

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 cloning⁵⁷ 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 drosophila⁵⁸ and was also found to be applicable to humans. The first demonstration of linkage in humans was done by Bell and Haldane⁵⁹, 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 function⁶⁰: $\varpi = -1/2 \ln(1-2\theta)$, where ϖ is the map distance and θ the recombination 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 θ is the recombination fraction, and by construction $Z(1/2)=0$. Two point lod scores are usually estimated at θ values 0, 0.01, 0.05, 0.1, 0.2, 0.3, and 0.4. Generally, lod scores are calculated at fixed values of θ 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 genes⁶³. 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 Retinoblastoma⁶⁴.

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 maps⁶⁵, restriction maps⁶⁶, maps of overlapping, bacteriophage P1⁶⁷, cosmid⁶⁸, BAC⁶⁹, PAC⁷⁰, YAC⁷¹ 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 map⁷² 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 cloning⁷³. 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 possible 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* locus⁸⁸ 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) locus⁸⁹ 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 dehydrogenase⁵⁴, steroid 5 α -reductase⁹¹ and TGF- β 1 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 peptide 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/SSCP⁹⁸ 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 electrophoresis⁹⁹ (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 tetroxide. 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 product¹⁰¹. For the diseases, such as Duchenne muscular dystrophy, adenomatous polyposis 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 scanning¹⁰². 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 transcripts¹⁰³. 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* studies¹⁰⁴. 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 pathogenic⁷³. 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 proposed³⁶. 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 nucleus¹⁰⁵. 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 Sin1p¹⁰⁶. 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 neuron degenerative 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 lymphocytes³⁶. 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 degeneration of neurons. Selective vulnerability is a common phenomenon in many human neurodegenerative diseases¹¹⁰. For example, Huntington'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 parkinsonism. 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 pedigrees 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 ADHSP¹¹⁵. 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 explored¹¹⁶.

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, destabilization 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 possibly 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 family¹¹⁴. 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 below¹¹⁸.

- (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 syndrome¹¹⁹ 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 molecule¹²⁰. 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 males¹²⁴. Increased ectopic expression of a chimerical mRNA encoding a normal protein causes the lethal yellow mutant at the mouse *agouti* locus¹²⁵. 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 protein¹²⁷. 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 mRNA¹²⁸.

(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 suppressor p53

oligomerizes *in 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 protein¹³¹, altered substrate specificity in mutant protein¹³², recessive antioncogene in retinoblastoma¹³³, and parental genomic imprinting¹³⁴ 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.