyze large fragments (over 1 kb) and the location of mismatch can be detected by the size of the fragments. The disadvantage is however the use of very toxic chemicals , such as osmium tetraoxide. An alternative is to use T4 phage resolvase or endonuclease VII for cutting heteroduplexes. This method however is not easily optimized.[100]

(3) The protein truncation test (PTT) is a specific test for frameshifts, splice site or nonsense mutations that truncate a protein product[101]. For the diseases, such as Duchenne muscular dystrophy, adenomatous polyposs coli or BRCA1-related breast cancer in which missense mutation is infrequent, PTT is useful because it ignores silent or missense mutation, and also reveals the approximate location of the mutation.

(4) High-density oligonucleotide arrays which are based on oligonucleotide hybridization can be used for a high throughput mutation scanning [102]. This technique is applied to the mutation scanning of the known genes. Two basic designs are as follows: Hybridization chips contain oligonucleotides matching all wild-type and single nucleotide substitution sequences in a gene. Test DNA is PCR-amplified, fluorescently labeled and hybridized to the array. Generally, homozygous base substitutions are easily detected, but heterozygous substitution detection can be difficult and detection of insertion mutations is almost impossible. Minisequencing chips use arrayed oligonucleotide primers with free 3'-OH groups. Unlabeled PCR-amplified test DNA is hybridized to the array, and DNA polymerase plus four differently labeled dideoxynucleotides are added. The test DNA acts as template for addition of a single labeled dideoxynucleotide to each array primer. The additions will occur only if the 3' end of the primer exactly matches the template. The array

can be made with primers specific not only for the wild-type sequence but also for all possible mutations.

(5) RT-PCR analysis: In addition to testing genomic DNA, mutation detection by RT-PCR allows study of aberrant splicing and also base changes in exons. However, not all genes are expressed in readily accessible tissues and RNA is much less stable; in particular mRNA. Transcripts which exist in cells that are not expected to express such mRNA are so-called ectopic or illegitimate transcprits[103]. Caution should be taken in interpretation of aberrant splicing product of ectopic transcripts , because the structure of ectopic transcripts are not always faithful. In addition, many mutations result in unstable mRNA so that only the wild-type transcript of a heterozygous patient can be detected by RT-PCR. An aberrant RT-PCR product lacking exon 8 in *SPAST* from a patient in that large SPG4-linked family was detected. In our lab-control samples, the same splicing artifact was observed after blood samples were placed for 10 hr at room temperature.

As shown in this dissertation, a single patient of a small pedigree was detected a mutation in *SPAST* but the mutation of the large SPG4-linked family was not found. The same phenomenon was also reported in other *SPAST* studies[104]. The reason may be that the mutations in these SPG4-linked family are gross deletions or another atypical intronic mutation. Sequencing is the direct and also still the most reliable method when both strands are sequenced to detect mutations. However, it is expensive and time-consuming to scan many exons of a gene and exclude artifacts.

Although sequence changes can be detected by many techniques as described above, not every sequence variant of an affected person is necessarily pathogenic. There are some criteria for deciding whether a DNA sequence change is pathogenic[73]. These are described as follows:

(1)Deletions of an entire gene, as well as nonsense and frameshift mutations probably abolish the gene function.

(2) Mutations in the conserved GT..AG nucleotides flanking most introns affect splicing, and usually alter the function of the gene.

(3)A missense mutation is likely to be pathogenic if it alters the domain of protein proved to be functionally important.

(4) Change of amino acids in the evolutionarily conserved domain across species or between members of a gene family is possible to be pathogenic.

(5) Amino acid substitutions where a polar amino acid is replaced by a nonpolar amino acid; or where an acidic amino acid is replaced by a basic amino acid are likely to alter gene function

(6) A sequence change in a disease gene of a *de novo* affected patient that is not in the unaffected parents is likely to be pathogenic.

### 5.3 Hypothesis of Spastin Function

On the basis of sequence homology, several hypotheses of spastin function have been proposed[36]. The high homology with the 26S proteasome subunit as well as the presence of two leucine-zipper domains and a coiled-coil dimerization suggest that spastin participates in protein complexes. Proteasomes are large multi-catalytic protease complexes involved in major proteolytic pathways in both the cytoplasm and the nucleus[105]. They function in cellular processes such as cell differentiation, adaptation to environmental stress, cell cycle control

and transcription regulation. There is also striking homology between spastin and Sap1p, a yeast nuclear AAA protein thought to interact with the chromatin protein Sin1p[106]. The change in yeast mating type occurs via a cassette mechanism in which one of the silent copies replaces the active gene at the MAT (mating type) locus changing the mating type. This process is controlled by an endonuclease which is the product of the HO (homothallism) gene and which makes a double-stranded cut at the MAT locus initiating switching. Sin 1p is one of SIN loci which encode transcription factors that repress HO gene expression. Nuclear AAA proteins exhibiting homology with 26S proteasome subunits have been suggested to play an indirect role in gene expression by mechanisms that include proteolytic activation or degradation of transcription factors.[107,108]

An other proposed function of spastin is that of a chaperone in corticospinal tract preservation. This idea comes from the comparison with the *Paraplegin*[25] gene at chromosome 16q24.3. Mutations in paraplegin cause autosomal recessive SPG7, resulting from typical mitochondrial oxidative phosphorylation (OXPHOS) impairments and result in both pure and complicated forms of AR-HSP. The *Paraplegin* gene encodes a protein, paraplegin, that is highly homologous to the yeast mitochondrial ATPases, AFG3, RCA1, and YME1. These ATPases have both proteolytic and chaperone-like activities at the inner mitochondrial membrane. Another autosomal dominant neurondegenerative disease, early onset torsion dystonia, is caused by a mutation in the *DYT1* gene which encodes torsinA. TorsinA is also a member of the AAA family and functions as a chaperone protein in the transport of dopamine-containing membranous vesicles.[109] Thus the functions of paraplegin and torsinA indicate that spastin may play a chaperone-like role.

The *SPAST* gene is expressed ubiquitously in brain, heart, liver, kidney, pancreas, skeletal muscle, and lymphocytes[36]. However, the primary pathology is the degeneration of the neurons. This indicates that mutations may have different effects in different cell types in which the gene is expressed and a tissue-specific vulnerability in SPG4 which results in the degeration of neurons. Selective vulnerability is a common phenomenon in many human neurondegenerative diseases[110]. For example, Huntigton's disease results in cell death in the caudate and causes in abnormal movement. Parkinson's disease destroys cells in the substantia nigra and causes rigidity and tremor of movement. Amyotrophic lateral sclerosis destroys the lower motor and pyramidal neurons and results in weakness and spasticity. However, the selective vulnerability is not absolutely sustained during the process of disease. For example, patients with Huntington's disease, Parkinson's disease, and amyotrophic lateral sclerosis can develop dementia, implying cortical pathology involved in late stage of the disease. Patients with Alzheimer's disease frequently develop parkinsomism. The selective vulnerability might be due to the variability of the gene expression or interaction with tissue-specific polypeptides, however, the mechanism is still unclear.

Heinzlef *et al.*[111] mapped a complicated familial spastic paraplegia to locus SPG4 on chromosome 2p, although SPG4 locus is frequently associated with pure HSP. Recently, another three reports of the pedgrees from Ireland, France, and England have described the association of SPG4-linked autosomal dominant pure HSP with dementia , epilepsy, or both[112,113,114], although SPG4 occurs frequently in a pure form. In addition, a study on five Danish SPG4-linked families revealed bowel, bladder and sexual dysfunction suggesting that autonomic system was involved in SPG4-linked ADHSP[115]. Pure HSP is thought as a disease of the CNS long tracts. The reason why an alteration in Spastin leads to specific axonal degeneration and more complicated pathology remains to be explored[116].

### 5.4 Hypothesis of mutation effect in *SPAST*

Hazan *et al.*[36] and Bürger *et al.*[93] suggested that SPG4 is caused by a loss of function due to mutations. This implied that a threshold dosage of spastin was critical for neuronal preservation and SPG4 exerted the dominant effect by haploinsufficiency. The phenotype of point mutations in a gene is likely the same pathological change as deletions which imply the loss of function. However, many different changes in a gene can lead to loss of function, for examples, the structural abolishment in a gene (including deletion, insertion, translocation, inversion), disturbance in the promoter region, destablization of the mRNA, and errors in splicing and post-transcriptional processing. Haploinsufficiency refers to the case where a 50% reduction in the level of gene function causes an abnormal phenotype. This posssibly explains why SPG4 has incomplete penetrance and wide variation of phenotype between affected individuals, even within families. Recently, a nonpenetrance at the SPG4 locus was reported where the carrier was 26 years older than the maximal age of onset for her family[114]. This suggests that gene modifiers are involved in the process of the neuronal degeneration resulting from mutation of *SPAST*.

Human genes which show haploinsufficiency effect that means two alleles are needed for a normal phenotype fall into the three categories: [117]

(1) Transcriptional regulators: *PAX3* (Waardenburg syndrome), *PAX6* (aniridia), *GLI3* (Greig cephalopolysyndactyly), *ZNF141* (Wolf-Hirschhorn syndrome), *TUPLE-1* (CATCH22 syndromes), and *WT1* (Wilms's tumor).

(2) Structural molecules: elastin (supravalvular aortic stenosis), ankyrin (spherocytosis type II), and type 1 collagen (osteogenesis imperfecta). The abnormal heterozygous phenotypes result in the imbalance with a matched component protein.

(3) Receptors and signal transduction molecules: *RET* (Hirschprung disease) and *LIS-1* (Miller-Dieker lissencephaly) result in the false stoichiometry of intermolecular complexes in the abnormal heterozygote.

Except haploinsufficiency described above, dominant mutations can exert their effects in various ways. Some mechanisms of dominant mutations which present gain of function are discussed below[118].

(1) Increased gene dosage: Trisomy in humans is usually associated with phenotypic abnormality which may be relevant when the increase in dosage at the mRNA and protein level exceeds the expected factor of 1.5. Studies of Down syndrome[119] with translocations show the critical region is in chromosome 21q22.2. At least two candidate genes for mental retardation have been identified from this region: *DYRK*, a gene whose drosophila and mouse homologues (*minibrain*) produces dosage sensitive learning defects, and *DSCAM*, a brain-specific cell adhesion molecule[120]. Gene amplification in somatic cells to much higher copy numbers frequently occurs in cancer. For example, a dominant phenotype is caused by the amplification of the *MDM2* gene in sarcomas. MDM2 protein binds and inactivates the tumor suppresser gene *P53*, leading to escape from normal p53 regulated cellular growth control.[121]

(2) Ectopic or temporally altered mRNA expression: The exquisite controls of mRNA expression that dictate the normal cellular distribution, temporal restriction, and absolute levels of mRNA are disturbed in some disorders. The disease phenotype may reflect a combination of alterations in the temporal specificity, tissue distribution, and absolute level of mRNA expression. In this case, the primary abnormality usually lies at the level of

transcription, but sometimes mRNA processing may be affected. chromosomal translocations resulting from errors in recombinase-mediated gene rearrangement in lymphocytes activate expression of transcription factors like MAC, causing B and T cell neoplasms[122,123]. Promoter mutations in the *Caenorhabditis* sex determining gene, *her-1,* increase expression levels and result in partial transformation of XX worms into phenotypic males[124]. Increased ectopic expression of a chimerical mRNA encoding a normal protein causes the lethal yellow mutant at the mouse *agouti* locus[125]. A point mutation of the γ globin promoter, which alters binding of the erythroid transcriptional factor GATA-1, results in hereditary persistence of fetal haemoglobin (HPFH). This mutation blocks the normal switch from expression of γ to δ and β globin that occurs around the time of birth.[126]

(3) Increased or constitutive protein activity: At the protein level, increased activity may be caused by increased half life or by loss of normal inhibitory regulation. PEST sequences which are rich in proline, glutamic acid, serine, and threonine can act as recognition signals for proteolytic degradation and loss of these sequences by C-terminal truncation stabilizes the protein[127]. The *glp-1* gene of *C elegans* is required for induction of germline proliferation and embryogenesis. The *glp-1* point mutation is particularly instructive, as it causes both semidominant (multivulva) and recessive (sterility/embryonic lethality) phenotypes which result from the destabilization of the mutant mRNA[128].

(4) Dominant negative mutations: In the heterozygous state, these mutants antagonise the activity of the remaining wild type allele. The mutant polypeptide loses not only its own function but also interferes with the product of the normal allele in a heterozygote. Dominant negative effects are very important in neoplasia. The tumor suppresser p53

oligomerizes i*n vitro* and can adopt two conformations, one active and the other inactive; wild type protein is normally in the active state. Cotranslation with certain missense mutants results in mixed oligomers that adopt the inactive conformation. Although p53 is conventionally considered as a recessive tumor suppresser gene, some mutants can deregulate p53 function in a dominant negative pattern.[129,130]

In addition, the disruptive interaction in toxic protein[131], altered substrate specificity in mutant protein[132], recessive antioncogene in retinoblastoma[133], and parental genomic imprinting[134] provide a variety of examples of genetic dominance.

**5.5 Prospective View**:

The isolation of the *SPAST* gene and the establishment of mutation spectrum have led the molecular diagnosis of SPG4 available. However, the functional study of *SPAST*, such as with transgenic models, is expected to illustrate the genetic and molecular pathology of SPG4. There are several fundamental questions that need to be explored.

(1) How is the mutant mRNA expressed in different cell types?

(2) How does the tissue specific vulnerability occur in SPG4?

(3) Is there a *SPAST* modifier gene involved in the gene expression?

(4) Spastin should be studied as well as to establish rapid detection methods like ELISA or RIA .

In this dissertation, the mutation of this large SPG4-linked family could not be detected because of methodological limitation. However, a gross deletion or other intronic mutations should be considered. In future, new techniques such as nucleic acid chips and microarrays are expected to make a more rapid screening and mutation detection possible. Finally, it would be interesting to trap some genes from the BAC and PAC contig in which *SPAST* is located.

# 6 Conclusion

Pure hereditary spastic paraplegias (pHSP) are relatively mild neurodegenerative disorders of the spinal cord. The predominant clinical sign is a slowly progressive gait anomaly due to spasticity of the legs. Age of onset is usually during adulthood, but varies greatly both within and between families. Autosomal dominant hereditary spastic paraplegia (ADHSP) are the major forms of pHSP and are genetically heterogeneous. The disease loci of pure ADHSP have been assigned to chromosomes 2 (*SPG4* in 2p 21-24, *SPG13* in 2q24-34), 8 (*SPG8* in 8q23-24), 12 (*SPG10* in 12p13), 14 (*SPG3* in 14q11.2-q24.3), 15 (*SPG6* in 15q11.1), and 19 (*SPG12* in 19q13). To date, the SPG4 gene *SPAST* has been identified. The purpose of this dissertation is to (1) determine the disease locus of a large family with pure ADHSP, (2) construct the physical map of the candidate region as the first step to identify the disease gene, (3) investigate whether CAG repeat expansions were involved in the mechanism of disease, and (4) screen the mutation of patients with the *SPAST* gene. The results are summarized below.

(1) The large family with pure ADHSP was analyzed by linkage test and a significant linkage of the disease loci to *SPG4* with lod scores = 3.63 and 3.39 at loci *D2S367* and *D2S352,* respectively was found.

(2) A yeast artificial chromosomes (YAC) contig which comprised 10 additional YACs and 13 expressed sequence tags (ESTs) on the basis of the previously published YAC contig of SPG4 critical region was constructed.

(3) A total of six CAG/CTG containing fragments were obtained from cosmid libraries within the SPG4 candidate region between loci *D2S367* and *D2S400*. Of the six CAG containing sequences, a highly polymorphic STRP (*D2S3008*) was serendipitously identified. It is composed of (GAAA)n(AAGG)n. A total of 28 alleles were identified in 50 Caucasian DNA samples. The marker has a heterozygosity of 0.81 and a PIC value of 0.78. Loci *D2S352* and *D2S2283* flank *D2S3008* distally and *D2S2351* flank it proximally. A 254 bp allele at locus *D2S3008* was segregated in the affected individuals of the SPG4-linked family.

(4) Based on the sequences of the remaining five CAG positive clones, five new STSs were generated and their dbSTS_ID are 98663, 98664, 98809, 98810, 98811, respectively. Two of the six fragments showed the maximum of repeat units was $(CTG)_6$. However, no CAG/CTG repeat expanded alleles were detected in the affected individuals of the SPG4-linked family.

(5) A contig which comprised 33 bacteriophage P1-derived artificial chromosomes (PACs) and 15 bacterial artificial chromosomes (BACs) was established in the SPG4 candidate region. The sizes of PACs and BACs were determined by pulse field gel electrophoresis and the average size of PACs and BACs was 120 kb and 127 kb, respectively. In addition, a total of 41 STSs comprising 26 ESTs were integrated into this contig. Of the 26 ESTs, SHGC12567 was found to be part of human xanthine dehydrogenase, N52847 was identical with human latent transforming growth factor beta binding protein 1 (LTBP1), and stSG 4150 and stSG 1707 were a part of human EH domain containing protein 3 (EHD3). The remaining ESTs had no striking homology with the known genes.

(6) A novel mutation 1206-1209delCCTT in the exon 7 of *SPAST* was identified (Human Gene Mutation Data Accession Number: H971836) in one patient. In addition, one polymorphism 1298+17A>C in the intron 8 of *SPAST* was detected and the same polymorphism was found in one of 100 unrelated controls. There was no mutation found in the SPG4-linked family when all 17 exons, 5'UTR, 3'UTR, and introns 6,10,11 were sequenced. The mechanism of mutation in the SPG4-linked family has to be further investigated for gross deletion or atypical intronic mutation.

## Zusammenfassung

Reine hereditäre spastische Paraplegie (pHSP) bezeichnet eine Gruppe von relativ gutartigen neurodegenerativen Erkrankungen des Rückenmarks. Das hervorstechende klinische Symptom ist eine langsam fortschreitende Gang-Anomalität, die durch eine Spastik der Beine begründet ist. Autosomal dominante hereditäre spastische Paraplegien (ADHSP) sind die Hauptformen von pHSP und sind genetisch heterogen. Als Genorte für die reine ADHSP wurden die Chromosomen 2 (SPG4 auf 2p 21-24, SPG13 auf 2q24-34), 8 (SPG8 auf 8q23-24), 12 (SPG10 auf 12p13), 14 (SPG3 auf 14q11.2-q24.3), 15 (SPG6 auf 15q11.1), und 19 (SPG12 auf 19q13) identifiziert. Bis zum heutigen Tag wurde allein das in SPG4-Patienten mutierte SPAST-Gen kloniert und analysiert.
Die Ziele dieser Dissertation sind: (1) Bestimmung des Krankheitslocus einer großen Familie mit reiner ADHSP; (2) Erstellung einer physikalischen Karte der Kandidaten-Gen-Region, um Kandidaten-Gene untersuchen zu können; (3) Untersuchung, ob CAG-Repeat-Expansionen für die Krankheitsentwicklung verantwortlich sind und (4) Suche nach neuen Mutationen im SPAST-Gen. Die Ergebnisse sind hier kurz zusammengefasst:

(1) Bei einer großen Familie mit reiner ADHSP wurde eine Linkage-Analyse durchgeführt und eine signifikante Assoziation der SPG4 Krankheitsloci mit lod score Werten von 3.63 am Locus D2S367 und 3.39 am Locus D2S352 erhalten.

(2) Eine Yeast Artificial Chromosome (YAC) Contig, die 10 zusätzliche YAC's und 13 EST's beinhaltet, wurde auf der Basis der zuvor veröffentlichten YAC Contig der SPG4-kritischen Region erstellt.

(3) Sechs Fragmente mit CAG/CTG Sequenzen wurden aus einer Cosmid-Genbank aus der SPG4 Kandidatengen-Region zwischen den Loci D2S367 und D2S400 isoliert. Bei der Analyse dieser sechs CAG-Sequenzen beinhaltenden Fragmente wurde ein hoch polymorpher STRP (D253008) zufällig identifiziert. Dieser STRP hat die Struktur (GAAA)n(AAGG)n und es konnten 28 Allele in 50 kaukasischen DNA-Proben identifiziert werden. Dieser Marker hat eine Heterozygosität von 0.81 und einen PIC Wert von 0.78.

Die Loci D2S352 und D2S2283 flankieren den Marker D2S3008 distal und D2S2351 flankiert ihn proximal. Ein 254 bp Allel am Locus D2S3008 segregiert mit den betroffenen Individuen aus der SPG4 zugeordneten Familie.

(4) Die verbleibenden fünf CAG-positiven Klone konnten genutzt werden, um fünf neue STSs (98663, 98664, 98809, 98810 und 98811) zu generieren. Zwei der sechs Fragmente beinhalteten sechs CTG-Repeat-Einheiten. In den Patienten der SPG4 assoziierten Familie konnte jedoch keine Expansion eines CAG/CTG Repeat-Allels nachgewiesen werden.

5) Über die SPG4 Kandidatenregion wurde ein Contig gelegt das aus 33 P1-Bakteriophagen Artifiziellen Chromosomen (PAC's) und 15 Bakteriellen Artifiziellen Chromosomen (BAC's) besteht. Die Längen der BAC's und PAC's wurden über Pulsfeld Gelelektrophorese bestimmt, die Durchschnittsgröße der PAC's belief sich auf 120kB, die der BAC's auf 127 kB. Zusätzlich wurde eine Summe von 41 STS's, die 26 EST's beinhalteten, in diesem Contig lokalisiert. Einige der 26 EST's wurden Genen zugeordnet: SHGC12567 wurde als Teil der humanen Xanthin Dehydrogenase erkannt; N52847 war identisch mit dem Human Latent Transforming Growth Factor Beta Binding Protein 1 (LTBP1); und sowohl stSG 4150 als auch stSG 1707 wurden als Teile des Humanen EH Domain Containing Protein 3 (EHD3) erkannt. Die weiteren EST's zeigten keine deutlichen Homologien zu bekannten Genen.

(6) In einem Patienten wurde im Exon 7 von SPAST eine neue Mutation 1206-1209(CCTT identifiziert (Human Gene Mutation Data Accession Number: H971836). Zusätzlich wurde in Intron 8 ein Polymorphismus 1298+17A>C gefunden, der aber auch in einer der 100 nicht verwandten Kontrollpersonen vorkam. In der Familie mit SPG4-Link konnte in allen 17 Exonen, sowie in 5'UTR, 3'UTR und Introns 6, 10, 11 von SPAST über DNA-Sequenzierung keine Mutation gefunden werden. Die Art der Mutation in der Familie mit SPG4-Link bleibt also noch zu bestimmen. Es könnte sich um eine große Deletion oder um eine atypische intronische Mutation handeln.

# 7 References:


1. Mckusick VA. Spastic paraplegia 3A. Online mendelian Inheritance in Man, OMIM (TM). McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, MD)2000. World Wide Web URL: Http://www.ncbi.nlm.nih.gov/omim/

2. Fink JK, Heiman-Patterson T, the Hereditary Spastic Paraplegia Working Group. Hereditary spastic paraplegia: advances in genetic research. *Neurology* 1996; **335**: 1169-1175.

3. Baraitser M. Spastic paraplegia/HSP. In: Motulsky AG, bobrow M, Harper PS, Scriver C, eds. *The genetics of neurological disorders*. 2nd ed. New York: Oxford University Press, 1990:275-290.

4. Harding AE. Hereditary spastic paraplegias. *Semin Neurol* 1993; **13**: 333-336.

5. Werderlin L. Hereditary ataxias. Occurrence and clinical features. *Acta Neurol Scand* 1986; **73** (suppl. 106): 124.

6. Skre H. Hereditary spastic paraplegia in western Norway. *Clin Genet* 1974; **6**: 165-183.

7. Polo JM, Calleja J, Combarros O. Hereditary atacisas and paraplegias in Cantabria, Spain. An epidemiological and clinical study. *Brain* 1991; **114**: 855-866.

8. Reid E. Pure hereditary spastic paraplegia. *J Med Genet* 1997; **34**: 499-503.

9. Harding AE. Hereditary ´´pure´´ spastic paraplegia: a clinical and genetic study of 22 families. *J Neurol Neurosurg Psychiatry* 1981; **44**: 871-883.

10. Keppen LD, Leppert MF, O`Connell PO. Etiological heterogeneity in X-linked spastic paraplegia. *Am J Hum Genet* 1987; **41**: 933-943.

11. Strümpell A. Beiträge zur Pathologie des Rückenmarks. *Arch Psychiatr Nervenkr* 1880; **10**: 676-717.

12. Dürr A, Brice A, Serdaru M, Rancurel G, Derouesne C, Lyon-Caen O, Agid Y, Fontaine B. The phenotype of ´´pure`` autosomal dominant spastic paraplegia. *Neurology* 1994; **44**: 1274-1277.

13. Behan W, Maia M. Strumpell´s familial spastic paraplegia genetics and neuropathology. *J Neurol Neurosurg Psychiatry* 1974; **37**: 8-20.

14. Kenwrick S, Ionasescu V, Ionasescu G, Searby C, King A, Dubowitz M, Davies KE. Linkage studies of X-linked recessive spastic paraplegia using DNA probes. *Hum Genet* 1986; **73**: 264-266.

15. Winter RM, Davies KE, Bell MV, Huson SM, Patterson MN. MASA syndrome: further clinical delineation and chromosomal localisation. *Hum Genet* 1989; **82**: 367-370.

16. Schrander-Stumpel C, Legius E, Fryns JP, Cassiman JJ. MASA syndrome: new clinical features and linkage analysis using DNA probes. *J Med Genet* 1990; **27**: 688-692.

17. Jouet M, Rosenthal A, Armstrong G, MacFarlane J, Stevenson R, Paterson J, Metzenberg A, Ionasescu V, Temple K, Kenwrick S. X-linked spastic paraplegia (SPG1), MASA syndrome and X-linked hydrocephalus result from mutations in the L1 gene. *Nature Genet* 1994; **7**: 402-407.

18. Bonneau D, Rozet JM, Bulteau C, Berthier M, Mettey R, Gil R, Munnich A, Le Merrer M. X linked spastic paraplegia (SPG2): clinical heterogeneity at a single gene locus. *J Med Genet* 1993; **30**: 381-384.

19. Saugier-Veber P, Munnich A, Bonneau D, Rozet JM, Le Merrer M, Gil R, Boespflug-Tanguy O. X-linked spastic paraplegia and Pelizaeus-Merzbacher disease are allelic disorders at the proteolipid protein locus. *Nature Genet* 1994; **6**: 257-262.

20. Kobayashi H, Hoffman EP, Marks HG. The rumpshaker mutation in spastic paraplegia. (Letter) *Nature Genet* 1994; **7**: 351-352.

21. Schneider A, Montague P, Griffiths I, Fanarraga M, Kennedy P, Brophy P, Nave KA. A mutation of the proteolipid protein gene uncoupling hypomyelination from glial cell death. *Nature* 1999; **358**: 758-761

22. Steinmuller R, Lantigua-Cruz A, Garcia-Garcia R, Kostrzewa M, Steinberger D, Müller U. Evidence of a third locus in X-linked recessive spastic paraplegia. *Hum Genet* 1997; **100**: 287-289.

23. Hentati A, Pericak-Vance MA, Hung WY, Belal S, Laing N, Boustany RM, Hentati F, Ben Hamida M, Siddique T. Linkage of 'pure' autosomal recessive familial spastic paraplegia to chromosome 8 markers and evidence of genetic locus heterogeneity. *Hum Mol Genet* 1994; **3**: 1263-1267.

24. De Michele G, De Fusco M, Cavalcanti F, Filla A, Marconi R, Volpe G, Monticelli A, Ballabio A, Casari G, Cocozza S. A new locus for autosomal recessive hereditary spastic paraplegia maps to chromosome 16q24.3. *Am J Hum Genet* 1998; **63**: 135-139.

25. Casari G, De Fusco M, Ciarmatori S, Zeviani M, Mora M, Fernandez P, De Michele G, Filla A, Cocozza S, Marconi R, Dürr A, Fontaine B, Ballabio A. Spastic paraplegia and OXPHOS impairment caused by mutations in paraplegin, a nuclear-encoded mitochondrial metalloprotease. *Cell* 1998; **93**: 973-983.

26. Confalonieri F, Duguet M. A 200-amino acid ATPase module in search of a basic fucntion. *Bioessays* 1995; **17**: 639-650.

27. Patel S, Latterich M. The AAA team: related ATPases with diverse functions. *Trends Cell Biol* 1998; **8**: 65-71.

28. Neuwald AF, Aravind L, Spouge JL, Koonin EV. $AAA^{+}$: a class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res* 1999; **9**: 27-43.

29. Martinez Murillo F, Kobayashi H, Pegoraro E, Galluzzi G, Creel G, Mariani C, Farina E, Ricci E, Alfonso G, Pauli RM, Hoffman EP. Genetic localization of a new locus for recessive familial spastic paraparesis to 15q13-15. *Neurology* 1999; **53**: 50-56.

30. Vazza G, Zortea M, Boaretto F, Micaglio GF, Sartori V, Mostacciuolo, ML. A new locus for autosomal recessive spastic paraplegia associated with mental retardation and distal motor neuropathy, SPG14, maps to chromosome 3q27-q28. *Am J Hum Genet* 2000; **67**: 504-509.

31. Hazan J, Lamy C, Melki J, Munnich A, de Recondo J, Weissenbach J. Autosomal dominant familial spastic paraplegia is genetically heterogeneous and one locus maps to chromosome 14q. *Nature Genet* 1993; **5**: 163-167.

32. Hentati A, Pericak-Vance MA, Lennon F, Wasserman B, Hentati F, Juneja T, Angrist MH, Hung WY, Boustany RM, Bohlega S, *et al*.. Linkage of a locus for autosomal dominant familial spastic paraplegia to chromosome 2p marker. *Hum Mol Genet* 1994; **3**: 1867-1871.

33. Gispert S, Santos N, Damen R, Orozco G, Ha-Hao D, Ratzlaff T, Aguiar J, Torrens I, Heredero L, Brice A, *et al*.. Autosomal dominant familial spastic paraplegia: reduction of the FSP1 candidate region chromosome 14q to 7cM and locus heterogeneity. *Am J Hum Genet* 1995; **56**: 183-187.

34. Hazan J, Fontaine B, Bruyn RPM, Lamy C, van Deutekom JCT, Rime CS, Dürr A, Melki J, Lyon-Caen O, Agid Y, Munnich A, Padberg GW, de Recondo J, Frants RR, Brice A, Weissenbach J. Linkage of a new locus for autosomal dominant familial spastic paraplegia to chromosome 2p. *Hum Mol Genet* 1994; **3**: 1569-1573.

35. Scott WK, Gaskell PC, Lennon F, Wolpert CM, Menold MM, Aylsworth AS, Warner C, Farrell CD, Boustany RMN, Albright SG, Boyd E, Kingston HM, Cumming WJK, Vance JM, Pericak-Vance MA. Locus heterogeneity, anticipation and reduction of the chromosome 2p minimal candidate region in autosomal dominant familial spastic paraplegia. *Neurogenetics* 1997; **1**: 95-102.

36. Hazan J, Fonknechten N, Mavel D, Paternotte C, Samson D, Artiguenave F, Davoine CS, Cruaud C, Dürr A, Wincker P, Brottier P, Cattolico L, Barbe V, Burgunder JM, Prud'homme JF, Brice A, Fontaine B, Heilig R, Weissenbach J. Spastin, a new AAA protein, is altered in the most frequent form of autosomal dominant spastic paraplegia. *Nature Genet* 1999; **23**: 296-303.

37. Hazan, J, Davoine C-S, Mavel D, Fonknechten N, Paternotte C, Fizames C, Cruaud C, Damson D, Muselet D, Vega-Czarny N, Brice A, Gyapay G, Heilig R, Fontaine B, Weissenbach J. A Fine intergarted map of the SPG4 locus excludes an expanded CAG repeat in chromosome 2p-linked autosomal dominant spastic paraplegia. *Genomics* 1999; **60**: 309-319.

38. Fink JK, Wu CB, Jones SM, Sharp GB, Lange BM, Lesicki A, Reinglass T, Varvil T, Otterud B, Leppert M. Autosomal dominant familial spastic paraplegia: tight linkage to chromosome 15q. *Am J Hum Genet* 1995 ; **56**: 188-192.

39. Hedera, P, Rainier S, Alvarado D, Zhao X, Williamson J, Otterud B, Leppert M, Fink JK. Novel locus for autosomal dominant hereditary spastic paraplegia, on chromosome 8q. *Am J Hum Genet* 1999 ; **64**: 563-569.

40. Seri M, Cusano R, Forabosco P, Cinti R, Caroli F, Picco P, Bini R, Morra VB, De Michele G, Lerone M, Silengo M, Pela I, Borrone C, Romeo G, Devoto M. Genetic mapping to 10q23.3-q24.2, in a large Italian pedigree, of a new syndrome showing bilateral cataracts, gastroesophageal reflux, and spastic paraparesis with amyotrophy. *Am J Hum Genet* 1999; **64**: 586-593.

41. Reid E, Dearlove AM, Rhodes M, Rubinsztein DC. A new locus for autosomal dominant 'pure' hereditary spastic paraplegia mapping to chromosome 12q13, and evidence for further genetic heterogeneity. *Am J Hum Genet* 1999; **65**: 757-763.

42. Reid E, Dearlove AM, Osborn O, Rogers MT, Rubinsztein DC. A locus for autosomal dominant 'pure' hereditary spastic paraplegia maps to chromosome 19q13. *Am J Hum Genet* 2000; **66**: 728-732.

43. Reid E, Grayson C, Rogers MT, Rubinzstein DC. Locus-phenotype corelations in autosomal dominant pure hereditary spastic paraplegia: a clinical and molecular genetic study of 28 United kingdom families. *Brain* 1999; **122**: 1741-1755.

44. Fontaine B, Davoine CS, Dürr A, Paternotte C, Feki I, Weissenbach J, Hazan J, Brice A. A new locus for autosomal dominant pure spastic paraplegia, on chromosome 2q24-q34. *Am J Hum Genet* 2000; **66**: 702-707.

45. Dracopoli NC, Haines JL, Korf BR, Moir DT, Morton CC, Seidman CE, Seidman JG, Smith DR. *Current Protocol in Human Genetics.* John Wiley & Sons, Inc. 1994.

46. Ausbubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. *Current Protocol in Molecular Biology*. John Wiley & Sons, Inc. 1993.

47. Dubé MP, Rouleau, GA, Kibar Z, Farlow MR, Ebers G, Harper P, Kolodny EH, Baumbach L, Figlewicz DA. Autosomal dominant hereditary spastic paraplegia: Linkage analysis of a heterogeneous trait. *Am J Hum Gent* 1995; **57**: A190.

48. Hudson TJ, Stein LD, Gerety SS, Ma J, Castle AB, Silva J, Slonim DK, Baptista R, Kruglyak L, Xu SH, *et al.* An STS-Based Map of the Human Genome. *Science* 1995; **270**: 1945-1954.

49. De Jonghe P, Krols L, Michalik A, Hazan J, Smeyers G, Löfgren A, Weissenbach J, Martin JJ, Van Broeckhoven C. Pure Familial Spastic Paraplegia: Clinical and Genetic Analysis of Nine Belgian Pedigrees. *Eur J Hum Genet* 1996; **4**: 260-266.

50. Lau EL, Kostrzewa M, Müller U. Transcript map of the chromosome 2 linked autosomal dominant spastic paraplegia (SPG4) critical region and identification of a highly informative STRP. *Neurogenetics* 1998; **2**: 75-76.

51. Nielsen JE, Koefoed P, Abell K, Hasholt L, Eiber H, Fenger K, Niebuhr E, Sørensen SA. CAG repeat expansion in autosomal dominant pure spastic paraplegia linked to chromosome 2p21-p24. *Hum Mol Genet* 1997; **6**: 1811-1816.

52. Schalling M, Hudson TJ, Buetow KH, Housman DE. Direct detection of novel expanded trinucleotide repeats in the human genome. *Nature Genet* 1993; **4**: 135-139.

53. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997; **25**: 3389-3402.

54. Ichida K *et al.* Cloning of the cDNA encoding human xanthine dehydrogenase (oxidase): structural analysis of the protein and chromosomal location of the gene. *Gene* 1993; **133**: 279-284.

55. Kanzaki, T. *et al*. TGF-β 1 binding protein: a component of the large latent complex of TGF-β 1 with multiple repeat sequences. *Cell* 1990; **61**: 1051-1061.

56. Pohl U, Smith JS, Tachibana I, Ueki K, Lee HK, Ramaswamy S, Wu Q, Mohrenweiser HW, Jenkins RB, Louis DN. EHD2, EHD3, and eHD4 enocode novel members of highly conserved family of EH domain-containing proteins. *Genomics* 2000; **63**: 255-262.

57. Collins FS. Positional cloing moves from preditional to traditional. *Nature Genet*. 1995; **9**: 347-350.

58. Morgan TH. Random segregation versus coupling in Mendelian inheritance. *Science* 1911; **34**: 384.

59. Bell J, Haldane JBS. The linkage between the genes for color-blindness and haemophilia in men. *Proc R Soc Ser [Biol]* 1937; **123**: 119-150.

60. Haldane JBS. The combination of linkage values and the calculation of distances between the loci of linked factors. *J Genet* 1919; **8**: 299-309.

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

62. Ott J. Analysis of human genetic linkage, 3rd ed. 1999. Johns Hopkins University Press, Baltimore.

63. Koenig M, Hoffman EP, Bertelsom CJ, Monaco AP, Feener C, Kunkel LM. Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell* 1987; **50**: 509-517.

64. Lee WH, Bookstein R, Hong F, Young LJ, Shew JY, Lee EY. Human retinoblastoma susceptibility gene: Cloning, identification, and sequence. *Science* 1987; **235**: 1394-1399.

65. Walter MA, Spillett DJ, Thomas P, Weissenbach J, Goodfellow PN. A method for constructing radiation hybrid maps of whole genomes. *Nature Genet* 1994; **7**: 22-28.

66. Krawczak M, Schmidtke J. *DNA Fingerprinting*, 2nd ed. BIOS Scientific Publishers, Oxford.

67. Sternberg N. Bacteriophage P1 cloning system for the isolation, amplification, and recovery of DNA fragments as large as 100 kilobase pairs. *Proc Natl Acad Sci USA* 1990; **87**: 103-107.

68. Dilella, AG, Woo SLC. Cloning large segments of genomic DNA using cosmid vecotrs. *Methods Enzymol*. 1987; **152**: 199-212.

69. Shizuya H, Birren B, Kim UJ, Mancino V, Slepak T, Tachiiri Y, Simon M. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci USA* 1992; **89**: 8794-8797.

70. Ioannou Pa, Amemiya CT, Garnes J, Kroisel PM, Shizuya H, Chen C, Batzer MA, de Jong PJ. P1-derived vector for the propagation of large human DNA fragments. *Nature Genet* 1994; **6**: 84-89.

71. Burke DT, Carle GF, Olson MV. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* 1987; **236**: 806-812.

72. Cohen D, Chumakov I, Weissenbach J. A first-generation map of the human genome. *Nature* 1993; **366**: 698-701.

73. Strachan T, Read AP. *Human Molecular Genetics*. BIOS Scientific Publishers, 1999.

74. Raskind WH, Pericak-Vance MA, Lennon F, Wolff J, Lipe HP, Bird TD. Familial Spastic Paraparesis: Evaluation of locus heterogeneity, anticipation, and Haplotype mapping of the SPG locus on the short arm of chromosome 2. *Am J Med Genet* 1997; **74**: 26-36.

75. Kremer E, Pritchard M, Lynch M, Yu S, Holman K, Baker E, Warren ST, Schlessinger D, Sutherland GR, Richards RI. Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Science* 1991; **252**: 1711-1714.

76. Knight SJL, Flanner AV, Hirst MC, Campbell L, Christodoulou Z, Phelps SR, Pointon J, Middleton-Price HR, Barnicoat A, Pembrey ME, Holland J, Oostra BA, Bobrow M, Davies KE. Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. *Cell* 1993; **74**: 127-134.

77. Gacy AM, Goellner GM, Spiro C, Chen X, Gupta G, Bradbury EM, Dyer RB, Mikesell MJ, Yao JZ, Johnson AJ, Richter A, Melancon SB, McMurray CT. GAA instability in Friedreich's Ataxia shares a common, DNA-directed and intraallelic mechanism with other trinucleotide diseases. *Mol Cell* 1998; **1**: 583-93.

78. Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion JP, Hudson T, sohn R, Zemelman B, Snell RG, Rundle SA, Crow S, Davies J, Shelbourne P, Buxton J, Jones C, Juvonen V, Johnson K, Harper PS, Shaw DJ,

Housman DE. Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG)repeat at the 3'end of a transcript encoding a protein kinase family member. *Cell* 1992; **68**: 799-808.

79. The Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 1993; **72**: 971-983.

80. La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 1991; **352**: 77-79.

81. Orr HT, Chung M, Banfi S, Kwiatkowski TJ Jr, Servadio A, Beaudet AL, McCall AE, Duvick LA, Ranum LPW, Zoghbi HY. Expansion of an unstable trinucleotide (CAG) repeat in spinocerebellar ataxia type 1. *Nature Genet* 1993; **4**: 221-226.

82. Pulst SM, Nechiporuk A, Nechiporuk T, Gisper S, Chen XN, Lopes-Cendes I, Pearlman S, Starkman S, Orozco-Diaz G, Lunkes A, De Jong P, Rouleau GA, Auburger G, Korenber JR, Figueroa C, Sahba S. Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nature Genet* 1996; **14**: 269-276.

83. Kawaguchi Y, Okamota Z, Zaniwaki M, Aizawa M, Inoue M, Katayama S, Kawakami H, Nakamura S, Nishimura M, Akiguchi I, Kimur J, Narumiya S, Kakizuka A. CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nature Genet*; **8**: 221-228.

84. Zhuchnko O, Bailey J, bonnen P, Ashizawa T,Stockton DW, Amos C, Dobyns WB, Subramony SH, Zoghbi HY, Lee CC. Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the $\alpha_{1A}$-voltage-dependent calcium channel. *Nature Genet* 1997; **15**: 62-69.

85. David G, Abbas N, Stevanin G, Dürr A, Yvert G, Cancel G, Wever C, Imbert G, Saudou F, Antoniou E, Drabkin H, Gemmil R, Giunti P, Benomar A, Wood N, Buber M, Agid Y, Mandel JL, Brice A. Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nature Genet* 1997; **17**: 65-70.

86. Koide R, Ikeuchi T, Onodera O, Tanaka H, Igarashi S, Endo K, Takasashi H, Kondo R, Ishikawa A, Hayashi T, Saito M, Tmoda A, Miike T, Naito H, Ikuta F, Tsuji S. Unstable expansionof CAG repeat in hereditary dentatorubral-pallidouluysian atrophy (DRPLA). *Nature Genet* 1994; **6**: 9-13.

87. Benson KF, Horwitz M, Wolff, J, Griend K, Thompson E, White S, Richards RI, Raskind WH, Bird TD. CAG repeat expansion in autosomal dominant familial spastic paraparesis: novel expansion in a subset of patients. *Hum Molec Genet* 1998; **7**: 1779-1786.

88. Breshel TS, McInnis MG, Margolis RL, Sirugo G, Corneliussen B, Simpson SG, McMahon FJ, Mackinnon DF, Xu JF, Pleasant N, Huo Y, Ashworth rG, Grundstrom C, Grundstrom T, Kidd KK, DePaulo JR, Ross CA. A novel, heritable expanding CTGrepeat in an intron of the SEF2-1 gene on chromosome 18q21.1. *Hum Mol Genet* 1997; **6**: 1855-1863.

89. Nakamoto M, Takebayashi H, Kawaguch Y, NarumiyaY, Taniwaki M, Nakamura Y, Ishikawa Y, Akiguchi I, Kimura J, Kakizka A. A CAG/CTG expansion in the normal population. *Nature Genet* 1997; **17**: 358-386.

90. Del-Favero J, Goossens D, De Jongh P, Benson K, Michalik A, Ban den Bossche D, Horwitz M, Van Broeckhoven C. Isolation of CAG/CTG repeats from within the chromosome 2p21-p24 locus for autosomal dominant spastic paraplegia (SPG4) by YAC fragmentation. *Hum Genet* 1999; **105**: 217-225.

91. Andersson S, Berman DM, Jenkins EP, Russel DW. Deletion of steroid 5'reductase 2 gene in male pseudohermaphroditism. *Nature* 1991; **354**: 159-161.

92. Fonknechten N, Mavel D, Byrne P, Davvoine CS, Cruaud C, Boentsch D, Samson D, Coutinho P, Hutchinson M, McMonagle P, Burgunder JM, Tartaglione An, Heinzlef O, Feki I, Deufel T, Parfrey N, Brice A, Fontaine B, Prud'homme JF, Weissenbach J, Dürr A, Hazan J. Sepectrum of SPG4 mutations in autosomal dominant spastic paraplegia. *Hum Mol Genet* 2000; **9**: 637-644.

93. Bürger J, Fonknechten N, Hoeltzenbein M, Neumann L, Bratanoff E, Hazan J, Reis A. Hereditary spastic paraplegia cased by mutations in the SPG4 gene. *Eur J Hum Genet*. 2000; **8**: 771-776.

94. Lindsey JC, Lusher ME, McDermott CJ, White KD, Reid E, Rubinsztein DC, Bashir R, Hazan J, Shaw PJ, Bushby KMD. Mutation analysis of the spastin gene (*SPG4*) in patients with hereditary spastic paraparesis. *J Med Genet* 2000; **37**: 759-765.

95. Santorelli FM, Patrono C, Fortini D,Tessa A, Comanducci G, Bertini E, Pierallini A, Nmabile GA, Casali C. Intrafamilial variability in hereditary spastic paraplegia associated with an *SPG4* gene mutation. *Neurology* 2000; **55**: 702-705.

96. Hentati A, Deng HX, Zhai H, Chen W, Yang Y, Hung WY, Azim AC, Bohlega S, Tandan R, Warner C, Laing NG, Cambi F, Mitsumoto H, Roos RP, Boustany RM Ben Hamida M, Hentati F, Siddique T. Novel mutations in spastin gene and absence of correlation with age at onset of symptoms. *Neurology* 2000; **55**: 1388-1390.

97. Lau EL , Müller U. 2000. *Human Mutation database*: accession number H971836.

98. Sheffield VC, Beck JS, Kwitek AE, Sandstrom DW, Stone EM. The sensitiveity of single strand conformaltional polymorphism analysis for the dtection of single base substitutions. *Genocmics* 1993; **16**: 325-332.

99. Cotton RGH, Edkins E, Forrest S. 1998 *Mutation Detection: a Pratical Approach*. IRL Press. Oxford.

100. Mashal RD, Koontz J, Sklar J. Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases. *Nature Genet* 1995; **9**: 177-183.

101. van der Luijt R, Khan PM, Vaen H, van Leeuwen C, tops C, Roest PA et al. Rapid detection of translation-terminating mutations at the adenomatous polyposis coli (APC) gene by direct protein truncation test. *Genomics* 1994; **11**: 365-368.

102. Hacia JG. Resequencing and mutational analysis using oligonucleotide microarrays. *Nature Genet* 1999; **21** (suppl.): 42-47.

103. Chelly J, Kaplan JC, Maire P, Gautron S, Kahn A. Transcription of the dystrophin gene in human muscle and non-muscle tissue. *Nature* 1988; **333**: 858-860.

104. McMonagle P, Byrne PC,Fitzgerald B, Webb S, Parfrey NA, Hutchinson M. Phenotype of AD-HSP due to mutations in the SPAST gene. *Neurology* 2000; **55**: 1794-1800.

105. Hilt W, Wolf DH. Proteassomes of the yeast *S.cerevisiae*: genes, structure and functions. *Mol biol Reports* 1995; **21**: 3-10.

106. Liberzon A, Shpungin S, Bangion H, Yona E, Katcoff DJ. Association of yeast SAP1, a novel member of the ``AAA´´ ATPase family of proteins, with the chromatin protein SIN1. *FEBS Lett* 1996; **388**: 5-10.

107. Andrews BJ, Mason SW. Gene expression and the cell cycle; a family affair. *Science* 1993; **261**: 1543-1544.

108. Andrews BJ, Donoviel MS. A heterodimeric transcriptional repressor becomes clear. *Science* 1995; **270**: 251-253.

109. Ozelius LJ, Heweet JW, Page CE, Bressman SB, Krmer PL, Shalish C, de Deon D, Brin MF, Raymond D, Corey DP, Fahn S, Risch NJ, Buckler AJ, Gusella JF, Breakefield XO. The early-onset torsion dystonia gene (DYT1) encodes an ATP-binding protein. *Nature Genet*. 1997; **17**: 40-48.

110. Hardy J, Gwinn-Hardy K. Genetic classification of primary neurodegenerative disease. *Science* 1998; **282**: 1075-1079.

111. Heinzlef O, Paternotte C, Jean-Franҫois Prud'homme F M, Dien J, Madigand M, Puget J, Weissenbach J, Roullet E, Hazan J. Mapping of a complicated familial spastic paraplegia to locus SPG4 on chormosome 2p. *J Med Genet* 1998; **35**: 89-93.

112. Webb S, Coleman D, Byrne P, Parfrey N, Burke T, Hutchinson J, Hutchinson M. Autosomal dominant hereditary spastic paraparesis with cognitive loss linked to chromosome 2p. *Brain* 1998; **121**: 601-609.

113. White KD, Ince PG, Lusher M, Lindsey J, Cookson M, Bahir R, Shaw PJ, Bushby KMD. Clinical and pathologic findings in hereditary spastic paraparesis with spastin mutation. *Neurology* 2000; **55**: 89-94.

114. Byrne PC, Mc Monagle P, Webb S, Fitzgerald B, Parfrey NA, Hutchinson M. Age-related cognitive decline in hereditary spastic paraparesis linked to chromosome 2p. *Neurology* 2000; **54**: 1510-1517.

115. Neerup JL, Gerstenberg T, Kallestrup EB, Koefoed P, Nordling J, Nielsen JE. Urodynamic evaluation of patients with autosomal dominant pure spastic paraplegia linked to chromosome 2p21-2p24. *J Neurol Neurosurg Psychiatry* 1998; **65**: 693-696.

116. Figlewicz D.A. and Bird T.D.. ´´Pure`` hereditary spastic paraplegia-The story becomes complicated. *Neurology* 1999; **53**: 5-7.

117. Fisher E . and Scambler P. Human haploinsufficiency-one for sorrow, two for joy. *Nature Genet*. 1994; **7**: 5-7.

118. Wilkie AOM . The molecular basis of genetic dominace. *J Med Genet*. 1994; **31**: 89-98.

119. Holtzman DM, Epstein CJ. The molecular genetics of Down syndrome. *Mol Genet Med* 1992; **2**: 105-120.

120. Yamakawa K, Huo Yk, Haendel MA, Huber R, Chen XN, Lyons GE, Korenberg JR. DSCAM: a novel member of the immunoglobulin superfamily maps in a Down syndrome region and is involved in the development of the nervous system. *Hum Mol Genet* 1998; **7**: 227-237.

121. Oliner JD, Kinzler KW, Meltzer PS, George DL, Vogelstein B. Amplification of a gene encoding a p53-associated protein in human sarcomas. *Nature* 1992; **358**: 80-83.

122. Cleary ML. Oncogenic conversion of transcription factors by chromosomal translocations. *Cell* 1991; **66**: 619-622.

123. Rabbitts TH. Translocations, master genes, and differences between the origins of acute and chronic leukemias. *Cell* 1991; **67**:641-644.

124. Trent C, Wood WB, Horvitz HR. A novel dominant transformer allele of the sex-determining gene *her-1* of *Caenorhabditis elegans*. *Genetics* 1988; **120**: 145-157.

125. Miller MW, Duhl DM, Vrieling H, Cordes SP, Ollmann MM, Winkes BM, Barsh GS. Cloning of the mouse *agouti* gene predicts a secreted protein ubiquitously expressed in mice carring the *lethal yellow* mutation. *Genes Dev* 1993; **7**: 1203-1213.

126. Feingold EA, Forget BG. The breakpoint of a large deletion causing hereditary persistence of fetal hemoglobin occurs within an erythroid DNA domain remote from the β-globin gene cluster. *Blood* 1989; **74**: 2178-2186.

127. Rogers S, Wells R, Rechsteiner M. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* 1986; **234**: 364-368.

128. Mango SE, Maine EM, Kimble J. Carboxy-termianl truncation activates *glp-1* protein to specify vulval fates in *Caenorhabditis elegans*. *Nature* 1991; **352**: 811-815.

129. Vogelstein B. A deadly inheritance. *Nature* 1990; **348**: 681-682.

130. Vogelstein B, Kinzler KW. P53 function and dysfunction. *Cell* 1992; **70**: 523-526.

131. Monplaisir N, Merault G, Poyart C, Rhoda MD, Craescu C, Vidaud M, Galacteros F, Blouquit Y, Rosa J. Hemoglobin S Antilles: a variant with lower solubility than hemoglobin S and producing sickle cell disease in heterozygotes. *Proc Natl Acad Sci USA* 1986; **83**: 9363-9367.

132. Owen MC, Brennan So, Lewis JH, Carrell RW. Mutation of antitrypsin to antithrombin α1-antitrypsin Pittsburgh (358 Met→Arg), a fatal bleeding disorder. *N Engl J Med* 1983; **309**: 694-698.

133. Knudson AG. Mutation and cancer: statistical study of retinoblastom. *Proc Natl Acad Aci USA* 1971; **68**: 820-823.

134. Cavenee WK, Dryja TP, Phillips RA, Benedict WF, Godbout R, Gallie BL, Murphree
135. AL, Strong LC, White RL. Expression of recessive alleles by chromosomal mechanisms
136. in retinoblastoma. *Nature* 1983; **305**: 779-784.

137. DeChiara TM, Roberston EJ, Efstratiadis A. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* 1991; **64**: 849-859.

## Acknowledgment

This dissertation could not have been finished if it were not for the understanding and support of the following people. I would like to thank my supervisor Prof.Dr. Müller for his intensive supervision and critical discussion during my four-years of study in Giessen and especially throughout this dissertation writing period. I would also like to thank my co-advisor Prof.Dr. Renkawitz for his critical reading and useful suggestions to my dissertation.

Thanks also go to Frau.Dr. Anders for kindly taking care of me in the past four years. I especially would like thank my colleagues in the Institute of Human Genetics, Giessen: Dr. Kostrzewa, Dr. Nolte, Dr. Meyer-Rogge, Dr. Stahl, Dr. Joos, Frau Winter, Frau Seimm, and the secretaries, Frau Reichmann, Frau Schmid, and Frau Walldorf. Without them, this dissertation-writing process could not have gone so smoothly.

I would also like to thank Prof.Dr. Pingoud for allowing me to use the Light cycler. And my thanks extend to Drs. Hahn and Wilhem for their direction in quantitative PCR. Thanks also go to the originators of the PAC and BAC libraries, Drs. Pieter de Jong, Kazutoyo Osoegawa, Minako Tateno, and J Gingrich and the library providers, German Resource Center and UK HGMP Resource Center for their very kind supports of materials.

I would also like to thank my English tutor, Dr. Burke, and my cousin Janice for their gratuitous time and help in proofreading my dissertation. Many thanks go to my overseas Taiwanese friends, Chi-ming, Ju-jen, Su-shia, Mr. and Mrs. Mu, Tsai, Lien, and Liu for their encouragement and friendship.

I also would like to thank my family, especially my parents, for their full support during my four years of study. Without them, this dissertation could not have even existed.

Finally, I would like to dedicate this thesis to Prof. em. Dr. Anders. I sincerely hope that he could share this delight in heaven.

## Danksagung

Mit Hilfe der verständnisvollen Unterstützung durch folgende Personen, konnte diese Dissertation abgeschlossen werden.

Ich möchte dem Betreuer meiner Arbeit Prof. Dr. Müller für seine intensive Betreuung während meiner vier Jahre in Gießen und seine Bereitschaft zu kritischer Diskussion, vor allem in der Phase des Schreibens, danken.

Weiterhin möchte ich meinem Betreuer aus dem Fachbereich Biologie, Herrn Prof. Dr. Renkawitz für hilfreiche Vorschläge zur Dissertation, sowie für die kritische Lektüre des Manuskripts danken.

Ein weiterer Dank geht an Frau Dr. Anders, die mir nicht nur fachliche Unterstützung zukommen ließ. Besonders möchte ich mich bei meinen Kollegen am Institut für Humangenetik in Gießen bedanken: Dr. Kostrzewa, Dr. Nolte, Dr. Meyer-Rogge, Dr. Stahl, Dr. Joos, Frau Winter, Frau Seim sowie die Sekretärinnen Frau Reichmann, Frau Schmid und Frau Walldorf. Ohne deren Unterstützung hätte die Arbeit nicht so reibungslos fertiggestellt werden können.

Auch möchte ich mich bei Herrn Prof. Dr. Pingoud bedanken für die Möglichkeit seinen Light-Cycler zu nutzen. In diesem Rahmen möchte ich mich auch bei den Drs. Hahn und Wilhelm bedanken für ihre Einweisung in die quantitative PCR mit dem oben genannten Gerät.

Weiterhin bedanke ich mich bei den Wissenschaftlern und Institutionen von denen ich verschiedene BAC- und PAC-Genbanken erhielt: Drs. de Jong, Osoegawa, Tateno und Gingrich sowie die Institutionen Deutsches Resourcen Zentrum und UK HGMP Resource Center.

Ein weiterer Dank geht an Dr. Burke und meine Cousine Janice, für ihre Zeit und Mühe, die sie auf die Fehlerkontrolle in der englischen Version der Dissertation verwendet haben. Vielmals möchte ich mich noch meinen taiwanesischen Freunden in Übersee Chi-ming Ju-jen, Su-shia, Mr. und Mrs. Mu, Tsai, Lien und Liu für ihre Freundschaft und Ermutigung bedanken.

Vor allem aber auch möchte ich meiner Familie danken, die mich die ganze Zeit meiner vierjährigen Doktorarbeit in Gießen voll unterstützt hat. Ohne diese Unterstützung wäre die Arbeit niemals möglich gewesen.

Endlich möchte ich diese Arbeit Herrn Prof. em. Dr. Anders widmen. Ich hoffe, dass er von höherer Warte den Abschluß dieser Arbeit mit Wohlwollen betrachtet.