uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Fabin Han

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Microbiology and Immunology)

GRADE / DEGREE

Department of Biochemistry, Microbiology and Immunology

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Characterization of the Genes for Myoclonus-Dystonia

TITRE DE LA THÈSE / TITLE OF THESIS

Dennis Bulman

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

John Bell

Peter Ray

Robert Korneluk

Catherine Tsilfidis

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Characterization of the Genes for Myoclonus-Dystonia

By

Fabin Han

Thesis
Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of
Doctorate of Philosophy

Department of Biochemistry, Microbiology and Immunology,
Graduate Collaborative Program in Human and Molecular Genetics,
Faculty of Medicine
University of Ottawa

Ottawa, Canada

Your file Votre référence
ISBN: 978-0-494-49354-0
Our file Notre référence
ISBN: 978-0-494-49354-0

Abstract: Myoclonus-dystonia (M-D; OMIM 159900) is an autosomal dominant movement disorder characterized by alcohol responsive myoclonic jerks typically affecting the arms and axial muscles. Mutations in the epsilon-sarcoglycan (*SGCE*) gene on chromosome 7q21 were shown to cause M-D but not in all cases. Currently, in the literature mutations in *SGCE* are responsible for 20-50% of all cases of M-D. By direct sequencing, we have identified mutations in four families and we were able to find large deletions in two additional families using a semi-quantitative PCR-based DHPLC method. One of these families (family 2) has an 18 kb deletion that includes exons 2 and 3 while another family (family 6) has a 13 kb deletion covering exons 2 through 5. We have confirmed these large deletions by identifying the exact deletion breakpoints. In our cohort of 17 unrelated M-D families, 35% of patients with M-D have mutations in the *SGCE* gene. Previously, we characterized a large five-generation Canadian family with M-D (family 1) that did not have a mutation in *SGCE* and mapped a novel locus for this disorder to a 4 Mb region between the markers D18S1132 and D18S843 on 18p11 (OMIM number: 607488). Since additional informative STS markers were not found between the flanking recombinant and non-recombinant markers, we utilized single nucleotide polymorphisms (SNP) for fine-mapping. As a result, we identified three recombinant SNPs (rs385769, rs727951 and rs727952) in the proximal flanking region between markers D18S1163 and D18S843 in family 1. Our previous 4 Mb critical region has been reduced to 3.18 Mb, flanked by markers D18S1132 and rs385769. There are seven known genes and eight predicted genes within or very close to this region. We have sequenced all the exons and exon-intron boundaries of these seven known genes (*MGC17515, ZFP161, EPB41L3 (KIAA0987), L3MBT4, ARHGAP28 (FLJ10312), LAMA1, and PTPRM*) and four predicted genes (*LOC388459, LOC388460,*

LOC400643 and *LOC388461*) without identifying the causative mutations. Future studies should be directed to the investigation of the functional effects of the known *SGCE* mutations using *in vitro* and *in vivo* models. With respect to the identification of a novel gene in family 1, we will focus on the novel transcripts in the critical region based on gene prediction and EST (expressed sequence tag) mapping. Once the second M-D causing gene is identified, we will perform gene expression and cellular localization analyses of the mutated and wild type protein in tissues and generate a mouse model.

Acknowledgements

After four years of hard work and diligent study, I have finished my graduate study with the submission of my PhD thesis. I would like first to thank my supervisors, Dr. Dennis E. Bulman and Dr. David A. Grimes for their careful mentorship, their unlimited patience and their enthusiastic encouragement in each step of my research project. I really appreciate that for all the experiments I did in this lab, they gave me the right direction, they carefully reviewed my research results, and they patiently listened to my explanations and discussed with me the critical issues on my work. Most importantly, they carefully corrected my abstracts, PowerPoint slide and poster presentations, and my paper drafts by showing me the basic word skills, presentation skills, and the skills to answer the questions asked by the audience to intensify my research abilities. In addition, they have provided me a lot of knowledge in molecular genetics, molecular biology, and bioinformatics as well as experimental skills in genotyping, genetic linkage analysis, PCR, RT-PCR, DNA sequencing, gene dosage analysis and so on. I will always remember their strict scientific attitudes and kindness in my future research no matter where I will go. Secondly, I would like to thank my thesis committee members, Dr. Rashmi Kothary and Dr. Ken Dimock for their professional advices, critical comments and constructive suggestions for the success of my research project. Thirdly, I would like to thank all the lab mates in Dr. Bulman's lab and my colleagues in other labs of the OHRI as well as in the Microbiology and Immunology Program at the University of Ottawa. Fourthly, I would like to thank all the clinicians who identified the families with myoclonus-dystonia and shipped blood samples of patients to our lab. Without their contributions, this thesis would never have been possible. These collaborators are: Dr. Antony E. Lang, Dr. Oksana Suchowersky, Dr. Daune McGregor, Dr.

Galit Kleiner-Fisman, Dr. Peter Humphreys and Dr. Alan Fryer. Fifthly, I should thank all the M-D family members for their participation in this research project. I would also thank Canadian Institutes of Health Research (CIHR), Ontario Research Development Fund and The Dystonia Medical Research Foundation for their funding to support my research. Finally, I am greatly indebted to my beloved parents, my wife, my son and my other family members for their emotional support, concern and continuous encouragement during my PhD study.

Table of Contents

Page

List of Figures

List of Tables

List of Abbreviations

M-D	myoclonus-dystonia
cM	centimorgans
chr	chromosome
LOD	logarithm of the odds ratio
Θ	(theta) recombination fraction
SD	standard deviation
DNA	deoxyribonucleic acid
RNA	ribonucleic acid
cDNA	complementary deoxyribonucleic acid
SNP	single nucleotide polymorphism
STS	sequence tagged site
PCR	polymerase chain reaction
RT	reverse transcriptase
DHPLC	denaturing high performance liquid chromatography
SGCE	ε-sarcoglycan
LAMA1	Laminin, alpha 1
7q21	chromosome 7, long arm, region 2, band 1
18p11	chromosome 18, short arm, region 1, band 1
BAC	bacterial artificial chromosome
FISH	fluorescence in situ hybridization
bp	base pair
Kb	kilobase

Mb	megabase
MgCl2	magnesium chloride
dNTP	deoxynucleoside triphosphate
EST	expressed sequence tag
M	moles per litre
mM	millimoles per litre
ml	millilitre
µl	microlitre
mg	milligram
ng	nanogram
nm	nanometer
NCBI	National Centre for Biotechnology Information (USA)
UCSC	University of California Santa Cruz
CIHR	Canadian Institutes of Health Research
OHRI	Ottawa Health Research Institute

Chapter 1.

Introduction: Primary Dystonia and Myoclonus-Dystonia

1.1. The Clinical manifestation and classification of dystonia

Dystonia is a clinically and genetically heterogeneous type of abnormal movement characterized by involuntary sustained muscle contractions, producing twisting or repetitive movements or abnormal postures (Geyer and Bressman, 2006). Dystonia can affect the arms, legs, trunk, neck, eyelids, face, or vocal cords with all gender and ethnic backgrounds being affected. While dystonia tends not to be fatal, it is chronic and prognosis is difficult to predict. The total prevalence of dystonia is approximately 127-329 per million of which focal dystonia accounts for 61-300 per million (Nutt *et al.*, 1988). If movement disorders are classified into those with reduced movement (akinetic-rigid syndromes) versus those with excessive movement (hyperkinetic syndromes), dystonia falls into the hyperkinetic disorder category. The clinical characteristic of dystonia that helps differentiate it from other hyperkinetic movement disorders is that dystonic movements are repetitive with involvement of the same group of muscles. For instance, a patient with idiopathic cervical dystonia or spasmodic torticollis with involuntary right-head rotation will tend to have repetitive right-head rotation throughout the duration of his/her disorder. By comparison, choreic movements are brief, continuous, non-stereotyped and random. Tics involve brief and intermittent movements or sounds. Compared to dystonia, tics may be suppressed or more abrupt, and are often preceded by a subjective compulsive urge or sensation, which is temporarily relieved after the tic occurs. Dystonia is often associated with tremor, which may make the diagnosis somewhat more difficult. Essential tremor and dystonia have been observed to coexist frequently, although tremors associated with dystonia may represent a forme-fruste of dystonia (Rivest and Marsden, 1990). Electromyogram (EMG) studies on

patients with dystonia have shown that there is excessive co-contraction of antagonist muscles and there is overflow activation of extraneous muscles (Cohen LG, 1988).

Dystonia can be classified in several ways based on the clinical characteristics and the etiology. One way is according to the etiology where dystonia is classified into primary dystonia, secondary dystonia, and dystonia-like conditions. Primary dystonia (primary torsion dystonia) occurs either in a familial or sporadic pattern with dystonia as the sole or major symptom. In contrast, secondary dystonia refers to dystonia in the context of a neurological disease in which dystonia is usually one of several symptoms of the disease. It can be caused by peripheral or central trauma, cerebrovascular diseases (e.g. stroke), neurodegenerative diseases (e.g. Huntington's disease and Parkinson's disease), infections, and environmental toxins (e.g. MPTP and manganese). Other dystonia-like conditions such as psychogenic dystonia and pseudodystonia are not true dystonia but may mimic dystonia and should be differentiated from dystonia. In addition, paroxysmal dystonia is a distinct group of dystonia which is different from other forms of dystonia by their periodic occurrence and the patient being free of symptoms in most of the time. This group of dystonia includes paroxysmal kinesigenic dyskinesia (PKD) (Bruno *et al.*, 2004), paroxysmal non-kinesigenic dyskinesia (Bressman *et al.*, 1988) and other paroxysmal dyskinesias (hypnogenic paroxysmal dyskinesias, paroxysmal exertional dyskinesia, infantile paroxysmal dystonias, Sandifer's syndrome and symptomatic paroxysmal dyskinesias) (Dressler and Benecke, 2005).

Another way to classify dystonia is simply by using clinical criteria (Bressman, 2004). If the symptoms of dystonia begin at the age of less than 26 years, they are classified as early-onset dystonia whereas the symptoms begin after this age, they are termed late-

onset dystonia. The distribution of the affected body parts can be used. For example when one area of the body is affected it is a focal dystonia such as a writer's cramp (arm), cervical dystonia (neck, also called torticollis) or blepharospasm (eyes). When two or more continuous regions are affected it is a segmental dystonia (eg. eyes and face, head and arm), hemidystonia when only one side of the body is affected, or generalized dystonia when the legs plus other body parts are involved (Bressman *et al.*, 1989).

1.2. Genetics of dystonia

Over the past decade, it has been recognized that genetics plays a role in primary dystonia (idiopathic). As more and more loci for various categories of primary dystonia, secondary dystonia and paroxysmal dystonia have been identified, a new molecular classification of dystonia has been widely recognized. Currently 15 different loci (*DYT1* to *DYT15*) have been linked to various forms of dystonia and mutations have been found in six genes: *DYT*-1, -3, -5, -8 , -11 and -12 as shown in Table 1-1 (Bressman, 2004; de Carvalho Aguiar *et al.*, 2004; de Carvalho Aguiar and Ozelius, 2002; Grimes *et al.*, 2002; Grotzsch *et al.*, 2002; Rainier *et al.*, 2004).

Table 1-1. Genetic classification of dystonia. AD, autosomal dominant; AR, autosomal recessive; OMIM, On line Mendelian Inheritance in Men; PTD, primary torsion dystonia.

Locus	Designation	Mode of inheritance	OMIM number	Chromosome location	Protein	Clinical characteristics
DYT1/TOR1A	Dystonia1, early-onset PTD	AD	12810	9q34	TorsinA	Early-onset PTD starts in a limb in most cases, and spreads to other body parts
DYT2	Dystonia 2, autosomal recessive	AR	224500	Unknown	Unknown	Early-onset, generalized, or segmental PTD
DYT3	Dystonia3, X-linked dystonia-parkinsonism	X-linked	314250	Xq13.1	Unknown	Segmental or generalized dystonia with concurrent or subsequent parkinsonism in about 50% of the cases
DYT4	Dystonia 4, "non-DYT1" PTD	AD	128101	Unknown	Unknown	Whispering dystonia
DYT5/GCH1	Dystonia 5, dopa-responsive dystonia,	AD	128230	14q22.1-22.2	GCH1	Dystonia with concurrent or subsequent parkinsonism, diurnal worsening of symptoms, and a dramatic response to levodopa
DYT6	Dystonia 6, adolescent-onset	AD	602629	8p21-q22	Unknown	Adolescent-onset, mostly segmental PTD, rarely generalizes
DYT7	Dystonia 7, adult-onset	AD	602124	18p11.3	Unknown	Adult-onset focal dystonia (torticollis, writer's cramp, focal PTD dysphonia, or blepharospasm)
DYT8/PNKD	Dystonia 8, paroxysmal	AD	118800	2q25–q33	MR-1	Attacks of dystonia/choreoathetosis, precipitated by stress, fatigue, alcohol, and chocolate
DYT9/ CSE	Dystonia 9, paroxysmal choreoathetosis with episodic ataxia and spasticity	AD	601042	1p13.3–p21	Unknown	Attacks of dystonia, parasthesias, double vision, precipitated by exercise, stress, and alcohol
DYT10	Dystonia 10, paroxysmal choreoathetosis	AD	128200	16p11.2–q12.1 16q13–q22.1	Unknown Unknown	Attacks of dystonia/choreoathetosis, brought on by sudden movements
DYT11	Dystonia 11, myoclonus- dystonia	AD	159900	7q21–q31	SGCE	Rapid, jerk-like movements, responsive to alcohol; combined with variable degrees of dystonia
DYT12	Dystonia 12, rapid-onset dystonia-	AD	128235	19q13	ATP1A3	Acute or subacute onset of (generalised) dystonia in combination with parkinsonism

	parkinsonism					
DYT13	Dystonia 13, Early and late-onset focal Focal or segmental dystonia	AD	607671	1p36.13	Unknown	Focal or segmental dystonia with onset either in the crania-cervical region in the or in the upper limbs; mild course
DYT14	Dystonia 14, Dopa-responsive, dystonia.	AD	607195	14q13	Unknown	Dopa-responsive dystonia
DYT15	Dystonia 15, myoclonus-dystonia	AD	607488	18p11	Unknown	Rapid, jerk-like movements, responsive to alcohol; combined with variable degrees of dystonia

1.2.1. *Autosomal dominant, early-onset dystonia (DYT1* or *TOR1A)*

The *DYT1* gene, which codes for the protein torsin A, maps to 9q34 and causes most cases of child-onset dystonia in both Jewish and non-Jewish families (Kramer *et al.*, 1990; Kramer *et al.*, 1994; Ozelius *et al.*, 1992) It is the most common and severe form of hereditary dystonia (Kruger *et al.*, 2003). Symptoms usually begin in an arm or leg in the first decade and the highest prevalence occurs in Ashkenazi-Jewish populations from Eastern Europe (Korczyn *et al.*, 1980). *DYT1* was shown to be inherited in an autosomal dominant mode with 30-40% penetrance (Bressman *et al.*, 1989). Linkage disequilibrium analysis among 174 Ashkenazi Jewish affected individuals with idiopathic torsion dystonia suggested this high incidence was caused by the founder mutation, which was introduced into the Ashkenazi-Jewish population about 20-30 generations ago (Bressman *et al.*, 1994). Haplotype analysis involving 68 Ashkenazi Jewish families with early onset torsion dystonia was responsible for mapping the gene within a 150 kb region between D9S2161 and D9S63 on chromosome 9q34 (Ozelius *et al.*, 1997a; Ozelius *et al.*, 1992). Mutation analysis of the five candidate genes in this region identified a GAG deletion from the carboxyl terminal of torsin A (Ozelius *et al.*, 1997b). Subsequently this GAG deletion had been found in families of diverse origins (Ikeuchi *et al.*, 1999; Im *et al.*, 2004; Lebre *et al.*, 1999; Major *et al.*, 2001).

Torsin A is a widely expressed member of the AAA protein family (ATPase associated with different cellular activities). These AAA proteins use ATP binding and hydrolysis to induce conformational changes in substrate molecules via a repetitive cycle of substrate binding and release. The in-frame GAG deletion (ΔE302/3) in exon 5 of *DYT1* (*TOR1A* gene) was found to re-localize the protein from the endoplasmic reticulum (ER) to

the nuclear envelope (NE) in both patient tissue and in neurons from ΔE302/3 –torsinA transgenic mice (Goodchild and Dauer, 2004). As the GAG deletion in *DYT1* accounts for most of the Jewish and non-Jewish patients, genetic testing for *DYT1* has been developed and is commercially available (Klein *et al.*, 1999b).

1.2.2. *X-linked dystonia-parkinsonism syndrome (DYT3)*

DYT3 is an adult-onset, severe and progressive X-linked recessive dystonia parkinsonism (XDP). It is characterized by combined dystonia and parkinsonism in 50% of the cases (Graeber *et al.*, 1992). XDP originated from a genetic founder effect in the Philippine island of Panay and subsequently spread across the major Philippine islands. This disorder has also been diagnosed in Canada and in the United States. To date, however, XDP is known only in patients of Filipino descent, supporting genetic homogeneity of the disease (Kupke *et al.*, 1990b). *DYT3* locus was mapped to the proximal long arm of the X chromosome (Kupke *et al.*, 1992; Kupke *et al.*, 1990a). A 48-bp deletion has been identified within a novel multiple transcript system which is composed of at least 16 exons spanning more than 109 kb of genomic DNA (Nolte *et al.*, 2003). This novel gene is called DSC (disease-specific single nucleotide change) which contains at least three independent transcription start sites coding for four distinct transcripts (transcripts1-4). Both transcripts 1 and 2 initiate at exon 1 of the novel gene and include the distal exons 36 and 37 of *TAF1* (TATA-box-binding protein-associated factor) gene while transcript 3 initiates transcription from exon 2. The transcript 4 initiates transcription from exon 3 of this novel *DYT3* gene.

1.2.3. *Dopa-responsive dystonia (DYT5)*

DYT5 is an autosomal-dominant inherited L-dopa-responsive dystonia (DRD) which is caused by a genetic mutation in the gene that encodes guanosine 5-triphosphate

cyclohydrolase I (*GCH1* gene). This disorder is also called autosomal guanosine triphosphate cyclohydrolase (GCH-I) deficiency (Segawa disease, MIM 128230). DRD is characterized by childhood-onset dystonia with marked diurnal fluctuation, aggravation of symptoms toward evening and alleviation in the morning after sleep, and a dramatic response to low doses of levodopa (L-DOPA) (Ichinose *et al.*, 2000; Segawa *et al.*, 1976). Linkage analyses in *DYT5* families had mapped the disease locus to chromosome 14 (14q11-q24.3) (Nygaard *et al.*, 1993). Subsequently, analysis of candidate genes in patients revealed mutations in *GCH1* located in 14q22.1-q22.2 as the cause for DRD in three of four families examined and in one sporadic case (Ichinose *et al.*, 1994). GCH-1 is a rate-limiting enzyme for the biosynthesis of tetrahydrobiopterin (BH4). BH4 is an essential cofactor for the three amino acid monooxygenases, phenylalanine-, tyrosine-, and tryptophan hydroxylase. Phenylalanine hydroxylase catalyzes the conversion of phenylalanine to tyrosine. Tyrosine hydroxylase (TH) is required for the synthesis of L-dopa and dopamine and for the subsequent production of adrenaline and noradrenaline whereas tryptophan hydroxylase is essential for the synthesis of serotonin. Because TH is the rate-limiting enzyme in dopamine synthesis (Nagatsu *et al.*, 1964), a deficiency of BH4 can lead to decreased dopamine production and symptoms of dopamine deficiency. Reduced activity of GCH-1 in DRD patients is thought to cause symptoms by depletion of dopamine. This assumption is consistent with the observed pronounced therapeutic effect of L-dopa, even in patients who had been severely incapacitated for years (Ichinose *et al.*, 2000; Segawa *et al.*, 2004). More than 85 different mutations, including large deletions, have been found in the coding region, the introns and the 5'-untranslated region (5'-UTR) of the *GCH-1* gene as

well as numerous mutations in the introns and 5`-regulatory region (Bandmann *et al.*, 1998; Furukawa, 2003; Furukawa *et al.*, 2000; Tassin *et al.*, 2000).

1.2.4. *Autosomal dominant focal dystonia (DYT7)*

A large German family was reported with primary late-onset focal cervical dystonia in addition to the mild cranial and arm involvement. In thirteen members of this family, seven were definitely affected, six were possibly affected and three deceased individuals were obligate carriers. The *DYT7* locus was mapped to a 30-cM region on chromosome 18p between the centromeric flanking marker D18S1153 and the telomere of 18p (Leube *et al.*, 1996). The presumptive disease haplotype (*D18S1153*-18pter) was shared among all definitely affected family members; all obligate carriers and five out of six possibly affected family members. A dystonia gene located on this locus was also supported by the identification of an 18p deletion in a family with dystonia (Klein *et al.*, 1999c). The deletion was localized between the telomere of 18p and D18S1104 and spanned 49.6 cM. Although this finding did not refine the portion of the *DYT7* locus, the dystonia in these patients was suggested to be caused by haploinsufficiency of the *DYT7* gene. Coincidently, our lab had been studying a large five generation Canadian family with myoclonus-dystonia and identified a 17 cM region between D18S1132 and D18S843 on 18p11, that was in the same region which would further supporting the existence of a new gene for dystonia on 18p (Grimes *et al.*, 2002). However, clinically our family manifests with an earlier onset and predominant myoclonus which is the same as that of families with *SGCE* mutations (*DYT11*), and different from that of families with *DYT7* (Grimes *et al.*, 2001).

1.2.5. Paroxysmal non-kinesigenic dyskinesia (PNKD or DYT8)

PKND is a rare hyperkinetic movement disorder characterized by episodic attacks of dystonia, chorea and athetosis. The attacks often begin in childhood and can last for a few minutes or up to several hours (Demirkiran and Jankovic, 1995; Richards and Barnett, 1968). Most of the families seem to inherit it as an autosomal dominant trait and the disease gene was eventually mapped to 2q35 (Chen *et al.*, 2005; Fink *et al.*, 1996; Fouad *et al.*, 1996; Lee *et al.*, 2004; Matsuo *et al.*, 1999; Rainier *et al.*, 2004). Recently, sequence analysis identified different missense mutations in the myofibrillogenesis regulator 1 gene (*MR-1*) as the causative gene for PNKD (Chen *et al.*, 2005; Lee *et al.*, 2004; Rainier *et al.*, 2004).

1.2.6. *Inherited myoclonus-dystonia or myoclonus-dystonia (DYT11)*

Mutations in ε-sarcoglycan gene (*SGCE*) were identified for the *DYT11* locus on chromosome 7q21 (Zimprich *et al.*, 2001). Patients with *SGCE* mutations are affected with a variant of dystonia called inherited myoclonus-dystonia (IMD) or myoclonus-dystonia (M-D). More detailed description on this locus will be discussed in section 1.4.

1.2.7. Rapid-onset dystonia-parkinsonism (RPD, DYT12)

RDP (OMIM 128235) is a rare movement disorder with autosomal dominant inheritance and characterized by sudden onset of dystonic spasms and parkinsonism (Dobyns *et al.*, 1993). Sudden onset or worsening of symptoms occurs after stresses such as high fever, exercise or emotional stress. Most affected individuals have limb and cranial dystonia accompanied by bradykinesia, slow gait and postural instability (Pittock *et al.*, 2000). The disease gene was originally mapped to a 8 cM region on chromosome19q13 (Kramer *et al.*, 1999) which was then refined to a 5.9 cM region (Kamm *et al.*, 2004). Mutations in the *DYT12* gene have recently been identified in the gene for the Na+/K+-

ATPase α3 subunit (ATP1A3) in seven families with RDP (de Carvalho Aguiar *et al.*, 2004). Functional analysis of the protein indicated the mutations impaired enzyme activity or stability.

1.3. Phenotypic characteristics of myoclonus-dystonia (M-D)

Myoclonus-dystonia (M-D) is an autosomal dominant neuromuscular movement disorder characterized, in addition to dystonia, by a predominance of alcohol responsive myoclonic jerks mostly in the arms and axial muscles (Asmus and Gasser, 2004). Onset of this disorder is usually in the first or second decade, however a few cases were described with the onset age of as late as 38 years old (Gasser, 1998; Quinn, 1996). The diagnostic criteria were first reported by Mahloudji and Pikielny in 1967 to describe a condition called essential myoclonus (Mahloudji and Pikielny, 1967). These initial criteria were: onset of myoclonus in the first or second decade; males and females equally affected; a benign course, often variable, but compatible with an active life of normal span; dominant mode of inheritance with variable severity; absence of seizures, dementia, gross ataxia, and other neurologic deficits; normal electroencephalogram. Families with these clinical features had been described as having essential myoclonus (Przuntek and Muhr, 1983), however it was recognized even then that many of the individuals had not only the myoclonus but also dystonia to varying degrees (Kyllerman *et al.*, 1990; Lang, 1997; Quinn *et al.*, 1988). Many different names were used in the literature causing significant confusion with the condition. At the Third International Dystonia Symposium a broader, more neutral term, "inherited myoclonus-dystonia syndrome" was suggested to designate an autosomal dominant disorder characterized by myoclonus or dystonia or a combination of both without other signs of neurologic dysfunction (Gasser, 1998). The clinical spectrum of M-D can be investigated in

detail, based on the analysis of patients from families with proven *SGCE* mutations. The predominant symptom of most patients is the myoclonic jerks which affect mostly axial muscles (neck and trunk) but also muscles of upper and lower extremities, with proximal muscles being more affected than distal ones. Jerks are usually very brief, lighting-like, and are aggravated by action and psychological stress. In some circumstances, jerks may also be present at rest. In about two thirds of patients, more sustained dystonic movements are observed with torticollis and writer's cramp being the most common manifestations. Occasionally, the dystonia is seen in the lower limbs leading to dystonic gait disturbances. Dystonia may also be the predominant clinical symptom of M-D but in rare young patients. The detailed clinical characteristics and diagnostic criteria of M-D are described here (Asmus and Gasser, 2004): brief, "lightning-like" myoclonus as primary feature with or without focal or segmental dystonia of subtle to marked severity may be also seen but is rarely sole feature; autosomal-dominant inheritance with reduced penetrance and variable expressivity; in SGCE-mutation positive cases suppression of phenotype upon maternal transmission or "pseudo-sporadic" inheritance; onset usually in childhood; exclusion of additional neurological features, such as cerebellar ataxia, spasticity, dementia, and seizures; no pathological abnormalities in brain imaging, normal EEG, and somatosensory evoked potentials; and usually a benign clinical course without continuous progression of symptoms and a normal life expectancy. The optional diagnostic criteria are a positive response of symptoms to alcohol and various personality disorders and psychiatric disturbances.

Following the identification of mutations in the *SGCE* gene (*DYT11*) in six families with M-D (Zimprich *et al.*, 2001), Asmus *et al* evaluated 24 affected patients from an

additional nine pedigrees with *SGCE* mutations and summarized these typical manifestations of M-D families with *SGCE* mutations (Asmus *et al.*, 2002). The mean age at onset for myoclonus was 5.4 years (0.5-20) in 23 cases with myoclonus. Myoclonus predominantly affected neck, arms and trunk. The mean age at onset of dystonia was 8.8 years (1–38 years) in 13 cases with dystonia, often occurring in parallel with worsening of myoclonus and therefore most likely reflecting progression of M-D. Dystonia presented mostly in a focal distribution, such as cervical dystonia and/or writer's cramp. Myoclonus was improved by alcohol ingestion in 21 cases. Notably this effect is dose dependent with improvement in some cases only after heavy drinking. Five affected individuals from three families had a history of psychological symptoms such as panic attacks, depression, and agoraphobia. No other neurological abnormalities were found. Electrophysiological studies of seven affected patients did not show any abnormality.

1.4. Myoclonus-dystonia and mutations in the *SGCE* gene

1.4.1. Linkage analysis localized the DYT11 locus to chromosome 7q21

Gasser and colleagues performed two-point and multipoint linkage analyses in a large German family with alcohol responsive myoclonic dystonia and proposed the existence of a novel locus for M-D by excluding the reported dystonia loci (Gasser *et al.*, 1996). Eleven affected members and six obligate carriers were identified in this five generation family. They excluded *DYT1* (chromosome 9q34), the 13 genes for subunits of gamma-aminobutyric acid A receptor (GABA-A receptor), as well as the alpha subunit of the glycine receptor (chromosome 5q31). Therefore, they proposed a novel locus for M-D. This locus was first supported in 1999 when Nygaard *et al* characterized a large family with essential familial myoclonus-dystonia and mapped the locus to a 28-cM region of

chromosome 7q21-q31 (DYT11) (Nygaard *et al.*, 1999). Subsequently, Klein *et al* evaluated eight families with myoclonus-dystonia and found that all eight families were linked to chromosome 7, with the maximum combined two-point LOD score (Zmax) at D7S646 of 8.96. Using the multipoint linkage analysis with five markers (D7S2212, D7S2410, D7S646, D7S1820 and D7S821) spanning this region, they showed a maximum combined multipoint LOD score of 11.71 at D7S646. Critical recombination events in some key affected individuals of the eight families defined the candidate gene region to a 14-cM interval between D7S2212 and D7S821 (Klein *et al.*, 2000) and further reduced the initially reported region of 28 cM (Nygaard *et al.*, 1999). However, within the linked region, no haplotype was shared across the families, excluding a common founder mutation for these eight families. Asmus *et al* evaluated four German families with M-D and found linkage to the reported locus on chromosome 7q21-q31 with the combined multiple maximum LOD score of 5.99 (Asmus *et al.*, 2001). In one of the four families, cross over events were identified in the proximal region between markers D7S652 and D7S1489 and in the distal region between markers D7S476 and D7S2480. The candidate region was thus refined to a 7.2 cM interval between markers D7S652 and D7S2480.

1.4.2. Identified mutations in the ε-sarcoglycan (SGCE) gene for M-D (DYT11)

Following the localization and refined mapping of the *DYT11* locus on 7q21 for M-D, Zimprich *et al* first identified mutations in the *SGCE* gene in six M-D families (Zimprich *et al.*, 2001). They screened more than 10 genes by PCR and direct sequencing and identified five different heterozygous loss-of-function mutations in the *SGCE* gene in six families with M-D (MD2, 6, 8, 9, 7 and 10). Families MD7 and MD10 have the same R102X mutations (304C→ T), while family MD6 has the R97X mutation (289C→T).

Family MD2 showed a deletion (565delA) leading to a frameshift of the coding region, with a premature stop at codon 169. They also identified a splice-site mutation at the 3′ exon–intron junction of exon 6 (907+1G→A) in family MD9 and a 97-bp deletion, affecting intron 3 and 15 bp of exon 4 in family MD8 (488–97del). Subsequently, several other groups including our lab reported additional mutations in *SGCE* gene for M-D (Han *et al.*, 2003; Hjermind *et al.*, 2003; Klein *et al.*, 2002). Asmus and colleagues reported seven mutations in *SGCE* gene for nine M-D families, of which three families have the same mutation (R102X) as previously reported (Zimprich *et al.*, 2001). They also found six novel point mutations and small deletions in exons 4, 5, 6, and 7 (Asmus *et al.*, 2002). Both of these point mutations and small deletions are predicted to produce a truncated protein with a premature stop code. Since mutations in the *SGCE* gene were reported to cause M-D in 2001, more than 30 different mutations have been identified in this gene by different research groups (Asmus and Gasser, 2004; Asmus *et al.*, 2005; Han *et al.*, 2003; Schule *et al.*, 2004a; Schule *et al.*, 2004b; Tezenas du Montcel *et al.*, 2006). The mutation spectrum of the *SGCE* gene includes nonsense mutations, missense mutations, deletions, insertions and splicing mutations. Most of these mutations are loss-of-function mutations which produce short truncated proteins. The identified mutations in the *SGCE* gene are summarized in Table 1-2.

Once mutations in *SGCE* were identified it was questioned whether other neurological conditions that had similar clinical symptoms could also be caused by mutations in *SGCE*. Tezenas du Montcel *et al* screened the *SGCE* gene for mutations in 76 French patients with M-D, essential myoclonus, primary myoclonic dystonia, generalized dystonia, dystonia with tremor and benign hereditary chorea. Eleven of the *SGCE* mutation

positive cases had M-D and five had E-M but no *SGCE* mutations were detected in patients with other phenotypes (Tezenas du Montcel *et al.*, 2006).

Table 1-2. Characteristics of patients with *SGCE* mutations. N = neck, A = arm, H = hand, L = leg, T = trunk, F = face, S = shoulders, V = voice, G = generalized, NK = not known
*: range of age at onset. When only one value is given, it is the mean age at onset in the family.

Nucleotide or amino acid change	Exons of *SGCE* gene	Affected (n)	Origin	Age at onset (year)*	Myoclonus (distribution)	Dystonia (distribution)	Alcohol sensitivity	References
Truncating mutations								
p. Arg97X	exon 3	3	German	3.5	NK	H, N	NK	Zimprich, et al, 2001
p. Arg97X	exon 3	1	NK	0.3	A, L, N	N	Yes	Valente, et al, 2005
p. Arg97X	exon 3	1	NK	3	A, N	A, N	No	Valente, et al, 2005
p. Arg102X	exon 3	7	German	4	Nnk	H, N	NK	Zimprich, et al, 2001
p. Arg102X	exon 3	8	German	8	NK	N	NK	Zimprich, et al, 2001
p. Arg102X	exon 3	5	German	5-18	A, N, T	H, N	Yes	Asmus, et al, 2002
p. Arg102X	exon 3	3	German	2-6	A, L, N, T	H, L	Yes	Asmus, et al, 2002
p. Arg102X	exon 3	3	French	0.5-6	A, F, N, T	A, H, N	Yes	Asmus, et al, 2002
p. Arg102X	exon 3	13	Canadian	Childhood	A, F, N, T	N	Yes	Han, et al, 2003
p. Arg102X	exon 3	2	Canadian	1-14	A, N	none	NK	Han, et al, 2003
p. Arg102X	exon 3	1	German	7	A	H	Yes	Hedrich, et al, 2004
p. Arg102X	exon 3	1	German	13	A, N	segmental	NK	Hedrich, et al, 2004
p. Gln286X	exon 7	3	French	Childhood	A, N, T	A, L, T	NK	Asmus, et al, 2002
p. Arg372X	exon 9	1	NK	3	A, L, N, T	A, L, N	Yes	Valente, et al, 2005
p. Arg372X	exon 9	1	NK	11	A, L, N	A, L	Yes	Valente, et al, 2005
p. Arg372X	eon 9	3		Adolescence	F, A, T		NK	Tezenas du Montcel, 2005
p. Trp100X	exon 3	3		2	A, L	L	NK	Tezenas du Montcel, 2005
p. Glu70X	exon 2	3		6	N, A	N, A, T	Yes	Tezenas du Montcel, 2005
Missense mutations								
p. His60Pro	exon 2	1	German	7	A, F, T	N	Yes	Hedrich, et al, 2004
p. His60Pro	exon 2	5	Serbian	7-15	A, H, L	T, H, L;	Yes	Schule, et al, 2004
p. Leu196Arg	exon 5	2	German	2-10	A, F, L, V	A, H	Yes	Klein, et al, 2002
p. Cys271Tyr	exon 6	1		1.5	F, A, L, T	A, L	NK	Tezenas du Montcel, 2005
p. Tyr115Cys	exon 3	8		1-12	N, A, T	N, F, A	NK	Tezenas du Montcel, 2005
p. Met92Thr	exon 3	1		1.75	L	A, L	NK	Tezenas du Montcel, 2005
Splicing mutations								
c. 233-1G>A	intron 2	3	French	3-10	A, N	G	Yes	Asmus, et al, 2002

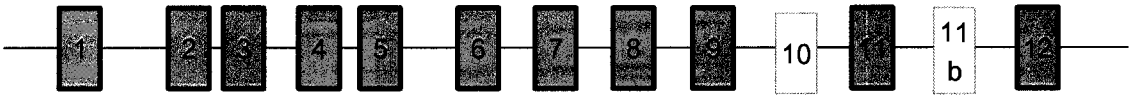
c.391-3T>C	intron 3	1	NK	30	A, N	none	NK	Valente, et al, 2005
c. 463+6T>C	intron 4	1	UK	3	A, N T	A, L	Yes	Asmus, et al, 2002
c. 662+5G>A	intron 5	3	German	1.5-12	N, A or G	H, N	Yes	Asmus, et al, 2002
c. 825+1G>A	intron 6	4	German	4	NK	N	Yes	Zimprich, et al, 2001
c. 232+2T>C	exon 2	5		4-24	N, A, T	NK	NK	Tezenas du Montcel, 2005
c. 233-1G>T	exon 3	3		4	N, A	A, T	NK	Tezenas du Montcel, 2005
c.233+1G>A	exon 2	1		6	N, A	A		
Deletions								
c. 164delG	exon 2	5	Welsh/Czech	<6	A, F, N	N	Yes	Hedrich, et al, 2004
c. 276delG	exon 3	1	German	4	A, L, N	none	Yes	Asmus, et al, 2002
c. 391_405del15	exon 4	4	German	7	NK	N	NK	Zimprich, et al, 2001
c. 402-405del4	exon 4	6	French	2-30	NK	A, H, L	Yes	Marechal, et al, 2003
c. 483del	exon 5	11	German	4-12	NK	H, N	Yes	Zimprich, et al, 2001
c. 564_576del3	exon 5	1	NK	2	A, N	A, N	Partial	Asmus, et al, 2002
c. 795delA	exon 6	1	Canadian	14	N, T, A, H	A, H	NK	Han, et al, *unpublished*
c. 832_836del5	exon 7	6	Canadian	Childhood	A, F, L, N	A	Yes	Han, et al, 2003
c. 835_839del5	exon 7	8	German/English	2-16	A, F, H, L, T	A, L, N, T	NK	Klein, et al, 2002
c. 966delT	exon 7	2	German	4-9.5	NK	NK	NK	Muller, et al, 2002
c. 966delT	exon 7	6	Serbian	4-15	A, H, L	T, L,	NK	Schule, et al, 2004
c. 974delC	exon 7	9	Danish	1-4	L, N, T, V	A, H, L	Yes	Hjermind, et al, 2003
c. 832_836del5	exon 7	2		4.5	N, A, L	N, A	NK	Tezenas du Montcel, 2005
c. 444_447del4	exon 4	3		6-18	N, T	F, N, T	Yes	Tezenas du Montcel, 2005
c. 221delA	exon 2	1		3	A	N, A, L	NK	Tezenas du Montcel, 2005
Insertions								
c. 625_626insG	exon 5	1	German	2	NK	NK	NK	Muller, et al, 2002
c. 885_886insT	exon 7	4	German	15-17	A, F, H, N	N, T	NK	Foncke, et al, 2003
c. 745_746insTGTA	exon 6	1		<6	L,	N, A	Yes	Tezenas du Montcel, 2005

1.4.3. Genomic structure of the *SGCE* gene on chromosome 7q21

Sequence analysis of *SGCE* cDNAs suggested that the *SGCE* gene is homologous to α-sarcoglycan and its genomic structure is shown in Figure 1-1. Human *SGCE* is composed of 11 common exons plus alternatively spliced exon 10 (also called exon 9a) and exon 11b. The *SGCE* gene spans 71 kb genomic region on 7q21. The sequences within exon 10 have homology to the Alu class of repetitive sequences indicating that inclusion of exon 10 in the transcripts may result from Alu-repeat induced splicing (McNally *et al.*, 1998). The primary transcript is composed of 11 exons without exons 10 and 11b in central nervous system and muscle tissues whereas the minor transcript contains 12 exons including exon 10 in muscle tissue(Figure 1-1, A and C). Recently, a new exon of mouse *Sgce*, exon 11b, was identified from the mouse brain (Nishiyama *et al.*, 2004). Nishiyama *et al* used RT-PCR to show that two major transcripts are present in the mouse brain. One form (*Sgce1*) is a conventional form with 11 exons and the other form (*Sgce2*) is a novel form including exon 11b, making the protein have 18 more amino-acid at the C-terminal compared to the conventional form. Immunoblot analysis indicated that mouse *Sgce1* is widely expressed in different tissues whereas *Sgce2* is exclusively expressed in the brain. Similarly, the same transcripts were also detected in human brain including the secondary transcript of the *SGCE* gene (Figure 1-1, B).

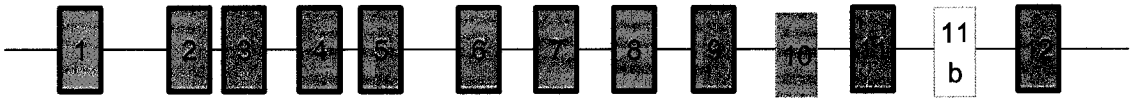
Figure 1-1. Genomic structure and alternatively spliced transcripts of the human *SGCE* gene. The human *SGCE* gene includes 11 common exons and two alternative splicing exons (exons 10 and 11b), spanning a 71 kb genomic region on 7q21. **A.** Primary transcript of *SGCE* codes for a protein of 437 amino acids expressed in human brain tissue and muscle tissue. **B.** *S*econdary transcript of *SGCE* codes for a protein of 448 amino acids expressed in human brain tissue. **C.** Minor transcript of *SGCE* codes for a protein of 462 amino acids expressed in human skeletal muscle tissue.

A. Primary transcript (Without exon 10 and 11b) of the *SGCE* gene



B. Secondary transcript (Without exon 10) of the *SGCE* gene

C. Minor transcript (Without exon 11b) of the *SGCE* gene



1.4.4. Expression patterns of the SGCE mRNA and protein

The sarcoglycans are a family of transmembrane proteins. They form a multiplex complex as a component of the dystrophin-glycoprotein complex (DGC) which was originally reported to be responsible for the integrity of skeletal and smooth muscles (Campbell, 1995) and are now known to exist in the peripheral nervous system (Imamura *et al.*, 2000). The DGC complex is composed of dystrophin, dystroglycans, sarcoglycans, syntrophins and sarcospan (Lim and Campbell, 1998). ε-sarcoglycan is the fifth member of sarcoglycan protein family which includes α-, β-, γ-, δ-sarcoglycans (SGCA, SGCB, SGCG, and SGCD). Using the rabbit α –sarcoglycan sequence to search the EST databases, Ettinger *et al* identified an *SGCE* transcript which resembles the mouse sequence of α-sarcoglycan (Ettinger *et al.*, 1997). Northern blot analysis confirmed the mRNA expression in several adult mouse tissues including brain, heart, lung, muscle, kidney, liver and E14 embryo. Subsequently, human *SGCE* was cloned and mapped to 7q21-22. Its cDNA is 1.7 kb and encodes a 437 amino acid transmembrane protein with an amino-terminal signal sequence and a transmembrane domain. The carboxyl-terminus is predicted to reside in the cytoplasm. Human *SGCE* is 62% homologous to α-sarcoglycan at the protein level (45% identical) and 47% homologous to *SGCA* at the nucleotide level (Hack *et al.*, 2000; McNally *et al.*, 1998). Unlike other sarcoglycans, which are largely restricted to muscle and heart, the *SGCE* mRNA is widely expressed in human non-muscle tissues including the central nervous system.

The mouse ε-sarcoglycan protein was predicted to begin with a 22-aa signal sequence, followed by a 262-aa extracellular domain, a single 23-aa hydrophobic transmembrane domain and a 98-aa intracellular domain. Human ε-sarcoglycan has a

similar predicted hydrophobic signal sequence followed by a substantial extracellular domain with a conserved consensus site for glycosylation as α -sarcoglycan. The glycosylation site is located within a highly conserved region of exon 5. There is a hydrophobic transmembrane domain followed by a short cytoplasmic domain. The predicted molecular mass of ε-sarcoglycan is ~45 kDa, and the pI is 5.78, close to the predicted pI (5.44) of α-sarcoglycan.

ε -sarcoglycan is expressed in many mouse tissues such as cerebrum, cerebellum, peripheral nerve, heart, lung, liver, kidney, skeletal muscle, intestine, spleen, and testis (Imamura *et al.*, 2000). Recently, the fluorescence in situ hybridization (FISH) and immunohistochemistry (IHC) analyses of the distribution patterns of *SGCE* mRNA and protein in different mouse brain sections showed a wide distribution of the *SGCE* mRNA and protein throughout the mouse brain. FISH analysis detected high levels of *SGCE* mRNA expression in substantia nigra (SN), olfactory bulb of midbrains, hippocampus, neocortex, and cerebellum (Chan *et al.*, 2005). *SGCE* polyclonal antiserum raised against a part close to the C-terminal of the human protein (amino acid 396-412) indicated protein expression of *SGCE* in the same brain regions. *SGCE* mRNA expression was most intense in cytoplasm rather than the nuclei of individual neurons. In 2002, the sixth sarcoglycan, ζ-sarcoglycan (ζ -SG) was cloned (Wheeler *et al.*, 2002) Its cDNA is 1.7 kb long and encodes for 299 amino acids. ζ-SG is shown 72% and 74% homologous to SGCG and SGCD at the protein level (56% and 57% identical to SGCG and SGCD). No mutations have been found in this gene for any human disease.

1.4.5. Mutated SGCE in M-D families is maternally imprinted

In the human genome, an allele from each parent usually contributes to the phenotype of the cell under normal conditions. However, there are about 30 genes in the human genome, which can only express the paternal allele or the maternal allele whereas the other allele is repressed because of unbalanced methylation of their promoters. If the non-repressed allele has a mutation to cause a disease, the disease will be inherited in an autosomal dominant mode. Maternally imprinted mouse *Sgce* was reported by Piras *et al* (Piras *et al.*, 2000). They described a procedure by which androgenetic (AG) and parthenogenetic (PG) mouse embryonic fibroblast (MEF) lines stably retain the parent-of-origin pattern of imprinted gene expression. Paternally expressed genes could be expressed in both AG MEF and WT MEF while the maternally expressed genes could be expressed in both PG MEFs and WT MEFs. *Sgce* was shown to be expressed in AG and WT MEFs but not in PG MEFs, suggesting paternal expression of *Sgce* gene. Mouse *Sgce* was also shown to be transcribed from the paternal allele in the mouse embryo and the majority of adult tissues, further supporting the maternal imprinting of the *Sgce* gene. Grabowski and colleagues evaluated the origins of the *SGCE* mutated alleles in M-D pedigrees and found that 49 affected individuals inherited the disease allele from their father and only four from their mother. In 17 clinically asymptomatic individuals, 14 inherited the maternally mutated allele whereas 3 cases inherited the paternal mutated alleles (Grabowski *et al.*, 2003). The methylation pattern of CpG islands within the promoter region and first exon of *SGCE* in fresh blood samples from patients with M-D was studied using bisulphate genomic sequencing for the generation of methylation maps with single-base resolution. The maternal allele was shown to be methylated in peripheral blood leukocytes whereas the paternal allele is unmethylated. This imprinting mechanism is in general agreement with the

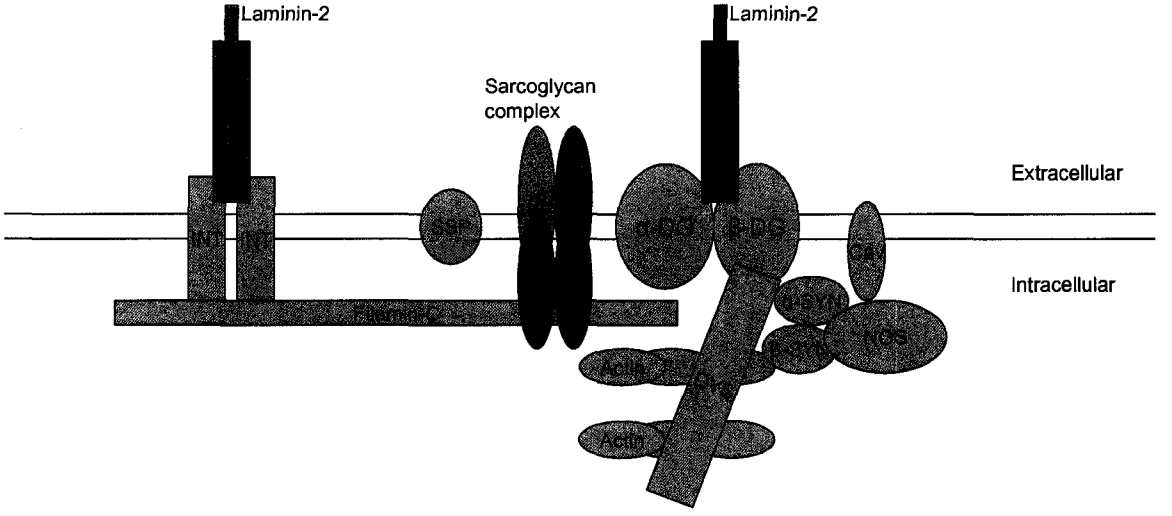
inheritance patterns of M-D families except for two affected of 29 cases, who inherited *SGCE* mutations from their mother (Grunau *et al.*, 2001). Furthermore one female patient of these two cases, who inherited the disease from the affected mother, was found to have a normal paternal allele expressed in peripheral blood leukocytes. These findings suggested incomplete maternal imprinting in this patient, which is in accordance with the incomplete imprinting of mouse *Sgce* gene in brain tissue (Piras *et al.*, 2000). In contrast to this partial expression of the *Sgce* gene, it has been reported that mouse *Sgce* was exclusively expressed in paternal origin, suggesting that maternal inherited M-D may not result from maternal expression of *SGCE*, but from other mechanisms (Yokoi *et al.*, 2005).

1.4.6. Functional studies of the *SGCE* gene and knockout mouse models

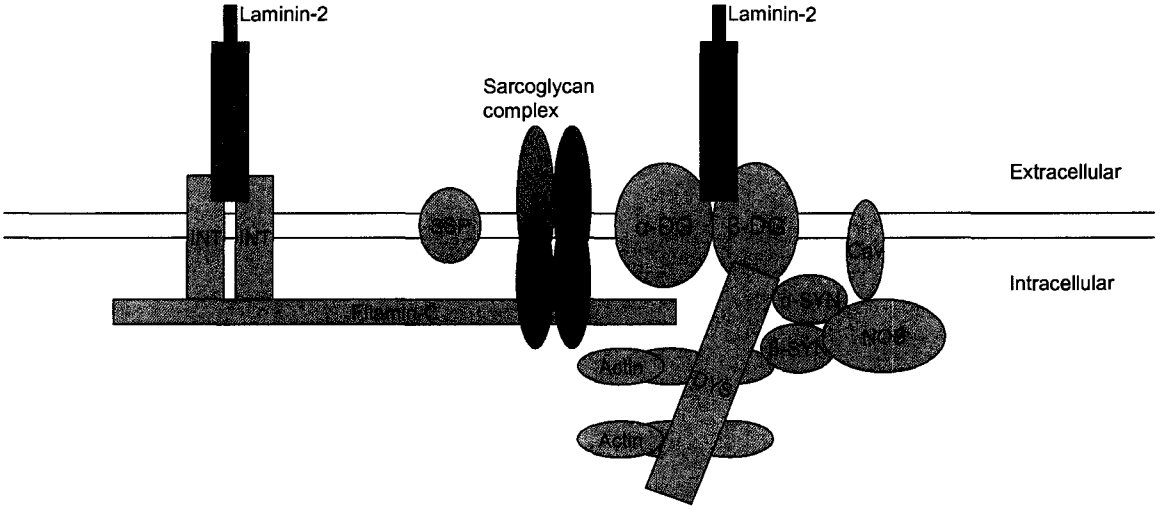
The sarcoglycan (SG) complex is a component of the dystrophin glycoprotein complex (DGC) in cell membranes. In muscle cells, the DGC is composed of dystrophin, dystroglycan, syntrophin, sarcospan, dystrobrevins and laminin-2 in addition to the sarcoglycans. Although the biological roles of the DGC are not completely understood, the DGC is thought to provide a mechanical signaling transduction between the outside of muscle cells and the intracellular cytoskeletons and stabilize the muscle cell membranes (Ozawa *et al.*, 1998). The SG complex is considered to reinforce the molecular interactions within the DGC to protect muscle cell membrane from the contraction-induced mechanical stress (Araishi *et al.*, 1999). Through the heavily glycosylated transmembrane protein-dystroglycan, dystrophin and the associated proteins are stablely anchored to laminin-2 of the extracellular matrix in Figure 1-2 (Ozawa *et al.*, 2005; Wheeler and McNally, 2003).

Figure 1-2. The dystrophin–glycoprotein complex (DGC) and the sarcoglycan subcomplex types at the cell membrane. The mechanic linkages from the muscle cell cytoskeleton to the extracellular matrix include dystroglycan bound to laminin 2 and associated proteins, and binding of integrins to laminin 2. **A.** The sarcoglycan complex interacts with dystroglycan on the extracellular surface, and with filamin C and dystrophin on its intracellular face. Dystrophin and filamin C have a number of additional binding partners. The sarcoglycan complex is composed of α-, β-, δ -, γ -sarcoglycan in skeletal and cardiac muscle. **B.** ε-sarcoglycan may substitute for α-sarcoglycan in a subset of striated muscle complexes and smooth muscle complexes. α-DG, α-dystroglycan; β-DG, β-dystroglycan; Cav, Caveolin; α-SYN, α-syntrophin; β-SYN, β-syntrophin; DYS, dystrophin; SSP, sarcospan; INT, integrins. Modified from:***Wheeler MT, McNally EM.*** *Sarcoglycans in vascular smooth and striated muscle. Trends Cardiovasc Med. 2003; 13(6):238-43 and* ***Ozawa E, et al.*** *Molecular and cell biology of the sarcoglycan complex. Muscle Nerve. 2005; 32(5):563-76.*

A.


Laminin-2
Laminin-2
Sarcoglycan complex
Extracellular
Intracellular
INT
INT
SSPG
α-DG
β-DG
Caᵛ
α-SYN
β-SYN
nNOS
Actin
Dys
Actin
Filamin-C

B.


Laminin-2
Laminin-2
Sarcoglycan complex
Extracellular
Intracellular
INT
INT
SSPG
α-DG
β-DG
Caᵛ
α-SYN
β-SYN
nNOS
Actin
Dys
Actin
Filamin-C

In humans, mutations in the genes encoding the components of DGC have been found to cause Duchenne muscular dystrophy and milder Becker's muscular dystrophy and X-linked cardiomyopathy (dystrophin gene) (Bulman *et al.*, 1991; Muntoni *et al.*, 2003), congenital muscular dystrophy (*LAMA2* gene) (Guicheney *et al.*, 1997), and limb-girdle muscular dystrophy and cardiomyopathy (sarcoglycan genes) (Bonnemann, 2005; Hack *et al.*, 2000; Van den Bergh *et al.*, 1995). The sarcoglycan complex is typically composed of at least six different sarcoglycans, α-, β-, γ-, δ-, ε -, and ζ -sarcoglycans. In general, two types of sarcoglycan complex exist in the muscle cell membranes; one containing α-, β-, γ-, δ-sarcoglycans, and the other containing β-, δ-, ζ-, and ε-sarcoglycans. In skeletal and cardiac muscle cells, α-, β-, γ-, δ-sarcoglycans form a complex while in vascular smooth muscle, the sarcoglycan complex are composed of β-, δ-, ζ-, and ε-sarcoglycans. α-sarcoglycan and ε-sarcoglycan are closely related to each other at the amino acid sequence level. However, their tissue and cell localization differs (Ettinger *et al.*, 1997; McNally *et al.*, 1998). Whereas α-sarcoglycan is only expressed in striated muscle, ε-sarcoglycan is expressed in a variety of tissues (Imamura *et al.*, 2000; Nishiyama *et al.*, 2004). In addition to the α-, β-, γ-, δ-sarcoglycan major complex, Liu *et al* used the α-sarcoglycan-deficient mice to study the localization and associations of sarcoglycans in skeletal muscles and demonstrated the existence of the ε -, β-, γ-, δ-sarcoglycan minor complex with ε-sarcoglycan replacing α-sarcoglycan in sarcoglycan complex in skeletal muscle. It was found that in C2C12 myocytes, α-sarcoglycan and ε-sarcoglycan form separate complexes with β-, γ-, δ-sarcoglycan (Liu and Engvall, 1999). However, low levels of the SGCE protein in skeletal muscle can not prevent the development of muscular dystrophy in α-sarcoglycan-deficient mice (Duclos *et al.*, 1998). A sarcoglycan complex of β-, γ-, δ-

sarcoglycan in peripheral nervous systems was found to form a single unit similar to the SG complex in muscle cells. Imamura and colleagues studied the overexpression of the *SGCE* gene as a means of compensation for α-sarcoglycan dysfunction and found that α-sarcoglycan-deficient mice with overexpression of *SGCE* did not show abnormal contraction function and muscle cell membrane damage. It was postulated that the increased *SGCE* level prevented the development of muscular dystrophy in mice (Imamura *et al.*, 2005). In addition to being a component of DGC, ε-sarcoglycan may play different roles in the CNS as mutations in *SGCE* are associated with M-D and *SGCE* is highly expressed in the nervous system.

Mouse knockout models (KO mice) have been widely used to understand the expression patterns and functions of a gene, to confirm the association between genes and diseases and to explore therapeutic strategies for inherited diseases. To further study the functions of the *SGCE* gene and the genetic mechanisms of *SGCE* mutations leading to M-D in humans, Yokoi *et al* created the first *Sgce* knockout mice by homologous recombination gene targeting to delete exon 4 of *Sgce*. Deletion of exon 4 would produce a truncated protein with premature stop in exon 5 of the *Sgce* gene. Phenotypically, these paternally inherited *Sgce* heterozygous knockout (KO) mice showed myoclonus and deficient movements of motor coordination and balance. They also evaluated the effects of the loss of ε-sarcoglycan on the mice's motor function and behavior and investigated the pathophysiological mechanism in myoclonus-dystonia (Yokoi *et al.*, 2006). They performed a series of behavioral tests and analyzed the levels of dopamine (DA), serotonin (5-HT), and their metabolites in the striatum. In their study, they have shown that the KO mice can mimic the phenotypes of M-D in that these mice develop myoclonus, deficits in motor

coordination, balance, learning, and psychiatric alterations that were consistent with anxiety and depression. Their results suggested that the typical M-D symptoms can be attributed to the effects of a single *SGCE* gene mutation. In addition, the high serotonin turnover associated with the hyperactivity of *SGCE* KO mice supports a potential treatment option of serotonin reuptake inhibitors for M-D. Our lab has made KO mice with a deletion of exon 1 of *Sgce* which will provide us more data for understanding M-D.

1.5. Genetic heterogeneity of myoclonus-dystonia

Although mutations in the *SGCE* gene are responsible for M-D, *SGCE* mutations can only be found in 20-50% of families with M-D, suggesting genetic heterogeneity. Our lab has identified a large five-generation Canadian family with M-D, which has the identical phenotype to that of families with mutations in the *SGCE* gene. Individuals in this family were examined, videotaped and had venous blood draw and DNA extracted. Originally, twelve individuals were clearly affected with M-D with a mean age of onset of 9.6 years. After excluding mutations in the *SGCE* gene, we mapped a novel locus for M-D to chromosome 18p11 with a maximum LOD score of 3.96 at GATA185C06 (D18S1363) (Grimes *et al.*, 2002). This locus was designated DYT15 (OMIM number: 607488) in accordance with other previous classifications of dystonia. A large German family with idiopathic torsion dystonia (ITD) had also been linked to chromosome 18p (*DYT7*) (Leube *et al.*, 1996). In this family, seven definitely affected, six possibly affected members and three obligate carriers were observed. Six of the seven definitely affected patients suffer from torticollis, the remaining one from typical spasmodic dysphonia. Five of the six possibly affected patients suffered from less obvious forms of torticollis while the sixth patient only presented with postural hand tremor. Symptoms of focal dystonia had been seen in all cases

over nine years of disease duration. Age of onset ranged between 28 and 70 years, with an average age of 58. After linkage to *DYT1* locus on 9q34 was excluded, genome-wide screening for linkage had detected one marker with LOD score of greater than 2.0 (D18S62). Treating the disease haplotype of D18S967/S62/S471/S458/S452 as one marker showed a maximal LOD score of 3.17 at θ=0.00 with the highest lod score of an individual marker being 2.68 for D18S452 at θ=0.00. This disease haplotype was shared in all the definitely affected members, all the obligate carriers. Three recombination events in definitely affected individuals of the family had suggested the locus between D18S843 and D18S52, which is located in the same, but larger region as our *DYT15* locus.

Genetic heterogeneity was also suggested in other studies. In a study of 6 familial and 10 sporadic families with M-D, symptoms for M-D started in the arm or neck in most cases with average onset at the age of 15 years. All patients showed a combination of myoclonus and dystonia. Eight of 10 patients who had tried alcohol showed positive response. These patients were excluded for the GAG and 18-bp deletions in the *DYT1* gene (Leung *et al.*, 2001). No mutations in *SGCE* gene were identified in any of these patients, indicating that one or more unidentified genes could be responsible for causing the M-D phenotype in some of these families (Valente *et al.*, 2003). Genetic heterogeneity of M-D was further implied in a study of 10 families. Schule *et al* identified 26 cases in seven familial M-D and 3 sporadic M-D cases. Phenotype investigations showed that the affected individuals inherited their disease in an autosomal dominant mode in six families and possibly in a seventh family. Five families showed reduced penetrance while paternal expression was consistent within two of these families. 15 of 17 (88%) affected index patients with available DNA samples were alcohol-responsive in eight of the ten families.

The average age of onset was 11.9 years (range 3–24 years) for all examined familial and sporadic patients, with 12.0 years (range 3-24) in the familial, and 11.0 years (range 5–19 years) in the sporadic cases. Clinically, all the 29 familial and sporadic patients were compatible with the diagnostic standards of M-D. Sequencing screening of the *SGCE* gene identified mutations in family 1 and 2, but no mutations were found in the other eight families. Interestingly, haplotype analysis showed that two families (family 3 and 4) shared the same haplotypes at the DYT15 locus on 18p11 as we reported previously (Schule *et al.*, 2004a).

Recently, Tezenas du Montcel and collaborators sequenced all 12 exons of the *SGCE* gene in 76 French origin patients with myoclonus-dystonia (M-D) (n=49), essential myoclonus (E-M) (n=5), primary myoclonic dystonia (n=13), as well as dystonia associated with tremor (n=5), generalized dystonia (n=3) and benign hereditary chorea (BHC) (n=1) to broaden the phenotypic spectrum. Investigations for secondary causes of the diseases were performed and a primary cause was suspected in all cases. All cases had been tested for the GAG deletion in the *DYT1* gene. Out of these 76 index patients, they found only 16 patients had mutations in the *SGCE* gene. Eleven index cases had M-D and five had E-M, two of whom had at least one relative with M-D. The phenotypes of 38 of the 60 patients without *SGCE* mutations were very similar to those of patients with *SGCE* mutations (i.e. M-D and E-M). The *SGCE* mutation carriers had significantly earlier onset, more frequent responses to alcohol and lower limb dystonia. *SGCE* mutation screening in this large cohort of patients showed that *SGCE* mutations were exclusively found in patients with typical myoclonus-dystonia or essential myoclonus, but only in a subset (30%), indicating further genetic heterogeneity of M-D (Tezenas du Montcel *et al.*, 2006).

1.6. Rationale, Hypothesis, and Aims

1.6.1. Rationale

Our current understanding and treatment of M-D and other abnormal involuntary movements remains poor as the causes of these diseases are rarely known. Genetic studies of the dystonias have resulted in a marked improvement in understanding the basic pathophysiological process of these movement disorders. The identification of mutations in *SGCE* gene for M-D and the recent generation of mouse models with *Sgce* knockout have greatly improved our understanding of M-D. However, mutations in *SGCE* gene had been reported to account for only a subset of the families with M-D (20-50%). Even including the recent identification of large deletions in the *SGCE* gene, the *SGCE* portion of the population with M-D can only account for 50% at best of cases. Our lab has collected 17 familial and sporadic families with M-D and identified a large five-generation Canadian family with M-D. With the rapid progress in genetic approaches and technologies, different groups across the world have also concluded that one or more novel loci exist for M-D families without *SGCE* mutations.

1.6.2. Hypothesis

We have collected DNA samples from 17 familial and sporadic families with M-D. We hypothesized that the *SGCE* gene or other genes are responsible for some of these families and we would use direct sequencing, gene dosage analysis and positional cloning approaches to identify these disease-causing genes.

1.6.3. Aims

In order to identify mutations in the *SGCE* gene and other novel genes for families with M-D in this research project, we would attain the following specific aims:

1) To screen all 17 M-D families for point mutations or small gene deletions/insertions in the *SGCE* gene by direct sequencing. As most mutations occur in exons of the genes for autosomal dominant diseases and the introns of the genes are usually too big to sequence, we would directly sequence the 12 exons of *SGCE* and the exon-intron boundaries in these 17 families.

2) As direct sequencing could miss the large deletions or mutations in introns of the *SGCE* gene, we will screen each exon of the *SGCE* gene for dosage alterations for those families who were mutation negative as a result of PCR and direct sequencing. Our laboratory has modified the existing multiplex PCR-based DHPLC procedures for quantifying gene copy numbers. We will use this method to screen the *SGCE* gene for dosage alterations.

3) We have identified a large five-generation Canadian family with M-D. The phenotypes of affected individuals in this family are indistinguishable from that of the families with *SGCE* mutations. Our lab has mapped a novel locus for M-D to a 17 cM region on 18p11 in this family. In order to identify the disease-causing gene for this family, we will use the SNP mapping and haplotype analysis to refine this locus and screen the candidate genes for mutations using direct sequencing and gene dosage approaches.

Once we identify the mutations in *SGCE* or other genes, we would investigate the functional roles of mutations in these genes and create a mouse model to explore the pathological mechanisms of mutations of the genes and the potential molecular therapies for M-D.

Chapter 2.

Subjects and General Methodology

2.1. Phenotypic characteristics of the patients in our cohort of families with M-D

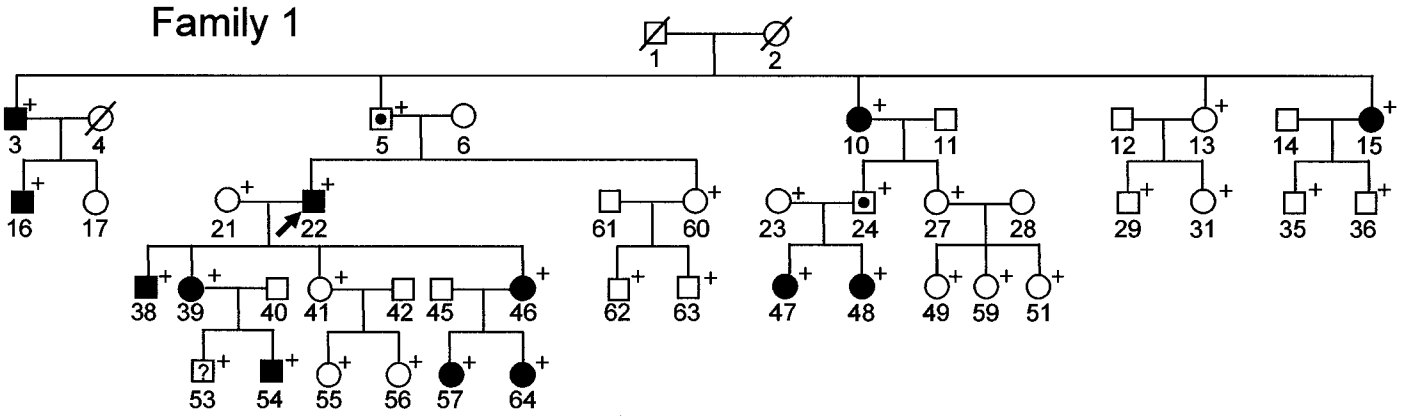
We obtained informed consent from all the participants of 17 familial and sporadic families with M-D and one family with focal dystonia (Families 1-18) after approval by the institutional review board at The Toronto Hospital and The Ottawa Hospital. Individuals from family 1 had a detailed history, physical examination, and videotape recorded using a standard questionnaire and videotape protocol. All patients and their relatives in the eighteen families were examined by neurologists specialized in movement disorders. The diagnosis of M-D was made according to the recently established criteria (Asmus and Gasser, 2004). The phenotypic characteristics of each family are listed in Table 2-1 and were reported previously (Grimes *et al.*, 2001; Grimes *et al.*, 2002; Han *et al.*, 2003). After obtaining informed consent, a blood sample was collected from all available members of 18 families and pedigrees were generated using Cyrillic software (version 2.1) and are shown in Figure 2-1.

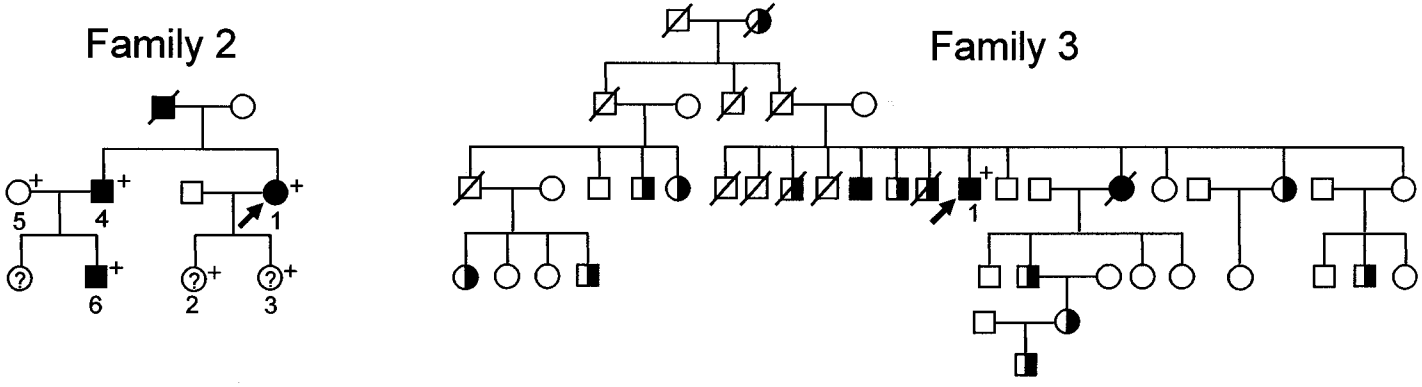
Table 2-1. Clinical characteristics of 17 families with myoclonus-dystonia and one family with focal dystonia (family 10). UE, upper extremities; LE, lower extremities; L, left; R, right.

Family/ Pedigree Number	Number of affected Individuals	Mean age of onset in years (range)	Myoclonus Location	Dystonia Location	Alcohol responsive	Unusual features	*SGCE* Mutation status
1	14	9.1 (3-15)	Mainly UE, face, trunk	UE, LE, axial	Yes		None
2	4	7.5 (1-17)	Face, neck UE, trunk	None	Yes		Exon 2 and 3 deleted
3	13	Childhood	Face, neck UE,	Neck	Yes		Exon 3 (R102X)
4	6	Childhood	Face, neck UE, LE	UE	Yes		Exon 7 (5-bp deletion)
5	2	15	Face, neck UE	None	Yes		None
6	3	2	Face, neck UE, LE	UE, LE	Unknown		Exons 2-5 deleted
7	3	7.5 (1-14)	Neck, UE	None	Unknown		exon 3 (R102X)
8	1	1	More mild Trunck, UE	Generalized dystonia	Unknown	More prominent dystonia vs myoclonus	None
9	1	5	Left arm, UE Trunk/axial	Left arm, subtle	Unknown	Kinetic Tremor L>R, UE	None
10	4			Focal Dystonia			None
11	1	15	Neck, shoulders Upper Trunk	Neck, trunk and Right UE	Unknown	Irregular postural and kinetic tremor, trunk muscles, R>L UE,	None

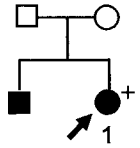
12	1	5	L>R UE	Neck	Yes	postural tremor in R LE, Dysarthria and irregular "tremor"	None
13	1	15	Both LE	no	Yes	"Tremor"	None
14	1	13	Face and UE	Neck and R arm	None	Postural and kinetic tremor of U and LE	None
15	1	14	Neck, axial>UE	UE bilateral	Unknown		exon 6 (846A deletion)
16	5	25	generalized	Eyes, Neck	Yes	FMHx of dystonia, vocal tremor	None
17	1	30	UE	None	Yes		None
18	1	36	Face, L>R UE	Both UE	Mild birth anoxia		None

Figure 2-1. The pedigrees of 17 families with myoclonus-dystonia and one family with focal dystonia. Filled symbols represent affected individuals; half-filled symbols represent those affected by history only; question mark represents possibly affected; Dot represents obligate carrier; "+" indicates those examined and who had DNA samples collected; arrows represent probands.

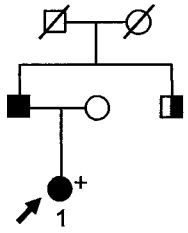

Family 1
1
2
3
4
5
6
10
11
12
13
14
15
16
17
21
22
61
60
23
24
27
28
29
31
35
36
38
39
40
41
42
45
46
62
63
47
48
49
59
51
53
54
55
56
57
64


Family 2
5
4
1
6
2
3
Family 3
1

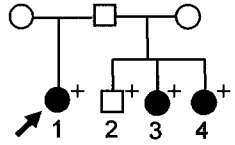
Family 4



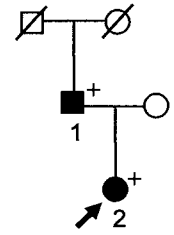
Family 5



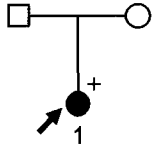
Family 6



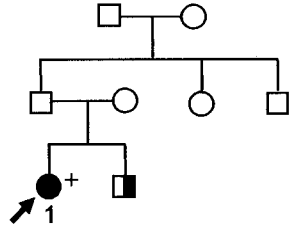
Family 7



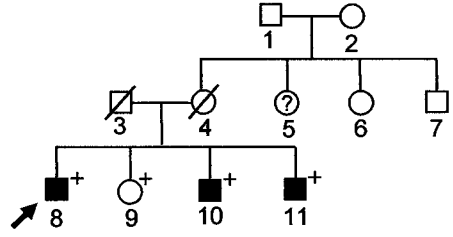
Family 8



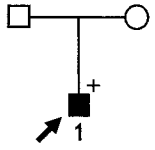
Family 9



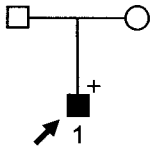
Family 10



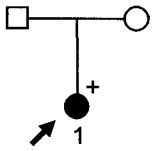
Family 11



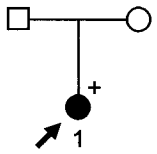
Family 12



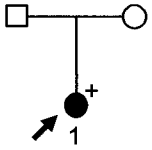
Family 13



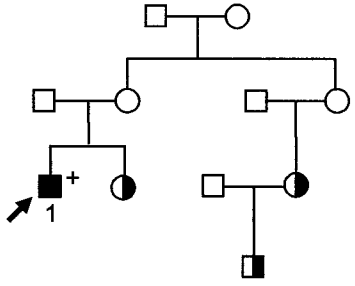
Family 14



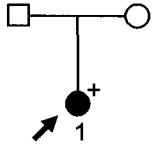
Family 15



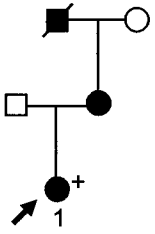
Family 16



Family 17



Family 18



2.2. Isolation of genomic DNA from whole blood

Genomic DNA was extracted from venous blood samples of available members of the eighteen families using the standard protocols of the QIAGEN kit (QIAGEN, Mississauga, Canada). In brief, 20 µl QIAGEN Protease (or Proteinase K) and 200 µl Buffer AL were added to a 200 µl whole blood sample in a 1.5 mL microcentrifuge tube. After mixing by pulse-vortexing, the tube was incubated at 56°C for 10-15 min. The 1.5 ml microcentrifuge tube was briefly centrifuged to remove drops from the inside of the lid and 200 µl of ethanol (96–100%) was added to the sample, and was mixed again by pulse-vortexing. The mixture was carefully applied to the QIAamp Spin Column in a 2 ml collection tube and was centrifuged at 6000 x *g* (8000 rpm) for 1 min. The column was placed in a clean 2 ml collection tube and the tube containing the filtrate was discarded. The QIAamp Spin Column was washed with 500 µl Buffer AW1, centrifuged at 6000 x *g* (8000 rpm) for 1 min and washed again with 500 µl Buffer AW2 at full speed (20,000 x *g*; 14,000 rpm) for 3 min. The QIAamp Spin Column was then placed in a clean 1.5 ml microcentrifuge tube, and 200 µl Buffer AE or distilled water was added to dissolve the DNA at room temperature (15–25°C) for 1 min. The sample was then centrifuged at 6000 x *g* (8000 rpm) for 1 min. DNA was stored in Buffer AE at 4° C, or at –20° C for long-term use. Using this approach, a sample of 200 µl whole human blood (~5 x 10^6 leukocytes/ml) typically can yield 6 µg of DNA in 200 µl Buffer AE or water with an A_{260}/A_{280} ratio of 1.7–1.9.

2.3. Genotyping and linkage analysis methods

2.3.1. Genotyping

A genomewide screening was performed on patients from family 1 initially using microsatellite markers with an average spacing of 25 cM between markers. Later, markers with an average spacing of 10 cM were used to fill the gaps between the 25 cM markers. The marker order and positions were obtained from the Marshfield Centre for Medical Genetics (Marshfield Medical Research Foundation). The DNA samples were genotyped using primer set 8 from MAPPAIR^{TM} microsatellite markers (Research Genetics, Huntsville). The PCR amplification was performed on PTC-225 Peltier Thermal Cycler (MJ Research, Watertown, MA) in 10 µl PCR tubes containing 2 pmol of reverse primer and M13-tailed forward primer each, 0.1 µM fluorescence-labeled IRD-700, 0.2 U Taq DNA polymerase, 0.2 mM dNTPs, and 2.5 mM $MgCl_2$. The PCR conditions were set at 94° C for 2 min, then 30 cycles of 94° C for 45 s, 55-61° C for 45 s, and 72° C for 45 s, followed with incubation at 72° C for 10 min. Following PCR amplification, products were denatured at 95° C for 5 min, separated by size on a 6% Sequengel-6 and detected on a Li-Cor automated sequencer (Lincoln, NE) and analyzed using RFLPscan software (version 3.0).

2.3.2. Linkage analysis

Two-point and multipoint linkage analyses were performed using MLINK and LINKMAP from the LINKAGE program package (version 5.1) specifying an autosomal dominant model of transmission (Lathrop *et al.*, 1985; Ott, 1999). Two-point lod scores were calculated for each marker by using an "affected-only analysis" in which only those clearly affected with M-D were considered affected. All others were coded as "unknown" for disease status. The multipoint linkage analysis was performed taking intermarker distances from the Marshfield map. The disease frequency was estimated at 0.0001 (Caviness *et al.*, 1999) with the recombination fractions assumed to be equal for men and

women. Marker allele frequencies were taken from Marshfield/Centre d'Etude du Polymorphisme if available; otherwise they were assumed to be equal. Haplotypes were established under the assumption of minimal recombination.

2.4. DNA sequencing

2.4.1. *The design of primers and optimization of PCR conditions*

We used the Primer 3 program to design oligonucleotide primers to sequence the exons, the exon-intron boundaries, 5'-untranslated region (5'-UTR) and 3'-untranslated region (3'-UTR) of a gene. To optimize the PCR conditions for each pair of primers, we usually choose three different annealing temperatures (57, 59 and 61° C) and three different concentrations (1.0, 1.5 and 2.0 mM) of $MgCl_2$ to make nine PCR combined reactions. If PCR amplification was sub-optimal, we used different annealing temperatures, or adjusted the concentrations of $MgCl_2$ or added specific reagents such as DMSO to obtain the optimized PCR conditions.

2.4.2. *Sequence analysis by P^{33}-radiolabled method*

Sequencing was performed using Thermo Sequenase Radiolabled Terminator Cycle Sequencing Kit (USB Corp., Cleveland, OH). In brief, the PCR products amplified from the whole blood genomic DNA of the patients and normal controls were treated with EXO-SAP (EXOnuclease I and Shrimp alkaline phophatase, USB Corp., Cleveland, OH, USA). Sequencing reactions were set in 20 µl reaction volumes containing 2 µl of EXO-SAP treated DNA, dNTP mix, 0.5 µl of forward/reverse primers (17 µM), 10X thermosequenase buffer, thermosequenase and P^{33}-labeled ddNTP* (radiolabeled ddGTP, ddATP, ddTTP, or ddCTP). The PCR cycling program was set for: 94° C for 2 min, then 30 cycles of 94° C for 45 s, 55-61° C for 45 s, and 72° C for 45 s, 72° C for 10 min. After PCR amplification, 4 µl

of Sequencing Stop Buffer was added to each reaction and denatured for 10 min. Three µl of the sequencing products per lane was loaded on Acrylamide Sequagel-6 gel and was electrophoresed for 2-4 hours depending on the size of the PCR products. The gels were then dried (BIO-RAD, Model583) and were exposed for 2-4 days at –80 °C.

2.4.3. *Sequence analysis using Applied Biosystems automatic sequencer*

Sequencing was performed on ABI3130x*l* Genetic Analyzer using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystem, Foster City, CA). In brief, for each exon and the intron-exon boundaries of the genes to be sequenced, the PCR products were amplified from genomic DNA from the patients and treated with EXO-SAP at 37 °C for 20 min. The sequencing reactions were set up in 10 µl reaction volumes containing 2 µl of EXO-SAP treated DNA, 4 µl of Ready Reaction premix (RR-100), 2 µl of BigDye Sequencing Buffer, 3.2 µM primer. The thermocycling conditions used on the gradient PCR equipment (MBS SATELLITE 0.2G Thermal Cycler, Thermo Electron Corporation, Milford, MA) were set: 96°C for 1 min, then 25 cycles at 96°C 10 sec, 50°C 5 sec, 60°C for 4 min, 4°C to hold. The sequencing products were purified using the modified Ethanol/EDTA precipitation protocol. To each sample, 5 µl of 125 mM EDTA and 60 µl of 100% ethanol were added to the microcentrifuge tubes and the tubes were incubated at -20°C for 15 min and then at room temperature for 15 min. Samples were centrifuged at 12000xg for 15 min and the supernatant was removed. Then, 60 µl of 70% ethanol was added to each tube and centrifuged at 12000xg for 8 min. After all liquids were removed from the tubes, the tubes were left in air to dry for 2 hr. The sequencing reactions were re-suspended in 15 µl of ddH2O for sequencing analysis using the Applied Biosystems automatic sequencer.

2.5. Mutation analysis using DHPLC (denaturing high-performance liquid chromatography)

For mutation screening of a specific fragment of a gene by DHPLC, PCR was performed on a thermal cycler at 94 °C for 2 min, then 30 cycles at 94 °C for 45 s; 55-63 °C for 45 s; 72 °C for 45s; and a final extension at 72 °C for 10 min. After DNA amplification, PCR products were denatured at 95°C for 5 min and cooled down to 25°C. DHPLC analysis of the PCR product was carried out on a Wave DNA Fragment Analysis System (Transgenomics, Omaha, NE) containing a DNASep column held at one of three oven temperatures. All analyses were performed at the following elution conditions: the acetonitrile and triethylammonium acetate were used to constitute the mobile phase. The mobile phase was formed by mixing buffer A (0.1 M triethylammonium acetate, 0.1 % acetonitrile) and buffer B (0.1 M triethylammonium acetate, 25 % acetonitrile). Each gradient was composed of 30%-60% eluent B. The PCR fragments were eluted at a flow rate of 0.9 ml/min. Each injection took 5 µl of PCR product and was run about 8-15 min depending on different lengths of PCR products. The eluted products were detected by using a UV detector set at 260 nm. After each run, the column was cleaned with 100% buffer B (0.5 min) and equilibrated (1.2 min) at the starting conditions. The melting temperature (Tm) required for each of the heteroduplex products was determined by the DHPLC melting algorithm. Variants found on DHPLC profiles were then directly sequenced to determine the mutation status.

2.6. Gene dosage analysis using DHPLC (denaturing high-performance liquid chromatography)

2.6.1. Amplification of each exon of a specific gene and internal control by multiplex PCR

To detect dosage alterations for any exon of a specific gene by DHPLC, the exons of an interested gene and an internal control gene were co-amplified in the multiplex PCR reactions. Amplification was performed in 50 μl reaction volumes containing 0.2-0.5 U of Taq DNA polymerase (5U/ul), 5 μl 10X buffer with concentration of $MgCl_2$ at 1.0, or1.5 or 2.5 mM, 2 pmol of each primer, 0.2 mM dNTPs, and 2-4 μl DNA (100-200 ng). We had determined that amplification with 25 cycles would significantly detect dosage alteration and our multiplex PCR conditions were set at 94° C for 2 min, then 25 cycles of 94° C for 45 s, 55-61° C for 45 s, and 72° C for 45 s, 72° C for 10 min. All reactions were carried out in a 96 well plate or tubes and were performed in triplicate for each individual DNA sample.

2.6.2. Quantification of PCR products using DHPLC and Wave-Maker Software

The multiplex PCR products (5 μl) were injected into the DHPLC system (Transgenomics). PCR products were eluted under non-denaturing conditions (50° C) and at a flow rate of 0.9 ml/min by a linear acetonitrile gradient. The buffer gradients were determined by the WAVEMAKER software. The PCR products were eluted at two different times based on the sizes of the products. The DHPLC profile of each individual is usually composed of two peaks with one peak representing the product to be analyzed and another peak representing the internal control.

2.6.3. Calculation of gene dosage expressed as the ratio of peak area for each exon of a gene and the internal control

Each peak area of the DHPLC profiles was read through the WAVE-MAKER software. The mean *gene/internal control* ratio for each exon of the gene was calculated

from 10 unaffected controls and was used to normalize each exon of the *gene/internal control* ratios of each patient based on the formula proposed by Naimi and colleagues (Naimi *et al.*, 2005):

$$\text{Normalized Ratio} = \frac{\text{Gene to be analyzed/internal control [Patients]}}{\text{Mean gene to be analyzed/internal control [Normal Controls]}}$$

2.6.4. Assessment of the gene dosage based on the normalized dosage ratio for each of the affected individuals

After calculating the normalized ratio for each patient, we assessed the gene dosage according to the criteria suggested by Naimi and colleagues (Naimi *et al.*, 2005): the normalized ratios below 0.75 indicate deletions; from 0.75 to 1.25 are normal; above 1.25 reflect duplications.

Chapter 3.

Screening for Mutations in the *SGCE* Gene Using Direct Sequencing

3.1. Introduction

Myoclonus-dystonia (M-D) is an autosomal dominant movement disorder characterized by alcohol responsive myoclonic jerks typically affecting the arms and axial muscles with reduced penetrance. Mutations in the epsilon-sarcoglycan (*SGCE*) gene on the chromosome 7q21 have been shown to cause M-D but not in all cases (Zimprich *et al.*, 2001). Since the loss-of-function mutations in the *SGCE* gene (*DYT11* locus) were reported to cause M-D in 2001, more than 30 different mutations in *SGCE* have been identified (Asmus and Gasser, 2004; Asmus *et al.*, 2005; Han *et al.*, 2003; Schule *et al.*, 2004a; Schule *et al.*, 2004b; Tezenas du Montcel *et al.*, 2006). The *SGCE* mutation spectrum includes nonsense, missense, and splicing mutations with most of the mutations producing non-functional truncated proteins (nonsense mutations). Currently we estimate that mutations in the *SGCE* gene account for approximately 20 to 50% of M-D. This means that in our cohort of 17 families with M-D, we would expect to find mutations in 3-8 of 17 families.

3.2. Subjects and Methods

3.2.1. *Subjects*

Informed consent was obtained from all the participants of the 18 familial and sporadic families (Families 1-18, **Chapter 2**) after approval by the institutional review board at The Toronto Hospital and The Ottawa Hospital. Individuals had a detailed history, physical examination using a standard questionnaire. All patients and their relatives in the eighteen families who had blood drawn were examined by neurologists and the clinical characteristics of these families are shown in **Chapter 2**. The diagnosis of M-D was made according to established criteria in **Chapter 1** (Asmus and Gasser, 2004). Family 10 had a

focal dystonia without myoclonus but was included in the *SGCE* screen because it mapped to a similar location as family 1.

3.2.2. Sequence analysis of the SGCE gene in 18 families

Genomic DNA was extracted from the venous blood of patients using the standard protocols of the QIAGEN kit as described in Chapter 2 (QIAGEN Inc. Mississauga, Canada). Two microlitres of DNA was used to amplify each of the 12 exons of the *SGCE* gene for the affected individuals from the 18 unrelated families. DNA sequencing was performed using the Thermo Sequenase Radiolabled Terminator Cycle Sequencing Kit (USB Corporation, Cleveland, Ohio) or the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystem, Foster City, CA) described in **Chapter 2.**

3.2.3. Mutation confirmation using allele-specific oligonucleotide amplification and sequencing

For the confirmation of 5-bp deletion in family 4, one hundred unrelated controls were screened for the five base pair deletion of *SGCE* exon 7 using allele-specific primers (left primer, 5'-TGAACCAAATGAAACTTTGCT-3', right primer, 5'-CAGGTTGATAAAGCAAGTG-3') that can specifically bind the deleted allele without binding the wild type allele. After PCR amplification of genomic DNA from the patient with 5-bp deletion and the normal individuals, the PCR products were separated by size on a 1% agarose gel to differentiate the mutant allele from the wild type allele. For the confirmation of the one base pair deletion of *SGCE* exon 6 in family 15, we directly sequenced exon 6 in one hundred normal controls using the BigDye Terminator v3.1 Cycle Sequencing Kit.

3.3. Results

3.3.1. *Mutation analysis of the SGCE gene in 18 families*

We had used the Primer 3 program to design oligonucleotide primers to amplify the previously reported 11 exons of *SGCE* gene plus an additional newly identified exon (exon 11b). The alternative spliced exon 10 is absent from the majority of transcripts, and therefore was not screened. All the primer sequences and PCR conditions for each pair of primers are listed in Table 3-1.

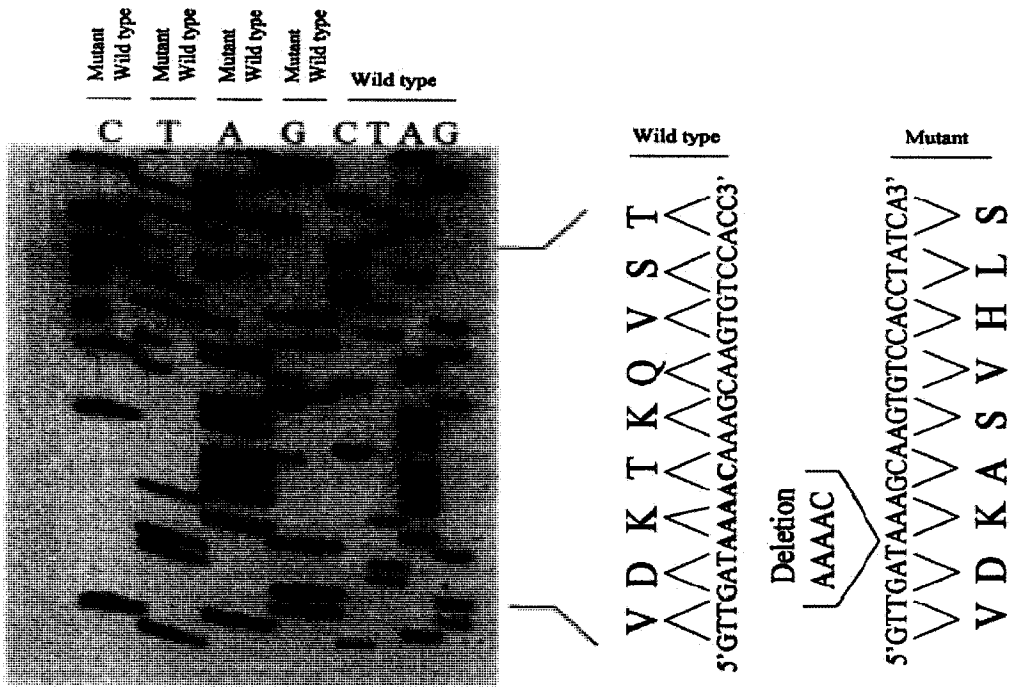
Table 3-1. Primers and PCR conditions to amplify all 12 exons of the *SGCE* gene.
F: Forward primer, R: Reverse primer, Temp: temperature.

Exons	Primers	Primer Sequences	Annealing Temp(C°)	Mg2+ mM	Amplimer Size (bp)
1	F1	5'-GTCCGGGACAGAAAGAGAGG-3'	57	1.0	357
	R1	5'-GCTAGAGGGAGTACGGGATT-3'			
2	F2	5'-TCCAGAAATATGCTTTCCTTAACA-3'	57	1.0	400
	R2	5'-GCTGAATTATCAAGGGCGTATC-3'			
3	F3	5'-CTGTGTTTCCATGCACAAAAT-3'	57	1.0	358
	R3	5'-CCAAAGCAACATGTGTGAAAA-3'			
4	F4	5'-TTTCACCATCATAACACACCAT-3'	57	2.0	417
	R4	5'-CTCATTGCCCAGAGAAGGAA-3'			
5	F5	5'-TGCAATAGGCCATCTTCCAT-3'	57	1.0	404
	R5	5'-CTTCCTGCTGCCAGGATTAT-3'			
6	F6	5'-AAACGTTAACTCCAGCCACA-3	59	1.0	302
	R6	5'-TCCTGCTTTTAAGGTGGATTG-3'			
7	F7	5'-TGAACCAAATGAAACTTTGCT-3'	57	2.0	391
	R7	5'-CAAAGAATGCTTTAGTGTATCCAG-3			
8	F8	5'-CACATGTATGGAGCATGATGG-3'	57	2.0	185
	R8	5'-TGCAAGATCCCATGAGCATA-3'			
9	F9	5'-TTTCAGGAGAGTAGGGGTGA-3'	57	1.0	411
	R9	5'-TTGATGACCCATCAGGCTAA-3'			
11	F11	5'-TGTTTTGCCTTATTTGGTGAA-3'	57	1.0	285
	R11	5'-GGGGAATTGTGCTGGATTAC-3'			
11b	F11b	5'-AATAGCGGCATTGTGGTAGG-3'	57	1.0	499
	R11b	5'-GCGAGATGGAAGCTACAAGG-3'			
12	F12	5'-TTGTGGAATGAGAATGAACACAT-3'	61	1.0	439
	R12	5'-TCCAATAATGCATGAGCTTTTC-3'			

By sequencing all 12 exons and the exon-intron boundaries of the *SGCE* gene in the affected individuals of our initial seven M-D families using the Thermo Sequenase Radiolabled Terminator Cycle Sequencing Kit, we first identified *SGCE* mutations in three families (Families 3, 4, and 7) (Han *et al.*, 2003). Family 4 had been found to carry a 5 base pair (5'-AAAAC-3') deletion in exon 7 of the *SGCE* gene (Figure 3-1, A). Sequence alignment analysis showed that this 5 bp deletion leads to a frameshift change in the *SGCE* protein sequence, resulting in a truncated protein with a premature stop at codon 294 (Figure 3-1, B). This truncated protein did not contain the transmembrane domain (Amino Acid 317-Amino Acid 339) by alignment analysis with ClustalW program (http://www.ebi.ac.uk/clustalw/).

Figure 3-1. A novel 5-bp deletion in exon 7 of the *SGCE* gene identified in family 4 with M-D. A. Partial nucleotide sequence of exon 7 of *SGCE* in the affected proband from family 4 and a normal control. The mutation is a heterozygous 5-bp deletion present in the affected patient of family 4. The five deleted nucleotides (AAAAC) are in boldface. **B.** Protein sequence alignment shows that the 5-bp deletion produced a truncated protein of 294 amino acids, which lost the transmembrane domain (red) of the wild type SGCE protein.

A.



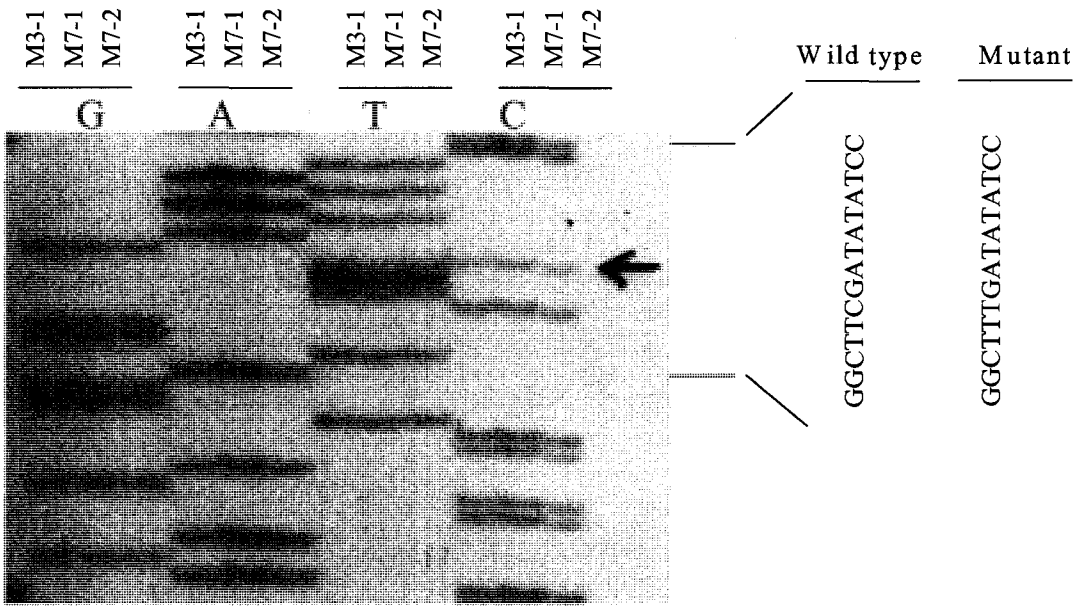
B.

```
SGCE-wild             MQLPRWWELGDPCAWTGQGRGTRRMSPATTGTFLLTVYSIFSKVHSDRNVYPSAGVLFVH 60
SGCE-5-bp-deletion    MQLPRWWELGDPCAWTGQGRGTRRMSPATTGTFLLTVYSIFSKVHSDRNVYPSAGVLFVH 60

SGCE-wild             VLEREYFKGEFPPYPKPGEISNDPITFNTNLMGYPDRPGWLRYIQRTPYSDGVLYGSPTA 120
SGCE-5-bp-deletion    VLEREYFKGEFPPYPKPGEISNDPITFNTNLMGYPDRPGWLRYIQRTPYSDGVLYGSPTA 120

SGCE-wild             ENVGKPTIIEITAYNRRTFETARHNLIINIMSAEDFPLPYQAEFFIKNMNVEEMLASEVL 180
SGCE-5-bp-deletion    ENVGKPTIIEITAYNRRTFETARHNLIINIMSAEDFPLPYQAEFFIKNMNVEEMLASEVL 180

SGCE-wild             GDFLGAVKNVWQPERLNAINITSALDRGGRVPLPINDLKEGVYVMVGADVPFSSCLREVE 240
SGCE-5-bp-deletion    GDFLGAVKNVWQPERLNAINITSALDRGGRVPLPINDLKEGVYVMVGADVPFSSCLREVE 240

SGCE-wild             NPQNQLRCSQEMEPVITCDKKFRTQFYIDWCKISLVDKTKQVSTYQEVIRGEGILPDGGE 300
SGCE-5-bp-deletion    NPQNQLRCSQEMEPVITCDKKFRTQFYIDWCKISLVDKASVHLSGSDSWRG*          294

SGCE-wild             YKPPSDSLKSRDYYTDFLITLAVPSAVALVLFLILAYIMCCRREGVEKRNMQTPDIQLVH 360

SGCE-wild             HSAIQKSTKELRDMSKNREIAWPLSTLPVFHPVTGEIIPPLHTDNYDSTNMPLMQTQQNL 420

SGCE-wild             PHQTQIPQQQTTGKWYP 437
```

Families 3 and 7 were found to have the same C→ T transition (R102X) mutation in exon 3 of *SGCE* as previously reported (Figure 3-2, A). This R102X nonsense mutation has also been predicted to produce a truncated protein of 101 amino acids (Figure 3-2, B). Alignment analysis of the protein sequence showed the R102X mutated protein lost most of the SGCE protein including the transmembrane domain (Amino Acid 317-Amino Acid 339). However, we did not find any point-mutation or small deletions in the *SGCE* gene for families 1, 2, 5, and 6.

Figure 3-2. A reported Arg102X mutation in exon 3 of the *SGCE* gene identified in two families with M-D. **A.** One affected from family 3 and two affected from family 7 have a 304C→T (Arg102 X) mutation in exon 3 of the *SGCE* gene. Black arrow indicates the C→T heterozygous mutation. **B.** Protein sequence alignment shows this mutation was predicted to produce a truncated protein of 101 amino acids (blue), and does not include the transmembrane domain (red) of wild type SGCE protein (black).

A.



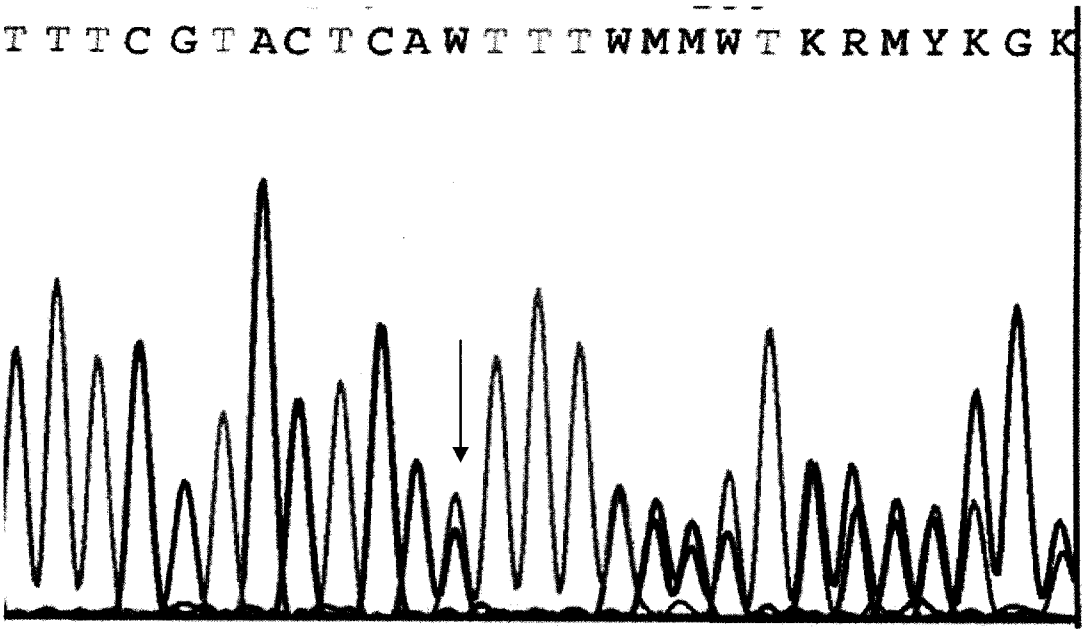
B.

```
SGCE-Wild       MQLPRWWELGDPCAWTGQGRGTRRMSPATTGTFLLTVYSIFSKVHSDRNVYPSAGVLFVH 60
SGCE-R102X      MQLPRWWELGDPCAWTGQGRGTRRMSPATTGTFLLTVYSIFSKVHSDRNVYPSAGVLFVH 60

SGCE-Wild       VLEREYFKGEFPPYPKPGEISNDPITFNTNLMGYPDRPGWLRYIQRTPYSDGVLYGSPTA 120
SGCE-R102X      VLEREYFKGEFPPYPKPGEISNDPITFNTNLMGYPDRPGW*                     101

SGCE-Wild       ENVGKPTIIEITAYNRRTFETARHNLIINIMSAEDFPLPYQAEFFIKNMNVEEMLASEVL 180
SGCE-R102X

SGCE-Wild       GDFLGAVKNVWQPERLNAINITSALDRGGRVPLPINDLKEGVYVMVGADVPFSSCLREVE 240
SGCE-R102X

SGCE-Wild       NPQNQLRCSQEMEPVITCDKKFRTQFYIDWCKISLVDKTKQVSTYQEVIRGEGILPDGGE 300
SGCE-R102X

SGCE-Wild       YKPPSDSLKSRDYYTDFLITLAVPSAVALVLFLILAYIMCCRREGVEKRNMQTPDIQLVH 360
SGCE-R102X

SGCE-Wild       HSAIQKSTKELRDMSKNREIAWPLSTLPVFHPVTGEIIPPLHTDNYDSTNMPLMQTQQNL 420
SGCE-R102X
SGCE-Wild       PHQTQIPQQQTTGKWYP 437
SGCE-R102X
```

We have been collecting additional families with M-D (families 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, and 18) and have screened these families for mutations in the *SGCE* gene. Out of these 11 families, one sporadic family (family 15) was found to have a one-base pair deletion (846delA) in exon 6 (Figure 3-3, A). This deletion was predicted to produce a truncated protein of 287 amino acids. Protein alignment shows that this truncated protein lacks the transmembrane domain (Amino Acid 317-Amino Acid 339) of the wild type SGCE (Figure 3-3, B).

Figure 3-3. A novel 1-bp deletion in exon 6 of the *SGCE* gene identified in family 15 with M-D. A. Sequencing analysis of the *SGCE* gene in family 15 showed the sequence frameshift caused by 795A deletion (black arrow). **B.** A truncated protein of 287 amino acids was produced in family 15 (blue) compared to the wild type *SGCE* sequence. The mutant premature protein was shown to lack the predicted transmembrane domain (Amino Acid 317-Amino Acid 339, red) of the wild type SGCE protein.

A.



B.

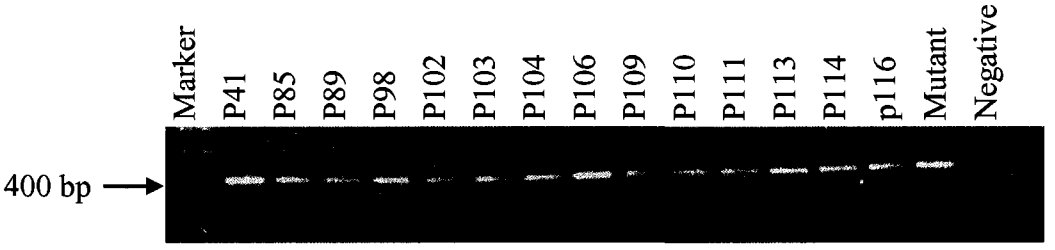
```
SGCE-wild       MQLPRWWELGDPCAWTGQGRGTRRMSPATTGTFLLTVYSIFSKVHSDRNVYPSAGVLFVH 60
SGCE-846del     MQLPRWWELGDPCAWTGQGRGTRRMSPATTGTFLLTVYSIFSKVHSDRNVYPSAGVLFVH 60

SGCE-wild       VLEREYFKGEFPPYPKPGEISNDPITFNTNLMGYPDRPGWLRYIQRTPYSDGVLYGSPTA 120
SGCE-846del     VLEREYFKGEFPPYPKPGEISNDPITFNTNLMGYPDRPGWLRYIQRTPYSDGVLYGSPTA 120

SGCE-wild       ENVGKPTIIEITAYNRRTFETARHNLIINIMSAEDFPLPYQAEFFIKNMNVEEMLASEVL 180
SGCE-846del     ENVGKPTIIEITAYNRRTFETARHNLIINIMSAEDFPLPYQAEFFIKNMNVEEMLASEVL 180

SGCE-wild       GDFLGAVKNVWQPERLNAINITSALDRGGRVPLPINDLKEGVYVMVGADVPFSSCLREVE 240
SGCE-846del     GDFLGAVKNVWQPERLNAINITSALDRGGRVPLPINDLKEGVYVMVGADVPFSSCLREVE 240

SGCE-wild       NPQNQLRCSQEMEPVITCDKKFRTQFYIDWCKISLVDKTKQVSTYQEVIRGEGILPDGGE 300
SGCE-846del     NPQNQLRCSQEMEPVITCDKKFRTHFTLTGAKFHWLIKQSKCPPIRK*              287

SGCE-wild       YKPPSDSLKSRDYYTDFLITLAVPSAVALVLFLILAYIMCCRREGVEKRNMQTPDIQLVH 360

SGCE-wild       HSAIQKSTKELRDMSKNREIAWPLSTLPVFHPVTGEIIPPLHTDNYDSTNMPLMQTQQNL 420

SGCE-wild       PHQTQIPQQQTTGKWYP 437
```

3.3.2. *Confirmation of novel mutations in the SGCE gene*

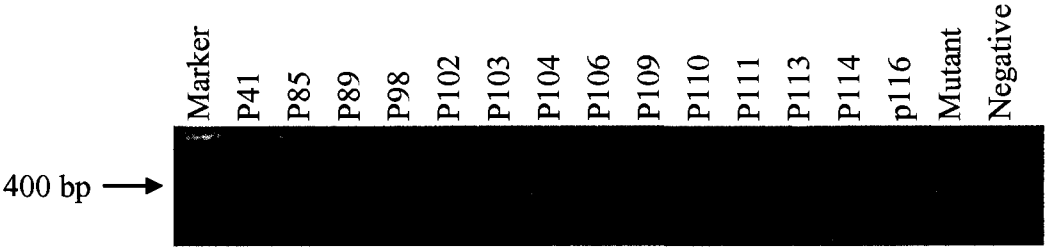
To confirm that the 5-bp deletion was a mutation, we designed an allele specific primer to bind only to the mutant deletion breakpoint. After PCR amplification with the specific primer pair, the mutant allele was only amplified in the affected individuals of family 4 and was not amplified in family members without M-D. The deletion was not seen in any of the 100 normal controls screened using allele-specific primer screening (Figure 3-4). For confirmation of the 1-bp deletion mutation in family 15, we did not find this variant in 100 normal controls nor in 20 affected individuals from 18 families. Out of 17 families with M-D we have now identified 4 families with *SGCE* mutations. The detected mutations in the *SGCE* gene by direct sequencing account for 24 % in our cohort of 17 M-D families.

Figure 3-4. Confirmation of the 5-bp deletion in exon 7 of *SGCE* by allele specific oligonucleotide hybridization (ASOH). A. PCR amplification of exon 7 of the *SGCE* gene using normal primers (*SGCE* F7/R7) showed the exon 7 fragments in normal controls (p41-p116) and the affected individual with a 5-bp deletion (M4-1). **B.** Allele specific primer only binds to the deletion breakpoint of mutant allele whereas the normal controls can not be amplified.

A.


Marker
P41
P85
P89
P98
P102
P103
P104
P106
P109
P110
P111
P113
P114
p116
Mutant
Negative
400 bp

B.


Marker
P41
P85
P89
P98
P102
P103
P104
P106
P109
P110
P111
P113
P114
p116
Mutant
Negative
400 bp

3.3.3. Phenotype analysis of patients with SGCE mutations

The examination of affected individuals of our four families with *SGCE* mutations (families 3, 4, 7, and 15) showed typical clinical manifestations of M-D compared to the affected individuals of families without *SGCE* mutations. These characteristics include early onset, responsiveness to alcohol (two of four families) and myoclonus in face, neck and upper extremities as the primary symptom. The age at onset for myoclonus is in the first decade while the dystonia is rarely manifested. These clinical characteristics are consistent to the observations reported by others (Asmus *et al.*, 2002; Tezenas du Montcel *et al.*, 2006).

3.4. Discussion

The α-, β-, γ-, δ-, ζ-, and ε-sarcoglycans are transmembrane components of the dystrophin–glycoprotein complex, which link the cytoskeleton to the extracellular matrix in muscle and cardiac tissue (Wheeler and McNally, 2003). Human *SGCE* gene is composed of 11 exons plus two alternatively spliced exons (exon 10 and 11b). The full-length sequence of ε-sarcoglycan is 437 amino acids long and shares 62% amino acid identity with α –sarcoglycan (Hack *et al.*, 2000). The *SGCE* gene is widely expressed in both muscle and non-muscle cells in embryos as well as adults. Recessive mutations in the α-, β-, γ-, or δ-sarcoglycan genes cause limb-girdle muscular dystrophy (Ozawa *et al.*, 2005). Mutations in *SGCE* are not associated with a muscular dystrophy or a muscle phenotype, but do cause myoclonus–dystonia (Asmus and Gasser, 2004; Asmus *et al.*, 2005; Han *et al.*, 2003; Schule *et al.*, 2004a; Schule *et al.*, 2004b; Tezenas du Montcel *et al.*, 2006). There does not seem to be any phenotypic distinguishing features in individuals with M-D who are negative for *SCGE* mutations. The psychiatric problems of alcohol abuse, obsessive–compulsive

disorder, and panic attacks that are being increasingly recognized in this disorder, but were not formally studied in our cohort because they do not seem to be predictive for mutation in the *SGCE* gene as previously reported (Doheny *et al.*, 2002). In the families where the *SGCE* mutations were identified along with mutation in the *DRD2* or *DYT1* genes, there did not seem to be any difference in the phenotype (Doheny *et al.*, 2002; Klein *et al.*, 2002). One individual with an M-D phenotype had a dramatic response to levodopa but was found to have a *GCH-1* mutation that is typically associated with dopa-responsive dystonia (Leuzzi *et al.*, 2002). To date, almost all families with the typical features of M-D have been reported to carry mutations in the *SCGE* gene by direct sequencing (Tezenas du Montcel *et al.*, 2006). In this study, we screened seventeen families with the typical M-D phenotype yet could identify mutations in only four of them using this technique. We did identify one novel 5-bp deletion (AAAAC) and one novel 1-bp deletion (795A deletion) both of which are predicted to produce a truncated protein lacking the transmembrane domain of *SGCE* gene needed to bind to the cell membrane. As with all the previous mutations in the *SCGE* gene, the associated phenotype did not involve any muscle disease. This 5-bp deletion in exon 7 has been recently found in a French family (Tezenas du Montcel *et al.*, 2006). Interestingly, Klein *et al* identified a similar 5-bp deletion (ACAAA) in exon 7 for a German family (Klein *et al.*, 2002). This deletion overlapped with our 5-bp deletion, indicating this region might be a hot spot for deletions. The 1-bp deletion in exon 6 has not been reported by other groups and suggests this mutation could be a rare mutation. Mutations in exon 3 of the *SCGE* gene appear to be common in M-D families. Including our 2 families, 20 families have mutations in exon 3 of the *SGCE* gene and 9 of these 20 families have the Arg102X mutation. None of these families is known to be related, so why this particular mutation

seems to be so common is unknown. If individuals with an M-D phenotype do not have a positive family history, they are much less likely to carry an *SGCE* mutation. However, seemly non-inherited (sporadic) cases have been found to harbor *SGCE* mutations (e.g. family 15). The limitation of sequencing exons in the *SGCE* gene using the genomic DNA is that intronic mutations causing alternate splicing or large deletions might be missed using this technique. However, at least for Family 1, linkage was identified to a distinct region on chromosome 18p11 as an explanation for their disease (Grimes *et al.*, 2002). Our lab has recently developed a semi-quantitative PCR-based DHPLC method to efficiently detect gene dosage alterations which was used to detect large deletions in some of our other M-D families and is the focus of **Chapter 4**.

Chapter 4.

Screening for Large Deletions/Duplications in the Epsilon-Sarcoglycan (*SGCE*) Gene

4.1. Introduction

The initial reports of screening for mutations in the epsilon-sarcoglycan (*SGCE*) gene on chromosome 7q21 in M-D used only direct sequencing (**Chapter 3**). Screening for gene dosage changes has only been reported to date by one group and it is unclear the number of families with M-D who might have larger deletions that were not identified by direct sequencing. We have collected 17 familial and sporadic families with M-D and one family with focal dystonia. By sequencing the 12 exons of the *SGCE* gene for these 18 families, we were able to identify 4 families with *SGCE* mutations. Two families had the same mutation in exon 3 (R102X) as reported previously while other two families were found to carry novel mutations. One family (family 4) had a 5-bp deletion in exon 7 and another family (family 15) had a 1-bp deletion in exon 6. All of these mutations are predicted to result in premature termination of translation.

Recently, large deletions in the *SGCE* gene have been reported in two of eight families with M-D (Asmus *et al.*, 2005). One family had a 15 kb deletion encompassing exon 5 of the *SGCE* gene while another family had a 6.8 kb deletion encompassing exon 6 of the *SGCE* gene, demonstrating that rearrangements within *SGCE* can also cause M-D. In order to explore the possibility of large deletions and duplications within the *SGCE* gene for our cohort of M-D families, we have developed a fast, sensitive and semi-quantitative PCR-based DHPLC method.

4.2. Subjects and Methods

4.2.1. *Subjects*

Information regarding informed consent, diagnosis and extraction of DNA is presented in **Chapter 2**. Fourteen *SGCE* mutation-negative familial and sporadic families were selected for dosage analysis (families 1, 2, 5, 6, 8, 9, 10, 11, 12, 13, 14, 16, 17, and 18).

4.2.2. Gene dosage analysis of the SGCE gene for fourteen families

Gene dosage analysis for all 12 exons of the *SGCE* gene was determined as described in section 2.6 of **Chapter 2**. In brief, each of 12 exons of the *SGCE* gene and an internal control (exon 7 of *PDZD2* gene on chromosome 5) were co-amplified in 50 µl reactions containing 0.2-0.5 U of Taq DNA polymerase (5U/µl), 5 µl 10X buffer with concentration of $MgCl_2$ at 1.0, or 1.5 or 2.5 mM, 2 pmol of each *SGCE* exon forward and reverse primer and 2 pmol of each of *PDZD2* exon 7 forward and reverse primers (Forward: 5'-ACTACAATGGGACAGATGCC-3'; Reverse: 5'-ACCAAGTTGACAAGATAGACA-3'), 0.2 mM dNTPs, and 2-4 µl genomic DNA (100-200 ng). As our multiplex PCR optimizations with 25 cycles were shown to detect the *SGCE* dosage alteration, we set our multiplex PCR conditions at 94° C for 2 min; then 25 cycles of 94° C for 45 s, 55-61° C for 45 s; and 72° C for 45 s, 72° C for 10 min. All of the reactions were performed in triplicate in a 96 well plate or PCR tubes. The oligonucleotide primers used to amplify each exon of *SGCE* are described in Table 3-1 of **Chapter 3**.

Quantification of multiplex PCR products was performed using DHPLC system and Wave-Maker Software (Transgenomics). The DHPLC profile of each patient or normal individual is composed of two peaks with one peak representing the PCR product of each of 12 *SGCE* exons and another peak representing the PCR product of *PDZD2* exon 7 (internal control). The *SGCE* exon dosage was expressed as the ratio of peak areas for each exon of the *SGCE* gene and *PDZD2* exon 7 (Internal control). Each peak area of the DHPLC

profiles was read through the WAVE-MAKER software. The mean *SGCE/PDZD2* ratio for each exon of *SGCE* gene was calculated from 10 unaffected controls and was used to normalize each exon of *SGCE/ PDZD2 exon 7* ratio of each patient based on the formula described in **Chapter 2**.

$$\text{Normalized Ratio} = \frac{SGCE \,/\, PDZD2\ [\text{Patient}]}{\text{Mean } SGCE \,/\, PDZD2\ [\text{Control}]}$$

After calculating the normalized ratio of each patient, we assessed the gene dosage based on the criteria described in **Chapter 2**: the normalized ratios below 0.75 indicate deletions; from 0.75 to 1.25 are normal; above 1.25 reflect duplications.

4.2.3. Mapping the deletion breakpoints of the SGCE gene

The deletion breakpoints were mapped in the same manner that *SGCE* exons were evaluated for copy number. For each side of the deletion, primer pairs were designed midway between the closest flanking PCR products exhibiting haploid (one copy) and diploid states (two copies). Once the genomic distances between two flanking non-deletion PCR primer sets were reduced to about 2 -3 kb, we amplified the intervening genomic DNA from patients of families with deletions. The resulting PCR products were then sequenced directly.

4.3. Results

4.3.1. *Gene dosage analysis of the SGCE gene in fourteen families*

In this study, we analyzed gene dosages in 12 exons of the *SGCE* gene (previously reported 11 exons of *SGCE* gene plus an additional newly identified exon 11b) in fourteen families using our semi-quantitative PCR-based DHPLC method. These families were shown to be *SGCE* mutation negative via PCR and direct sequencing (**Chapter 3**). To

detect dosage changes in the *SGCE* gene for these families, each exon of *SGCE* was co-amplified with an internal control (exon 7 of *PDZD2),* visualized on agarose gels (Figure 4-1) prior to quantifying the PCR products on DHPLC system (Figure 4-2).

By analyzing the copy number for all 12 exons of the *SGCE* gene in the affected individuals of these fourteen families, we identified large deletions in two of them (families 2 and 6). Family 2 had a deletion of *SGCE* exons 2 and 3 and family 6 had a deletion spanning exons 2 through 5 (Figure 4-3, 4, 5, and 6). We did not identify dosage alterations in any exons of the other 12 families. The calculations of normalized ratios are listed in Appendix 1.

Figure 4-1. Co-amplification of *SGCE* exon 2 and the *PDZD2* exon 7 (internal control) for patients with M-D and normal controls using multiplex PCR. Lanes N1 and N2 were normal controls. Lanes M1-16 through M18-1 were affected individuals with M-D except for M10-11 with focal dystonia. Black arrows indicate the affected individuals with deletions in *SGCE*. Red arrow indicates *PDZD2* exon 7 (top bands), Blue arrow indicates *SGCE* exon 2 (bottom bands).

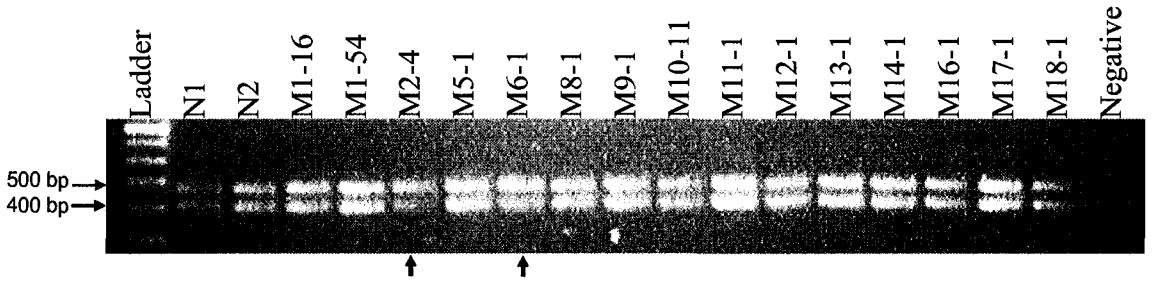

Ladder
N1
N2
M1-16
M1-54
M2-4
M5-1
M6-1
M8-1
M9-1
M10-11
M11-1
M12-1
M13-1
M14-1
M16-1
M17-1
M18-1
Negative
500 bp
400 bp

Figure 4-2. Chromatograms obtained by DHPLC analysis of a normal control (N2) and an M-D patient from family 2 (M2-4). *SGCE* exon 2 and the *PDZD2* exon 7 (internal control) fragments were co-amplified by multiplex PCR (25 cycles) and analyzed by DHPLC. *SGCE* and *PDZD2* amplicons were eluted at 8.5 and 9.0 min, respectively. The areas under the peaks, which were proportional to the amount of PCR products, were calculated automatically. Left panel, the profile for an affected individual of family 2 (M2-4); Right panel, the profile for a normal individual.

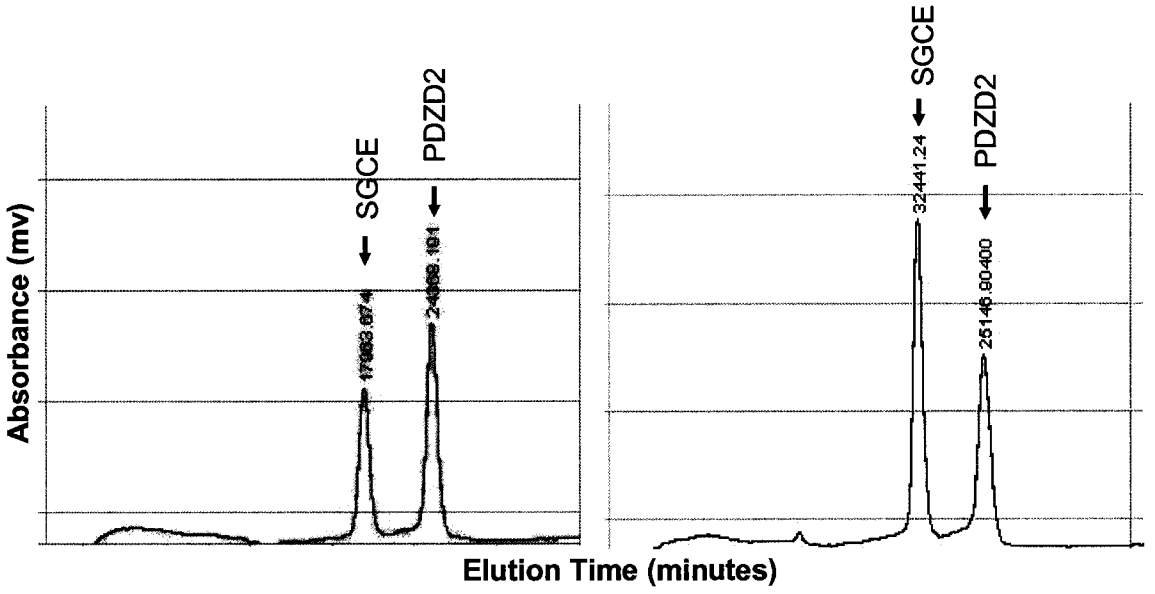
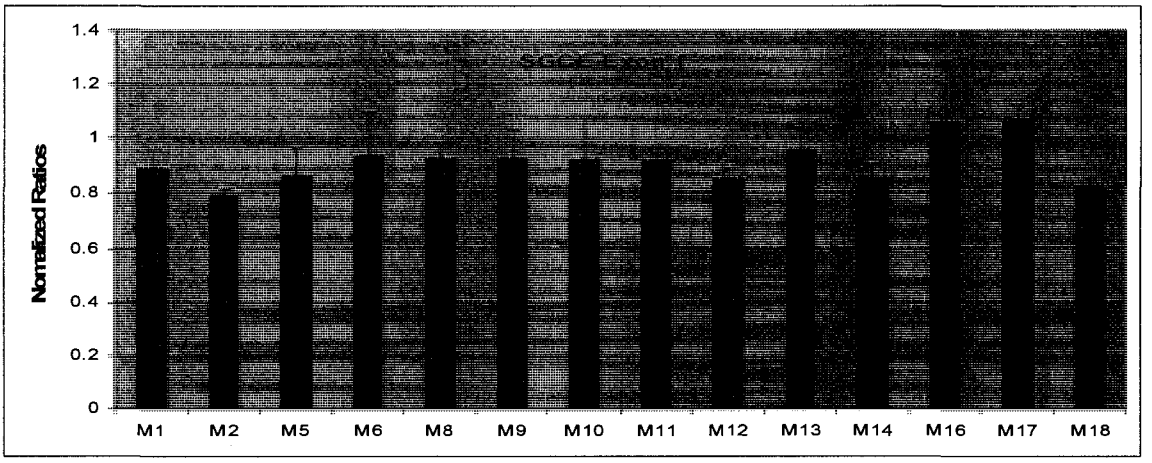
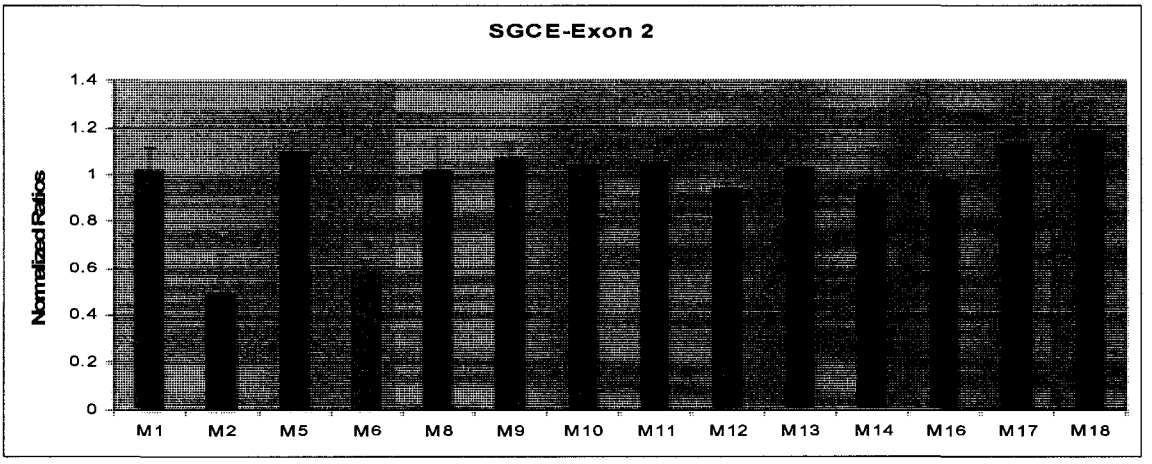

SGCE
PDZD2
Absorbance (mv)
32441.24
25146.90400
Elution Time (minutes)

Figure 4-3. Determination of deletions or duplications in exons 1, 2 and exon 3 of the *SGCE* gene for each affected individual from fourteen families. Upper panel, No patient from the14 families revealed a deletion or duplication in exon 1. Middle and Lower panels, two patients, M2-4 (red) and M6-1 (orange), revealed deletions in exon 2 and exon 3. * indicates the patients have a reduced copy number consistent with a deleted allele.


Normalized Ratios
1.4
1.2
1
0.8
0.6
0.4
0.2
0
M1 M2 M5 M6 M8 M9 M10 M11 M12 M13 M14 M16 M17 M18


SGCE-Exon 2
Normalized Ratios
1.4
1.2
1
0.8
0.6
0.4
0.2
0
M1 M2 M5 M6 M8 M9 M10 M11 M12 M13 M14 M16 M17 M18

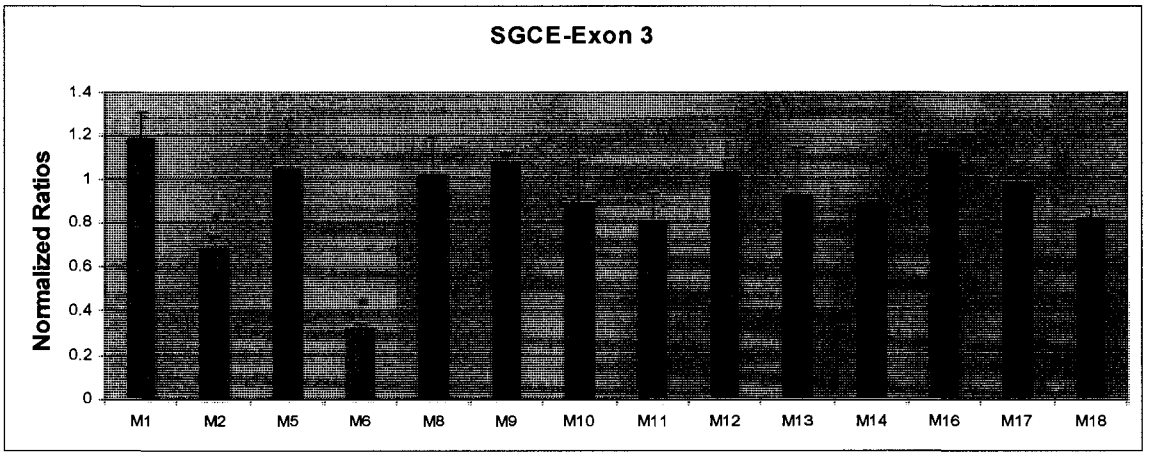
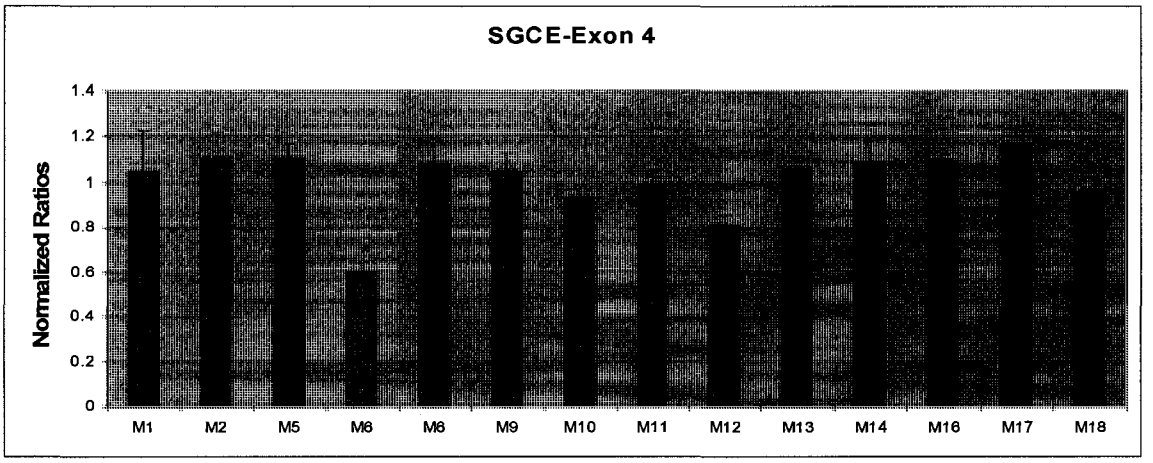
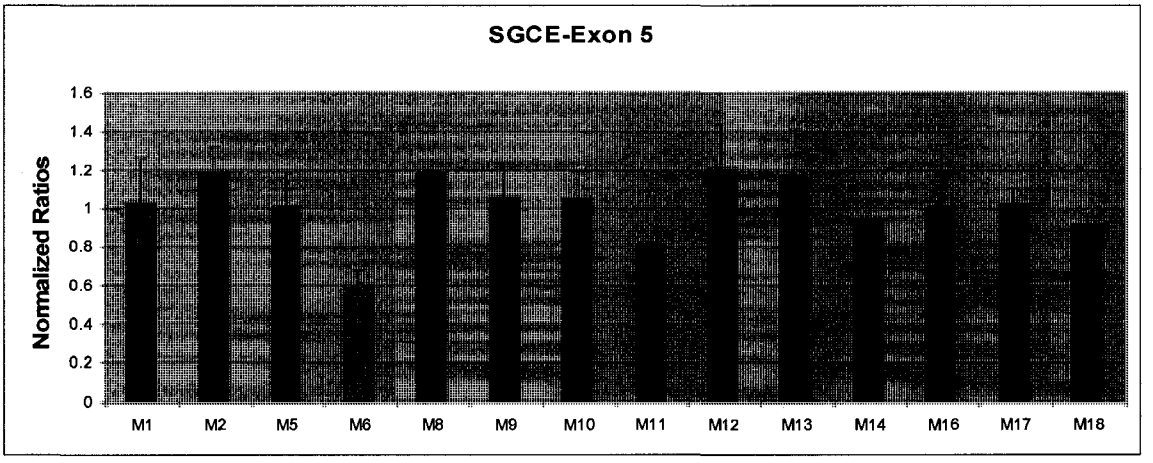

SGCE-Exon 3
Normalized Ratios
1.4
1.2
1
0.8
0.6
0.4
0.2
0
M1 M2 M5 M6 M8 M9 M10 M11 M12 M13 M14 M16 M17 M18

Figure 4-4. Determination of deletions or duplications in exons 4, 5, and 6 of the *SGCE* gene for each affected individual from fourteen families. Upper and middle panels, One patient (M6-1) revealed deletions in exons 4 and 5. Lower panel: No patient from 14 families revealed a deletion or duplication in exon 6.


SGCE-Exon 4
Normalized Ratios
1.4
1.2
1
0.8
0.6
0.4
0.2
0
M1 M2 M5 M6 M8 M9 M10 M11 M12 M13 M14 M16 M17 M18


SGCE-Exon 5
Normalized Ratios
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
M1 M2 M5 M6 M8 M9 M10 M11 M12 M13 M14 M16 M17 M18

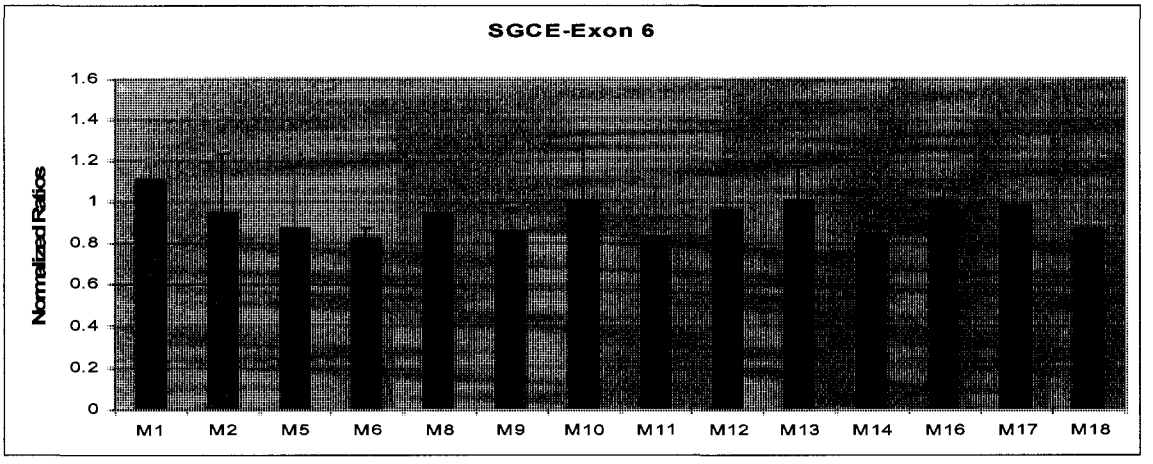
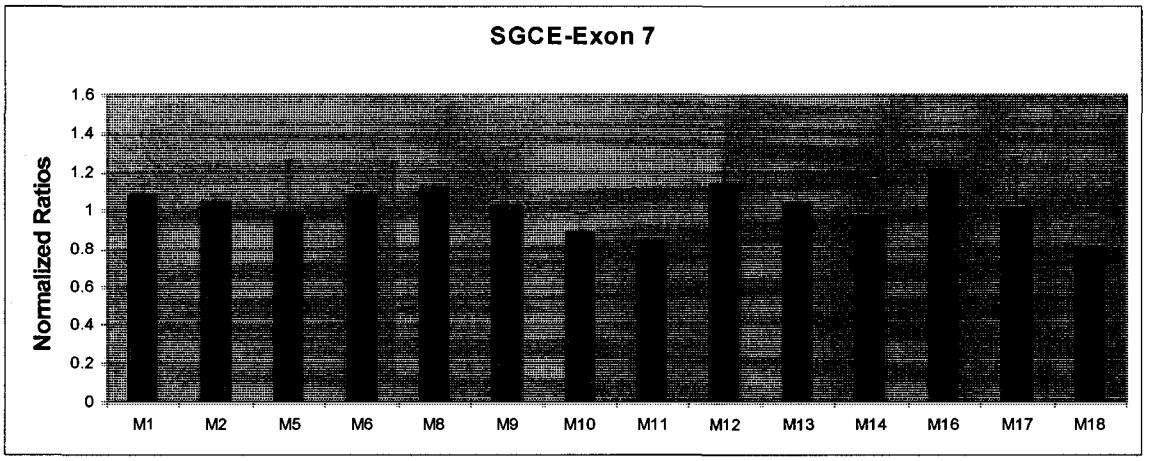
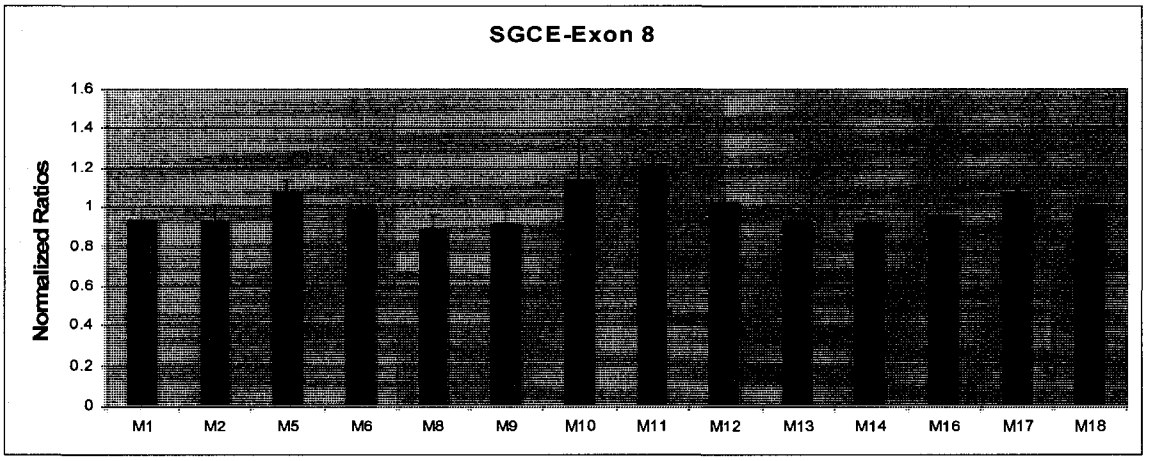

SGCE-Exon 6
Normalized Ratios
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
M1 M2 M5 M6 M8 M9 M10 M11 M12 M13 M14 M16 M17 M18

Figure 4-5. Determination of deletions or duplications in exons 7, 8 and 9 of the *SGCE* gene for each affected individual from fourteen families. No patient revealed deletions/duplications in exons 7 (upper panel), 8 (middle panel) and 9 (lower panel).


SGCE-Exon 7
Normalized Ratios
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
M1
M2
M5
M6
M8
M9
M10
M11
M12
M13
M14
M16
M17
M18


SGCE-Exon 8
Normalized Ratios
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
M1
M2
M5
M6
M8
M9
M10
M11
M12
M13
M14
M16
M17
M18

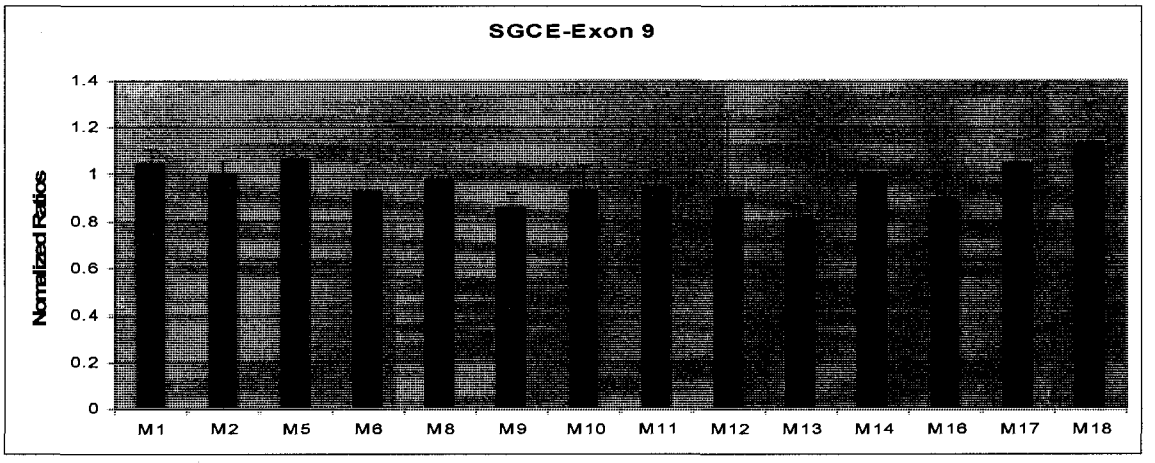
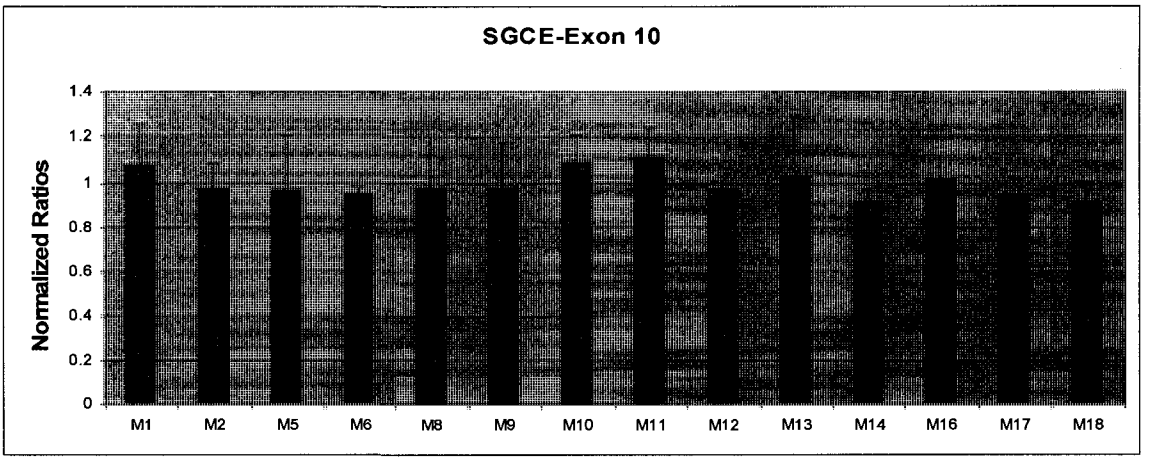
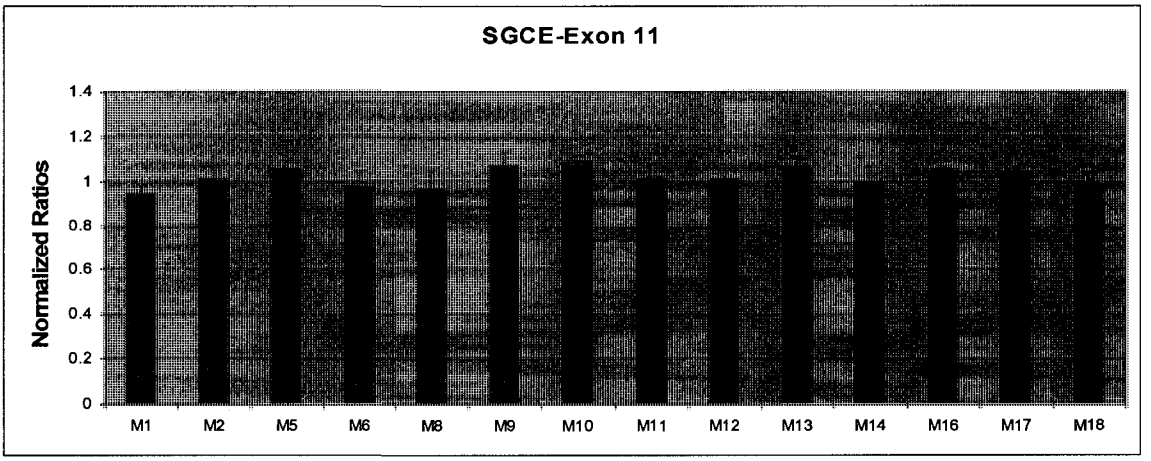
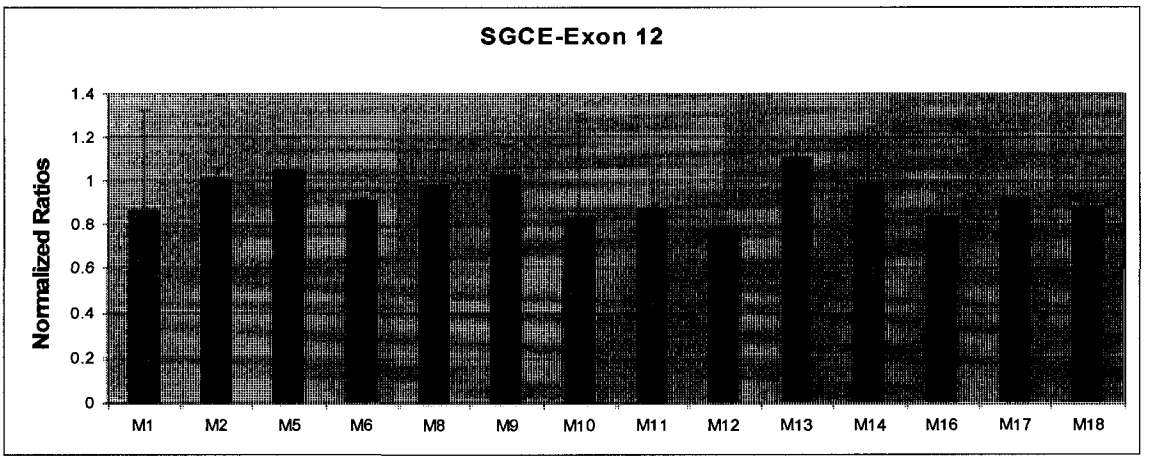

SGCE-Exon 9
Normalized Ratios
1.4
1.2
1
0.8
0.6
0.4
0.2
0
M1
M2
M5
M6
M8
M9
M10
M11
M12
M13
M14
M16
M17
M18

Figure 4-6. Determination of deletions or duplications in exons 10, 11 and 12 of the *SGCE* gene for each affected individual from fourteen families. No patient revealed deletions/duplications in exons 10 (upper panel), 11 (middle panel) and 12 (lower panel).


SGCE-Exon 10
Normalized Ratios
1.4
1.2
1
0.8
0.6
0.4
0.2
0
M1
M2
M5
M6
M8
M9
M10
M11
M12
M13
M14
M16
M17
M18

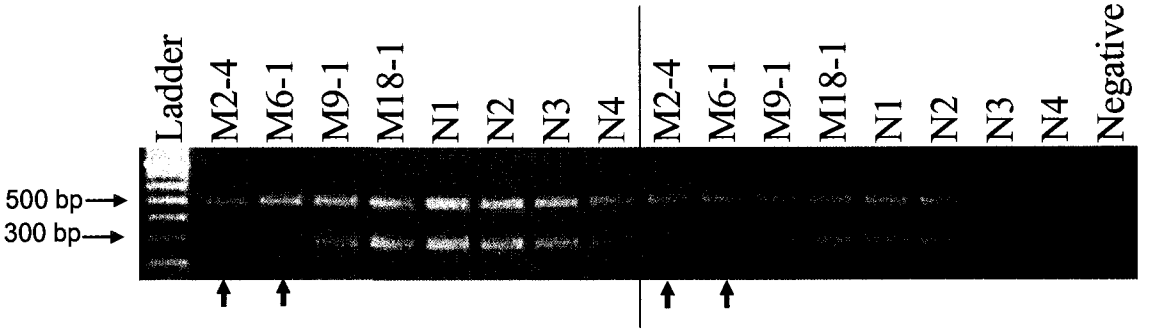

SGCE-Exon 11
Normalized Ratios
1.4
1.2
1
0.8
0.6
0.4
0.2
0
M1
M2
M5
M6
M8
M9
M10
M11
M12
M13
M14
M16
M17
M18


SGCE-Exon 12
Normalized Ratios
1.4
1.2
1
0.8
0.6
0.4
0.2
0
M1
M2
M5
M6
M8
M9
M10
M11
M12
M13
M14
M16
M17
M18

4.3.2. Identification of deletion breakpoints in the SGCE gene

To define the genomic breakpoints for the two families with deletions, we used the strategy as described in section 4.2 to narrow down the non-deleted regions flanking the breakpoints. As exon 1 is not deleted in two families, we postulated that the proximal breakpoints for these two families are located in intron 1. We know that the distal breakpoint of family 2 is located in intron 3 as exon 4 is not deleted in family 2. We also know that the distal breakpoint of family 6 is located in intron 5 as exon 6 is not deleted. Based on these observations, we first designed three sets of PCR primers which were equally distributed within intron 1. We also designed three sets of primers which were equally distributed in intron 3 and intron 5 for dosage analyses of distal breakpoints in family 2 and family 6 respectively. For the sake of speed we performed multiplex PCR and analyzed the products on agarose gels (Figure 4-7). We then continued to fine map the deletion junctions by repeating the process of amplifying fragments between the closest flanking diploid and haploid regions.

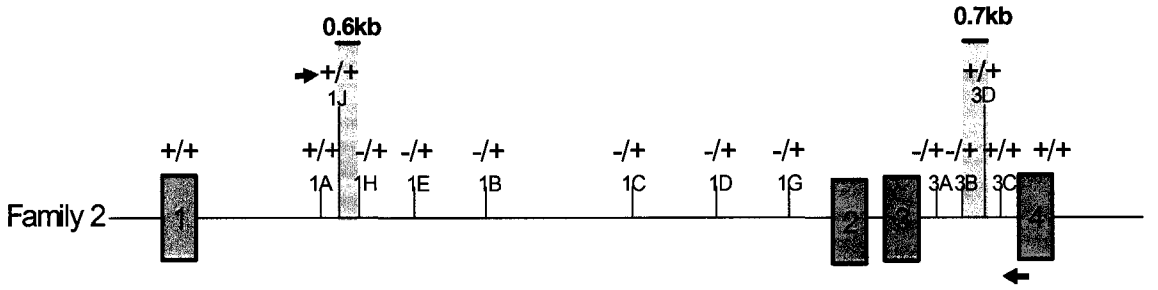
Figure 4-7. Multiplex PCR amplification of *SGCE* F3B/R3B (intron 3) and *PDZD2* exon 7 (internal control) for the patients with M-D and normal controls was performed in triplicate (only 2 replicates are shown here). N1, N2, N3 and N4 are normal controls. Lanes M2-4, M6-1, M9-1 and M18-1 are affected individuals with M-D. Black arrows indicate the affected individuals with deletions in *SGCE*. Red arrow indicates *PDZD2* exon 7 (top bands), Blue arrow indicates *SGCE* F3B/R3B of intron 3 (bottom bands).


Ladder
M2-4
M6-1
M9-1
M18-1
N1
N2
N3
N4
M2-4
M6-1
M9-1
M18-1
N1
N2
N3
N4
Negative
500 bp
300 bp

For the identification of the deletion encompassing exon 2 and exon 3 in family 2, we used 8 primers in intron 1 and 4 primers in intron 3 to determine that the proximal deletion breakpoint was located between primers for *SGCE* 1J and *SGCE* 1H (0.6 kb) and the distal deletion breakpoint was located between primers for *SGCE* 3B and *SGCE* 3D (0.7 kb) (Figure 4-8, A). We were able to amplify a 1.1 kb PCR product with primers Forward-1J/Reverse-3D (Figure 4-8, B). Direct sequencing of this 1.1 kb PCR product identified the deletion breakpoint junction between intron 1 and intron 3. Two affected individuals (M2-1 and M2-4) of family 2 have the deleted alleles with partial sequences of intron 1 and intron 3 (Figure 4-9).

Figure 4-8. Mapping the deletion breakpoint of the *SGCE* gene for family 2. A. 12 primers were used to narrow down the deletion flanking regions of affected individuals of family 2. **B**. PCR amplification of two affected individuals of family 2 (M2-1, M2-4) and two normal controls was performed using the primers *SGCE* F1J/R3D (red arrows). A 1.1 kb fragment was identified, confirming the deletion of *SGCE* exons 2 and 3 for family 2.

A.


0.6kb
0.7kb
+/+
1J
+/+
3D
+/+
+/+ -/+ -/+ -/+ -/+ -/+ -/+ -/+ -/+ +/+ +/+
1A 1H 1E 1B 1C 1D 1G 3A 3B 3C
Family 2
1
2
3
4

B.

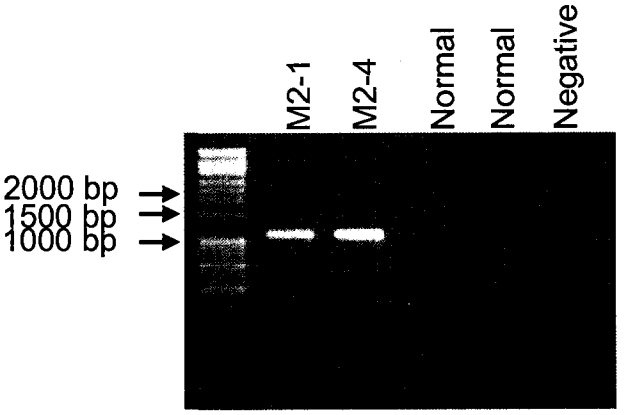
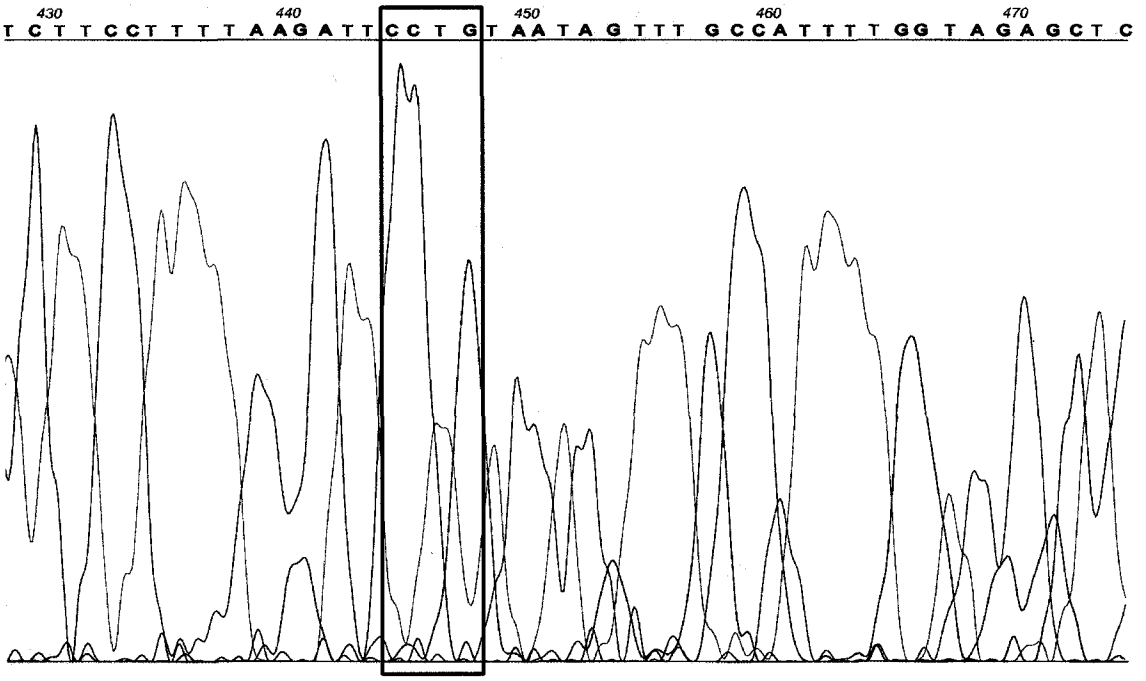

M2-1
M2-4
Normal
Normal
Negative
2000 bp
1500 bp
1000 bp

Figure 4-9. Sequencing of the deletion breakpoint of *SGCE* for family 2. Sequencing the 1.1 kb *SGCE* F1J/R3D PCR product of two affected individuals of family 2 (M2-1, M2-4) identified the deletion breakpoint between intron 1 and intron 3 of the *SGCE* gene (only the sequence of M2-4 was shown here). Black box indicates the 4-bp sequence (CCTG) was shared between intron 1 and intron 3 of the *SGCE* gene.

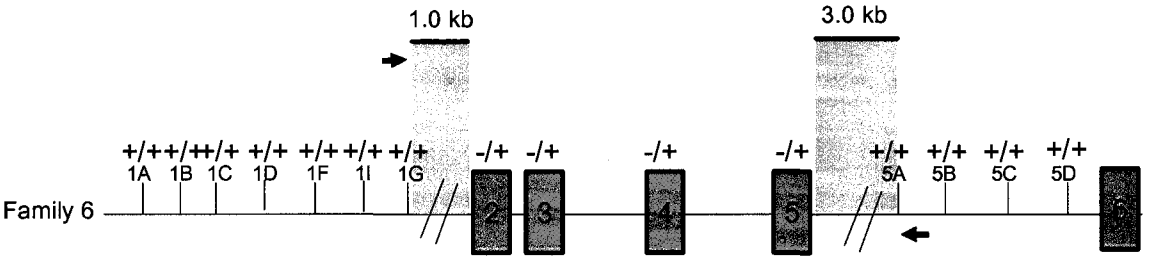
3'- Partial sequence of *SGCE* intron 1 5'- Partial sequence of *SGCE* intron 3



For the identification of the deletion covering exon 2 through exon 5 in family 6, we used 7 primers in intron 1 and 4 primers in intron 5 to determine that the proximal deletion breakpoint was located between primers SGCE1G and *SGCE* exon 2 (1.0 kb) and the distal deletion breakpoint was located between primers for *SGCE* exon 5 and *SGCE* 5A (3.0 kb) Figure 4-10, A. We were able to amplify a 1.6 kb PCR product with primers Forward-1G/Reverse-5A as shown in Figure 4-10, B. Direct sequencing of this 1.6 kb PCR product we identified the deletion breakpoint junction between intron 1 and intron 5. The affected individual (M6-1) of family 6 has the deleted allele with partial sequences of intron 1 and intron 5 of the *SGCE* gene (Figure 4-11).

Figure 4-10. Mapping the deletion breakpoint of the *SGCE* gene for family 6. A. 11 primers were used to determine the deletion boundaries of the *SGCE* gene for family 6. **B.** PCR amplification of one affected individual of family 6 (M6-1) and two normal controls was performed using the primers *SGCE* F1G/R5A (red arrows). 1.6 kb fragment was identified, confirming the deletion of *SGCE* exons 2 through 5 for family 6.

A.


1.0 kb
3.0 kb
+/+ +/+ +/+ +/+ +/+ +/+ +/+ -/+ -/+ -/+ -/+ +/+ +/+ +/+ +/+
1A 1B 1C 1D 1F 1I 1G 5A 5B 5C 5D
Family 6
2
3
4
5
6

B.

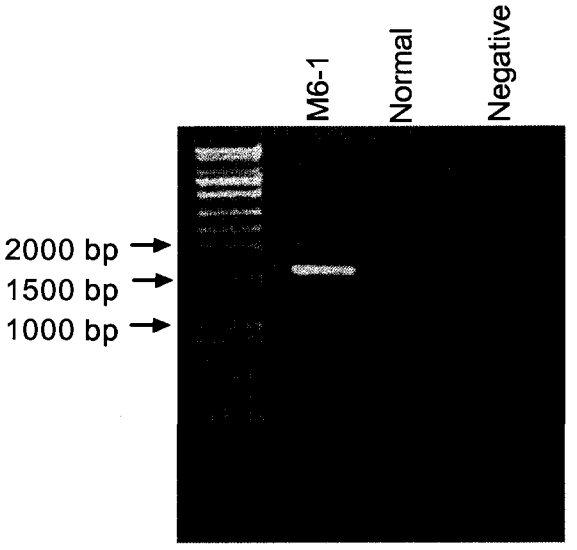
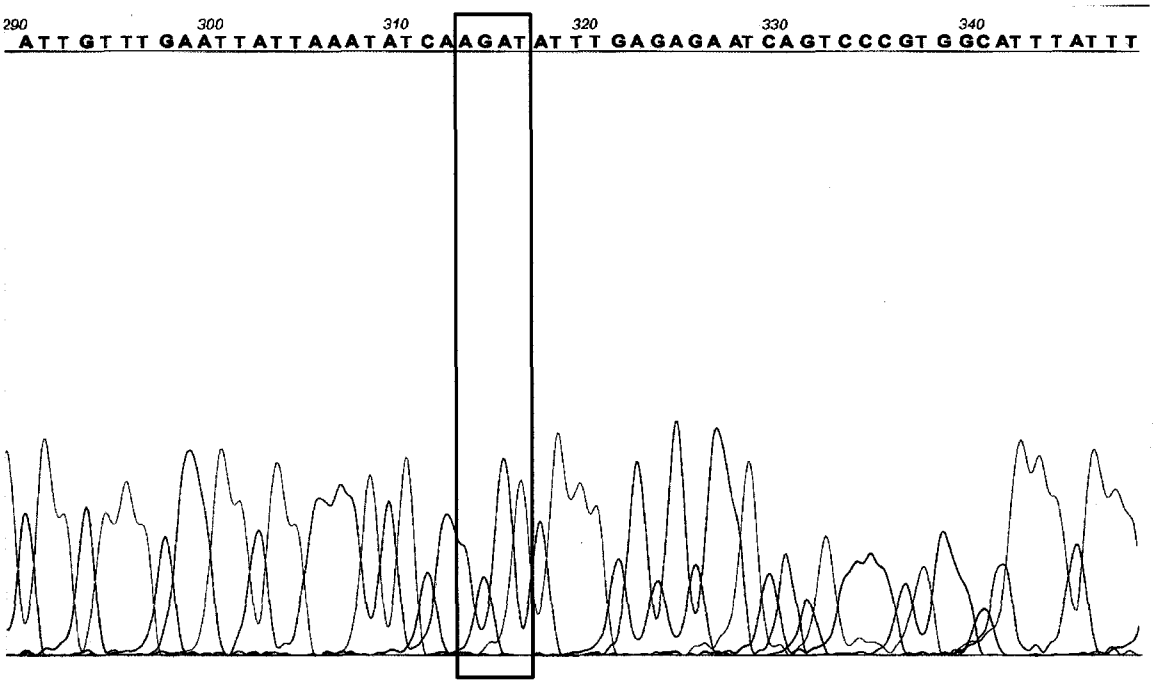

M6-1
Normal
Negative
2000 bp
1500 bp
1000 bp

Figure 4-11. Sequencing of the deletion breakpoint of the *SGCE* gene for family 6. Sequencing the 1.6 kb *SGCE* F1G/R5A PCR product of one affected individual of family 6 (M6-1) identified the deletion breakpoint between intron 1 and intron 5 of the *SGCE* gene. Black box indicates the 4-bp sequence (AGAT) was shared between intron 1 and intron 5 of the *SGCE* gene.

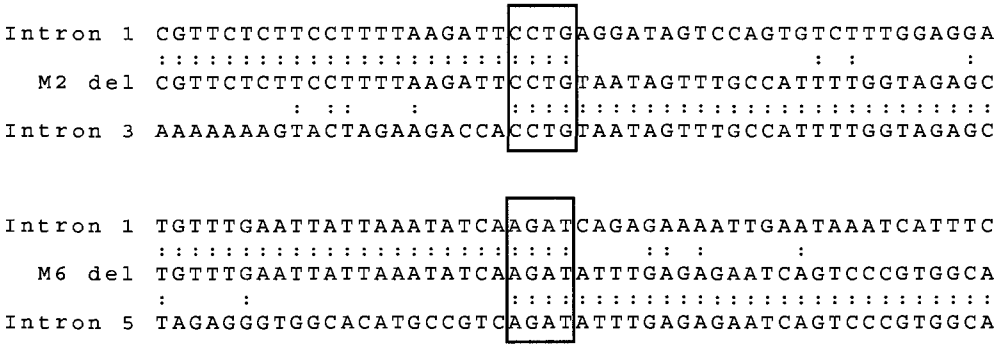
3'- Partial sequence of *SGCE* intron 1 5'- Partial sequence of *SGCE* intron 5



4.3.3. *Analysis of the mechanism for large deletions in the SGCE gene*

The nucleotide sequences surrounding the breakpoints were investigated for the purpose of identifying a mechanism for the deletions. The RepeatMasker program did not detect repetitive elements within 25 bases flanking the junctions of either deletion, nor were there any direct or inverted repeats present. Both deletions appear to have occurred as a result of non-homologous recombination. Comparing the sequences around the breakpoint revealed a common 4 base sequence. For family 2 the sequence was "CCTG", while for family 6 the sequence was "AGAT" (Figure 4-12).

Figure 4-12. Alignment of DNA sequences around the large *SGCE* deletions in families 2 and 6. A common 4 base tract "CCTG" (black box) was identified to flank the deletion breakpoint for family 2 (upper panel) while another 4 base tract "AGAT" (black box) was found to flank the deletion breakpoint for family 6. Note: M2 and M6 indicate family 2 and family 6, respectively; del indicates the sequence with deletion.


Intron 1 CGTTCTCTTCCTTTTAAGATTCCTGAGGATAGTCCAGTGTCTTTGGAGGA
M2 del CGTTCTCTTCCTTTTAAGATTCCTGTAATAGTTTGCCATTTTGGTAGAGC
Intron 3 AAAAAAAGTACTAGAAGACCACCTGTAATAGTTTGCCATTTTGGTAGAGC
Intron 1 TGTTTGAATTATTAAATATCAAGATCAGAGAAAATTGAATAAATCATTTC
M6 del TGTTTGAATTATTAAATATCAAGATATTTGAGAGAATCAGTCCCGTGGCA
Intron 5 TAGAGGGTGGCACATGCCGTCAGATATTTGAGAGAATCAGTCCCGTGGCA

4.4. Discussion

Point mutations and small deletions/insertions in *SGCE* gene have been reported to cause M-D in 20-50% of cases (Asmus and Gasser, 2004; Asmus *et al.*, 2005; Asmus *et al.*, 2002; Han *et al.*, 2003; Schule *et al.*, 2004a; Schule *et al.*, 2004b; Tezenas du Montcel *et al.*, 2006)). We had identified point mutations and small deletions in four of 17 families with M-D, accounting for 24% of the cases (**Chapter 3**). As PCR and direct sequencing of exons of any autosomal gene could miss intronic mutations and large genomic rearrangements, we used a sensitive semi-quantitative PCR-based DHPLC method to screen for large deletions/duplications in 14 families without *SGCE* mutations. Dosage analysis by DHPLC identified large deletions of the *SGCE* gene in two families. This number represents 12% (2/17) of our cases with M-D. Taken together with the point mutations, small deletions and large deletions, *SGCE* mutations were found in 35% of our M-D, supporting the fact that the *SGCE* gene plays an important role in pathogenesis of M-D, but other genes must also be involved.

In addition to the fact that point mutations and small deletion of a gene can cause disease, genomic rearrangements such as the large deletions/duplications and translocations are well known as causes for a wide range of human diseases such as Duchenne and Becker muscular dystrophies (Hu *et al.*, 1992; Kunkel *et al.*, 1986; Worton *et al.*, 1984), familial breast cancer (Puget *et al.*, 1999), velocardiofacial /VCFS (Driscoll *et al.*, 1992) and Parkinson's disease (Clarimon *et al.*, 2005).

Several mechanisms have been proposed to induce gene rearrangements. The most common mechanism of genomic rearrangements is nonallelic homologous recombination (NAHR) through repetitive elements (RE) or transposable elements (TE). These repetitive

elements include short interspersed nuclear elements (SINEs) and long interspersed nuclear elements (LINEs) and have been implicated in chromosome rearrangements and inherited disorders (Lander *et al.*, 2001; Pentao *et al.*, 1992). Because LINEs and SINEs are present in multiple copies in the human genome, mis-alignment during meiosis would result in homologous recombination. Unequal homologous recombination between dispersed LINEs or SINEs can result in deletions in one copy of a gene, whereas duplications would be the reciprocal event occurring on the other allele (Deininger and Batzer, 1999). Alu elements of the SINE family have been shown to mediate gene rearrangements in human diseases (Batzer and Deininger, 2002; Kolomietz *et al.*, 2002). Recently, a 270 bp Alu element has been found to result in a 2.3 Mb deletion of 22q11 for VCFS (Uddin *et al.*, 2006). We searched the repeat sequences using the RepeatMasker tool (RepeatMasker website) and identified 23 repeat sequences in intron 1, 9 repeat sequences in intron 3 and 22 repeat sequences in intron 5. However, using Blastn (NCBI Blastn program) analysis of the repeat sequences at the proximal and distal breakpoints, we did not find a homologous sequence between the deletion breakpoints of either family. Non-homologous end-joining (NHEJ) was regarded the mechanism for *SGCE* large deletions in two families reported by Asmus *et al* (Asmus *et al.*, 2005). In one family with a large *SGCE* deletion, a 6-bp (-TGTTAA-) sequence close to the breakpoint was found to carry a topo II cleavage site which could be recognized and cut by DNA topoisomerase II to induce the deletion (Spitzner and Muller, 1988). In another family with a large *SGCE* deletion, a 3-bp sequence (-CTA-) at the upstream breakpoint which was found to be inverted in the downstream sequence was assumed as the mechanism to induce deletion by a single stranded DNA looping mechanism (Bacolla *et al.*, 2004). We did not detect a similar situation with our two families with

deletions. However, both deletions appear to have involved a common 4-bp sequence. A common "CCTG" was identified to flank the deletion for family 2 while an "AGAT" sequence was found to flank the deletion for family 6.

Many methods have been developed to determine the gene deletions/ duplications. The hybridization-based techniques including Southern blotting (Ikegami *et al.*, 1997; Petrij-Bosch *et al.*, 1997) and fluorescence in situ hybridization (FISH) (Lupski *et al.*, 1991; Pemble *et al.*, 1994) have several common problems, such as the need to use labeled probes and large amounts of DNA together with time-consuming methods. Quantitative PCR has widely been used to detect the gene dosage alterations, but is costly (Glockner *et al.*, 2000). Comparative genomic hybridization (CGH) had also been used but has limited resolution (Pinkel *et al.*, 1998). Denaturing high performance liquid chromatography (DHPLC) is a low-cost method developed to detect gene dosage (Xiao and Oefner, 2001). Naimi and colleagues have recently validated this method by testing 90 patients with inherited peripheral neuropathies, who were diagnosed by molecular methods. This DHPLC method was reported to have 100% specificity and sensitivity to detect the deletions in patients with hereditary neuropathy with liability to pressure palsy (HNPP), and have 95% sensitivity to detect duplications in patients with Charcot-Marie-Tooth type 1A (CMT1A) (Naimi *et al.*, 2005). We have developed a semi-quantitative PCR-based DHPLC method to detect the gene dosage alterations of *SGCE* gene by including PDZD2 exon 7 as an internal control and optimizing multiplex PCR conditions. Using this DHPLC method, we detected the *SGCE* gene deletions in two families with M-D and confirmed these deletions by identifying the genomic breakpoints in these two families. As we did not compare our PCR-based DHPLC method to other methods for detecting the gene rearrangements and we only

evaluated 14 patients using our method, we can not conclude that our PCR-based DHPLC method has 100% sensitivity and specificity even though both potential deletions in two families were detected through our method and were subsequently confirmed by the identification of the deletion breakpoints. However, these findings demonstrate that screening for large gene deletions/duplications in *SGCE* is necessary in M-D families and that our method using genomic DNA and DHPLC can be used to identify these changes.

Chapter 5.

A 10 cM Genomewide Linkage Analysis in a Large Canadian Family with Myoclonus-Dystonia

5.1. Introduction

M-D has been shown to be inherited as an autosomal dominant trait, and characterized by non-progressive form of myoclonus (involuntary rapid jerking) often seen in combination with dystonia (Asmus and Gasser, 2004; Mahloudji and Pikielny, 1967). Onset is usually in childhood with a wide spectrum of disease severity. Some individuals have only subtle symptoms that do not interfere with daily living, whereas others are markedly impaired from their myoclonus or dystonia. Physical examinations are typically normal and no other neurologic symptoms develop. The clinical diagnostic criteria, genetic components and genetic heterogeneity of M-D have been described in **Chapter 1**. Genetic linkage studies had suggested that three chromosomal regions are associated with M-D. One locus was mapped to chromosome 11q23, where two separate missense mutations in the D2 dopamine receptor (*DRD2*) gene were identified in two unrelated families (Garavagilia *et al.*, 2000; Klein *et al.*, 1999a). As the families with *DRD2* mutations were also found to have mutations in the *SGCE* gene, the involvement of the *DRD2* gene has been brought into question (Klein *et al.*, 2002). A second locus for M-D was then mapped on chromosome 7q21 (*DYT11*) and subsequently narrowed down to a 7-cM critical region (Asmus *et al.*, 2001; Klein *et al.*, 2000; Vidailhet *et al.*, 2001). Soon after, different loss-of-function mutations in the *SGCE* gene were identified for 6 families of 12 families with M-D (Zimprich *et al.*, 2001). Currently, more than 30 M-D families have been shown to have mutations in the *SGCE* suggesting that the *SGCE* gene plays an important role in the development of M-D (**Chapter 1**). However, mutations in *SGCE* gene can only account for 20-50% of the families with M-D, suggesting other loci are responsible for causing M-D in

the remaining families (Asmus and Gasser, 2004; Asmus *et al.*, 2005; Han *et al.*, 2003; Schule *et al.*, 2004b).

In our cohort of 17 families with M-D that we have collected, we have identified a large five-generation Canadian family (family 1). The phenotypes of affected individuals in this family are indistinguishable from that of the families with *SGCE* mutations which we described in **Chapter 2** (Table 2-1). In a previous study, Grimes *et al.*, excluded linkage to the 11q23 region and ruled out mutations in *DRD2* gene for this family (Grimes *et al.*, 2001). In another study, we genotyped this family with markers in the *DYT11* locus (7q21), excluded linkage to this locus and performed a genome-wide screen with a microsatellite marker density of 25 cM. As no markers showed LOD scores greater than 3.0, we genotyped this family with additional markers surrounding markers which generated LOD score equal to or greater than 0.5 in our 25 cM genome-wide screen. We obtained a LOD score of 2.44 at marker D18S391. Subsequently, we identified a set of 6 markers linked to this family, with the maximum LOD score of 3.50 at GATA185C06 and mapped a novel locus for M-D to a 17 cM region (3.99 Mb) flanked by markers D18S1132 and D18S843 on 18p11 (*DYT15*) (Grimes *et al.*, 2002). In order to further narrow down this linked region, we collected additional individuals of this family including one affected individual (#67). Here we report the results from linkage analysis and haplotype analysis for the additional two individuals with linked markers on 18p11 and the result for a 10 cM genome-wide linkage study of this family with 15 affected individuals (originally 13 affected, with two more affected confirmed later).

5.2. Subjects and Methods

5.2.1. *Subjects*

Informed consent was obtained from all participants in this study after approval by the institutional review board at The Toronto Hospital and the Ottawa Hospital. Individuals had a detailed history, physical examination, and videotape recorded using a standardized questionnaire and videotape protocol. A movement disorder expert (Anthony.E.Lang, MD, FRCPC) who was blinded to the patient's status or relationship reviewed the videotape. The clinical diagnosis of M-D was made according to established criteria as described in **Chapter 1**.

5.2.2. *Genomewide genotyping*

Genotyping was performed on DNA extracted from patient's whole blood using standard PCR protocols (specified by Research Genetics, Huntsville, AL) as described in **Chapter 2**. Genotyping of the additional two individuals of the family was initially performed with the linked markers on 18p11. A 10 cM genomewide screen with 144 more markers was completed to narrow the gaps from the 25 cM screen. The genetic distances were based on the Marshfield Research Foundation Map.

5.2.3. Linkage analysis and haplotype analysis

The two-point linkage analysis was described as in **Chapter 2**. For the calculation of LOD score, allele frequencies were assumed to be equal, the gene frequency for M-D was assumed to be 0.0001, and male and female recombination rates were assumed to be equal. Haplotypes were built for all the individuals of the family based on minimum recombination events. The penetrance was set to 70%.

5.3. Results

5.3.1. *Clinical features*

The clinical features and summary of most of the family members are described in Table 2-1 of **Chapter 2** and have previously been reported. Thirty-eight individuals including the two new individuals were examined and videotaped and had venous blood drawn and DNA extracted. Originally, M-D was diagnosed in 13 individuals in this family, with two more (#53 and #67) diagnosed as affected after the 25 cM genome screen had been completed and linkage to chromosome 18 was found. When individual #53 was initially seen it was felt he was possibly affected with myoclonus of his right arm with writing but his movements were very subtle. When he was seen in follow-up at the age of 19 he now had developed definite myoclonus especially involving his upper extremities and trunk. Individual #67 was seen initially at the age of 4 years and he was not born when the family was first being investigated. His parents reported they had first noticed spontaneous jerks of his arms and trunk at the age of two. These involuntary movements did interfere with his ability to perform many tasks including eating or drinking. He has slight dystonic posturing of his feet. One unusual feature is his excessive blinking that has not been seen in other family members nor is it a typical feature of myoclonus dystonia. A movement disorder neurologist following him (Anthony E. Lang) feels the excessive eye blinking is more of a tic than dystonia. However, overall his clinical features have been felt to be compatible with a diagnosis of myoclonus-dystonia.

5.3.2. *Linkage analysis*

Linkage analysis was initially performed by genotyping the new individuals of the family with linked markers on 18p11. Potential allele calling errors were checked for by the UNKNOWN program in the LINKAGE package. Unexpectedly, the recalculated LOD scores for the previously linked markers on 18p11 were reduced when the two new affected

individuals were added. Compared to our previous LOD scores of 2.78, 3.50, 2.44 and 2.03 , the maximum LOD scores with these two additional individuals were decreased to 1.60, 1.19, 1.36, and 1.05 at θ=0.10 for our linked markers D18S976, Gata185C06, D18S391 and D18S967, respectively (Table 5-1).

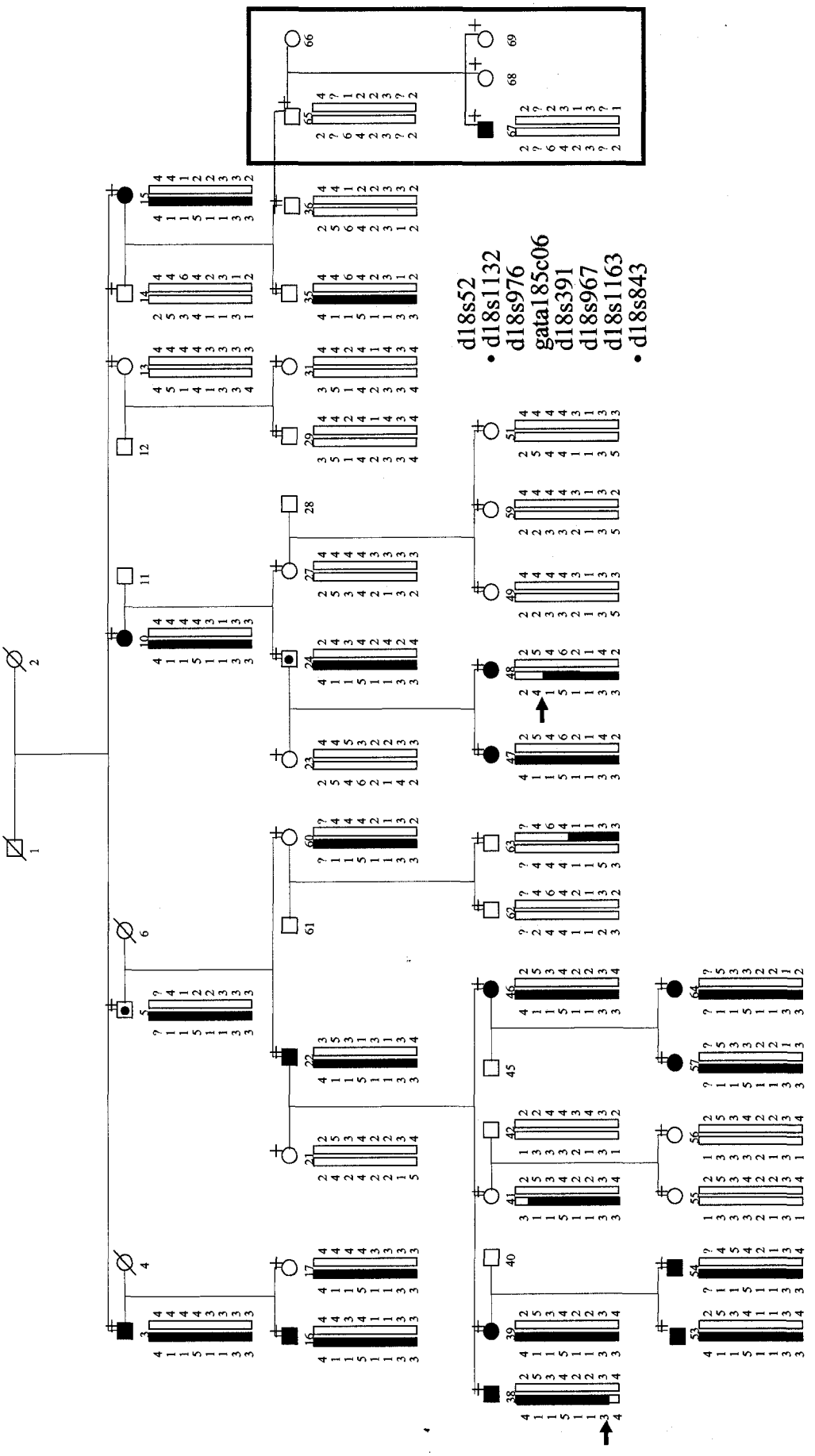
Table 5-1. Two-point LOD scores for the markers on chromosome 18p11 in the large family with an additional affected individual (#67). * Map locations are based on the Marshfield Medical Research Foundation Map.

Markers	Distance* (Kosambi), cM	LOD Score at θ= 0.00	0.01	0.05	0.10	0.20	0.30	0.40
D18S976	12.81	-1.66	1.13	1.60	1.60	1.29	0.86	0.42
GATA185C06	16.54	-0.76	-0.22	0.86	1.19	1.19	0.86	0.43
D18S391	18.70	-1.46	0.68	1.24	1.36	1.20	0.85	0.42
D18S967	19.64	-1.89	0.37	0.94	1.05	0.93	0.65	0.33

5.3.3. *Haplotype analysis for the additional individuals of the large family with markers on 18p11*

Haplotype analysis of the two additional individuals (#65 and #67) with four linked markers (D18S976, GATA185C06, D18S391 and D18S967) on 18p11 revealed that the two individuals shared a 2/3/1/3 haplotype which is different from original disease haplotype (1/5/1/1) for these four linked markers (Figure 5-1, purple box). This result is consistent with the reduced LOD scores as shown in Table 5-1. But after these two additional individuals were added to the family, the lack of a common disease haplotype suggested that the disease gene may be located somewhere other than on 18p11. Our original critical meiotic recombination events between distal markers D18S1132 and D18S976 in individual #48, and between proximal markers D18S1163 and D18S843 in individual #38, had placed the disease gene in a 17-cM region between D18S1132 and D18S843. Two obligate carriers #5 and #24 and all affected individuals carried the same locus haplotype. However, five other unaffected individuals #17, #60, #35, #41, and #63 also carried at least part of the haplotype, suggesting a reduced penetrance for M-D in this family (Figure 5-1, Black box). Because of the controversy between the original linkage analysis with 36 individuals and the current linkage analysis with 38 individuals including the two new individuals, we performed a 10 cM genomewide screen for this enlarged family.

Figure 5-1. Partial Pedigree and haplotype analysis for the two additional individuals of the large Canadian family (family 1) with the linked markers on 18p11. Blackened bars represent original disease-carrying haplotype. Filled symbols indicate individuals with definite M-D. "+" indicates individuals who have been assessed and had a DNA sample taken. Black arrows indicate the distal and proximal recombination events to define the *DYT15* locus. Individuals in red box were added after the original 25 cM screen.


d18s52
• d18s1132
d18s976
gata185c06
d18s391
d18s967
d18s1163
• d18s843

5.3.4. *A 10 cM genomewide linkage analysis for newly recruited family members of the expanded pedigree*

We performed the 10 cM genomewide screening for this expanded family with an additional 144 markers. Reanalysis of the family using the LINKAGE programs demonstrated that most markers gave negative lod scores except for markers on chromosomes17 and 20 (Table 5-2). We found that one marker on each of these two chromosomes has a LOD score greater than 1.0 with D17S2196 having LOD score of 1.51 at θ=0.10. More markers around D17S2196 were then genotyped. Both markers, D17S740 and D17S1925 beside D17S2196 gave negative LOD scores of -12.29 and -10.33, respectively, excluding the linkage on chromosome17 in this large family (Family 1). The two-point LOD scores for 10 cM markers in the family are listed in Table 5-2.

Table 5-2. Two-point LOD scores from the 10 cM genome scan of the large family with M-D. * Map locations are based on the Marshfield Medical Research Foundation Map. The LOD scores of great than 1.0 are highlighted as yellow.

Markers	Distance (Kosambi), *cM	LOD Score at θ= 0.00	0.01	0.05	0.10	0.20	0.30	0.40
Chr1								
D1S468	0.00	-13.39	-6.10	-2.91	-1.53	-0.42	-0.06	0.00
D1S3669	28.00	-13.06	-11.09	-8.27	-6.05	-3.75	-2.06	-0.88
D1S1622	49.00	-18.18	-10.87	-6.85	-4.59	-2.36	-1.16	-0.43
D1S3721	68.00	-26.28	-15.38	-8.59	-5.56	-2.72	-1.28	-0.45
D1S3728	90.00	-13.31	-10.43	-6.02	-3.90	-1.94	-0.97	-0.39
D1S1728	110.00	-10.92	-5.83	-3.05	-1.89	-0.82	-0.32	-0.08
D1S551	113.00	-19.35	-10.42	-5.67	-3.71	-1.89	-0.96	-0.38
D1S1631	137.00	-9.00	-2.97	-0.98	-0.24	0.29	0.40	0.29
D1S3723	141.00	-12.53	-4.67	-2.03	-1.00	-0.20	0.04	0.07
GATA124F08	237.00	-15.29	-4.71	-2.30	-1.24	-0.34	-0.02	0.05
D1S2141	240.00	-15.11	-3.33	-1.12	-0.28	0.27	0.32	0.18
D1S3462	263.00	-10.01	-6.21	-3.20	-1.91	-0.75	-0.23	-0.01
D1S235	269.00	-19.46	-6.99	-3.04	-1.44	-0.19	0.18	0.16
Chr2								
GATA116B01	16.00	-9.97	-3.31	-1.57	-0.78	-0.09	0.15	0.16
D2S405	50.00	-8.91	-4.89	-3.95	-2.75	-1.38	-0.69	-0.28
D2S441	81.00	-15.53	-8.38	-4.60	-2.92	-1.34	-0.56	-0.16
D2S1777	108.00	-10.47	-5.03	-2.36	-1.28	-0.38	-0.04	0.06
D2S1790	113.00	-15.66	-5.73	-3.18	-1.98	-0.84	-0.30	-0.06
D2S1334	155.00	-16.78	-11.63	-7.23	-4.95	-2.66	-1.40	-0.58
D2S442	164.00	-14.68	-8.08	-4.65	-3.07	-1.55	-0.78	-0.31
D2S1353	187.00	-14.67	-9.76	-6.73	-4.59	-2.46	-1.28	-0.51
D2S1776	198.00	-9.25	-7.77	-4.56	-3.06	-1.60	-0.81	-0.31
GATA30E06	238.00	-19.49	-7.53	-4.27	-2.69	-1.14	-0.40	-0.06
D2S434	243.00	-10.30	-3.67	-2.16	-1.46	-0.75	-0.37	-0.14
D2S427	267.00	-12.75	-3.78	-1.76	-0.96	-0.32	-0.10	-0.02
Chr3								
D3S1304	20.00	-23.20	-10.81	-6.26	-4.18	-2.13	-1.06	-0.41
D3S1259	38.00	-25.81	-13.53	-7.65	-5.01	-2.45	-1.13	-0.40
D3S3038	48.00	-16.40	-9.96	-7.25	-4.91	-2.44	-1.16	-0.43
D3S4542	95.00	-12.91	-5.76	-2.91	-1.71	-0.67	-0.24	-0.07
D3S3045	133.00	-29.33	-12.21	-6.67	-4.31	-2.01	-0.89	-0.29
D3S2459	126.00	-32.30	-12.18	-6.75	-4.32	-2.05	-0.94	-0.31
D3S3023	148.00	-26.46	-13.15	-7.15	-4.54	-2.1	-0.89	-0.25
D3S1764	167.00	-22.42	-9.33	-5.23	-3.34	-1.52	-0.64	-0.18
D3S1763	186.00	-5.52	-1.36	-0.40	0.04	0.38	0.42	0.28
D3S3053	195.00	-15.79	-4.94	-2.19	-1.08	-0.25	-0.01	0.03
D3S1262	217.00	-24.09	-11.47	-6.33	-4.02	-1.81	-0.72	-0.17
D3S2398	226.00	-15.53	-5.25	-3.03	-2.04	-1.04	-0.49	-0.16
D3S1311	241.00	-11.48	-4.96	-2.45	-1.33	-0.39	-0.06	0.01
Chr4								
D4S2639	21.00	-24.52	-8.01	-4.63	-2.97	-1.31	-0.49	-0.10
D4S2397	32.00	-11.49	-3.71	-2.01	-1.22	-0.51	-0.19	-0.04
D4S1647	101.00	-10.77	-2.17	-0.20	0.49	0.83	0.69	0.35
D4S2623	110.00	-20.93	-8.08	-4.01	-2.36	-0.95	-0.36	-0.09

D4S1644	137.00	-17.94	-7.22	-4.33	-2.82	-1.27	-0.49	-0.10
D4S1629	156.00	-13.74	-4.35	-1.67	-0.65	0.12	0.32	0.25
D4S2417	179.00	-5.40	-2.38	-1.48	-1.00	-0.50	-0.23	-0.08
D4S2367	70.00	-17.50	-6.85	-3.60	-2.08	-0.66	-0.06	0.12
D4S3248	63.00	-10.76	-3.88	-1.98	-1.14	-0.44	-0.17	-0.07
Chr5								
D5S2849	6.00	-10.38	-4.05	-2.04	-1.26	-0.62	-0.35	-0.19
D5S2505	13.00	-18.83	-11.99	-6.99	-4.69	-2.43	-1.21	-0.46
D5S2848	42.00	-15.48	-7.91	-4.44	-2.90	-1.42	-0.67	-0.24
D5S1501	92.00	-31.62	-12.50	-6.73	-4.25	-1.95	-0.83	-0.25
D5S1462	113.00	-17.66	-7.31	-3.83	-2.33	-0.97	-0.36	-0.08
D5S1453	123.00	-28.44	-10.71	-5.39	-3.24	-1.35	-0.51	-0.12
D5S816	149.00	-16.31	-6.21	-3.42	-2.22	-1.05	-0.46	-0.14
D5S1480	158.00	-19.12	-10.87	-8.00	-5.72	-2.98	-1.50	-0.58
D5S820	173.00	-15.90	-9.02	-5.39	-3.64	-1.89	-0.94	-0.35
D5S1471	188.00	-10.83	-5.39	-2.72	-1.63	-0.70	-0.31	-0.12
D5S211	204.00	-16.28	-7.99	-4.60	-2.92	-1.36	-0.60	-0.19
D5S408	221.00	-23.17	-9.07	-4.37	-2.40	-0.72	-0.07	0.09
Chr6								
D6S2427	49.00	-19.92	-5.83	-2.48	-1.20	-0.25	-0.01	0.00
D6S1009	137.00	-17.70	-7.56	-4.30	-2.78	-1.29	-0.55	-0.16
GATA184A08	147.00	-16.55	-7.59	-4.06	-2.55	-1.16	-0.50	-0.17
D6S2436	156.00	-17.00	-10.22	-6.18	-4.01	-1.91	-0.86	-0.29
D6S305	168.00	-22.88	-9.88	-6.77	-4.56	-2.34	-1.18	-0.49
Chr8								
D8S277	9.00	-20.00	-10.51	-7.72	-6.10	-3.29	-1.71	-0.73
D8S1145	31.00	-27.59	-12.57	-7.18	-4.75	-2.40	-1.16	-0.42
D8S136	39.00	-25.23	-11.05	-6.27	-4.00	-1.96	-0.97	-0.40
D8S1110	70.00	-19.21	-6.74	-4.10	-2.76	-1.40	-0.69	-0.27
GATA1A4	111.00	-4.71	-0.72	-0.08	0.15	0.28	0.23	0.10
D8S592	128.00	-23.61	-10.06	-5.83	-3.77	-1.73	-0.71	-0.18
D8S256	155.00	-6.01	-3.00	-1.12	-0.29	0.34	0.42	0.23
Chr9								
D9S921	8.00	-13.43	-7.38	-4.28	-2.80	-1.35	-0.62	-0.21
D9S1121	33.00	-9.98	-2.88	-1.14	-0.47	-0.02	0.07	0.04
D9S301	53.00	-3.64	-0.12	0.53	0.72	0.69	0.47	0.20
D9S1122	63.00	0.88	0.86	0.78	0.68	0.47	0.26	0.08
D9S257	81.00	-7.45	-4.36	-2.40	-1.45	-0.58	-0.19	-0.04
D9S938	99.00	-11.42	-3.59	-1.06	-0.15	0.40	0.40	0.19
D9S930	109.00	-7.22	-4.81	-2.71	-1.67	-0.64	-0.19	-0.02
D9S282	126.00	-17.23	-7.28	-3.80	-2.24	-0.81	-0.18	0.05
D9S2157	137.00	-17.42	-4.96	-2.59	-1.54	-0.59	-0.17	0.01
Chr10								
D10S1423	41.00	-21.54	-10.52	-5.99	-3.93	-1.98	-0.99	-0.39
D10S1426	49.00	-20.37	-12.63	-7.17	-4.68	-2.32	-1.12	-0.42
D10S1208	57.00	-23.77	-7.51	-3.93	-2.43	-1.08	-0.46	-0.15
D10S1221	72.00	-20.00	-7.68	-4.56	-2.89	-1.39	-0.69	-0.28
D10S2327	100.00	-6.62	-2.51	-1.36	-0.78	-0.26	-0.06	0.01
D10S1419	113.00	-11.58	-4.54	-2.37	-1.43	-0.57	-0.19	-0.03
D10S1237	137.00	-17.86	-9.98	-5.62	-3.64	-1.71	-0.75	-0.25

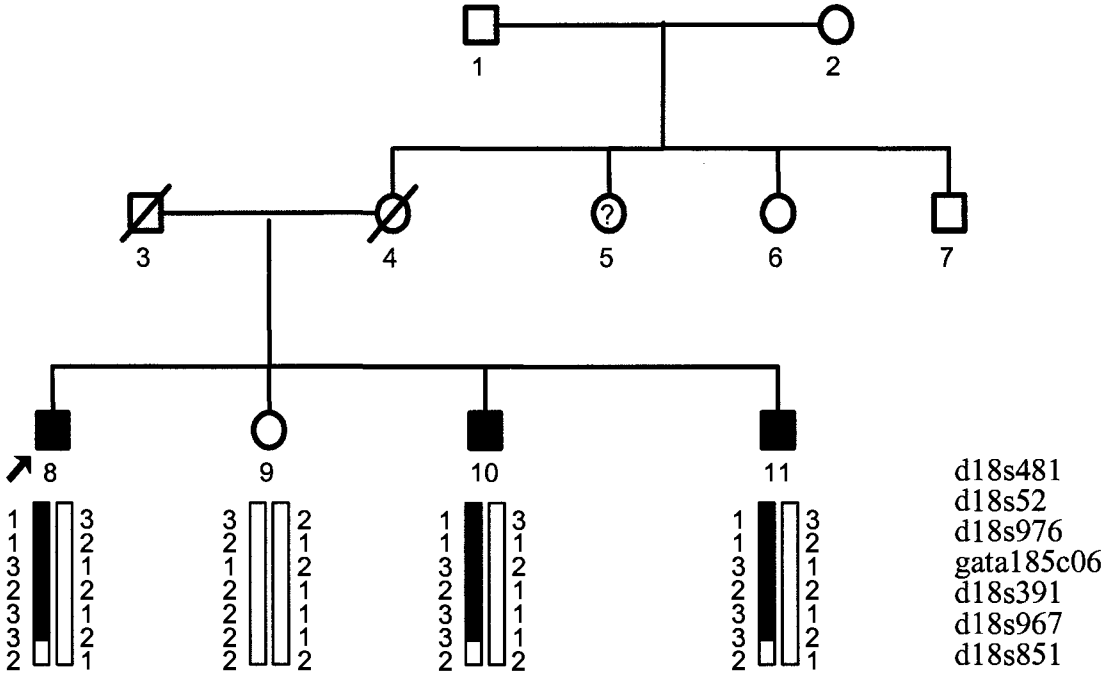
D10S1230	150.00	-26.89	-10.05	-5.09	-3.00	-1.15	-0.36	-0.04
D10S212	181.00	-17.06	-8.78	-4.57	-2.70	-1.09	-0.42	-0.14
Chr11								
D11S2362	6.00	-8.58	-5.54	-3.32	-1.85	-0.58	-0.07	0.08
D11S1981	26.00	-12.01	-8.78	-4.38	-2.40	-0.72	-0.07	0.09
GATA34E08	34.00	-21.29	-5.16	-1.83	-0.58	0.30	0.45	0.29
D11S2002	82.00	-25.44	-10.36	-5.04	-2.70	-0.78	-0.08	0.04
D11S912	130.00	-23.55	-7.08	-3.08	-1.55	-0.33	0.06	0.11
Chr12								
GATA49D12	11.00	-10.51	-2.58	-1.18	-0.60	-0.11	0.08	0.11
D12S373	26.00	-15.17	-3.40	-1.41	-0.65	-0.07	0.10	0.11
D12S398	57.00	-12.50	-7.27	-4.17	-2.72	-1.29	-0.57	-0.18
D12S1294	64.00	-10.28	-5.97	-3.10	-1.80	-0.69	-0.23	-0.04
D12S1052	74.00	-9.46	-1.90	-0.62	-0.17	0.12	0.15	0.09
D12S1300	99.00	-9.10	-3.66	-1.51	-0.62	0.07	0.24	0.17
D12S2070	127.00	-6.95	-1.45	-0.33	0.16	0.43	0.34	0.13
D12S395	136.00	-12.27	-7.24	-4.22	-2.50	-1.00	-0.38	-0.14
D12S2078	150.00	-11.26	-5.17	-2.80	-1.74	-0.77	-0.32	-0.11
Chr13								
D13S1493	19.00	-15.13	-6.92	-4.13	-2.43	-0.95	-0.35	-0.13
D13S894	28.00	-10.34	-8.16	-5.02	-3.24	-1.59	-0.76	-0.28
D13S325	34.00	-11.27	-4.00	-2.02	-1.07	-0.26	-0.01	0.03
D13S788	41.00	-25.44	-13.00	-6.88	-4.22	-1.80	-0.70	-0.17
D13S800	53.00	-16.69	-11.00	-6.40	-4.18	-2.13	-1.09	-0.44
D13S779	94.00	-22.56	-8.14	-3.80	-2.07	-0.67	-0.18	-0.03
Chr14								
D14S742	0.00	-11.98	-5.72	-2.44	-1.15	-0.14	0.18	0.18
D14S599	31.00	-0.72	-0.51	-0.05	0.21	0.37	0.32	0.18
D14S587	46.00	-20.91	-6.81	-3.59	-2.13	-0.77	-0.17	0.05
D14S1279	90.00	-16.42	-7.40	-4.43	-3.05	-1.63	-0.82	-0.31
D14S617	99.00	-21.79	-7.61	-4.26	-2.71	-1.24	-0.53	-0.18
D14S1434	108.00	-9.96	-5.00	-3.21	-2.25	-1.20	-0.59	-0.21
Chr15								
D15S165	5.00	-20.96	-11.30	-6.22	-3.88	-1.69	-0.67	-0.17
D15S1507	51.00	-11.54	-6.12	-3.86	-2.27	-0.91	-0.37	-0.13
D15S211	69.00	-20.67	-10.47	-5.96	-3.60	-1.48	-0.51	-0.08
D15S652	84.00	-13.18	-4.95	-2.43	-1.32	-0.36	-0.01	0.08
D15S816	92.00	-6.49	-6.33	-5.47	-3.82	-1.99	-0.98	-0.04
Chr16								
D16S748	11.00	-4.91	-2.34	-0.90	-0.31	0.12	0.19	0.11
D16S403	33.00	-29.68	-11.53	-5.77	-3.36	-1.26	-0.36	0.00
D16S3396	50.00	-12.48	-7.06	-3.51	-2.02	-0.72	-0.20	-0.03
D16S2620	71.00	-17.21	-9.68	-5.36	-3.41	-1.57	-0.67	-0.20
D16S516	92.00	-18.40	-11.29	-6.23	-3.94	-1.77	-0.71	-0.17
D16S402	106.00	-16.39	-4.90	-2.10	-0.96	-0.05	0.21	0.17
Chr17								
D17S1298	12.00	-15.54	-8.17	-4.72	-3.01	-1.43	-0.65	-0.2
D17S974	28.00	-17.7	-8.22	-4.89	-3.18	-1.54	-0.17	-0.25
D17S740	44.62	-12.29	-4.49	-1.19	-0.93	-0.2	0.03	0.06
D17S1925	44.62	-10.33	-4.21	-2.55	-1.66	-0.74	-0.28	-0.06

D17S2196	47.32	-3.10	0.85	1.41	1.51	1.29	0.86	0.37
D17S1294	58.00	-22.63	-8.92	-4.78	-2.89	-1.16	-0.42	-0.11
D17S1299	69.00	-14.74	-8.51	-6.11	-4.18	-1.95	-0.89	-0.33
D17S2180	73.00	-0.71	-0.43	0.01	0.20	0.26	0.20	0.10
Chr19								
D19S1034	13.00	-9.32	-4.36	-2.31	-1.47	-0.73	-0.38	-0.15
D19S714	36.00	-4.84	-1.13	-0.41	-0.14	0.03	0.03	0.00
D19S178	63.00	-8.89	-5.85	-3.30	-1.97	-0.81	-0.32	-0.10
D19S589	88.00	-17.33	-8.94	-4.81	-3.02	-1.37	-0.58	-0.18
Chr20								
D20S103	0.00	-6.18	-3.42	-1.85	-1.13	-0.49	-0.22	-0.10
D20S851	25.00	-9.70	-3.00	-0.04	0.52	1.03	0.90	0.45
D20S604	32.00	-4.56	-0.80	0.29	0.68	0.83	0.64	0.29
D20S477	45.00	-8.78	-2.89	-0.99	-0.31	0.13	0.18	0.11
Chr21								
D21S1437	12.00	-23.45	-7.26	-3.36	-1.71	-0.35	0.11	0.17
Chr22								
GCT10C10	19.00	-13.33	-8.69	-5.04	-3.32	-1.63	-0.76	-0.27
D22S445	49.00	-9.93	-2.78	-1.32	-0.71	-0.21	-0.05	-0.01

5.3.5. *Linkage analysis and haplotype analysis for a family with focal dystonia*

Linkages to 7q21 and mutations in *SGCE* have been excluded in a family with focal dystonia (family 10). Clinically, in this family individual #8 was a 69-year-old man who first noticed symptoms of writer's cramp at the age of 59. Neurological examination showed only a focal dystonia involving his right hand, brought out by attempts to write. Individual #10 was the 65-year-old brother of #8 who developed writer's cramp of right hand in his mid 50's. Six years later, he experienced episodes of postural tremor involving both arms. Individual #11 was the 75-year-old brother who developed resting tremor and difficulty writing with his right hand at the age of 68. His symptoms did not improve following the treatment with levodopa and did not progress over the seven year period. Individual #5 was a 93-year-old maternal aunt living in Europe who was diagnosed with idiopathic PD. Here we assumed she is possibly affected as other affected family members had rest tremor and writing difficulties that could be mistaken for Parkinson's disease. Genotyping was performed with linked markers on 18p11 in this family and disease haplotype with linked markers on 18p11 was found in three affected (Figure 5-2), providing additional evidence for our *DYT15* locus on 18p11. Haplotype analysis of this family showed a consistent region of 73 cM in the three affected siblings overlapping our 18p11 locus. As this family could have the same disease gene as our large family, we included one affected individual of this family in our sequencing set to search for mutations in candidate genes in the *DYT15* locus (**Chapter 6**).

Figure 5-2. Haplotype analysis for a focal dystonia family with linked markers on 18p11 (family10). Filled symbols indicate affected individuals. Linkage data generated with chromosome 18 markers were shown. Boxed black bars represent the disease-carrying haplotype. A73 cM region flanked by D18S481 and D18S851 was shared in three affected siblings that overlapped our *DYT15* locus. The question mark represents possibly affected. The unaffected individual (#9) did not carry the disease haplotype.


1
2
3
4
5
6
7
8
9
10
11
d18s481
d18s52
d18s976
gata185c06
d18s391
d18s967
d18s851

5.4. Discussion

Microsatellite DNA markers are small arrays of simple sequence repeats (usually less than 10 bp). They are interspersed throughout the genome, accounting for 2% of the human genome. Arrays of dinucleotide repeats such as CA, TG, AT, AG are the most common type with about 0.5% of the human genome. Trinucleotide repeats and tetranucleotide repeats are comparatively rare, but are often highly polymorphic markers. A marker is any polymorphic Mendelian character that can be used to follow a chromosomal segment through a pedigree. The critical requirement for a useful genetic marker is that it should be sufficiently polymorphic that a randomly selected person has a good chance of being heterozygous. The sequence of the human genome has produced more than 10,000 highly polymorphic microsatellite markers for mapping disease loci. Generally, a 10-cM genomewide genotyping set contains about 400 microsatellite markers including the di-, tri-, and tetra-nucleotide repeat markers.

Previously, using a 25 cM genetic linkage approach, we mapped a novel locus for M-D on 18p11 based on finding two-point and multipoint LOD scores of greater than 3.0. The highest two-point lod score of 3.5 at θ=0.00 was obtained with marker GATA185C06 (also known as D18S1363). Multipoint analysis of this family resulted in a plateau of 3.9 across the entire region from D18S976 and D18S967. Haplotype analysis revealed critical meiotic-recombination events between markers D18S1132 and D18S976 in individual #48, and between D18S1163 and D18S843 in individual #38 placing the disease gene in a 17 cM region between D18S1132 and D18S843. This novel locus has been designated *DYT15* and has the MIM reference number of 607488.

Phenotypic heterogeneity had already been recognized in this disorder. The clinical heterogeneity has resulted in numerous different terms being used to describe this condition (Asmus and Gasser, 2004; Kyllerman *et al.*, 1990; Mahloudji and Pikielny, 1967; Przuntek and Muhr, 1983; Quinn, 1996). Recognition that both myoclonus and dystonia may occur to varying degrees even within one kindred has brought about the current terminology of M-D. In this large Canadian family, all affected individuals had myoclonus and four (five if we include an additional individual #67) of whom also had dystonia involving their limbs.

There is other evidence to suggest our chromosome 18p11 linkage is correct in addition to our other linked family (Family 10) with focal dystonia. A family (Family K) described as having an autosomal dominant form of focal dystonia has their disease linked to a similar area on chromosome 18 (locus named *DYT7* [MIM 602124]) (Leube *et al.*, 1996). This might suggest that some cases of myoclonus-dystonia share a similar patho-physiologic abnormality to focal dystonia. In this German kindred, most of the affected individuals had a focal cervical dystonia, with one having a focal limb dystonia. A "postural hand tremor" was described in three definitely and three possibly affected persons. One other individual who carried the disease haplotype also was described as having a "postural hand tremor." Postural tremor has been described in other families with M-D (Kyllerman *et al.*, 1990; Vidailhet *et al.*, 2001) and more rhythmic myoclonus can be mistaken for postural tremor. Initial linkage and haplotype analyses placed the region telomeric to D18S843. Further studies in one possibly affected individual suggested that the region was centromeric to D18S52, which would overlap completely with the *DYT15* locus in our Canadian kindred. The *DYT7* gene was subsequently placed within a 6-cM region (roughly 3 cM telomeric to our region) by a combination of linkage analysis and allelic association studies using

additional individuals with dystonia (Leube *et al.*, 1997a; Leube *et al.*, 1997b). However, this allelic association has been disputed by the original authors (Leube and Auburger, 1998) and others (Klein *et al.*, 1998). If the two conditions do share a common gene locus, our data would suggest that the disease gene interval of Family K has been reduced from its original 30 cM region to 17cM. Other disorders that have myoclonus as a feature have not been associated with the 18p region, but dystonia has been reported in patients with the "18p deletion syndrome" (Klein *et al.*, 1999c; Tezzon *et al.*, 1998). These patients typically have mental retardation and obvious dysmorphic features that are not present in the family we studied. A common large 49-cM deleted area was found in three individuals with dystonia that incorporates our region (Klein *et al.*, 1999c; Tezzon *et al.*, 1998). Even if the two conditions share a similar etiologic basis for the abnormal movements, this knowledge does not narrow our region of interest. Another study showed that three families excluded from *SGCE* mutations on 7q21, also have an 18p11 haplotype, providing further evidence for the existence of the *DYT15* locus for M-D (Schule *et al.*, 2004a).

Unexpectedly, after we genotyped more family members of the expanded pedigree, the affected individual #67 and his father #65 shared a haplotype which is different from the 18p11 disease haplotype in other affected family members. There are four possible explanations for this controversy. 1. These two additional individuals are unrelated to the family. 2. Individual #67 is mis-diagnosed with M-D. 3. The affected individual #67 is a phenocopy. 4. The disease gene is located elsewhere.

We have excluded the possibility these new individuals are not related to the rest of the family when we performed the 10 cM genome screen. These individuals' alleles are consistent with those found in the rest of the family throughout the genome. The second

possibility is one of error in diagnosis. Unfortunately, no confirmatory or ancillary testing (except for those families with *SGCE* mutations) is helpful in diagnosing myoclonus-dystonia. His diagnosis was made by Anthony E. Lang, our collaborator from Toronto who is an expert in movement disorders and all of individual #67's clinical features are in keeping with this diagnosis except for his eye tics. There is a possibility that individual #67 is a phenocopy; however M-D is a rare disease which would make this unlikely. We did sequence the entire *SGCE* gene in this affected individual and we did not find a mutation. The final possibility is that the disease gene for the family may NOT be located at the 18p11 locus, however after a 10 cM genome scan we could not find evidence of a second locus other than 18p11.

Chapter 6.

Refinement of the Gene Locus *DYT15* for Myoclonus-Dystonia and Exclusion of the Candidate Genes

6.1. Introduction

M-D is an autosomal dominant neuromuscular disorder characterized, in addition to dystonia, by a predominance of alcohol responsive myoclonic jerks mostly in the arms and axial muscles. As the prevalence of M-D is unknown, it may be more common than is currently appreciated with more familial and sporadic cases being reported. Since the identification of mutations causing M-D in the epsilon-sarcoglycan (*SGCE*) gene on 7q21 in six Germany families, several other groups including our lab have reported additional mutations (**Chapter 3**).

Most investigators screening for mutations in the *SGCE* gene in families with M-D have identified mutations in less than half of the cases. Our lab originally screened the *DRD2* variants in a large five-generation Canadian family with M-D (Grimes *et al.*, 2001) and excluded mutations in the *SGCE* gene (*DYT11*) and linkage for *DYT11* on 7q21 using direct sequencing (**Chapter 3**), semi-quantitative gene dosage analysis (**Chapter 4**) and linkage analysis (Grimes *et al.*, 2002). Neither *DRD2* nor *SGCE* genes appeared to play a role in this Canadian family. Following a 25 cM genome screen and linkage analysis, we mapped a novel locus for M-D on chromosome 18p11 (Multipoint Max LOD score of 3.96). This locus is now designated *DYT15* (OMIM number: 607488). In this family, the clinical characteristics of M-D are identical to those whose disease is due to a mutation in *SGCE* gene. Haplotype analysis for this family had revealed critical meiotic-recombination events between distal markers D18S1132 and D18S976 in affected individual #48 and between proximal markers D18S1163 and D18S843 in affected individual #38, placing the disease gene to a 17 cM region (4 Mb) between markers D18S1132 and D18S843 as we reported (Grimes *et al.*, 2002). This locus has also been implicated in a family of German-

origin with idiopathic dystonia (Leube *et al.*, 1996). After a genome-wide linkage screening this group detected one marker with a LOD score of greater than 2.0 (D18S62) and a disease haplotype of D18S967/S62/S471/S458/S452 was identified with a maximal LOD score of 3.17 at D18S452. Three recombination events in definitely affected individuals of the family had shown that the locus was mapped to a region flanked by D18S843 and D18S52, which encompasses the *DYT15* locus on 18p11. We also have a family with focal dystonia (family 10) which was shown to be linked to our *DYT 15* locus, but the haplotype analysis with linked markers on *DYT15* in this family could not narrow down the *DYT15* locus (**Chapter 5**).

Recently, we determined genotypes for single nucleotide polymorphisms (SNPs) occurring between the flanking recombinant and non-recombinant markers of our critical region. As expected, we were able to identify three definite recombinant SNPs in the proximal arm between the markers D18S1163 and D18S843 and narrowed down the *DYT15* locus to a 3.18 Mb from the original region of 4 Mb. We report here our effort to narrow this critical region using single nucleotide polymorphisms (SNP) and the sequencing results of known and predicted genes within this region.

6.2. Subjects and Methods

6.2.1. *Subjects*

Informed consent was obtained from all the participants of 18 familial and sporadic cases (Families 1-18, **Chapter 2**) after approval by the institutional review board at The Toronto Hospital and The Ottawa Hospital. Individuals had a detailed history, physical examination, and videotape recorded using a standard questionnaire and videotape protocol. All patients and their relatives in the eighteen families were examined by neurologists

specialized in movement disorders and the clinical characteristics of these families were shown in **Chapter 2**. The diagnosis of M-D was made according to the established criteria (**Chapter 1**). Genomic DNA was extracted from whole blood of the patients and used for genotyping, sequencing, and 500K SNP array analysis.

6.2.2. Single nucleotide polymorphism (SNP) genotyping to narrow down the 18p11 region

SNPs were chosen to span the distal flanking region and the proximal flanking region of the disease haplotype from our large Canadian family (family 1). The physical position of each SNP was confirmed using The SNP Consortium database (http://snp.cshl.org). The SNPs for fine mapping were only chosen if their minor allele frequencies were equal to or greater than 30%. Primers for amplifying each SNP were designed by the Primer 3 program. After PCR amplification of the SNPs, we directly sequenced the PCR products to determine the genotype (G, or A, or T, or C) using the P^{33}-radiolabeled kit (**Chapter 2**). Six evenly distributed SNPs on each flanking arm were selected for the initial genotyping between the markers D18S1132 and D18S976 and between the markers D18S1163 and D18S843. Once a recombinant SNP had been identified, more SNPs were chosen to genotype the critical affected individuals to further refine the region. Totally we used 30 SNPs to genotype partial individuals of family 1 on the distal flanking arm and proximal flanking arm of the 18p11 region.

6.2.3. Sequence analysis of candidate genes in the critical region of chromosome 18p11

We prioritized the candidate genes within the critical region based on known expression in brain, known function, phenotype of known knockout mouse models of the candidate genes. All exons and exon-intron boundaries of the candidate genes were

sequenced. We have been using the Thermo Sequenase Radiolabled Terminator Cycle Sequencing Kit (USB Corp., Cleveland, OH) or the AB3130xl genetic analyzer using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystem, Foster City, CA) to sequence the PCR products of the SNPs and exons of candidate genes as described in **Chapter 2**. Our goal was to identify the mutation for family 1, but family 10 also showed support for linkage to *DYT15*. Therefore our panel of affected patients who were sequenced, included two individuals from family 1, two affected individuals of families 2, affected individuals from families 5, 6, 8, 9, 10 and 12 (UK) (families 2 and 6 had not been identified to have large deletions of the *SGCE* gene at that time). Families 3, 4, and 7 have been excluded from our sequencing set as mutations in the *SGCE* gene were identified in these three families (**Chapter 3**).

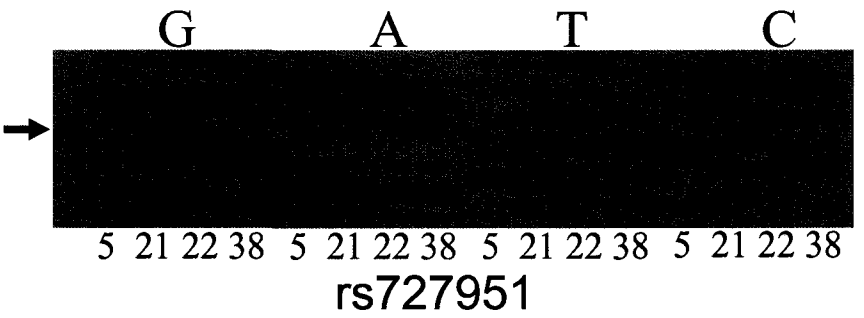
6.3. Results

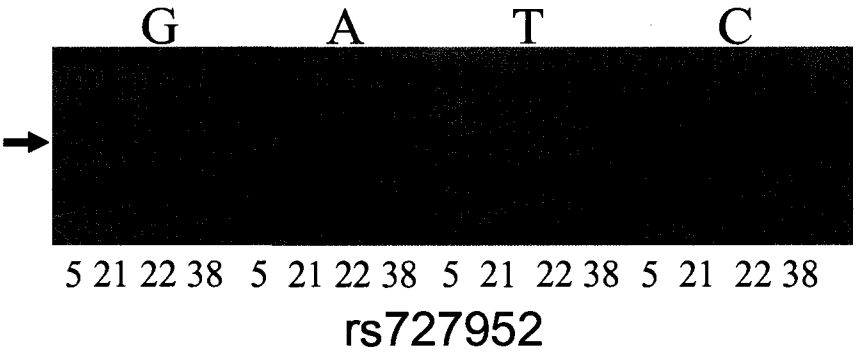
6.3.1. *Fine mapping of the DYT15 locus using SNPs*

With the high density of SNPs for gene mapping, we decided to refine the recombinational events by genotyping the key family members using SNPs spanning the distal flanking region between markers D18S1132 and D18S976 as well as the proximal flanking region between the markers D18S1163 and D18S843. As a result, out of the first six genotyped SNPs we identified two definite recombinant SNPs (rs727951, and rs727952) for family 1. Subsequently, another six SNPs between the recombinant SNP (rs727951) and the non-recombinant marker (D18S1163) were further genotyped and one additional recombinant SNP was identified (rs385769) (Figures 6-1 and 6-2). The proximal recombinant marker has now been designated as rs385769 with its physical position at 7790480 on18p11. We also genotyped SNPs within the distal flaking region without

identifying any recombinant SNP marker in family 1. Taken together, the *DYT15* locus has been reduced to a 3.18 Mb critical region flanked by rs385769 and D18S1132 (Figure 6-3).

Figure 6-1. Genotyping the proximal SNP markers for key individuals from family 1. Three recombinant SNP markers were identified (rs727951, rs727952, and rs385769). Black arrows indicate the nucleotide positions of rs727951, rs727952, and rs385769. Individuals 5, 21, 22 and 38 are indicated at the bottom of each panel.


G A T C
5 21 22 38 5 21 22 38 5 21 22 38 5 21 22 38
rs727951


G A T C
5 21 22 38 5 21 22 38 5 21 22 38 5 21 22 38
rs727952

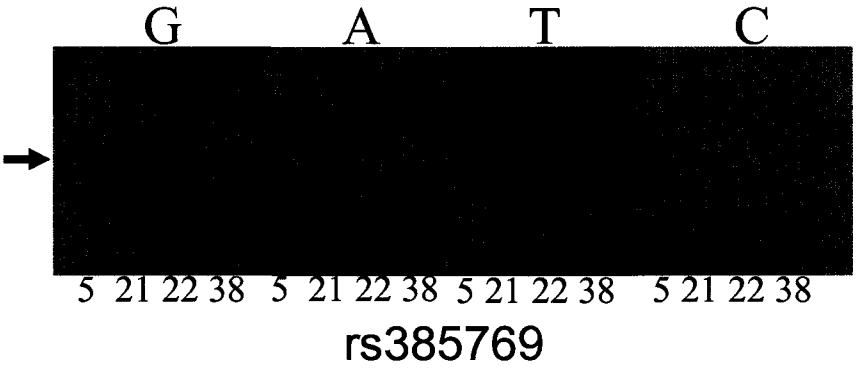

G A T C
5 21 22 38 5 21 22 38 5 21 22 38 5 21 22 38
rs385769

Figure 6-2. Haplotype analysis for key individuals of family 1 identified the proximal recombinant SNPs at *DYT15* locus. Black arrows indicate the previous proximal recombinant (R) and non-recombinant (NR) markers on chromosome 18p11 (grey bar) as shown in Figure 5-1. Red arrows indicate the current proximal boundary. Those SNPs showing a recombination are marked with an asterisk.

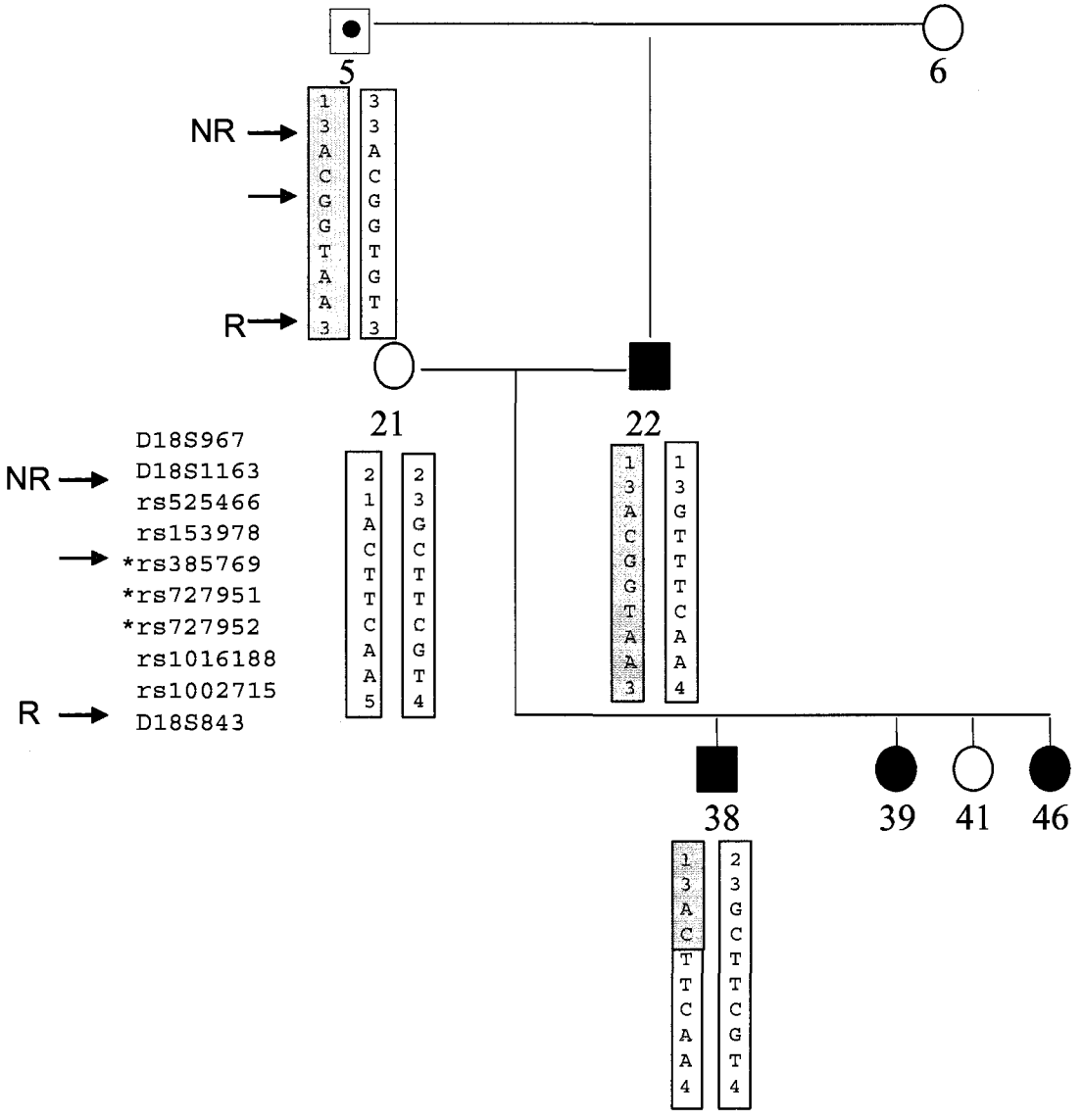

5
6
NR
R
21
22
D18S967
NR
D18S1163
rs525466
rs153978
*rs385769
*rs727951
*rs727952
rs1016188
rs1002715
R
D18S843
38
39
41
46

Figure 6-3. Schematic representation of candidate genes in the reduced 3.18 Mb region for *DYT15* locus on 18p11. The previous flanking recombinant (R), non-recombinant (NR) and the newly identified recombinant (red) are indicated as nucleotide positions of these markers. Physical distances are determined according to NCBI (http://www.ncbi.nlm.nih.gov/mapview). The previous proximal recombinant marker was D18S843. Additional recombinant SNPs (rs385769, rs727951 and rs727952) between D18S1163 and D18S843 were also detected.

Candidate genes
Markers
Genomic position
LOC390826
MGC17515
LOC390827
ZFP161
KIAA0987
(EPB41L3)
LOC388459
Ak093255
(L3MBTL4)
LOC388460
LOC388461
LOC400643
FLJ10312
(ARHGAP28)
LAMA1
LOC390828
LOC339291
PTPRM
D18S1132 R
4609908
D18S976 NR
5238944
D18S1363 NR
5327461
D18S391 NR
5771224
D18S458 NR
5931032
3.18 Mb
D18S967 NR
6585300
rs603258 NR
7005716
D18S1163 NR
7462347
rs153978 NR
7770445
rs385769 R
7790480
rs727951 R
7830668
rs727952 R
7830952
0.81Mb
centromere
D18S843 R
8603770

6.3.2. *Screening for mutations in candidate genes at the DYT15 locus*

With the complete sequence of the human genome, we can identify candidate genes within the *DYT15* locus. Based on the physical map of NCBI Build 36.2 and Genome Browser at UCSC, seven known candidate genes, *MGC17515, ZFP161, EPB41L3 (KIAA0987), L3MBT4 (*Homo sapiens l(3)mbt-like 4), *ARHGAP28 (FLJ10312), LAMA1, and PTPRM* and eight predicted genes, *LOC390826, LOC390827, LOC388459, LOC388460, LOC400643, LOC388461, LOC390828 and LOC339291* have been mapped in this region (Figure 6-3). We have sequenced all the known genes and four of the predicted genes in the region and have not detected any mutations; however we have identified numerous SNPs for each of the candidate genes (Table 6-1). The SNPs will be prefixed with rs# if those SNPs that we found were listed in dbSNP database of NCBI (http://www.ncbi.nlm.nih.gov/SNP/).

Table 6-1. The identified SNPs for the known and predicted genes at the *DYT15* locus for M-D. Notes: C=Confirmed gene-model based on alignment of mRNA, or mRNAs plus ESTs; P=Predicted only-model predicted by GenomeScan; E=EST evidence; I=Interim Locus ID, model based on alignment of mRNAs, or mRNAs plus ESTs, to the genome. SNPs=number of identified single nucleotide polymorphisms. ND= not done.

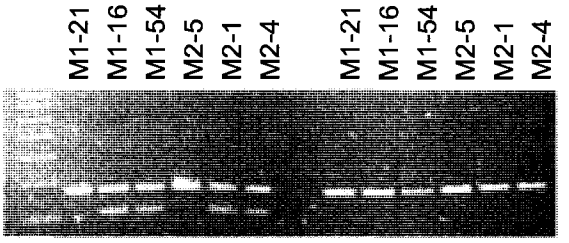
Genomic start	Genomic end	Gene name	Alternate name	Status	SNPs	Description of the genes
4992751	4993240	*LOC390826*		P	ND	
4992751	5123389	*LOC390827*		P	ND	similar to peptidyl-Pro cis trans isomerase
5221724	5233028	*MGC17515*	*C18orf18*	C	5	BC010538
5278979	5283713	*ZFP161*		C	1	zinc finger protein 161 homolog (mouse
5381988	5534386	*EPB41L3*	*KIAA0987*	C	3	erythrocyte membrane protein band 4.1-like 3
5615069	5882347	*LOC388459*		E	8	hypothetical LOC388459
6018986	6405301	*L3MBTL4*	*AK093255*	C	23	l(3)mbt-like 4 (Drosophila)
6451743	6453409	*LOC388460*		E	0	similar to 60S ribosomal protein L6
6718528	6742150	*LOC388461*		E	7	hypothetical LOC388461
6778093	6889118	*LOC388462*	*ARHGAP28*	C	16	similar to KIAA1314 protein, FLJ10312
6915077	6918951	*LOC400643*		I	3	hypothetical LOC400643
6931488	7108213	*LAMA1*		C	29	laminin, alpha 1
7124529	7165250	*LOC390828*		P	ND	similar to mitochondrial carrier triple repeat 1
7221137	7222042	*LOC339291*		P	ND	similar to RIKEN cDNA 2610040E16
7557417	8397254	*PTPRM*		C	5	protein tyrosine phosphatase, receptor type, M

6.3.3. *Sequencing of the candidate gene LAMA1*

LAMA1 (laminin, alpha 1 precursor) was considered the strongest candidate gene in our region based on its location, function and expressions in different human tissues such as brain and muscle. The *LAMA1* gene is composed of 62 exons spanning a genomic region of 176,725 bp on 18p11. Its mRNA is 9530 bp (NM_005559) and codes for a protein of 3075 amino acids. We have systematically sequenced the promoter, all 62 exons, the exon-intron boundaries, the 5'- and 3'-UTRs of *LAMA1*. In total we have identified 29 SNPs in the *LAMA1* gene by sequencing families 1, 2, 5, and 6. There were two SNPs that could be potential mutations as they were only found in affected individuals. One SNP located 7 bp downstream of exon 22 was first identified in both affected individuals of families 1 and 2 but not in normal controls. This A→G transition could create potentially an alternate splicing site for the expression of a different transcript of the *LAMA1* gene. We were able to find a restriction enzyme (AfeI) that can cut the G allele without cutting the A allele. We confirmed the results of AfeI digestion in the individuals we sequenced and then screened for this G allele in normal controls (Figure 6-4, A). We found that some unaffected individuals and normal controls also had the G allele, excluding this A→G change as a mutation (Figure 6-4, B). This change has recently been reported in the dbSNP and designated as rs603258. Another possible mutation was the SNP in exon 44 (exon 44B of Table 6-2). This C→T transition changes the amino acid threonine to isoleucine in an affected individual (M6-1 in Figure 6-4, C). This change destroys the restriction site for ApoI which can cut the C allele without cutting the T allele. We screened this change in normal controls and found that eight of twenty-four normal controls also had this T allele (Figure 6-4, D), excluding this change as a mutation.

Figure 6-4. Restriction enzyme digestions of two SNPs in the *LAMA1* gene. In panel **A**, G/A polymorphism (rs603258) in intron 22 of *LAMA1* were shown in affected individuals of families 1 and 2. In panel **B**, rs603258 polymorphism was also seen in unaffected controls (M1-49 and M1-62). In Panels **C** and **D**, a C →T transition mutation in exon 44 changes amino acid threonine to isoleucine in an affected individual (M6-1 in panel **C**) and normal controls (panel **D**). This change destroys a restriction site for ApoI which can cut the wild-type, but the mutant. Eight of the twenty-four normal individuals also have this T allele (panel **D**). M1-49 and M1-62 are unaffected individuals of family 1; 93991 is an affected individual with chin tremor; Individuals P41 to P137 are normal controls.

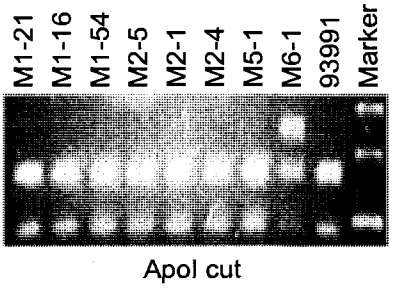
A.


M1-21
M1-16
M1-54
M2-5
M2-1
M2-4
M1-21
M1-16
M1-54
M2-5
M2-1
M2-4
AfeI cut
Uncut

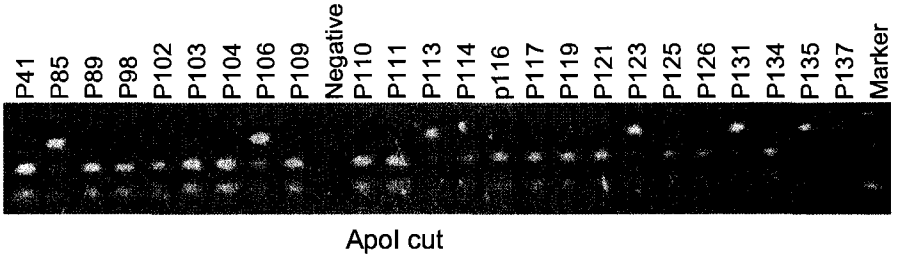
B.


Marker
M1-3
M1-22
M1-47
M1-39
M1-46
M1-13
M1-27
M1-35
M1-49
M1-53
M1-56
M1-62
M1-47(-AfeI)
M1-62(-AfeI)
AfeI cut
Uncut

C.


M1-21
M1-16
M1-54
M2-5
M2-1
M2-4
M5-1
M6-1
93991
Marker
ApoI cut

D.


P41
P85
P89
P98
P102
P103
P104
P106
P109
Negative
P110
P111
P113
P114
p116
P117
P119
P121
P123
P125
P126
P131
P134
P135
P137
Marker
ApoI cut

Interestingly, between exons 22 and 44 of *LAMA1*, there is a 38 kb region in which the affected individuals of families 1 and 2 are homozygous, implicating a possible large genomic deletion shared between the affected individuals of families 1 and 2 (Table 6-2). In fact this region of homozygosity may extend to exon 60 because no heterozygous SNPs were identified in exons 44-60 of the *LAMA1* gene. We have used a semi-quantitative DHPLC method to analyze the gene dosage of the homozygous region (Figure 6-5). We chose exon 60 of the *DMD* gene (dystrophin) located on the X chromosome and quantified the copy number of the *DMD* exon 60 in male and female individuals by comparing the DHPLC peak areas of male and female individuals. Upon confirming that we could detect a significant dosage difference between male and female *DMD* exon 60, we used a multiplex PCR to amplify the *DMD* exon 60 with *LAMA1* exons 2, 24, 32, and 42 and compared the dosage of *LAMA1* exon 2 outside the region of homozygosity to exons 24, 32 and 42 located inside of the homozygous region among the affected and unaffected individuals. No dosage difference was found in this region for the affected individuals (M1-22, M2-4), excluding the possibility of large deletion in this homozygous region. Subsequently, we have identified a large deletion in the *SGCE* gene in family 2, indicating the existence of this large region of homozygosity may not represent a rare event. All other SNPs identified in the *LAMA1* gene were summarized in Table 6-2 and Figure 6-6.

Table 6-2. Detected single nucleotide polymorphisms (SNPs) in the *LAMA1* gene for *SGCE*-negative families with M-D. Notes: M1-16 and M1-54 are affected individuals from family 1 while M1-21 is an unaffected individual who married into family 1. M2-1 and M2-4 are affected individuals from family 2 while M2-5 is an unaffected individual who married into family 2. M5-1 and M6-1 are affected individuals from families 5 and 6. N/N indicates unknown nucleotide. Homozygous regions are highlighted in yellow. The flanking heterozygous SNPs are highlighted in light blue.

			Affected and unaffected individuals								
Exons	Base Position	SNP#	M1-21	M1-16	M1-54	M2-5	M2-1	M2-4	M5-1	M6-1	Mutation status
Exon 8	7033457	novel	C/C	A/C	A/A	C/C	C/C	C/C	C/C	C/C	Intron 7, 60 bp upstream
Exon 14	7024510	reported	A/A	A/A	A/A	A/A	A/C	A/C	A/C	A/A	ACC(Threonine)/AAC(Asparagine), missense
Exon 15	7023039	reported	T/C	C/C	C/C	C/C	T/C	T/C	T/C	C/C	GCC/T(Alanine), Wobble
Exon 18	7014345	rs548577	T/C	C/C	C/C	C/C	C/C	C/C	T/C	C/C	Intron 18, 34 bp downstream
Exon 19	7013373	reported	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T	TGT/C (Cysteine), Wobble
Exon 20	7007324	novel	T/C	T/C	C/C	T/T	T/C	T/C	T/T	C/C	TGT/C(Cysteine), Wobble
[redacted]	7005716	rs603258	A/A	[redacted]		A/A	[redacted]		A/A	A/G	Intron 22, 7 bp downstream, AGgtgagca/g
Exon 24	7002125	novel	T/C	T/T	T/T	C/T	T/T	T/T	T/T	T/C	GGT/C (Glycine), Wobble,
Exon 25A	7001415	rs619106	C/T	C/C	C/C	C/T	C/C	C/C	C/C	C/C	ACC/T (Threonine), Wobble
Exon 25B	7001407	novel	T/G	T/T	T/T	T/G	T/T	T/T	T/T	T/T	GTT (Valine)/GGT (Glycine), missense
Exon 28A	6998593	rs662741	A/A	A/A	A/A	A/G	A/A	A/A	A/A	N/N	ATG(Methionine)/GTG(Valine), missense
Exon 28B	6998585	novel	G/T	T/T	T/T	T/T	T/T	T/T	T/T	X/X	GTT/G(Valine), Wobble
Exon 31	6989965	novel	C/C	C/C	C/C	C/C	C/C	C/C	C/C	T/C	CAC/T (Histidine), Wobble
Exon 32	6989628	rs625106	G/G	A/A	A/A	G/A	A/A	A/A	A/A	G/A	TCA/G (Serine), Wobble
Exon 33	6987820	novel	T/T	T/T	T/T	G/T	T/T	T/T	T/T	G/T	GCT (Alanine)/TCT(Serine), missense
Exon 35	6983675	reported	A/A	A/A	A/A	G/A	A/A	A/A	A/A	A/A	ATA (Isoleucine)/GTA (Valine), missense
Exon 36	6982509	rs7243349	G/G	A/A	A/A	G/G	A/A	A/A	A/G	G/G	Intron 36, 43 bp downstream
Exon 37A	6976261	reported	T/T	C/C	C/C	T/C	C/C	C/C	T/T	C/C	CAT/C (Histidine), Wobble
Exon 37B	6976227	rs12607841	C/C	C/C	C/C	T/C	C/C	C/C	C/C	C/C	GCG (Alanine)/GTG (Valine), missense
Exon 38A	6975633	reported	G/G	A/A	A/A	G/G	A/A	A/A	G/G	G/G	CTG/A (Leucine), wobble
Exon 38B	6975509	novel	C/C	T/T	T/T	C/C	T/T	T/T	C/C	C/C	Intron 38, 19 bp downstream
Exon 38C	6975508	novel	G/G	G/G	G/G	G/G	G/G	G/G	G/G	A/A	Intron 38, 20 bp downsream
Exon 38D	6975501	reported	G/G	G/G	G/G	A/G	G/G	G/G	G/G	G/G	Intron 38, 27 bp downstream
Exon 38E	6975470	reported	T/T	T/T	T/T	G/T	T/T	T/T	T/T	T/T	Intron 38, 96 bp downstream
Exon 42A	6970525	reported	A/A	G/G	G/G	G/A	G/G	G/G	A/A	A/A	GAA (Glutamic acid)/AAA (Lysine), missense
Exon 42B	6970489	novel	A/A	G/G	G/G	G/A	G/G	G/G	A/A	G/G	Intron 42, 33 bp downstream
Exon 44A	6967846	novel	T/T	T/T	T/T	T/T	T/T	T/T	T/T	C/T	ACT (Threonine)/ATT (Isoleucine), missense
Exon 44B	6967721	novel	T/T	T/T	T/T	T/T	T/T	T/T	T/T	C/T	Intron 44, 8 bp downstream, Ctgtgagtt/c
[redacted]	6933266	reported	C/C	[redacted]		C/C	[redacted]		C/C	C/T	3'-UTR

Figure 6-5. DHPLC analysis of multiplex PCR products of the *LAMA1* gene (M-D) and exon 60 of the *DMD* gene. M1-21 is a non-affected female of family 1, M1-22 is an affected male of family 1, M2- 5 is a non-affected female of family 2, M2- 4 is an affected male of family 2. The sizes of PCR products for exons 2, 22, 24, 42 and 60 are 371 bp, 327 bp, 286 bp, 243 bp and 139 bp respectively. There is no gene dosage change for the *LAMA1* homozygous region between unaffected and affected individuals, indicating that there is no large deletion in the *LAMA1* homozygous region in the affected individuals of families 1 and 2.

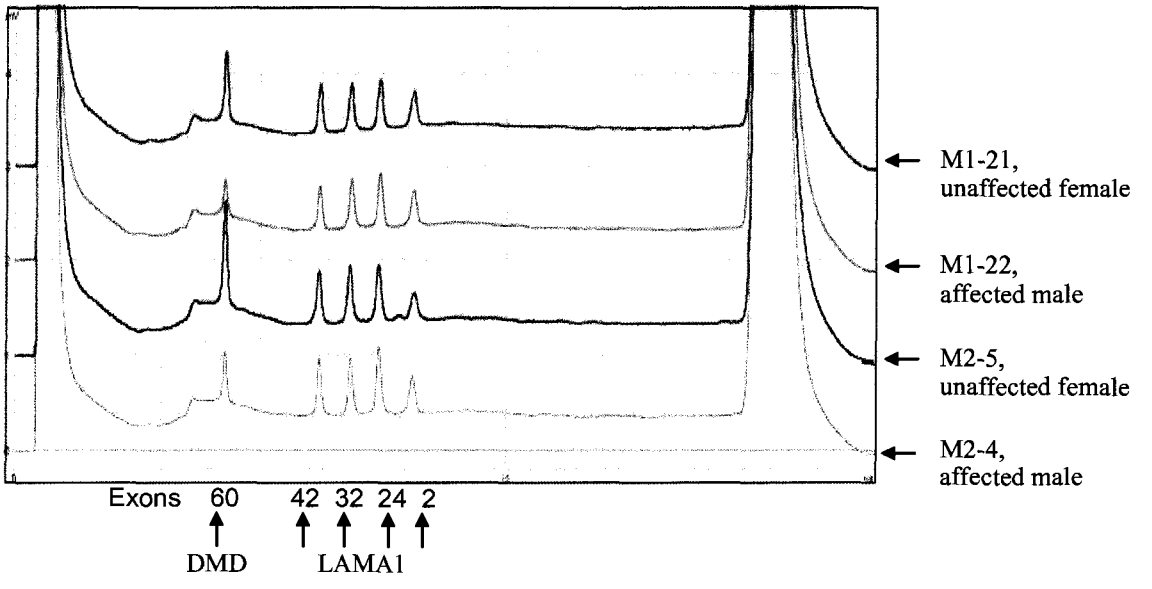
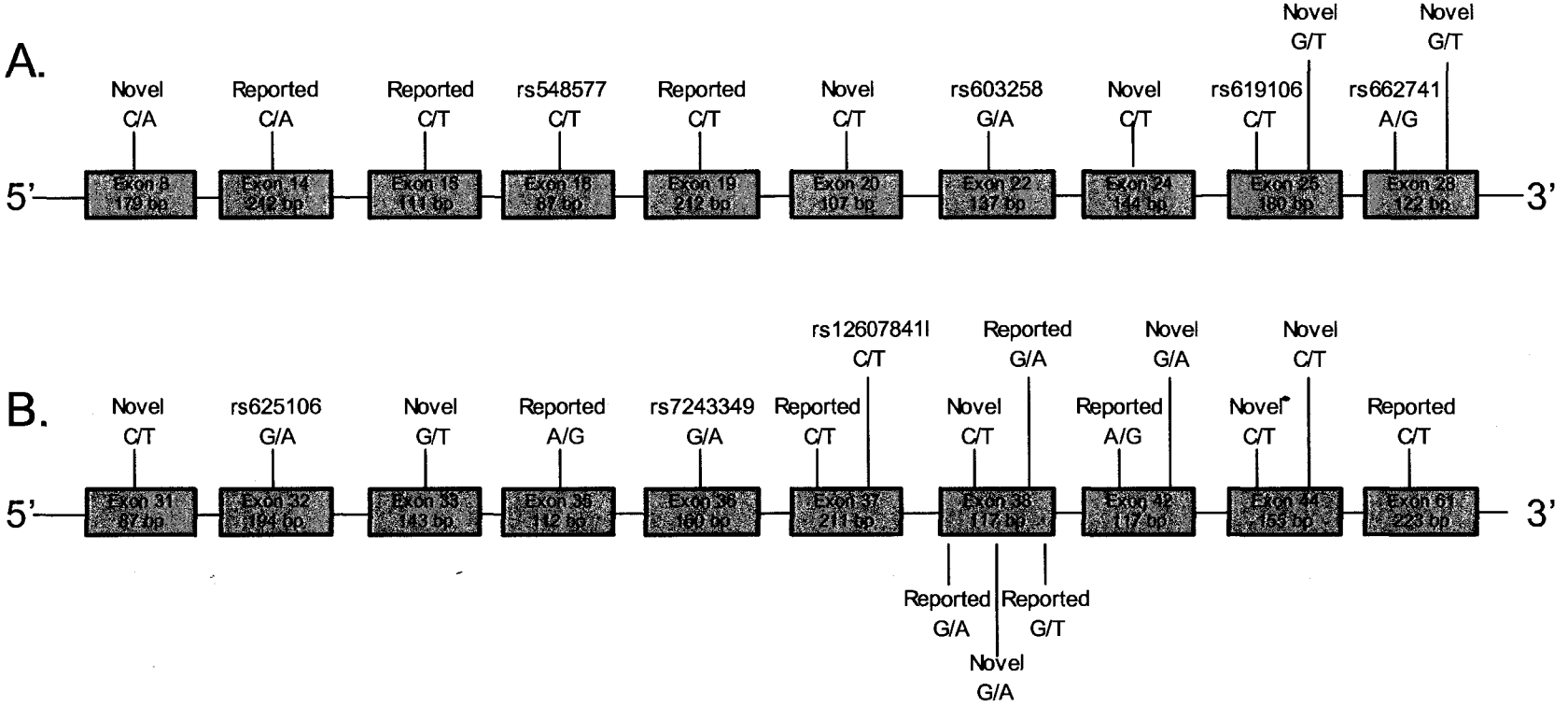

M1-21,
unaffected female
M1-22,
affected male
M2-5,
unaffected female
M2-4,
affected male
Exons 60 42 32 24 2
DMD
LAMA1

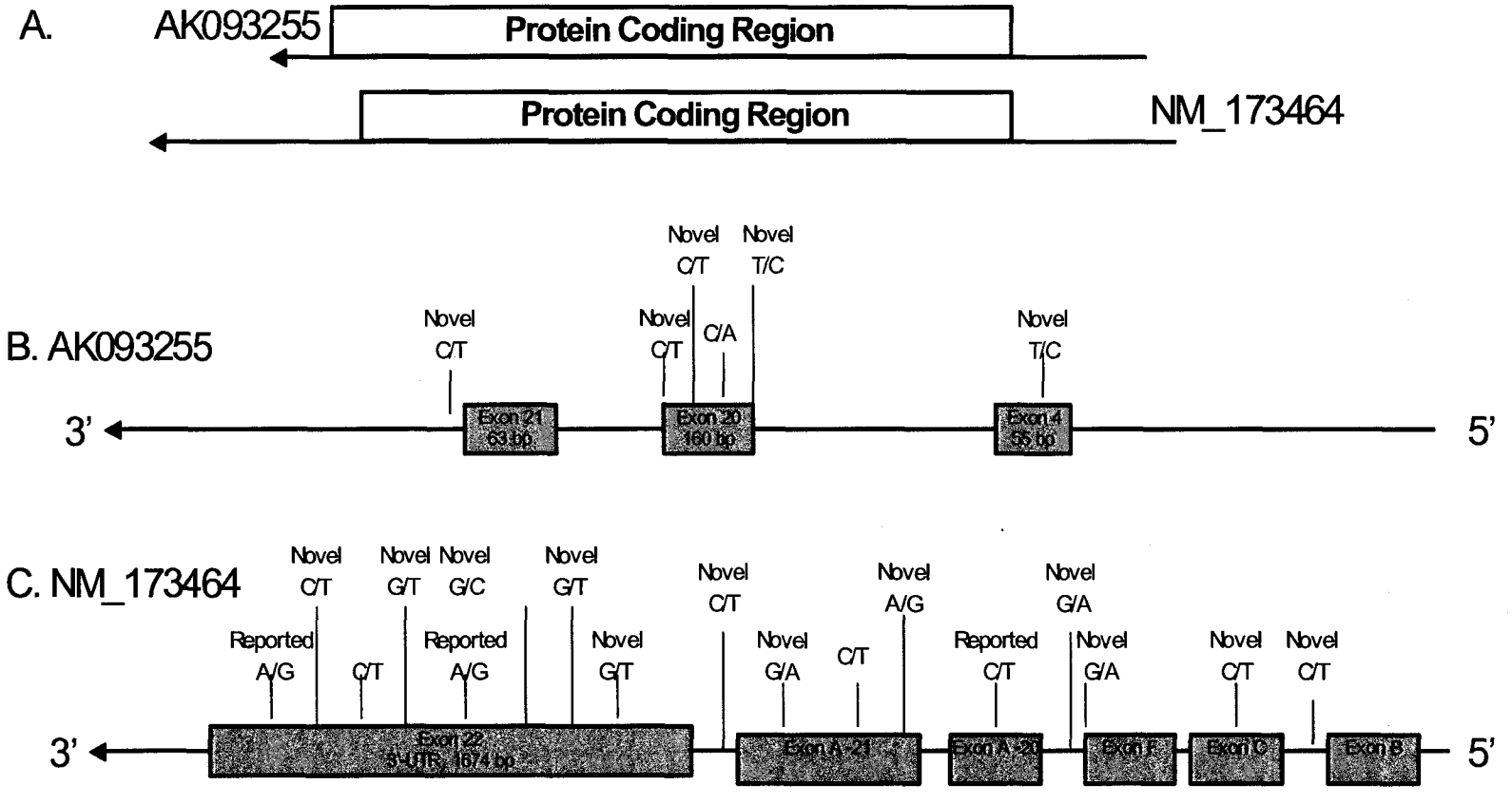
Figure 6-6. Schematic representation of the SNPs identified in exons and introns of the *LAMA1* gene. A. Linear representation of identified SNPs in exons 8 through 28 of *LAMA1* gene, their location, and the corresponding nucleotide. **B.** Linear representation of identified SNPs in exons 31 through 61 of the *LAMA1* gene, their location, and the corresponding nucleotide. Grey boxes indicate the exons of *LAMA1* which have SNPs identified by sequencing.


A.
Novel C/A
Reported C/A
Reported C/T
rs548577 C/T
Reported C/T
Novel C/T
rs603258 G/A
Novel C/T
rs619106 C/T
Novel G/T
rs662741 A/G
Novel G/T
5'
Exon 8 179 bp
Exon 14 212 bp
Exon 15 111 bp
Exon 18 87 bp
Exon 19 212 bp
Exon 20 107 bp
Exon 22 137 bp
Exon 24 144 bp
Exon 25 180 bp
Exon 28 122 bp
3'
B.
Novel C/T
rs625106 G/A
Novel G/T
Reported A/G
rs7243349 G/A
Reported C/T
rs12607841l C/T
Novel C/T
Reported G/A
Reported A/G
Novel G/A
Novel C/T
Novel C/T
Reported C/T
5'
Exon 31 87 bp
Exon 32 194 bp
Exon 33 143 bp
Exon 36 142 bp
Exon 36 160 bp
Exon 37 211 bp
Exon 38 117 bp
Exon 42 117 bp
Exon 44 153 bp
Exon 61 223 bp
3'
Reported G/A
Novel G/A
Reported G/T

6.3.4. *Sequencing of the candidate gene L3MBTL4*

Another good candidate gene was *L3MBTL4* (Homo sapiens l (3) mbt-like 4). It is located in the critical region and expressed in brain, muscle, liver and lungs. This gene has several alternate gene symbols (*FLJ35936, AK093255, and BC039316*) and there are two common transcripts of this gene. One transcript (*AK093255*) of 2099 bp mRNA is composed of 22 exons which codes for a protein of 623 amino acids. The second mRNA (*NM_173464*) of *L3MBTL4* is 4534 bp long and composed of 24 exons with several newly identified exons. The protein consists of 534 amino acids (NP_775735). These two transcripts share most of the first 2099 bp mRNA sequence. We initially sequenced the exons in the shared region without identifying mutations. Then we sequenced the newly identified exons, 5'-UTR and 3'-UTR of *L3MBTL4*. In total we identified 23 SNPs in the *L3MBTL4* gene and all these SNPs are listed in Figure 6-7 and Appendix 2-1.

Figure 6-7. Schematic representation of SNPs identified in the *L3MBTL4* gene. A. Alignment of two transcripts (AK093255 and NM_173464) of *L3MBT4* shows these two variants share most of the protein coding region or open reading frame (ORF). **B and C.** Linear representation of polymorphisms of the *L3MBTL4* variants (AK093255 and NM_173464), their location, and the corresponding nucleotides. Grey boxes indicate the exons which have SNPs identified by sequencing the *L3MBTL4* gene.

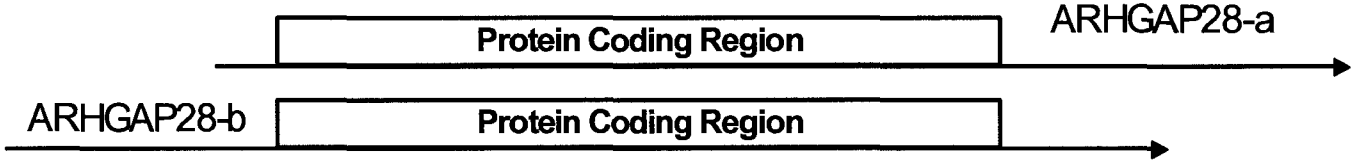

A.
AK093255
Protein Coding Region
Protein Coding Region
NM_173464
B. AK093255
Novel C/T
Novel C/T
Novel T/C
Novel C/T
C/A
Novel T/C
3'
Exon 21 63 bp
Exon 20 160 bp
Exon 4 35 bp
5'
C. NM_173464
Novel C/T
Novel G/T
Novel G/C
Novel G/T
Novel C/T
Novel A/G
Novel G/A
Reported A/G
C/T
Reported A/G
Novel G/T
Novel G/A
C/T
Reported C/T
Novel G/A
Novel C/T
Novel C/T
3'
Exon 22 5'-UTR, 1674 bp
Exon A -21
Exon A -20
Exon F
Exon C
Exon B
5'

6.3.5. *Sequencing of the candidate gene ARHGAP28*

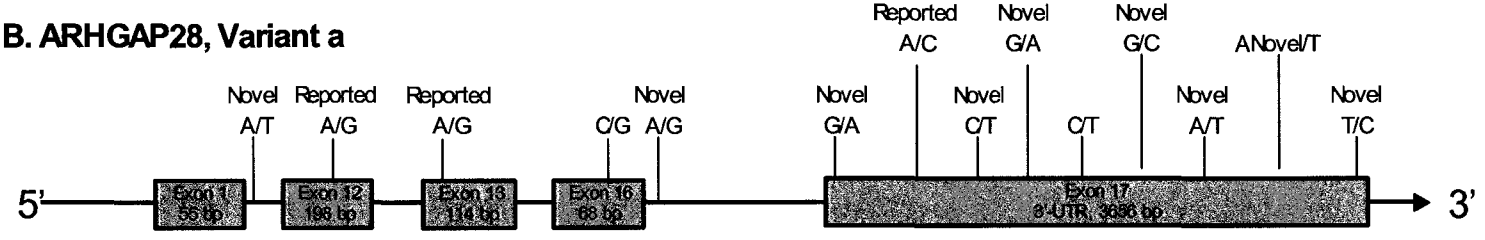
The candidate gene *ARHGAP28* (Rho GTPase activating protein 28) has also been given several gene symbols (*FLJ10312, LOC388462, AB037735, AK001174 and KIAA1314*) and has two common isoforms. One isoform (*ARHGAP28a*) has 17 exons to be transcribed to a 5482 bp mRNA (NM_001010000). Another isoform (*ARHGAP28b*) has 18 exons to form a 6308 bp mRNA (NM_030672). These two transcripts code for the same protein of 390 amino acids (Figure 6-8, A). We have sequenced all 17 exons of *ARHGAP28-a* and the extra exons of *ARHGAP28-b*. We have identified 16 SNPs in this gene (Figure 6-8 and Appendix 2-2). One of these SNPs (C→T missense mutation) located in exon 2 of *ARHGAP-b* isoform changes the amino acid proline to leucine and was found in two affected individuals of family 2 (Figure 6-9). As no restriction enzyme site was created or destroyed to differentiate the C allele from the T allele, we sequenced additional normal controls and found two of 12 normal controls also had this variant, excluding this SNP as the disease-causing mutation in family 2. Recent identification of *SGCE* large deletion in family 2 further supported this change as a SNP.

Figure 6-8. Schematic representation of SNPs identified in the *ARHGAP28* gene. A. Alignment of *ARHGAP28* variants a and b shows these two variants share the same protein coding region or open reading frame (ORF). **B** and **C.** Linear representation of polymorphisms of *ARHGAP28* variants a & b, their location, and the corresponding nucleotides. Grey boxes indicate the exons which have SNPs identified by sequencing the *ARHGAP28* gene.

A. Alignment of ARHGAP28 transcripts



B. ARHGAP28, Variant a



C. ARHGAP28,Variant b

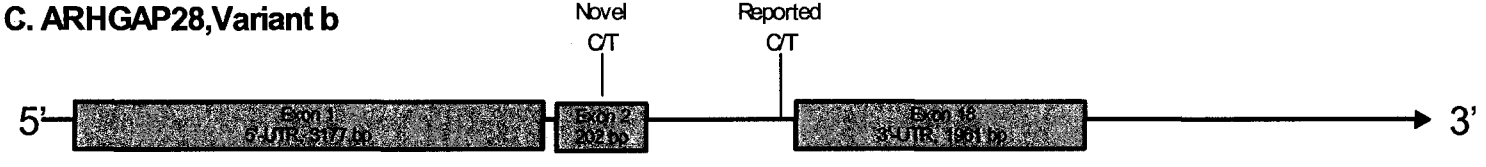
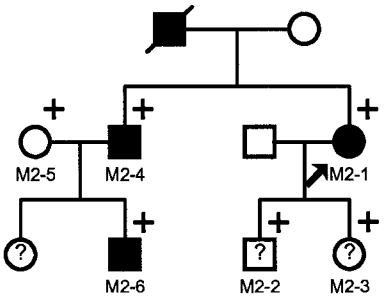
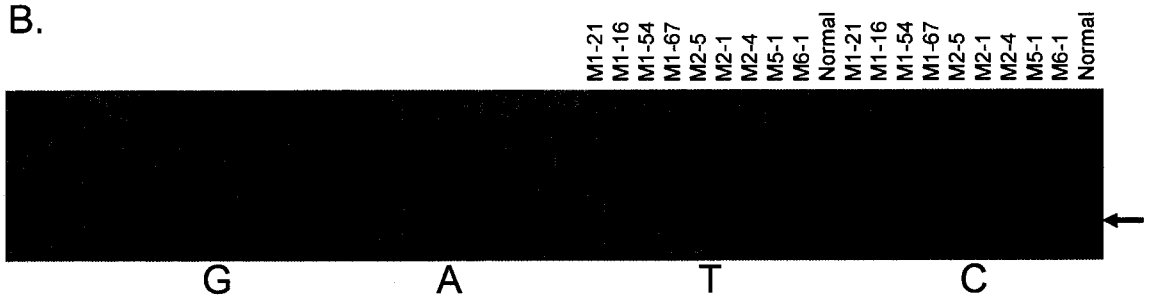


Figure 6-9. Sequence analysis of the candidate gene *ARHGAP28* (*FLJ10312, AB037735*) in *SGCE*-negative families with M-D. A. Pedigree of family 2. **B.** We identified a SNP in exon 2 for the affected (M2-1 and M2-4) of family 2. This C → T transition changes the amino-acid proline to leucine, but it was confirmed to be a polymorphism as normal controls also have this change. M1-21 and M2-5 are unaffected individuals.

A.


M2-5
M2-4
M2-1
M2-6
M2-2
M2-3

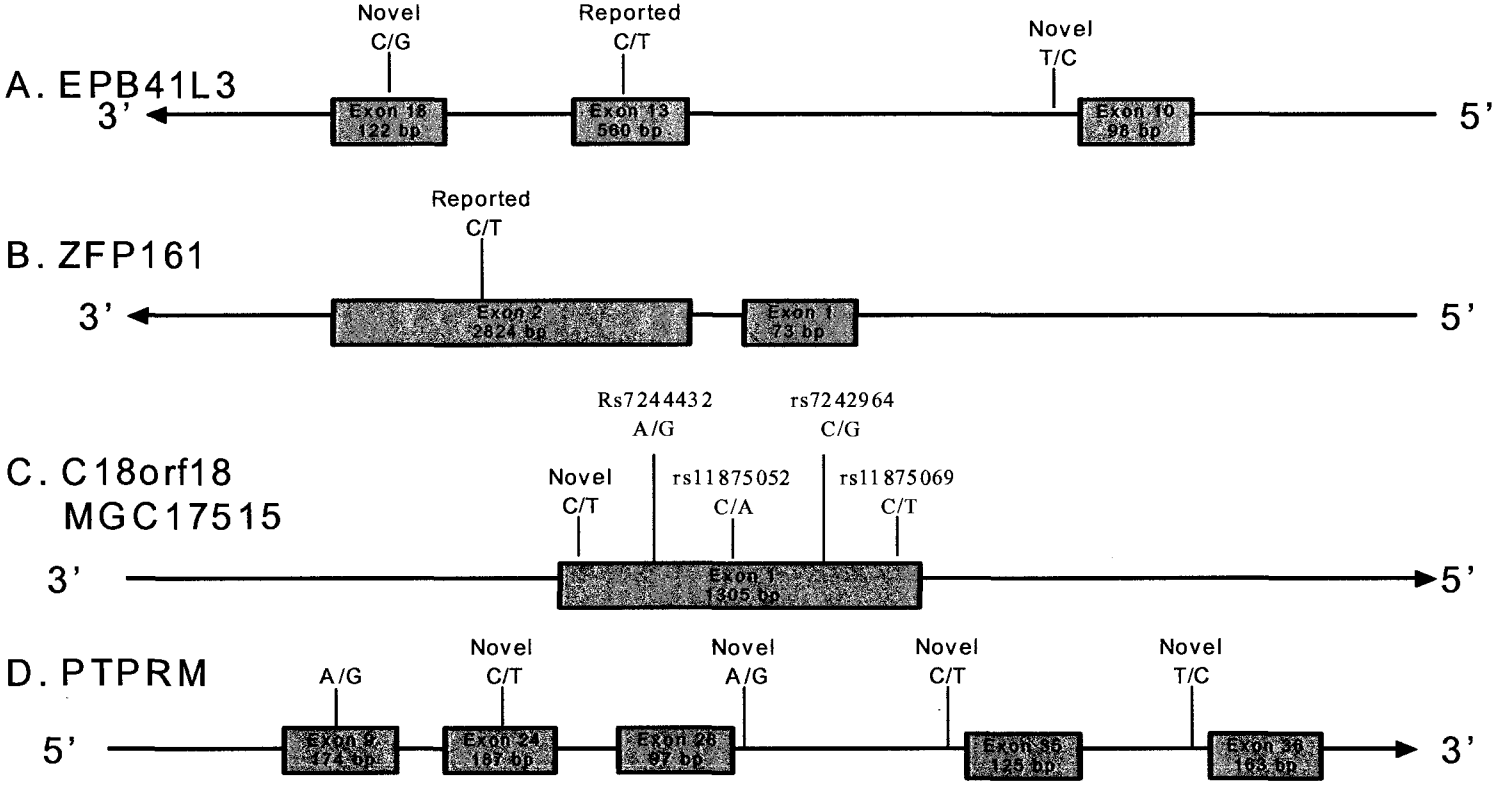
B.


M1-21
M1-16
M1-54
M1-67
M2-5
M2-1
M2-4
M5-1
M6-1
Normal
M1-21
M1-16
M1-54
M1-67
M2-5
M2-1
M2-4
M5-1
M6-1
Normal
G
A
T
C

6.3.6. Sequencing of the candidate gene EPB41L3 (KIAA0987)

Human erythrocyte membrane protein band 4.1-like-3(*EPB41L3*) contains 21 exons. Its mRNA (NM_012307) is 4446 bp long and codes for a 1087 amino acids (accession number: Q9Y2J2). This gene has been shown to be widely expressed in human tissues including the brain and muscle. We have sequenced all of the protein coding region, 5'-UTR and 3-UTR of EPB41L3 without identifying mutations in this gene for the *SGCE*-negative families. We identified three SNPs in this gene and listed them in Figure 6-10 and Appendix 2-3. A G→C transversion in exon 18 which changes amino acid glutamine to histidine was seen in one affected individual of family 1, but the other affected members of the family are G/G homozygotes, suggesting this change is a SNP.

Figure 6-10. Schematic representation of SNPs identified in the candidate genes, *EPB41L3, ZFP161, MGC17515, and PTPRM.* (**A**), SNPs in the *EPB41L3* gene. (**B**), SNPs in the *ZFP161* gene. (**C**), SNPs in the *MGC17515* gene. (**D**), SNPs in the *PTPRM* gene.


A. EPB41L3
Novel
C/G
Reported
C/T
Novel
T/C
3'
Exon 18
122 bp
Exon 13
560 bp
Exon 10
98 bp
5'
B. ZFP161
Reported
C/T
3'
Exon 2
2824 bp
Exon 1
73 bp
5'
C. C18orf18
MGC17515
Rs7244432
A/G
rs7242964
C/G
Novel
C/T
rs11875052
C/A
rs11875069
C/T
3'
Exon 1
1305 bp
5'
D. PTPRM
A/G
Novel
C/T
Novel
A/G
Novel
C/T
Novel
T/C
5'
Exon 9
174 bp
Exon 24
187 bp
Exon 28
97 bp
Exon 36
125 bp
Exon 38
103 bp
3'

6.3.7. *Sequencing of the candidate gene ZFP161*

ZFP161 (zinc finger protein 161 homolog) is only composed of two exons and has several alternate symbols: *Y12726; ZNF478*; Zinc finger protein 5 (*hZF5*) and; Zinc finger and BTB domain-containing protein 14 (*ZBTB14*). Its mRNA (NM_003409) is 2896 bp in length and codes for a protein of 449 amino acids. *ZFP161* has been shown to be widely expressed in human tissues and functions as a transcriptional activator of the dopamine transporter (DAT). We sequenced these two exons and identified one SNP in the gene, a heterozygous C/T which was found in the normal control. The affected individuals of family 1 were homozygous C/C, excluding this SNP as a mutation (Figure 6-10 and Appendix 2-3).

6.3.8. Sequencing of the candidate gene PTPRM

Human protein tyrosine phosphatase, receptor type, M (*PTPRM*) is composed of 39 exons. Its mRNA (NM_002845) is 5065 bp and codes for a protein of 1425 amino acids. As it is located in our original critical region of 4 Mb and half of the exons are still located in the reduced region, we sequenced all the 39 exons and identified numerous SNPs without finding any mutations in this gene (Figure 6-10 and Appendix 2-3).

6.3.9. Sequencing of the candidate gene MGC17515

MGC17515 is also named as *C18orf18*. The genomic structure of *MGC17515* only includes one exon. Its mRNA (BC010538) is 1322 bp and codes for a protein of 95 amino acids. We identified five SNPs in this gene. Four of them were reported in the dbSNP database (rs11875069, rs7242964, rs11875052, rs7244432). One novel variant (C→T) was identified in affected individuals of families 2, 5 and 6, but this variant codes for the same amino acid (leucine), indicating it is a SNP (Figure 6-10 and Appendix 2-4).

6.3.10. Sequencing of the candidate gene LOC400643

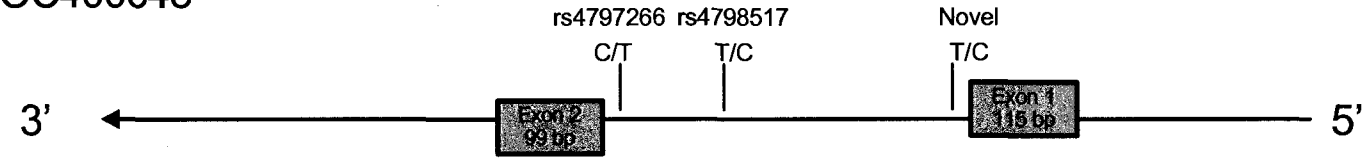
LOC400643 (also called AK095347) is another predicted gene in the candidate region for M-D. It is composed of 4 exons with 1778 bp of mRNA (BC035413). We sequenced these four exons and identified 3 SNPs in the noncoding region of this gene (Figure 6-11 and Appendix 2-4). Two of them have been reported in the dbSNP database (rs4797266 and rs4798517).

6.3.11. Sequencing of the candidate gene LOC388459

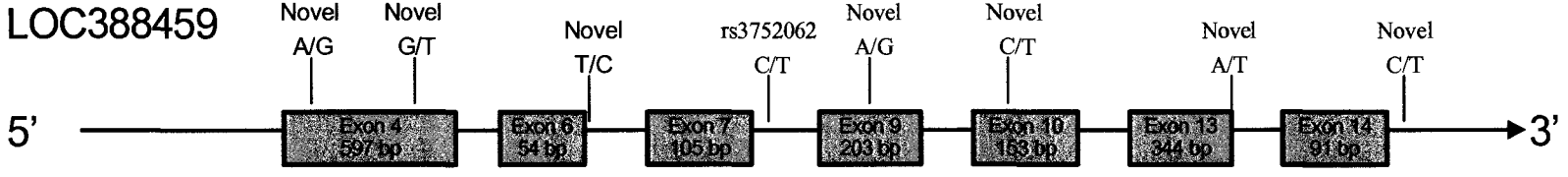
LOC388459 is a hypothetical (predicted) gene and has 18 predicted exons. However, it's predicted mRNA (XM_373772) is only 717 bp long and codes for a protein of 230 amino acids (XP_373772). As we did not know the exact size of its mRNA and protein, we designed primers to sequence all the 18 predicted exons and identified 8 SNPs (Figure 6-11 and Appendix 2-5). One of these SNPs was shown to have a G→T missense mutation which changes the amino acid alanine to threonine in two affected individuals of family 1 and one affected individual of family 8. We have sequenced 16 normal controls and found that two controls also have this heterozygous change (Figure 6-12).

Figure 6-11. Schematic representation of SNPs identified in the candidate genes, *LOC400643, LOC388459* and *LOC388461*. **A**. SNPs in the *LOC400643* gene. **B**. SNPs in the *LOC388459* gene. **C**. SNPs in the *LOC388461* gene.

A. LOC400643



B. LOC388459



C. LOC388461

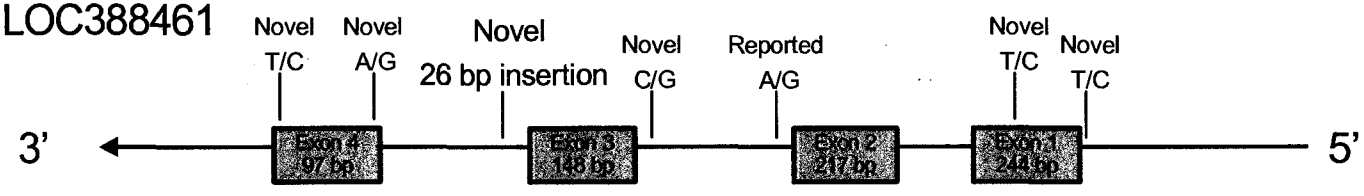
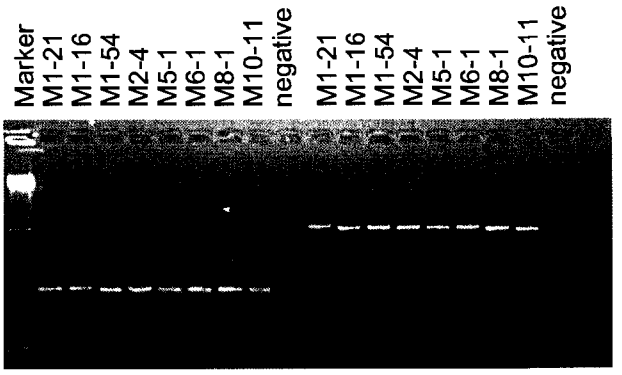
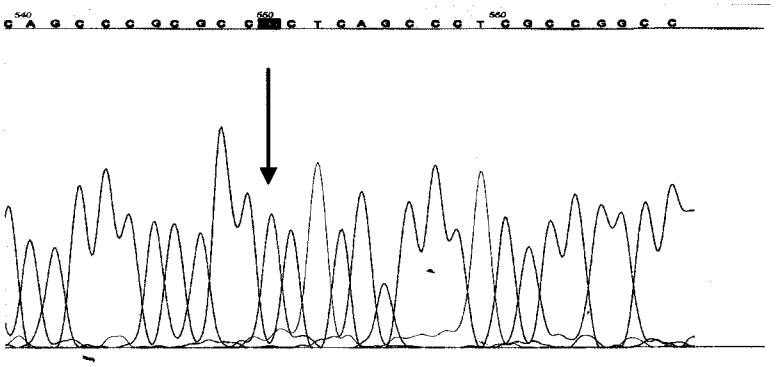


Figure 6-12. Sequence analysis of the candidate gene *LOC388459*. **A.** PCR products for exons 1 and 4 of *LOC388459*. **B.** Wild-type homozygous G/G for normal control M1-21 in exon 4 of *LOC388459*. **C.** This homozygous T/T changes the amino acid alanine to threonine in affected individuals M1-16 and M5-1. **D.** Heterozygous G/T was identified in affected individuals M1-54, M2-4, M6-1 and M8-1.

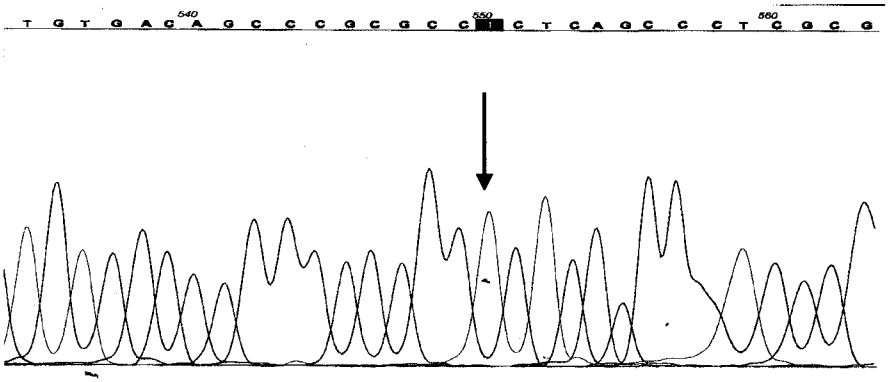
A.


Marker
M1-21
M1-16
M1-54
M2-4
M5-1
M6-1
M8-1
M10-11
negative
M1-21
M1-16
M1-54
M2-4
M5-1
M6-1
M8-1
M10-11
negative

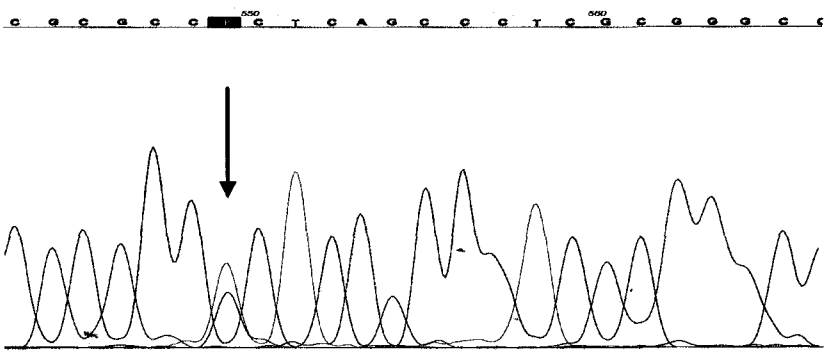
B.



C.



D.



6.3.12. Sequencing of the candidate genes LOC388461 and LOC388460

LOC388461 is a predicted gene which contains 4 exons. We sequenced all four exons and identified 6 SNPs in this gene. The SNPs identified in this gene were listed in Figure 6-11 and Appendix 2-6. We identified a 26 bp insertion polymorphism downstream of exon 3. As this insertion was only seen in the affected individuals of family 2 (we had not found the deletion in *SGCE* in this family at that time), we screened 50 normal controls and found two normal controls had this insertion. Another variant (genomic position: 6741746) was identified in affected individuals of families 1, 10 and 12. This C→T transition in exon1 of *LOC388461* changes the amino acid threonine to methionine (Figure 6-13). We first obtained the DHPLC profiles of those individuals we had sequenced and then screened normal controls using DHPLC. We determined that three of eleven normal controls (P41, P98, and P11) also had this change, confirming this variant is a SNP (Figure 6-14).

LOC388460 is a predicted gene which contains 2 exons. We sequenced both exons and did not identify any sequence variant in this gene.

Figure 6-13. Sequence analysis of exon 1 of *LOC388461* in *SGCE*-negative families with M-D. This C→T transition in exon1 of the predicted gene *LOC388461* changes the amino acid threonine to methionine in two affected individuals of family 1 (M1-16 and M1-54) and in the affected individual from each affected individual of families 10 and 12 (upper arrow). However, three normal controls (p41, p98, and p111) also had this change by DHPLC screening. Note: Affected individuals in family 1 and 2 had a different allele (T or C) for the SNP in exon 1 of *LOC388461*. The heterozygous C/T of M1-67 on the lower part of the gel is located 9 bp upstream of exon 1 (lower arrow). Except for the individual, M1-21, all other individuals were affected.

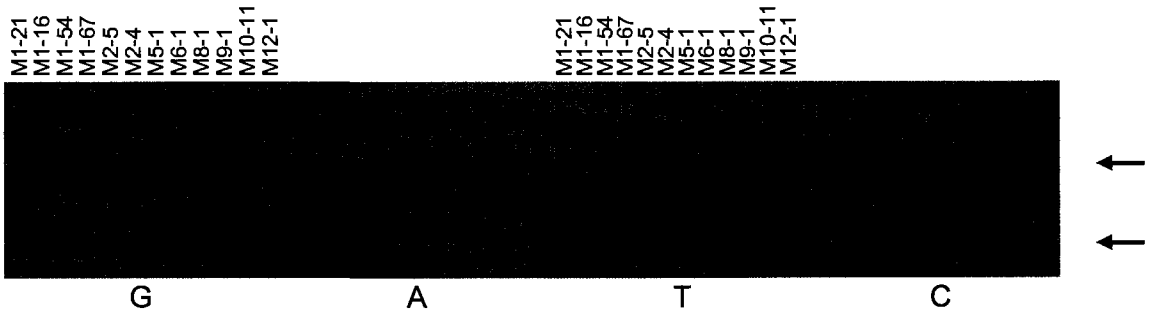
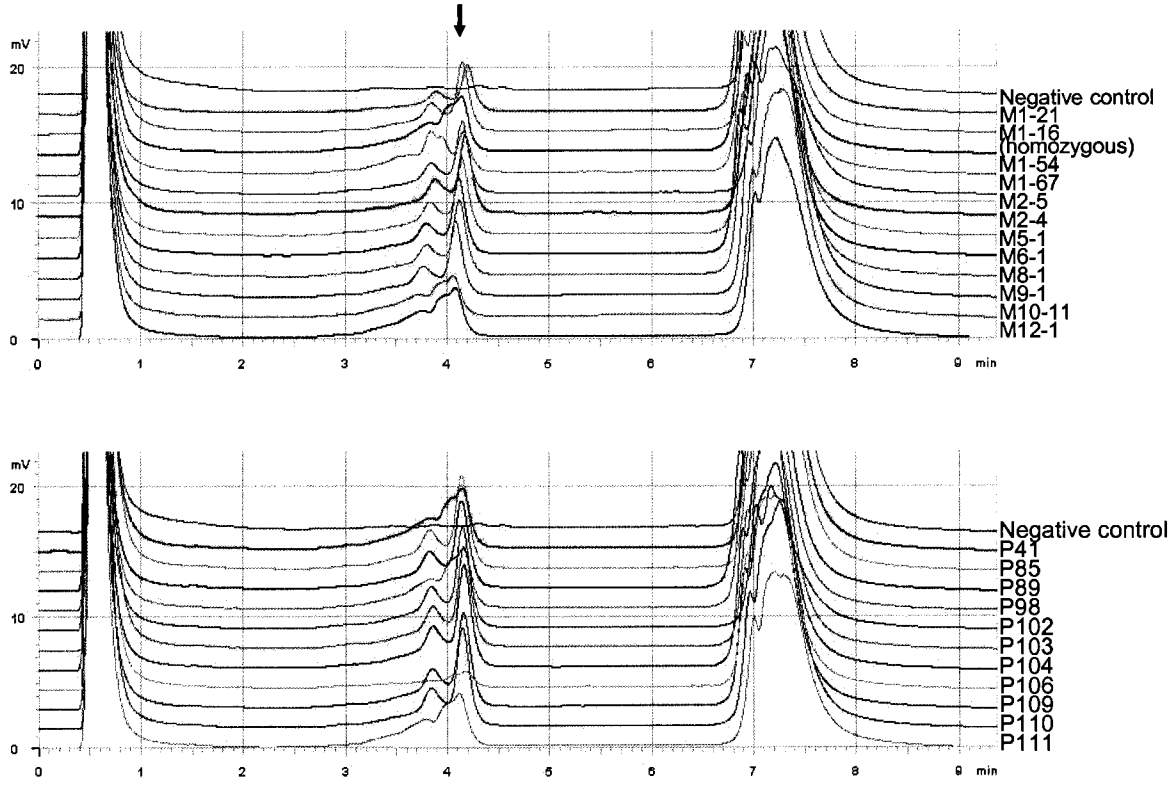

M1-21
M1-16
M1-54
M1-67
M2-5
M2-4
M5-1
M6-1
M8-1
M9-1
M10-11
M12-1
M1-21
M1-16
M1-54
M1-67
M2-5
M2-4
M5-1
M6-1
M8-1
M9-1
M10-11
M12-1
G
A
T
C

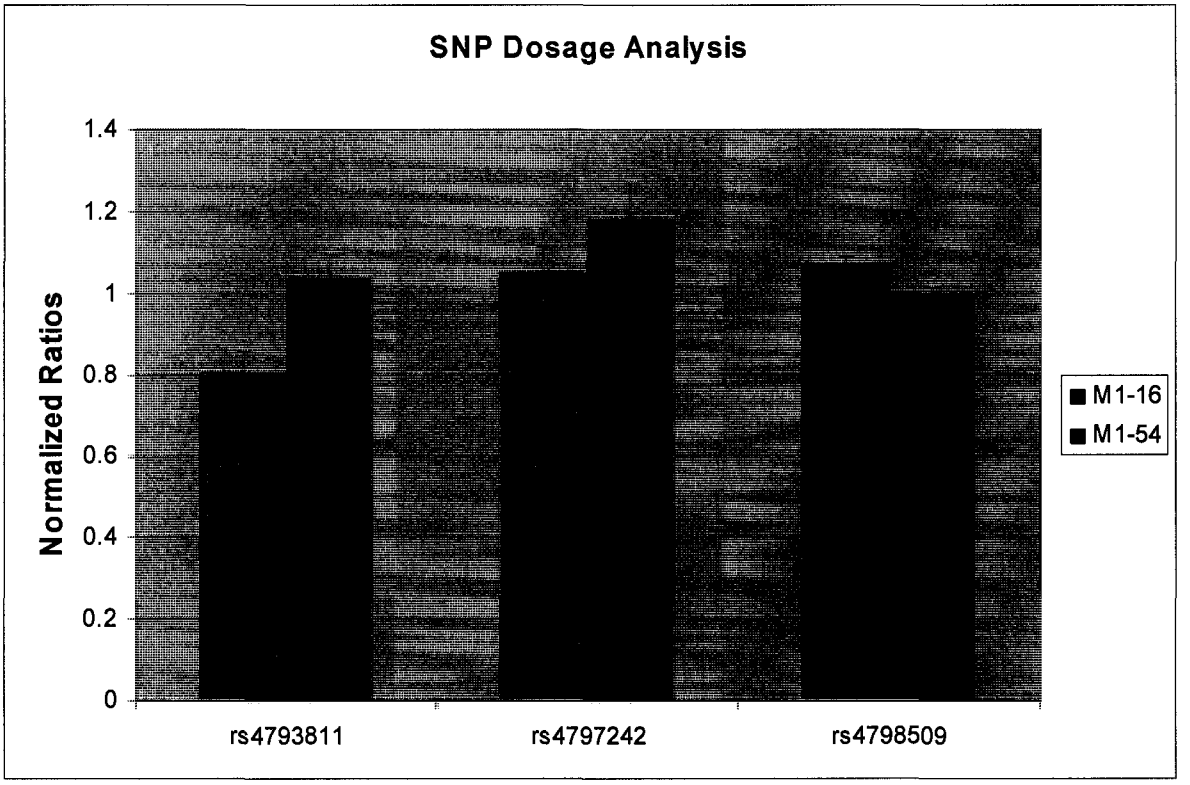
Figure 6-14. DHPLC analysis of *LOC388461* exon 1 for the affected individuals and normal controls. M1-54, M10-11 and M12-1 had the same heterozygous profile (upper panel). Some normal controls (P41, P98, and P111) also had the same profiles as M1-54, M10-11, and M12-1 had in exon1 of *LOC388461* (lower panel). M1-67 has a different profile as this individual has a heterozygous C/T located in a different position as shown in Figure 6-13. Black arrow indicates the position of elution peaks.


mV
20
10
0
0
1
2
3
4
5
6
7
8
9 min
Negative control
M1-21
M1-16
(homozygous)
M1-54
M1-67
M2-5
M2-4
M5-1
M6-1
M8-1
M9-1
M10-11
M12-1
mV
20
10
0
0
1
2
3
4
5
6
7
8
9 min
Negative control
P41
P85
P89
P98
P102
P103
P104
P106
P109
P110
P111

6.3.13. *Genomewide gene dosage analysis using a 500K SNP array*

Although we did not perform the 500K SNP array analysis, our lab also used a 500K SNP array to search for potential large deletions/duplications that could lead to M-D. A 500K SNP array analysis was performed on DNA from an affected individual M1-39 and an unaffected control. M1-39 has an affected father and affected children. Within our critical region on 18p11, we were able to identify 3 homozygous SNPs (rs4798311, rs4797242, and rs4798509) which had reduced signal intensity. The genomic positions of rs4798311, rs4797242, and rs4798509 are 5080563, 6397483 and 6878513, respectively. To search for the potential deletions in these homozygous regions, we designed primers to amplify these three SNPs and evaluated the gene dosages using our semi-quantitative DHPLC method. All the PCR for dosage analysis were performed in triplicate on the same plate, along with DNA from normal controls. However, no dosage alterations of these SNPs were found for the two affected individuals of family 1 (M1-16 and M1-54) as compared with the 10 normal controls, excluding the possibility of large deletions in these three SNP regions for families 1 (Figure 6-15).

Figure 6-15. Dosage analysis of three homozygous SNPs (rs4793811, rs4797242, and rs4798509) identified by a 500K SNP array analysis at *DYT15* locus. No significant dosage alteration were shown for these three SNPs in two affected individuals from family 1, M1-16 (blue bar) and M1-54 (red bar),.


SNP Dosage Analysis
Normalized Ratios
1.4
1.2
1
0.8
0.6
0.4
0.2
0
rs4793811
rs4797242
rs4798509
M1-16
M1-54

6.4. Discussion

We had previously mapped a novel locus for M-D to a 4 Mb region on chromosome18p11. As the critical region was still too large and we could not reduce the linked region by genotyping additional individuals of the family 1 and the individuals of family 10 with the linked markers on 18p11 (**Chapter 5**), we refined the region by narrowing the gap between the flanking recombinant and non-recombinant markers. SNPs have been widely used to map disease loci because of their high-density and accuracy (Fan and Xiong, 2002; Nairz *et al.*, 2002; Rannala and Reeve, 2001). We have refined the *DYT15* locus by identifying three definite recombinant SNPs in family 1 with M-D. The previous region of 3.99 Mb for DYT15 has been reduced to a 3.18 MB region. Based on the NCBI and genome browser of UCSC, seven known genes (*MGC17515, ZFP161, EPB41L3 (KIAA0987), L3MBT4, ARHGAP28 (FLJ10312), LAMA1, and PTPRM*) and eight predicted genes have been mapped in this critical region. *LOC 388459*, *LOC388460*, and *LOC388461* are predicted with EST support while *LOC390826*, *LOC390827*, *LOC400643*, *LOC390828* and *LOC339291* are predicted based on Genscan and alignment of the mRNAs on UCSC genome browser database.

Sequencing of the seven genes did not result in the identification of a mutation. Four of the known candidate genes (*LAMA1, L3MBTL4, ARHGAP28 and EPB41L3)* have similar expression patterns to the *SGCE* gene, suggesting that they would be good candidate genes for M-D. EPB41-related (protein 4.1) genes are a group of four paralogous genes (*EPB41*, *EPB41L2*, *EPB41L3*, and *EPB41L1*) that play important roles in the development of different tissues such as red blood cells, brain, skeletal muscle and the heart (Peters *et al.*, 1998). Defects in the *EPB41* gene has been reported to be associated with abnormal

morphology and unstable membranes leading to hereditary elliptocytosis and hemolytic anemia (Tchernia *et al.*, 1981). *Epb4.1* knockout mice demonstrated phenotypes which include the familiar erythroid disorders caused by *EPB41* deficiency and neuronal defects in fine motor coordination and spatial learning (Shi *et al.*, 1999; Walensky *et al.*, 1998). Both *L3MBTL4* and *ARHGAP28* (Rho GTPase activating protein 28) have several different transcripts because of the alternative splicing for their exons. They are expressed mostly in brain and muscle tissues, making them strong candidate genes for M-D even though their biological functions remain unknown (Ota *et al.*, 2004). The laminins consist of three subunit chains (α, β, and γ). These subunits are assembled into at least 15 distinct laminin isoforms. Even though Laminin α1 (*LAMA1*) is considered to have a restricted expression in fetal and adult human tissues, the major site of expression is the central nervous system (Miner *et al.*, 2004; Virtanen *et al.*, 2000) and it is the major component of early embryonic basement membrane. In the study of laminin knockout mice, Lamb1 knockout mice have shown that the gene is required for early development as Lamb1-/- embryos die at embryonic day 5.5. Even though Lama1 knockout (Lama1-/-) embryos can form the embryonic ectoderm cavity and undergo partial endoderm differentiation, they can not survive beyond embryonic day 7 (Miner *et al.*, 2004). The sarcoglycan complex is known to interact with *LAMA2* and mutations in *LAMA2* have been reported to cause congenital muscular dystrophy. Given that mutations in *SGCE* having been shown to cause myoclonus-dystonia, the mapping of *LAMA1* into our critical region make it the strongest candidate for M-D. We have sequenced all 62 exons and the regulatory elements (5'-UTR and 3'-UTR) of the *LAMA1* gene without identifying mutations. We noted a region of homozygosity between exon 22 and exon 44 and that this region co-segregated with the

affected individuals in both families 1 and 2. One possible explanation for this homozygous region is that there is a large deletion in the two families because deletions/duplications of multiple exons or genes are not detected by the conventional PCR sequencing when individual exons were screened for mutations. Our lab has developed a sensitive DHPLC-based quantitative multiplex PCR system to detect genomic deletions and duplications. Using this method no large deletions or duplications were detected for this homozygous region in *LAMA1*. Our identification of a large deletion of the *SGCE* gene in family 2 further confirmed that the existence of this homozygous region of the *LAMA1* gene is not caused by genomic deletions of the *LAMA1* gene, at least in family 2 (**Chapter 4**).

We have also sequenced the other known candidate genes and predicted genes with EST support in the reduced critical region. Even though we identified SNPs in each of these genes, we did not find any disease-causing mutations. DHPLC analysis of three homozygous SNPs (rs4793811, rs4797242, and rs4798509) did not identify gene dosage alterations, ruling out large deletion/duplication in these SNP regions as the cause of family 1. Since our initial sequencing only covered the exons and exon-intron boundaries of the candidate genes, the intronic mutations could also be missed via PCR and direct sequencing, we will use a reverse transcriptase PCR (RT-PCR) approach to search for intronic mutations in the candidate genes we sequenced, especially in the strong candidate genes, *LAMA1, L3MBTL4, ARHGAP28 and EPB41L3.* If we still can't find a mutation in the candidate genes in our region, we will work on the identification of novel transcripts (unknown genes) in the critical region which currently are not in the public resources such as the NCBI or genome browser of UCSC. For example, Cox and colleagues have used RT-PCR on human fetal brain poly(A) RNAs and 5' and 3' RACE on a Marathon-Ready human fetal brain

cDNA library to identify three novel genes in a locus for holoprosencephaly (HPE8) on human chromosome 14q13 between markers D14S49 and AFM205XG5 (Kamnasaran *et al.*, 2005). We can use this approach to characterize novel candidate genes in the *DYT15* locus and screen for mutations in these novel candidate genes for family 1.

Although M-D is an uncommon disorder, a better understanding of its pathogenesis will be important in helping to elucidate the cause of involuntary movements. Individuals with mutations in *SGCE* have a very narrow phenotype. Although this is the most common genetic cause, less than half of the individuals with the classic M-D phenotype harbor mutations in *SGCE*. Our *DYT15* family is clinically indistinguishable from those with *SGCE* mutations and would suggest that the *DYT15* gene product interacts within the same functional/biological pathway as that of *SGCE*. Prior to the identification of mutations in the *SGCE* gene as a cause for M-D, neurotransmitters and their receptors were felt to be the most likely cause of the abnormal movements seen in M-D. Despite the existence of several good candidate genes with similar potential roles within the central nervous system within our critical region, we were not able to identify any mutations within these genes.

Chapter 7.

Future Studies

7.1. Future studies

7.1.1. Alternative approaches for the identification of a novel gene at the DYT 15 locus

There are two kinds of mutations that could be missed by our sequencing methods. The first one is a large deletion/duplication. To overcome this problem, several mutation detection techniques have been used. Southern blots can be used to detect deletions and duplications by the identification of novel restriction fragments created by gained or lost restriction enzyme sites. But Southern blot analysis is time-consuming and requires substantial amounts of DNA for the identification of novel fragments. Quantitative real-time PCR (Q-PCR) can be used to detect deletions/duplications with relatively high cost, if one incorporates controls into every reaction. Q-PCR was used to detect homozygous tumor deletions and duplications and showed good correlation with Southern blot analysis (M'Soka T *et al.*, 2000) and used to detect the deletions in cystic fibrosis transmembrane regulator (CFTR) gene (Costes *et al.*, 2000). Array-based comparative genomic hybridization (CGH) has been widely used for the genome-wide detection of large deletions and duplications with a resolution of 40 kb genomic regions (Hodgson *et al.*, 2001; Le Caignec *et al.*, 2005; Pinkel *et al.*, 1998). CGH was first reported for investigating dosage alterations in breast cancer (Pinkel *et al.*, 1998) and was modified by degenerate oligonucleotide PCR (DOP-PCR) (Hodgson *et al.*, 2001). But CGH can not detect inversions and translocations because the copy number does not change. Our lab has developed a fast and sensitive semi-quantitative PCR-based DHPLC method for the detection of deletions and duplications and it has been successfully used to identify large deletions in the *SGCE* gene. We will use this method to analyze the copy number for

potential deletions which were implicated in the homozygous regions of the candidate genes we sequenced in **Chapter 6**.

Another possible mutation that could be missed in our sequencing of exons and flanking intronic sequence are intronic mutations which can create cryptic splice sites. Reverse-transcriptase PCR (RT-PCR) when combined with dosage analysis can be used to overcome this problem. We will use RT-PCR to search for intronic mutations in the strong candidate genes, *LAMA1*, *EPB41L3*, *ARHGAP28* and *L3MBTL4* which were directly sequenced in **Chapter 6.**

7.1.2. Characterization of novel transcripts as candidate genes for the DYT15 locus

If we don't find a mutation in the predicted genes in our region, we will work on the identification of novel transcripts (unknown genes) in the critical region which currently are not in public resources such as the NCBI gene bank or Genome browser of UCSC. For example, Kamnasaran *et al* have used RT-PCR on human fetal brain poly(A) RNAs and 5' and 3' RACE on human fetal brain cDNA library to identify three novel genes for a locus for holoprosencephaly (HPE8) on human chromosome 14q13 between markers D14S49 and AFM205XG5 (Kamnasaran *et al.*, 2005). We can use this approach to characterize the novel potential candidate genes in our *DYT15* locus.

7.1.3. Functional studies of the identified mutations in the SGCE gene

Not every sequence variant seen in an affected person is pathogenic. Expression analysis and functional studies usually follow to explore the pathological effects of the variant on cultured cells and in animal models (Dang *et al.*, 2005; Gonzalez-Alegre *et al.*, 2005; Goodchild and Dauer, 2004). A common experiment is to see if the variant has an effect in a knockout mouse model. In six of seventeen families with M-D, we identified

mutations which resulted in premature termination of protein translation. All these mutations are predicted to produce truncated proteins but the effect on the brain is unknown. *SGCE* is widely expressed in human and mice tissues but little is known of the role it plays in the brain. To understand the molecular mechanisms of myoclonus dystonia, it is important to realize that the majority of mutations result in a functional null. It would be interesting to investigate those rare missense mutations and determine the effects of these mutations on the localization of *SGCE* within the cell. We can also construct a knockout mouse model with deficient *Sgce* as a first step in determining the role of *SGCE* in the central nervous system.

7.1.4. Generation of a knockout mouse model with the novel gene

Mouse models allow detailed examination of the physical and pathological basis of diseases (Jinnah *et al.*, 2005; Shashidharan *et al.*, 2005). Successful mouse models with targeted genes can also be used to study novel therapeutic approaches. Once we have identified the disease causing gene in the *DYT15* locus, we will establish a mouse model. In addition to allowing us to examine the effects of the mutation in an *in vivo* model, this would also provide us with the ability to examine pathological mechanisms of the disease at the level of cell culture. Standard phenotypic characterization of a knockout/knockin or transgenic mouse would be performed and tissue and temporal expression analysis would be performed. Cellular localization would be determined by immunofluorescence microscopy using the primary and secondary antibodies for the mutated genes or by using a GFP-fusion protein.

7.2. Conclusion

In summary, we have identified point mutations and small deletions of *SGCE* gene in four of seventeen families with M-D using direct sequencing (families 3, 4, 7, and 15). Two families had the same mutation in exon 3 (R102X) as reported previously while other two families were found to carry novel mutations. One family (family 4) had a 5-bp deletion in exon 7 and another family (family 15) had a 1-bp deletion in exon 6. In addition, using a semi-quantitative PCR-based DHPLC method we were able to find large deletions of the *SGCE* gene in two of thirteen M-D families without other smaller *SGCE* mutations. One of these families (family 2) has an 18 kb deletion that includes exons 2 and 3 while another family (family 6) has a 13 kb deletion covering exons 2 through 5. We have confirmed these large deletions by identifying the exact breakpoints of the deletions. Therefore in our cohort of 17 unrelated M-D families, 35% of patients with M-D have mutations in the *SGCE* gene. These findings demonstrate that in addition to screening for point mutations and small deletions, screening for large gene deletions/duplications in *SGCE* is also necessary.

In a large five-generation Canadian family with M-D (family 1) that did not have *SGCE* mutations, we previously mapped a novel locus to a 4 Mb region between the markers D18S1132 and D18S843 on 18p11 (Max LOD score of 3.96 for GATA185c06). This locus is designated *DYT15* (OMIM: 607488). Since no further informative STS markers were found within the flanking regions, we utilized single nucleotide polymorphisms (SNP) for fine-mapping. As a result, we identified three recombinant SNPs (rs385769, rs727951 and rs727952) in the proximal flanking region between markers D18S1163 and D18S843 in family 1. Our previous 4 Mb linked region has been reduced to a 3.18 Mb critical region flanked by markers D18S1132 and rs385769. There are seven known genes and eight predicted genes within or very close to this region. We have

sequenced all the exons and exon-intron boundaries of these seven known genes (*MGC17515, ZFP161, EPB41L3 (KIAA0987), L3MBT4, ARHGAP28 (FLJ10312), LAMA1,* and *PTPRM*) and the four predicted genes with EST support (*LOC388459, LOC388460, LOC400643* and *LOC388461*) without identifying the causative mutations.

Future studies should be directed to the investigation of the functional effects of the known *SGCE* mutations using *in vitro* and *in vivo* models. For the identification of a novel gene, we will focus on the detection of gene rearrangements (deletions/duplications) and intronic mutations and search for mutations in novel transcripts in the critical reduced region based on gene prediction and EST mapping.

References

Araishi, K., Sasaoka, T., Imamura, M., Noguchi, S., Hama, H., Wakabayashi, E., Yoshida, M., Hori, T., and Ozawa, E. (1999). Loss of the sarcoglycan complex and sarcospan leads to muscular dystrophy in beta-sarcoglycan-deficient mice. Hum Mol Genet *8*, 1589-1598.

Asmus, F., and Gasser, T. (2004). Inherited myoclonus-dystonia. Adv Neurol *94*, 113-119.

Asmus, F., Salih, F., Hjermind, L. E., Ostergaard, K., Munz, M., Kuhn, A. A., Dupont, E., Kupsch, A., and Gasser, T. (2005). Myoclonus-dystonia due to genomic deletions in the epsilon-sarcoglycan gene. Ann Neurol *58*, 792-797.

Asmus, F., Zimprich, A., Naumann, M., Berg, D., Bertram, M., Ceballos-Baumann, A., Pruszak-Seel, R., Kabus, C., Dichgans, M., Fuchs, S., *et al.* (2001). Inherited Myoclonus-dystonia syndrome: narrowing the 7q21-q31 locus in German families. Ann Neurol *49*, 121-124.

Asmus, F., Zimprich, A., Tezenas Du Montcel, S., Kabus, C., Deuschl, G., Kupsch, A., Ziemann, U., Castro, M., Kuhn, A. A., Strom, T. M., *et al.* (2002). Myoclonus-dystonia syndrome: epsilon-sarcoglycan mutations and phenotype. Ann Neurol *52*, 489-492.

Bacolla, A., Jaworski, A., Larson, J. E., Jakupciak, J. P., Chuzhanova, N., Abeysinghe, S. S., O'Connell, C. D., Cooper, D. N., and Wells, R. D. (2004). Breakpoints of gross deletions coincide with non-B DNA conformations. Proc Natl Acad Sci U S A *101*, 14162-14167.

Bandmann, O., Valente, E. M., Holmans, P., Surtees, R. A., Walters, J. H., Wevers, R. A., Marsden, C. D., and Wood, N. W. (1998). Dopa-responsive dystonia: a clinical and molecular genetic study. Ann Neurol *44*, 649-656.

Batzer, M. A., and Deininger, P. L. (2002). Alu repeats and human genomic diversity. Nat Rev Genet *3*, 370-379.

Bonnemann, C. G. (2005). Limb-girdle muscular dystrophy in childhood. Pediatr Ann *34*, 569-577.

Bressman, S. B. (2004). Dystonia genotypes, phenotypes, and classification. Adv Neurol *94*, 101-107.

Bressman, S. B., de Leon, D., Brin, M. F., Risch, N., Burke, R. E., Greene, P. E., Shale, H., and Fahn, S. (1989). Idiopathic dystonia among Ashkenazi Jews: evidence for autosomal dominant inheritance. Ann Neurol *26*, 612-620.

Bressman, S. B., de Leon, D., Kramer, P. L., Ozelius, L. J., Brin, M. F., Greene, P. E., Fahn, S., Breakefield, X. O., and Risch, N. J. (1994). Dystonia in Ashkenazi Jews: clinical characterization of a founder mutation. Ann Neurol *36*, 771-777.

Bressman, S. B., Fahn, S., and Burke, R. E. (1988). Paroxysmal non-kinesigenic dystonia. Adv Neurol *50*, 403-413.

Bruno, M. K., Hallett, M., Gwinn-Hardy, K., Sorensen, B., Considine, E., Tucker, S., Lynch, D. R., Mathews, K. D., Swoboda, K. J., Harris, J., *et al.* (2004). Clinical evaluation of idiopathic paroxysmal kinesigenic dyskinesia: new diagnostic criteria. Neurology *63*, 2280-2287.

Bulman, D. E., Murphy, E. G., Zubrzycka-Gaarn, E. E., Worton, R. G., and Ray, P. N. (1991). Differentiation of Duchenne and Becker muscular dystrophy phenotypes with amino- and carboxy-terminal antisera specific for dystrophin. Am J Hum Genet *48*, 295-304.

Campbell, K. P. (1995). Three muscular dystrophies: loss of cytoskeleton-extracellular matrix linkage. Cell *80*, 675-679.

Caviness, J. N., Alving, L. I., Maraganore, D. M., Black, R. A., McDonnell, S. K., and Rocca, W. A. (1999). The incidence and prevalence of myoclonus in Olmsted County, Minnesota. Mayo Clin Proc *74*, 565-569.

Chan, P., Gonzalez-Maeso, J., Ruf, F., Bishop, D. F., Hof, P. R., and Sealfon, S. C. (2005). Epsilon-sarcoglycan immunoreactivity and mRNA expression in mouse brain. J Comp Neurol *482*, 50-73.

Chen, D. H., Matsushita, M., Rainier, S., Meaney, B., Tisch, L., Feleke, A., Wolff, J., Lipe, H., Fink, J., Bird, T. D., and Raskind, W. H. (2005). Presence of alanine-to-valine substitutions in myofibrillogenesis regulator 1 in paroxysmal nonkinesigenic dyskinesia: confirmation in 2 kindreds. Arch Neurol *62*, 597-600.

Clarimon, J., Johnson, J., Dogu, O., Horta, W., Khan, N., Lees, A. J., Hardy, J., and Singleton, A. (2005). Defining the ends of Parkin exon 4 deletions in two different families with Parkinson's disease. Am J Med Genet B Neuropsychiatr Genet *133*, 120-123.

Cohen LG, H. M. (1988). Hand cramps: clinical features and electromyographic patterns in a focal dystonia. Neurology *38*, 1005-1012.

Costes, B., Girodon, E., Vidaud, D., Flori, E., Ardalan, A., Conteville, P., Fanen, P., Niel, F., Vidaud, M., and Goossens, M. (2000). Prenatal detection by real-time quantitative PCR and characterization of a new CFTR deletion, 3600+15kbdel5.3kb (or CFTRdele19). Clin Chem *46*, 1417-1420.

Dang, M. T., Yokoi, F., McNaught, K. S., Jengelley, T. A., Jackson, T., Li, J., and Li, Y. (2005). Generation and characterization of Dyt1 DeltaGAG knock-in mouse as a model for early-onset dystonia. Exp Neurol *196*, 452-463.

de Carvalho Aguiar, P., Sweadner, K. J., Penniston, J. T., Zaremba, J., Liu, L., Caton, M., Linazasoro, G., Borg, M., Tijssen, M. A., Bressman, S. B., *et al.* (2004). Mutations in the Na+/K+ -ATPase alpha3 gene ATP1A3 are associated with rapid-onset dystonia parkinsonism. Neuron *43*, 169-175.

de Carvalho Aguiar, P. M., and Ozelius, L. J. (2002). Classification and genetics of dystonia. Lancet Neurol *1*, 316-325.

Deininger, P. L., and Batzer, M. A. (1999). Alu repeats and human disease. Mol Genet Metab *67*, 183-193.

Demirkiran, M., and Jankovic, J. (1995). Paroxysmal dyskinesias: clinical features and classification. Ann Neurol *38*, 571-579.

Dobyns, W. B., Ozelius, L. J., Kramer, P. L., Brashear, A., Farlow, M. R., Perry, T. R., Walsh, L. E., Kasarskis, E. J., Butler, I. J., and Breakefield, X. O. (1993). Rapid-onset dystonia-parkinsonism. Neurology *43*, 2596-2602.

Doheny, D., Danisi, F., Smith, C., Morrison, C., Velickovic, M., De Leon, D., Bressman, S. B., Leung, J., Ozelius, L., Klein, C., *et al.* (2002). Clinical findings of a myoclonus-dystonia family with two distinct mutations. Neurology *59*, 1244-1246.

Dressler, D., and Benecke, R. (2005). Diagnosis and management of acute movement disorders. J Neurol *252*, 1299-1306.

Driscoll, D. A., Spinner, N. B., Budarf, M. L., McDonald-McGinn, D. M., Zackai, E. H., Goldberg, R. B., Shprintzen, R. J., Saal, H. M., Zonana, J., Jones, M. C., and *et al.* (1992). Deletions and microdeletions of 22q11.2 in velo-cardio-facial syndrome. Am J Med Genet *44*, 261-268.

Duclos, F., Broux, O., Bourg, N., Straub, V., Feldman, G. L., Sunada, Y., Lim, L. E., Piccolo, F., Cutshall, S., Gary, F., *et al.* (1998). Beta-sarcoglycan: genomic analysis and identification of a novel missense mutation in the LGMD2E Amish isolate. Neuromuscul Disord *8*, 30-38.

Ettinger, A. J., Feng, G., and Sanes, J. R. (1997). epsilon-Sarcoglycan, a broadly expressed homologue of the gene mutated in limb-girdle muscular dystrophy 2D. J Biol Chem *272*, 32534-32538.

Fan, R., and Xiong, M. (2002). High resolution mapping of quantitative trait loci by linkage disequilibrium analysis. Eur J Hum Genet *10*, 607-615.

Fink, J. K., Rainer, S., Wilkowski, J., Jones, S. M., Kume, A., Hedera, P., Albin, R., Mathay, J., Girbach, L., Varvil, T., *et al.* (1996). Paroxysmal dystonic choreoathetosis: tight linkage to chromosome 2q. Am J Hum Genet *59*, 140-145.

Fouad, G. T., Servidei, S., Durcan, S., Bertini, E., and Ptacek, L. J. (1996). A gene for familial paroxysmal dyskinesia (FPD1) maps to chromosome 2q. Am J Hum Genet *59*, 135-139.

Furukawa, Y. (2003). Genetics and biochemistry of dopa-responsive dystonia: significance of striatal tyrosine hydroxylase protein loss. Adv Neurol *91*, 401-410.

Furukawa, Y., Guttman, M., Sparagana, S. P., Trugman, J. M., Hyland, K., Wyatt, P., Lang, A. E., Rouleau, G. A., Shimadzu, M., and Kish, S. J. (2000). Dopa-responsive dystonia due to a large deletion in the GTP cyclohydrolase I gene. Ann Neurol *47*, 517-520.

Garavagilia, B., Invernizzi, N., Nardocci, N., Zorzi, G., Girotti, F., and Zeviani, M. (2000). Molecular studies of childhood onset primary dystonias in Italy. Am J Hum Genet *67(Supplement 2)*, 375.

Gasser, T. (1998). Inherited myoclonus-dystonia syndrome. Adv Neurol *78*, 325-334.

Gasser, T., Bereznai, B., Muller, B., Pruszak-Seel, R., Damrich, R., Deuschl, G., and Oertel, W. H. (1996). Linkage studies in alcohol-responsive myoclonic dystonia. Mov Disord *11*, 363-370.

Geyer, H. L., and Bressman, S. B. (2006). The diagnosis of dystonia. Lancet Neurol *5*, 780-790.

Glockner, S., Lehmann, U., Wilke, N., Kleeberger, W., Langer, F., and Kreipe, H. (2000). Detection of gene amplification in intraductal and infiltrating breast cancer by laser-assisted microdissection and quantitative real-time PCR. Pathobiology *68*, 173-179.

Gonzalez-Alegre, P., Bode, N., Davidson, B. L., and Paulson, H. L. (2005). Silencing primary dystonia: lentiviral-mediated RNA interference therapy for DYT1 dystonia. J Neurosci *25*, 10502-10509.

Goodchild, R. E., and Dauer, W. T. (2004). Mislocalization to the nuclear envelope: an effect of the dystonia-causing torsinA mutation. Proc Natl Acad Sci U S A *101*, 847-852.

Grabowski, M., Zimprich, A., Lorenz-Depiereux, B., Kalscheuer, V., Asmus, F., Gasser, T., Meitinger, T., and Strom, T. M. (2003). The epsilon-sarcoglycan gene (SGCE), mutated in myoclonus-dystonia syndrome, is maternally imprinted. Eur J Hum Genet *11*, 138-144.

Graeber, M. B., Kupke, K. G., and Muller, U. (1992). Delineation of the dystonia-parkinsonism syndrome locus in Xq13. Proc Natl Acad Sci U S A *89*, 8245-8248.

Grimes, D. A., Bulman, D., George-Hyslop, P. S., and Lang, A. E. (2001). Inherited myoclonus-dystonia: evidence supporting genetic heterogeneity. Mov Disord *16*, 106-110.

Grimes, D. A., Han, F., Lang, A. E., St George-Hyssop, P., Racacho, L., and Bulman, D. E. (2002). A novel locus for inherited myoclonus-dystonia on 18p11. Neurology *59*, 1183-1186.

Grotzsch, H., Pizzolato, G. P., Ghika, J., Schorderet, D., Vingerhoets, F. J., Landis, T., and Burkhard, P. R. (2002). Neuropathology of a case of dopa-responsive dystonia associated with a new genetic locus, DYT14. Neurology *58*, 1839-1842.

Grunau, C., Clark, S. J., and Rosenthal, A. (2001). Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. Nucleic Acids Res *29*, E65-65.

Guicheney, P., Vignier, N., Helbling-Leclerc, A., Nissinen, M., Zhang, X., Cruaud, C., Lambert, J. C., Richelme, C., Topaloglu, H., Merlini, L., *et al.* (1997). Genetics of laminin alpha 2 chain (or merosin) deficient congenital muscular dystrophy: from identification of mutations to prenatal diagnosis. Neuromuscul Disord *7*, 180-186.

Hack, A. A., Groh, M. E., and McNally, E. M. (2000). Sarcoglycans in muscular dystrophy. Microsc Res Tech *48*, 167-180.

Han, F., Lang, A. E., Racacho, L., Bulman, D. E., and Grimes, D. A. (2003). Mutations in the epsilon-sarcoglycan gene found to be uncommon in seven myoclonus-dystonia families. Neurology *61*, 244-246.

Hjermind, L. E., Werdelin, L. M., Eiberg, H., Krag-Olsen, B., Dupont, E., and Sorensen, S. A. (2003). A novel mutation in the epsilon-sarcoglycan gene causing myoclonus-dystonia syndrome. Neurology *60*, 1536-1539.

Hodgson, G., Hager, J. H., Volik, S., Hariono, S., Wernick, M., Moore, D., Nowak, N., Albertson, D. G., Pinkel, D., Collins, C., *et al.* (2001). Genome scanning with array CGH delineates regional alterations in mouse islet carcinomas. Nat Genet *29*, 459-464.

Hu, X., Bulman, D. E., Ray, P. N., and Worton, R. G. (1992). Frameshift duplication resulting in truncated dystrophin in a patient with Duchenne muscular dystrophy. Hum Mutat *1*, 172-173.

Ichinose, H., Inagaki, H., Suzuki, T., Ohye, T., and Nagatsu, T. (2000). Molecular mechanisms of hereditary progressive dystonia with marked diurnal fluctuation, Segawa's disease. Brain Dev *22 Suppl 1*, S107-110.

Ichinose, H., Ohye, T., Takahashi, E., Seki, N., Hori, T., Segawa, M., Nomura, Y., Endo, K., Tanaka, H., Tsuji, S., and *et al.* (1994). Hereditary progressive dystonia with marked diurnal fluctuation caused by mutations in the GTP cyclohydrolase I gene. Nat Genet *8*, 236-242.

Ikegami, T., Ikeda, H., Chance, P. F., Kiyosawa, H., Yamamoto, M., Sobue, G., Ohnishi, A., Tachi, N., and Hayasaka, K. (1997). Facilitated diagnosis of CMT1A duplication in chromosome 17p11.2-12: analysis with a CMT1A-REP repeat probe and photostimulated luminescence imaging. Hum Mutat *9*, 563-566.

Ikeuchi, T., Shimohata, T., Nakano, R., Koide, R., Takano, H., and Tsuji, S. (1999). A case of primary torsion dystonia in Japan with the 3-bp (GAG) deletion in the DYT1 gene with a unique clinical presentation. Neurogenetics *2*, 189-190.

Im, J. H., Ahn, T. B., Kim, K. B., Ko, S. B., and Jeon, B. S. (2004). DYT1 mutation in Korean primary dystonia patients. Parkinsonism Relat Disord *10*, 421-423.

Imamura, M., Araishi, K., Noguchi, S., and Ozawa, E. (2000). A sarcoglycan-dystroglycan complex anchors Dp116 and utrophin in the peripheral nervous system. Hum Mol Genet *9*, 3091-3100.

Imamura, M., Mochizuki, Y., Engvall, E., and Takeda, S. (2005). Epsilon-sarcoglycan compensates for lack of alpha-sarcoglycan in a mouse model of limb-girdle muscular dystrophy. Hum Mol Genet *14*, 775-783.

Jinnah, H. A., Hess, E. J., Ledoux, M. S., Sharma, N., Baxter, M. G., and Delong, M. R. (2005). Rodent models for dystonia research: characteristics, evaluation, and utility. Mov Disord *20*, 283-292.

Kamm, C., Leung, J., Joseph, S., Dobyns, W. B., Brashear, A., Breakefield, X. O., and Ozelius, L. J. (2004). Refined linkage to the RDP/DYT12 locus on 19q13.2 and evaluation of GRIK5 as a candidate gene. Mov Disord *19*, 845-847.

Kamnasaran, D., Chen, C. P., Devriendt, K., Mehta, L., and Cox, D. W. (2005). Defining a holoprosencephaly locus on human chromosome 14q13 and characterization of potential candidate genes. Genomics *85*, 608-621.

Klein, C., Brin, M. F., Kramer, P., Sena-Esteves, M., de Leon, D., Doheny, D., Bressman, S., Fahn, S., Breakefield, X. O., and Ozelius, L. J. (1999a). Association of a missense change in the D2 dopamine receptor with myoclonus dystonia. Proc Natl Acad Sci U S A *96*, 5173-5176.

Klein, C., Friedman, J., Bressman, S., Vieregge, P., Brin, M. F., Pramstaller, P. P., De Leon, D., Hagenah, J., Sieberer, M., Fleet, C., *et al.* (1999b). Genetic testing for early-onset torsion dystonia (DYT1): introduction of a simple screening method, experiences from testing of a large patient cohort, and ethical aspects. Genet Test *3*, 323-328.

Klein, C., Liu, L., Doheny, D., Kock, N., Muller, B., De Carvalho Aguiar, P., Leung, J., De Leon, D., Bressman, S. B., Silverman, J., *et al.* (2002). varepsilon -sarcoglycan mutations found in combination with other dystonia gene mutations. Ann Neurol *52*, 675-679.

Klein, C., Ozelius, L. J., Hagenah, J., Breakefield, X. O., Risch, N. J., and Vieregge, P. (1998). Search for a founder mutation in idiopathic focal dystonia from Northern Germany. Am J Hum Genet *63*, 1777-1782.

Klein, C., Page, C. E., LeWitt, P., Gordon, M. F., de Leon, D., Awaad, Y., Breakefield, X. O., Brin, M. F., and Ozelius, L. J. (1999c). Genetic analysis of three patients with an 18p- syndrome and dystonia. Neurology *52*, 649-.

Klein, C., Schilling, K., Saunders-Pullman, R. J., Garrels, J., Breakefield, X. O., Brin, M. F., deLeon, D., Doheny, D., Fahn, S., Fink, J. S., *et al.* (2000). A major locus for myoclonus-dystonia maps to chromosome 7q in eight families. Am J Hum Genet *67*, 1314-1319.

Kolomietz, E., Meyn, M. S., Pandita, A., and Squire, J. A. (2002). The role of Alu repeat clusters as mediators of recurrent chromosomal aberrations in tumors. Genes Chromosomes Cancer *35*, 97-112.

Korczyn, A. D., Kahana, E., Zilber, N., Streifler, M., Carasso, R., and Alter, M. (1980). Torsion dystonia in Israel. Ann Neurol *8*, 387-391.

Kramer, P. L., de Leon, D., Ozelius, L., Risch, N., Bressman, S. B., Brin, M. F., Schuback, D. E., Burke, R. E., Kwiatkowski, D. J., Shale, H., and *et al.* (1990). Dystonia gene in Ashkenazi Jewish population is located on chromosome 9q32-34. Ann Neurol *27*, 114-120.

Kramer, P. L., Heiman, G. A., Gasser, T., Ozelius, L. J., de Leon, D., Brin, M. F., Burke, R. E., Hewett, J., Hunt, A. L., Moskowitz, C., and *et al.* (1994). The DYT1 gene on 9q34 is responsible for most cases of early limb-onset idiopathic torsion dystonia in non-Jews. Am J Hum Genet *55*, 468-475.

Kramer, P. L., Mineta, M., Klein, C., Schilling, K., de Leon, D., Farlow, M. R., Breakefield, X. O., Bressman, S. B., Dobyns, W. B., Ozelius, L. J., and Brashear, A. (1999). Rapid-onset dystonia-parkinsonism: linkage to chromosome 19q13. Ann Neurol *46*, 176-182.

Kruger, R., Fischer, C., Schulte, T., Strauss, K. M., Muller, T., Woitalla, D., Berg, D., Hungs, M., Gobbele, R., Berger, K., *et al.* (2003). Mutation analysis of the neurofilament M gene in Parkinson's disease. Neurosci Lett *351*, 125-129.

Kunkel, L. M., Hejtmancik, J. F., Caskey, C. T., Speer, A., Monaco, A. P., Middlesworth, W., Colletti, C. A., Bertelson, C., Muller, U., Bresnan, M., *et al.* (1986). Analysis of

deletions in DNA from patients with Becker and Duchenne muscular dystrophy. Nature *322*, 73-77.

Kupke, K. G., Graeber, M. B., and Muller, U. (1992). Dystonia-parkinsonism syndrome (XDP) locus: flanking markers in Xq12-q21.1. Am J Hum Genet *50*, 808-815.

Kupke, K. G., Lee, L. V., and Muller, U. (1990a). Assignment of the X-linked torsion dystonia gene to Xq21 by linkage analysis. Neurology *40*, 1438-1442.

Kupke, K. G., Lee, L. V., Viterbo, G. H., Arancillo, J., Donlon, T., and Muller, U. (1990b). X-linked recessive torsion dystonia in the Philippines. Am J Med Genet *36*, 237-242.

Kyllerman, M., Forsgren, L., Sanner, G., Holmgren, G., Wahlstrom, J., and Drugge, U. (1990). Alcohol-responsive myoclonic dystonia in a large family: dominant inheritance and phenotypic variation. Mov Disord *5*, 270-279.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Lang, A. E. (1997). Essential myoclonus and myoclonic dystonia [letter; comment]. Mov Disord *12*, 127.

Lathrop, G. M., Lalouel, J. M., Julier, C., and Ott, J. (1985). Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. Am J Hum Genet *37*, 482-498.

Le Caignec, C., De Mas, P., Vincent, M. C., Boceno, M., Bourrouillou, G., Rival, J. M., and David, A. (2005). Subtelomeric 6p deletion: clinical, FISH, and array CGH characterization of two cases. Am J Med Genet A *132*, 175-180.

Lebre, A. S., Durr, A., Jedynak, P., Ponsot, G., Vidailhet, M., Agid, Y., and Brice, A. (1999). DYT1 mutation in French families with idiopathic torsion dystonia. Brain *122 (Pt 1)*, 41-45.

Lee, H. Y., Xu, Y., Huang, Y., Ahn, A. H., Auburger, G. W., Pandolfo, M., Kwiecinski, H., Grimes, D. A., Lang, A. E., Nielsen, J. E., *et al.* (2004). The gene for paroxysmal non-kinesigenic dyskinesia encodes an enzyme in a stress response pathway. Hum Mol Genet *13*, 3161-3170.

Leube, B., and Auburger, G. (1998). Questionable role of adult-onset focal dystonia among sporadic dystonia patients. Ann Neurol *44*, 984-985.

Leube, B., Hendgen, T., Kessler, K. R., Knapp, M., Benecke, R., and Auburger, G. (1997a). Evidence for DYT7 being a common cause of cervical dystonia (torticollis) in Central Europe. Am J Med Genet *74*, 529-532.

Leube, B., Hendgen, T., Kessler, K. R., Knapp, M., Benecke, R., and Auburger, G. (1997b). Sporadic focal dystonia in northwest Germany: molecular basis on chromosome 18p. Ann Neurol *42*, 111-114.

Leube, B., Rudnicki, D., Ratzlaff, T., Kessler, K., Benecke, R., and Auburger, G. (1996). Idiopathic torsion dystonia: assignment of a gene to chromosome 18p in a German family with adult onset, autosomal dominant inheritance and purely focal distribution. Hum Mol Genet *5*, 1673-1677.

Leung, J. C., Klein, C., Friedman, J., Vieregge, P., Jacobs, H., Doheny, D., Kamm, C., DeLeon, D., Pramstaller, P. P., Penney, J. B., *et al.* (2001). Novel mutation in the TOR1A (DYT1) gene in atypical early onset dystonia and polymorphisms in dystonia and early onset parkinsonism. Neurogenetics *3*, 133-143.

Leuzzi, V., Carducci, C., Cardona, F., Artiola, C., and Antonozzi, I. (2002). Autosomal dominant GTP-CH deficiency presenting as a dopa-responsive myoclonus-dystonia syndrome. Neurology *59*, 1241-1243.

Lim, L. E., and Campbell, K. P. (1998). The sarcoglycan complex in limb-girdle muscular dystrophy. Curr Opin Neurol *11*, 443-452.

Liu, L. A., and Engvall, E. (1999). Sarcoglycan isoforms in skeletal muscle. J Biol Chem *274*, 38171-38176.

Lupski, J. R., de Oca-Luna, R. M., Slaugenhaupt, S., Pentao, L., Guzzetta, V., Trask, B. J., Saucedo-Cardenas, O., Barker, D. F., Killian, J. M., Garcia, C. A., *et al.* (1991). DNA duplication associated with Charcot-Marie-Tooth disease type 1A. Cell *66*, 219-232.

M'Soka T, J., Nishioka, J., Taga, A., Kato, K., Kawasaki, H., Yamada, Y., Yu, A., Komada, Y., and Nobori, T. (2000). Detection of methylthioadenosine phosphorylase (MTAP) and p16 gene deletion in T cell acute lymphoblastic leukemia by real-time quantitative PCR assay. Leukemia *14*, 935-940.

Mahloudji, M., and Pikielny, R. T. (1967). Hereditary essential myoclonus. Brain *90*, 669-674.

Major, T., Svetel, M., Romac, S., and Kostic, V. S. (2001). DYT1 mutation in primary torsion dystonia in a Serbian population. J Neurol *248*, 940-943.

Matsuo, H., Kamakura, K., Saito, M., Okano, M., Nagase, T., Tadano, Y., Kaida, K., Hirata, A., Miyamoto, N., Masaki, T., *et al.* (1999). Familial paroxysmal dystonic choreoathetosis: clinical findings in a large Japanese family and genetic linkage to 2q. Arch Neurol *56*, 721-726.

McNally, E. M., Ly, C. T., and Kunkel, L. M. (1998). Human epsilon-sarcoglycan is highly related to alpha-sarcoglycan (adhalin), the limb girdle muscular dystrophy 2D gene. FEBS Lett *422*, 27-32.

Miner, J. H., Li, C., Mudd, J. L., Go, G., and Sutherland, A. E. (2004). Compositional and structural requirements for laminin and basement membranes during mouse embryo implantation and gastrulation. Development *131*, 2247-2256.

Muntoni, F., Torelli, S., and Ferlini, A. (2003). Dystrophin and mutations: one gene, several proteins, multiple phenotypes. Lancet Neurol *2*, 731-740.

Nagatsu, T., Levitt, M., and Udenfriend, S. (1964). Tyrosine Hydroxylase. the Initial Step in Norepinephrine Biosynthesis. J Biol Chem *239*, 2910-2917.

Naimi, M., Tardieu, S., Depienne, C., Ruberg, M., Brice, A., Dubourg, O., and Leguern, E. (2005). Detection of genomic rearrangements by DHPLC: a prospective study of 90 patients with inherited peripheral neuropathies associated with 17p11.2 rearrangements. Am J Med Genet A *136*, 136-139.

Nairz, K., Stocker, H., Schindelholz, B., and Hafen, E. (2002). High-resolution SNP mapping by denaturing HPLC. Proc Natl Acad Sci U S A *99*, 10575-10580.

Nishiyama, A., Endo, T., Takeda, S., and Imamura, M. (2004). Identification and characterization of -sarcoglycans in the central nervous system. Brain Res Mol Brain Res *125*, 1-12.

Nolte, D., Niemann, S., and Muller, U. (2003). Specific sequence changes in multiple transcript system DYT3 are associated with X-linked dystonia parkinsonism. Proc Natl Acad Sci U S A *100*, 10347-10352.

Nutt, J. G., Muenter, M. D., Aronson, A., Kurland, L. T., and Melton, L. J., 3rd (1988). Epidemiology of focal and generalized dystonia in Rochester, Minnesota. Mov Disord *3*, 188-194.

Nygaard, T. G., Raymond, D., Chen, C., Nishino, I., Greene, P. E., Jennings, D., Heiman, G. A., Klein, C., Saunders-Pullman, R. J., Kramer, P., *et al.* (1999). Localization of a gene for myoclonus-dystonia to chromosome 7q21-q31. Ann Neurol *46*, 794-798.

Nygaard, T. G., Wilhelmsen, K. C., Risch, N. J., Brown, D. L., Trugman, J. M., Gilliam, T. C., Fahn, S., and Weeks, D. E. (1993). Linkage mapping of dopa-responsive dystonia (DRD) to chromosome 14q. Nat Genet *5*, 386-391.

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., *et al.* (2004). Complete sequencing and characterization of 21,243 full-length human cDNAs. Nat Genet *36*, 40-45.

Ott, J. (1999). Analysis of human genetic linkage, 3rd edn (Baltimore: John Hopkins University Press).

Ozawa, E., Mizuno, Y., Hagiwara, Y., Sasaoka, T., and Yoshida, M. (2005). Molecular and cell biology of the sarcoglycan complex. Muscle Nerve *32*, 563-576.

Ozawa, E., Noguchi, S., Mizuno, Y., Hagiwara, Y., and Yoshida, M. (1998). From dystrophinopathy to sarcoglycanopathy: evolution of a concept of muscular dystrophy. Muscle Nerve *21*, 421-438.

Ozelius, L. J., Hewett, J., Kramer, P., Bressman, S. B., Shalish, C., de Leon, D., Rutter, M., Risch, N., Brin, M. F., Markova, E. D., *et al.* (1997a). Fine localization of the torsion dystonia gene (DYT1) on human chromosome 9q34: YAC map and linkage disequilibrium. Genome Res *7*, 483-494.

Ozelius, L. J., Hewett, J. W., Page, C. E., Bressman, S. B., Kramer, P. L., Shalish, C., de Leon, D., Brin, M. F., Raymond, D., Corey, D. P., *et al.* (1997b). The early-onset torsion dystonia gene (DYT1) encodes an ATP-binding protein. Nat Genet *17*, 40-48.

Ozelius, L. J., Kramer, P. L., de Leon, D., Risch, N., Bressman, S. B., Schuback, D. E., Brin, M. F., Kwiatkowski, D. J., Burke, R. E., Gusella, J. F., and *et al.* (1992). Strong allelic association between the torsion dystonia gene (DYT1) andloci on chromosome 9q34 in Ashkenazi Jews. Am J Hum Genet *50*, 619-628.

Pemble, S., Schroeder, K. R., Spencer, S. R., Meyer, D. J., Hallier, E., Bolt, H. M., Ketterer, B., and Taylor, J. B. (1994). Human glutathione S-transferase theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism. Biochem J *300 (Pt 1)*, 271-276.

Pentao, L., Wise, C. A., Chinault, A. C., Patel, P. I., and Lupski, J. R. (1992). Charcot-Marie-Tooth type 1A duplication appears to arise from recombination at repeat sequences flanking the 1.5 Mb monomer unit. Nat Genet *2*, 292-300.

Peters, L. L., Weier, H. U., Walensky, L. D., Snyder, S. H., Parra, M., Mohandas, N., and Conboy, J. G. (1998). Four paralogous protein 4.1 genes map to distinct chromosomes in mouse and human. Genomics *54*, 348-350.

Petrij-Bosch, A., Peelen, T., van Vliet, M., van Eijk, R., Olmer, R., Drusedau, M., Hogervorst, F. B., Hageman, S., Arts, P. J., Ligtenberg, M. J., *et al.* (1997). BRCA1 genomic deletions are major founder mutations in Dutch breast cancer patients. Nat Genet *17*, 341-345.

Pinkel, D., Segraves, R., Sudar, D., Clark, S., Poole, I., Kowbel, D., Collins, C., Kuo, W. L., Chen, C., Zhai, Y., *et al.* (1998). High resolution analysis of DNA copy number

variation using comparative genomic hybridization to microarrays. Nat Genet *20*, 207-211.

Piras, G., El Kharroubi, A., Kozlov, S., Escalante-Alcalde, D., Hernandez, L., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., and Stewart, C. L. (2000). Zac1 (Lot1), a potential tumor suppressor gene, and the gene for epsilon-sarcoglycan are maternally imprinted genes: identification by a subtractive screen of novel uniparental fibroblast lines. Mol Cell Biol *20*, 3308-3315.

Pittock, S. J., Joyce, C., O'Keane, V., Hugle, B., Hardiman, M. O., Brett, F., Green, A. J., Barton, D. E., King, M. D., and Webb, D. W. (2000). Rapid-onset dystonia-parkinsonism: a clinical and genetic analysis of a new kindred. Neurology *55*, 991-995.

Przuntek, H., and Muhr, H. (1983). Essential familial myoclonus. J Neurol *230*, 153-162.

Puget, N., Stoppa-Lyonnet, D., Sinilnikova, O. M., Pages, S., Lynch, H. T., Lenoir, G. M., and Mazoyer, S. (1999). Screening for germ-line rearrangements and regulatory mutations in BRCA1 led to the identification of four new deletions. Cancer Res *59*, 455-461.

Quinn, N. P. (1996). Essential myoclonus and myoclonic dystonia. Mov Disord *11*, 119-124.

Quinn, N. P., Rothwell, J. C., Thompson, P. D., and Marsden, C. D. (1988). Hereditary myoclonic dystonia, hereditary torsion dystonia and hereditary essential myoclonus: an area of confusion. Adv Neurol *50*, 391-401.

Rainier, S., Thomas, D., Tokarz, D., Ming, L., Bui, M., Plein, E., Zhao, X., Lemons, R., Albin, R., Delaney, C., *et al.* (2004). Myofibrillogenesis regulator 1 gene mutations cause paroxysmal dystonic choreoathetosis. Arch Neurol *61*, 1025-1029.

Rannala, B., and Reeve, J. P. (2001). High-resolution multipoint linkage-disequilibrium mapping in the context of a human genome sequence. Am J Hum Genet *69*, 159-178.

Richards, R. N., and Barnett, H. J. (1968). Paroxysmal dystonic choreoathetosis. A family study and review of the literature. Neurology *18*, 461-469.

Rivest, J., and Marsden, C. D. (1990). Trunk and head tremor as isolated manifestations of dystonia. Mov Disord *5*, 60-65.

Schule, B., Kock, N., Svetel, M., Dragasevic, N., Hedrich, K., de Carvalho Aguiar, P., Liu, L., Kabakci, K., Garrels, J., Meyer, E.-M., *et al.* (2004a). Genetic heterogeneity in ten families with myoclonus-dystonia. J Neurol Neurosurg Psychiatry *75*, 1181-1185.

Schule, B., Kock, N., Svetel, M., Dragasevic, N., Hedrich, K., De Carvalho Aguiar, P., Liu, L., Kabakci, K., Garrels, J., Meyer, E. M., *et al.* (2004b). Genetic heterogeneity in ten families with myoclonus-dystonia. J Neurol Neurosurg Psychiatry *75*, 1181-1185.

Segawa, M., Hosaka, A., Miyagawa, F., Nomura, Y., and Imai, H. (1976). Hereditary progressive dystonia with marked diurnal fluctuation. Adv Neurol *14*, 215-233.

Segawa, M., Nomura, Y., Yukishita, S., Nishiyama, N., and Yokochi, M. (2004). Is phenotypic variation of hereditary progressive dystonia with marked diurnal fluctuation/dopa-responsive dystonia (HPD/DRD) caused by the difference of the locus of mutation on the GTP cyclohydrolase 1 (GCH-1) gene? Adv Neurol *94*, 217-223.

Shashidharan, P., Sandu, D., Potla, U., Armata, I. A., Walker, R. H., McNaught, K. S., Weisz, D., Sreenath, T., Brin, M. F., and Olanow, C. W. (2005). Transgenic mouse model of early-onset DYT1 dystonia. Hum Mol Genet *14*, 125-133.

Shi, Z. T., Afzal, V., Coller, B., Patel, D., Chasis, J. A., Parra, M., Lee, G., Paszty, C., Stevens, M., Walensky, L., *et al.* (1999). Protein 4.1R-deficient mice are viable but have erythroid membrane skeleton abnormalities. J Clin Invest *103*, 331-340.

Spitzner, J. R., and Muller, M. T. (1988). A consensus sequence for cleavage by vertebrate DNA topoisomerase II. Nucleic Acids Res *16*, 5533-5556.

Tassin, J., Durr, A., Bonnet, A. M., Gil, R., Vidailhet, M., Lucking, C. B., Goas, J. Y., Durif, F., Abada, M., Echenne, B., *et al.* (2000). Levodopa-responsive dystonia. GTP cyclohydrolase I or parkin mutations? Brain *123*, 1112-1121.

Tchernia, G., Mohandas, N., and Shohet, S. B. (1981). Deficiency of skeletal membrane protein band 4.1 in homozygous hereditary elliptocytosis. Implications for erythrocyte membrane stability. J Clin Invest *68*, 454-460.

Tezenas du Montcel, S., Clot, F., Vidailhet, M., Roze, E., Damier, P., Jedynak, C. P., Camuzat, A., Lagueny, A., Vercueil, L., Doummar, D., *et al.* (2006). Epsilon sarcoglycan mutations and phenotype in French patients with myoclonic syndromes. J Med Genet *43*, 394-400.

Tezzon, F., Zanoni, T., Passarin, M. G., and Ferrari, G. (1998). Dystonia in a patient with deletion of 18p. Ital J Neurol Sci *19*, 90-93.

Uddin, R. K., Zhang, Y., Siu, V. M., Fan, Y. S., O'Reilly, R. L., Rao, J., and Singh, S. M. (2006). Breakpoint Associated with a novel 2.3 Mb deletion in the VCFS region of 22q11 and the role of Alu (SINE) in recurring microdeletions. BMC Med Genet *7*, 18.

Valente, E. M., Misbahuddin, A., Brancati, F., Placzek, M. R., Garavaglia, B., Salvi, S., Nemeth, A., Shaw-Smith, C., Nardocci, N., Bentivoglio, A. R., *et al.* (2003). Analysis of the epsilon-sarcoglycan gene in familial and sporadic myoclonus-dystonia: evidence for genetic heterogeneity. Mov Disord *18*, 1047-1051.

Van den Bergh, P. Y., Tome, F. M., and Fardeau, M. (1995). Etiology and pathogenesis of the muscular dystrophies. Acta Neurol Belg *95*, 123-141.

Vidailhet, M., Tassin, J., Durif, F., Nivelon-Chevallier, A., Agid, Y., Brice, A., and Durr, A. (2001). A major locus for several phenotypes of myoclonus-dystonia on chromosome 7q. Neurology *56*, 1213-1216.

Virtanen, I., Gullberg, D., Rissanen, J., Kivilaakso, E., Kiviluoto, T., Laitinen, L. A., Lehto, V. P., and Ekblom, P. (2000). Laminin alpha1-chain shows a restricted distribution in epithelial basement membranes of fetal and adult human tissues. Exp Cell Res *257*, 298-309.

Walensky, L. D., Shi, Z. T., Blackshaw, S., DeVries, A. C., Demas, G. E., Gascard, P., Nelson, R. J., Conboy, J. G., Rubin, E. M., Snyder, S. H., and Mohandas, N. (1998). Neurobehavioral deficits in mice lacking the erythrocyte membrane cytoskeletal protein 4.1. Curr Biol *8*, 1269-1272.

Wheeler, M. T., and McNally, E. M. (2003). Sarcoglycans in vascular smooth and striated muscle. Trends Cardiovasc Med *13*, 238-243.

Wheeler, M. T., Zarnegar, S., and McNally, E. M. (2002). Zeta-sarcoglycan, a novel component of the sarcoglycan complex, is reduced in muscular dystrophy. Hum Mol Genet *11*, 2147-2154.

Worton, R. G., Duff, C., Sylvester, J. E., Schmickel, R. D., and Willard, H. F. (1984). Duchenne muscular dystrophy involving translocation of the dmd gene next to ribosomal RNA genes. Science *224*, 1447-1449.

Xiao, W., and Oefner, P. J. (2001). Denaturing high-performance liquid chromatography: A review. Hum Mutat *17*, 439-474.

Yokoi, F., Dang, M. T., Li, J., and Li, Y. (2006). Myoclonus, motor deficits, alterations in emotional responses and monoamine metabolism in epsilon-sarcoglycan deficient mice. J Biochem (Tokyo) *140*, 141-146.

Yokoi, F., Dang, M. T., Mitsui, S., and Li, Y. (2005). Exclusive paternal expression and novel alternatively spliced variants of epsilon-sarcoglycan mRNA in mouse brain. FEBS Lett *579*, 4822-4828.

Zimprich, A., Grabowski, M., Asmus, F., Naumann, M., Berg, D., Bertram, M., Scheidtmann, K., Kern, P., Winkelmann, J., Muller-Myhsok, B., *et al.* (2001).

Mutations in the gene encoding epsilon-sarcoglycan cause myoclonus- dystonia syndrome. Nat Genet *29*, 66-69.

Appendix 1. Calculation of normalized ratios to determine the gene dosage alterations for 14 families in 12 exons of the *SGCE* gene. Control indicates a PCR amplicon of *PDZD2* exon 7 (internal control); Ratio-A, -B, or -C indicate the triplicate ratios for the area of each *SGCE* exon divided by the area of *PDZD2* exon 7; The area represent the peak area under the curve; average ratio indicates the average of the triplicate ratios; normalized ratio indicates the average ratio of each affected individual for each exon of *SGCE* divided by the mean ratio of the *SGCE* exon. The individuals with reduced gene copy number are highlighted in yellow rows.

SGCE-exon 1

patients#	Exon 1-A	control	Ratio-A	Exon 1-B	control	Ratio-B	Exon 1-C	control	Ratio-C	Average Ratio	Standard Deviation	Normalized Ratio
M1-27	19643	20383	0.9637	25118	24409	1.0290	21143	19248	1.0985	1.0304	0.0674	0.8890
M2-4	22643	25495	0.8881	23707	26037	0.9105	15239	16219	0.9396	0.9127	0.0258	0.7875
M5-1	26895	29508	0.9114	31428	31972	0.9830	28299	25589	1.1059	1.0001	0.0984	0.8629
M6-1	16020	16650	0.9622	18344	17939	1.0226	19437	15433	1.2594	1.0814	0.1571	0.9330
M8-1	16079	15322	1.0494	18292	17829	1.0260	18183	16040	1.1336	1.0697	0.0566	0.9229
M9-1	25460	25832	0.9856	25011	24131	1.0365	23365	19681	1.1872	1.0698	0.1048	0.9230
M10-11	31606	33451	0.9448	42896	42781	1.0027	32542	26073	1.2481	1.0652	0.1610	0.9191
M11-1	34451	32828	1.0494	39841	38027	1.0477	42322	38874	1.0887	1.0619	0.0232	0.9163
M12-1	20727	20097	1.0313	21073	21997	0.9580	24232	25073	0.9665	0.9853	0.0401	0.8501
M13-1	21411	20965	1.0213	25526	22903	1.1145	26816	22812	1.1755	1.1038	0.0777	0.9524
M14-1	24848	26188	0.9488	23637	24103	0.9807	24855	23850	1.0421	0.9905	0.0474	0.8547
M16-1	20711	20283	1.0211	23446	20038	1.1701	19788	13430	1.4734	1.2215	0.2305	1.0540
M17-1	37134	31755	1.1694	43543	34548	1.2604	33494	26435	1.2670	1.2323	0.0546	1.0632
M18-1	25496	27580	0.9244	24839	25339	0.9803	24102	25092	0.9605	0.9551	0.0283	0.8241

SGCE exon 2

patients#	Exon 2-A	control	Ratio-A	Exon 2-B	control	Ratio-B	Exon 2-C	control	Ratio-C	Average Ratio	Standard Deviation	Normalized Ratio
M1-27	26289	19680	1.3358	24789	18912	1.3108	23345	20145	1.1588	1.2685	0.0958	1.0180
M2-4	4039	9645	0.4188	17903	24612	0.7274	10068	14814	0.6796	0.6086	0.1661	0.4884
M5-1	18201	12696	1.4336	33450	23910	1.3990	33060	26001	1.2715	1.3680	0.0854	1.0979
M6-1	6806	9402	0.7239	15951	20496	0.7782	10015	14750	0.6790	0.7270	0.0497	0.5835
M8-1	12908	9023	1.4306	25752	22081	1.1663	29660	24293	1.2209	1.2726	0.1395	1.0213
M9-1	15625	11090	1.4089	30695	23862	1.2864	30743	23437	1.3117	1.3357	0.0647	1.0720
M10-11	30433	24647	1.2348	38259	27136	1.4099	41311	33076	1.2490	1.2979	0.0973	1.0416
M11-1	37171	29048	1.2796	46385	33620	1.3797	46661	37516	1.2438	1.3010	0.0704	1.0442
M12-1	10758	9035	1.1907	26557	22522	1.1792	22758	19937	1.1415	1.1705	0.0257	0.9394
M13-1				29717	21689	1.3701	32462	27009	1.2019	1.2860	0.1190	1.0321
M14-1	28807	28685	1.0043	35832	27037	1.3253	28903	23911	1.2088	1.1794	0.1625	0.9466
M16-1	25321	21131	1.1983	24837	20952	1.1854	30188	23811	1.2678	1.2172	0.0443	0.9769
M17-1	44479	33695	1.3200	51045	33678	1.5157	45589	32667	1.3956	1.4104	0.0987	1.1320
M18-1	34322	24995	1.3732	33194	20466	1.6219	32180	22736	1.4154	1.4701	0.1331	1.1799

SGCE exon 3

patients#	Exon 3-A	control	Ratio-A	Exon 3-B	control	Ratio-B	Exon 3-C	control	Ratio-C	Average Ratio	Standard Deviation	Normalized Ratio
M1-27	3113	2045	1.5222	6616	4921	1.3444	9813	6206	1.5812	1.4826	0.1233	1.1837
M2-4	3060	3588	0.8528	3156	3494	0.9033	4804	6134	0.7832	0.8464	0.0603	0.6758
M5-1	2300	2128	1.0808	4904	3758	1.3049	7689	4990	1.5409	1.3089	0.2301	1.0450
M6-1	1521	3851	0.3950	1343	3546	0.3787	2110	5038	0.4188	0.3975	0.0202	0.3174
M8-1	2457	2224	1.1048	4891	3869	1.2642	7414	5115	1.4495	1.2728	0.1725	1.0162
M9-1	12236	15135	0.8085	10843	14141	0.7668	12325	14391	0.8564	0.8106	0.0449	1.0726
M10-11	6482	4974	1.3032	7199	6468	1.1130	13309	15002	0.8871	1.1011	0.2083	0.8791
M11-1	13009	12159	1.0699	12922	11989	1.0778	16115	18830	0.8558	1.0012	0.1260	0.7993
M12-1	6712	8288	0.8098	7292	10274	0.7098	10898	13656	0.7980	0.7725	0.0547	1.0223
M13-1				5638	4317	1.3060	15292	15401	0.9929	1.1495	0.2214	0.9177
M14-1				6878	5491	1.2526	14514	14889	0.9748	1.1137	0.1964	0.8892
M16-1	5555	4012	1.3846				8065	5666	1.4234	1.4040	0.0274	1.1210
M17-1	8254	6292	1.3118	16344	14538	1.1242	18751	15430	1.2152	1.2171	0.0938	0.9717
M18-1	14172	9640	1.4701	10778	7863	1.3707	11579	7835	1.4779	1.4396	0.0597	0.8174

SGCE exon 4

patients#	Exon 4-A	control	Ratio-A	Exon 4-B	control	Ratio-B	Exon 4-C	control	Ratio-C	Average Ratio	Standard Deviation	Normalized Ratio
M1-27	18334	14239	1.2876	19894	20887	0.9525	24755	20305	1.2192	1.1531	0.1771	1.0505
M2-4	15173	11364	1.3352	22487	20130	1.1171	24873	20574	1.2090	1.2204	0.1095	1.1119
M5-1	31178	26359	1.1828	28884	24369	1.1853	32829	25441	1.2904	1.2195	0.0614	1.1111
M6-1	10370	15724	0.6595	9190	14565	0.6310	12717	18022	0.7056	0.6654	0.0377	0.6062
M8-1	15985	12128	1.3180	23103	20774	1.1121	22712	19555	1.1614	1.1972	0.1075	1.0907
M9-1	33307	27856	1.1957	21838	20599	1.0601	25355	21045	1.2048	1.1535	0.0810	1.0510
M10-11	42995	41634	1.0327	36188	36990	0.9783	29284	27262	1.0742	1.0284	0.0481	0.9369
M11-1	48897	41953	1.1655	37295	36005	1.0358	33450	31555	1.0601	1.0871	0.0690	0.9905
M12-1	16232	18022	0.9007	12679	13679	0.9269	16691	19782	0.8437	0.8904	0.0425	0.8113
M13-1	26855	22215	1.2089	27200	24275	1.1205	28587	23988	1.1917	1.1737	0.0469	1.0693
M14-1	25899	21821	1.1869	33006	29631	1.1139	33570	25909	1.2957	1.1988	0.0915	1.0922
M16-1	15636	13025	1.2005	21749	20177	1.0779	15988	11851	1.3491	1.2092	0.1358	1.1016
M17-1	41322	33918	1.2183	50300	39332	1.2789	46346	34097	1.3592	1.2855	0.0707	1.1712
M18-1	37268	35945	1.0368	44251	41027	1.0786	67051	63110	1.0624	1.0593	0.0211	0.9651

SGCE exon 5

patients#	Exon 5-A	control	Ratio-A	Exon 5-B	control	Ratio-B	Exon 5-C	control	Ratio-C	Average Ratio	Standard Deviation	Normalized Ratio
M1-27	11598	5478	2.1172	22804	12541	1.8184	22311	13407	1.6641	1.8666	0.2303	1.0324
M2-4	9934	4892	2.0307	12697	5802	2.1884	13536	5967	2.2685	2.1625	0.1210	1.1961
M5-1	12194	6172	1.9757	20839	11093	1.8786	26613	16376	1.6251	1.8265	0.1810	1.0102
M6-1	4217	4040	1.0438				6696	5860	1.1427	1.0932	0.0699	0.6047
M8-1	9656	4602	2.0982	10466	4998	2.0940	8892	4029	2.2070	2.1331	0.0640	1.1798
M9-1	12966	6129	2.1155	19369	10442	1.8549	21537	12147	1.7730	1.9145	0.1788	1.0589
M10-1	22796	12093	1.8851	23687	11555	2.0499	22336	12528	1.7829	1.9060	0.1347	1.0542
M11-11				33854	26138	1.2952	38758	23613	1.6414	1.4683	0.2448	0.8121
M12-1	13747	6809	2.0189				10380	4422	2.3474	2.1831	0.2322	1.2075
M13-1	13646	6090	2.2407	25091	13756	1.8240	10550	4669	2.2596	2.1081	0.2462	1.1660
M14-1	22502	13085	1.7197	24183	13878	1.7425	27539	16329	1.6865	1.7162	0.0282	0.9492
M16-1	19728	11750	1.6790	12971	5910	2.1948	24839	15650	1.5872	1.8203	0.3275	1.0068
M17-1	30451	16144	1.8862	36568	19230	1.9016	25097	14318	1.7528	1.8469	0.0818	1.0215
M18-1	17827	11135	1.6010	23092	13838	1.6687	23694	13990	1.6936	1.6545	0.0479	0.9151

SGCE exon 6

patients#	Exon 6-A	control	Ratio-A	Exon 6-B	control	Ratio-B	Exon 6-C	control	Ratio-C	Average Ratio	Standard Deviation	Normalized Ratio
M1-27	2939	1631	1.8020	2951	1734	1.7018				1.7519	0.0708	1.1050
M2-4	6728	3946	1.7050				1610	1231	1.3079	1.5064	0.2808	0.9502
M5-1	3601	3122	1.1534				6398	4008	1.5963	1.3749	0.3132	0.8672
M6-1	3097	2469	1.2544	3322	2493	1.3325	3202	2407	1.3303	1.3057	0.0445	0.8236
M8-1	3220	2354	1.3679	2421	1530	1.5824	6111	3960	1.5432	1.4978	0.1142	0.9447
M9-1	3062	2247	1.3627	4372	3509	1.2459	6904	4707	1.4668	1.3585	0.1105	0.8569
M10-11	4373	2205	1.9832	3762	2611	1.4408	6627	4881	1.3577	1.5939	0.3397	1.0054
M11-1	7273	4628	1.5715	6492	5940	1.0929	9598	7483	1.2826	1.3157	0.2410	0.8299
M12-1	1515	1113	1.3612				3280	1930	1.6995	1.5303	0.2392	0.9653
M13-1	5181	3038	1.7054	7793	5273	1.4779				1.5917	0.1609	1.0039
M14-1	4025	2804	1.4354	3634	2815	1.2909	7121	5551	1.2828	1.3364	0.0859	0.8429
M16-1	6689	4082	1.6387				6143	3943	1.5580	1.5983	0.0571	1.0081
M17-1				6198	4094	1.5139	8418	5244	1.6053	1.5596	0.0646	0.9837
M18-1	3953	3493	1.1317	6588	3922	1.6798	8319	6224	1.3366	1.3827	0.2769	0.8721

SGCE-exon7

patients#	Exon 7-A	control	Ratio-A	Exon 7-B	control	Ratio-B	Exon 7-C	control	Ratio-C	Average Ratio	Standard Deviation	Normalized Ratio
M1-27	17501	11557	1.5143	12176	7578	1.6068	13986	9089	1.5388	1.5533	0.0479	1.0849
M2-4	17701	12176	1.4538	12033	7892	1.5247	14441	9558	1.5109	1.4965	0.0376	1.0452
M5-1	20880	14383	1.4517	14297	8666	1.6498	28934	26349	1.0981	1.3999	0.2795	0.9778
M6-1	13041	8147	1.6007	9288	6005	1.5467	10931	7199	1.5184	1.5553	0.0418	1.0863
M8-1	14330	8813	1.6260	7770	5063	1.5347	12642	7484	1.6892	1.6166	0.0777	1.1292
M9-1	18529	12728	1.4558	6485	3653	1.7753	24392	20912	1.1664	1.4658	0.3045	1.0238
M10-11	37554	27948	1.3437	26568	22813	1.1646	30928	23924	1.2928	1.2670	0.0923	0.8850
M11-1	35121	29646	1.1847	23547	20582	1.1441	31249	24224	1.2900	1.2062	0.0753	0.8425
M12-1	14720	9070	1.6229	10036	6539	1.5348	13933	8009	1.7397	1.6325	0.1028	1.1402
M13-1	15995	10824	1.4777	9158	5441	1.6831	28228	21840	1.2925	1.4845	0.1954	1.0368
M14-1	16287	11424	1.4257	14981	9698	1.5448	27256	23411	1.1642	1.3782	0.1946	0.9626
M16-1	15270	9358	1.6318	7717	4763	1.6202	10918	5626	1.9406	1.7309	0.1818	1.2090
M17-1	36560	28794	1.2697	31721	24446	1.2976	14849	8462	1.7548	1.4407	0.2724	1.0063
M18-1	31939	31185	1.0242	25469	22919	1.1113	20332	15335	1.3259	1.1538	0.1553	0.8059

SGCE-exon 8

patients#	Exon 8-A	control	Ratio-A	Exon 8-B	control	Ratio-B	Exon 8-C	control	Ratio-C	Average Ratio	Standard Deviation	Normalized Ratio
M1-27	18197	15949	1.1409	22353	18985	1.1774	17351	16714	1.0381	1.1188	0.0722	0.9266
M2-4	14997	14345	1.0455	17904	16397	1.0919	20365	16877	1.2067	1.1147	0.0830	0.9231
M5-1	22796	18410	1.2382	29410	22211	1.3241	35180	26329	1.3362	1.2995	0.0534	1.0762
M6-1	12695	11826	1.0735	15850	12728	1.2453	12729	10337	1.2314	1.1834	0.0954	0.9800
M8-1	14436	14537	0.9931	20773	18476	1.1243	21088	19413	1.0863	1.0679	0.0675	0.8844
M9-1	16503	16206	1.0183	22088	18978	1.1639	29193	26143	1.1167	1.0996	0.0743	0.9107
M10-11	34086	21292	1.6009	32082	26859	1.1945	32612	24871	1.3112	1.3689	0.2092	1.1336
M11-1	35767	23377	1.5300	48712	34101	1.4285	40853	29022	1.4077	1.4554	0.0655	1.2053
M12-1	20559	16722	1.2295	24789	18650	1.3292	25018	22004	1.1370	1.2319	0.0961	1.0202
M13-1	21180	19360	1.0940	21714	18651	1.1642	24661	22653	1.0886	1.1156	0.0422	0.9239
M14-1	22270	20252	1.0996	22675	19734	1.1490	21197	20030	1.0583	1.1023	0.0454	0.9129
M16-1	19081	17295	1.1033	18228	15803	1.1535	19508	16640	1.1724	1.1430	0.0357	0.9466
M17-1	38600	30080	1.2832	38009	28754	1.3219	32672	26400	1.2376	1.2809	0.0422	1.0608
M18-1	30853	25042	1.2321	27556	22488	1.2254	23552	21174	1.1123	1.1899	0.0673	0.9854

SGCE-exon 9

patients#	Exon 9-A	control	Ratio-A	Exon 9-B	control	Ratio-B	Exon 9-C	control	Ratio-C	Average Ratio	Standard Deviation	Normalized Ratio
M1-27	16368	8534	1.9180	14903	8191	1.8194	21727	11336	1.9166	1.8847	0.0565	1.0447
M2-4	16890	9300	1.8161	14743	8462	1.7423	14516	7870	1.8445	1.8010	0.0528	0.9983
M5-1	23734	12361	1.9201	18549	10147	1.8280	20174	10141	1.9894	1.9125	0.0809	1.0601
M6-1	14023	7900	1.7751	9227	5968	1.5461	10952	6502	1.6844	1.6685	0.1153	0.9248
M8-1	10918	5984	1.8245	10112	5980	1.6910	10676	6105	1.7487	1.7547	0.0670	0.9726
M9-1	12399	7715	1.6071	9603	6433	1.4928	11737	7593	1.5458	1.5486	0.0572	0.8584
M10-11	14424	9197	1.5683	11926	6749	1.7671	14350	8420	1.7043	1.6799	0.1016	0.9312
M11-1	21647	12161	1.7800	18166	10912	1.6648	19870	11815	1.6818	1.7089	0.0622	0.9472
M12-1	4783	2391	2.0004	4073	3235	1.2590	4277	2694	1.5876	1.6157	0.3715	0.8956
M13-1	12840	8487	1.5129	10784	7401	1.4571	13382	9338	1.4331	1.4677	0.0410	0.8135
M14-1	18238	10100	1.8057	19280	11180	1.7245	19728	10610	1.8594	1.7965	0.0679	0.9958
M16-1	11658	7764	1.5015	11350	7425	1.5286	12910	7212	1.7901	1.6067	0.1593	0.8906
M17-1	30637	16864	1.8167	25894	12986	1.9940	27768	15105	1.8383	1.8830	0.0967	1.0437
M18-1	20506	9494	2.1599	19208	9047	2.1231	19400	10485	1.8503	2.0444	0.1692	1.1332

SGCE-exon 10

patients#	Exon10-A	control	Ratio-A	Exon10-B	control	Ratio-B	Exon10-C	control	Ratio-C	Average Ratio	Standard Deviation	Normalized Ratio
M1-27	23829	14567	1.6358	19821	10421	1.9020	19465	12515	1.5553	1.6977	0.1814	1.0748
M2-4	19137	11514	1.6621	18145	12259	1.4801	16293	11139	1.4627	1.5350	0.1104	0.9717
M5-1	16829	9417	1.7871	24933	16768	1.4869	29713	22704	1.3087	1.5276	0.2418	0.9671
M6-1	10147	6650	1.5259	16300	11144	1.4627	12271	8230	1.4910	1.4932	0.0317	0.9453
M8-1	15189	8423	1.8033	17032	11899	1.4314	15874	11425	1.3894	1.5414	0.2278	0.9758
M9-1	18253	10418	1.7521	23609	15404	1.5327	22351	16388	1.3639	1.5495	0.1946	0.9810
M10-11	36851	19932	1.8488	32336	18974	1.7042	34622	21273	1.6275	1.7269	0.1124	1.0932
M11-1	36105	19273	1.8733	38312	21552	1.7777	36417	22428	1.6237	1.7582	0.1259	1.1131
M12-1	18920	11846	1.5972	17637	11452	1.5401	23335	15917	1.4660	1.5344	0.0657	0.9714
M13-1	13526	7998	1.6912	26881	20182	1.3319	13214	7132	1.8528	1.6253	0.2666	1.0289
M14-1	24923	16680	1.4942	20354	13935	1.4606	23311	16710	1.3950	1.4500	0.0504	0.9179
M16-1	15644	9801	1.5962	12727	7431	1.7127	15711	10470	1.5006	1.6031	0.1062	1.0149
M17-1	33121	20941	1.5816	31947	22246	1.4361	33212	22703	1.4629	1.4935	0.0775	0.9455
M18-1	26247	17036	1.5407	22527	16493	1.3659	24135	16819	1.4350	1.4472	0.0880	0.9162

SGCE-exon11

patients#	Exon11-A	control	Ratio-A	Exon11-B	control	Ratio-B	Exon11-C	control	Ratio-C	Average Ratio	Standard Deviation	Normalized Ratio
M1-27	18159	19617	0.9257	20987	24352	0.8618	23879	28301	0.8438	0.8771	0.0430	0.9423
M2-4	11468	10623	1.0795	21657	24872	0.8707	18914	21751	0.8696	0.9400	0.1209	1.0098
M5-1	14538	12969	1.1210	23324	24730	0.9431	26308	29678	0.8864	0.9835	0.1224	1.0566
M6-1	7685	7857	0.9781	20204	22676	0.8910	19916	22957	0.8675	0.9122	0.0583	0.9800
M8-1	10519	10958	0.9599				17679	20954	0.8437	0.9018	0.0822	0.9689
M9-1	20849	19590	1.0643	30485	31201	0.9771	28730	30469	0.9429	0.9947	0.0626	1.0687
M10-11	31969	31799	1.0053	37193	35425	1.0499	37230	37826	0.9842	1.0132	0.0335	1.0885
M11-1	31855	31293	1.0180	36772	39112	0.9402	39268	44635	0.8798	0.9460	0.0693	1.0163
M12-1	23662	24031	0.9846	23947	25451	0.9409	30979	34844	0.8891	0.9382	0.0478	1.0080
M13-1	24487	23606	1.0373	24203	25380	0.9536	33101	33220	0.9964	0.9958	0.0419	1.0698
M14-1	23049	22873	1.0077	21870	25247	0.8662	27192	29919	0.9089	0.9276	0.0726	0.9966
M16-1	18152	18618	0.9750	20412	20562	0.9927	23751	24474	0.9705	0.9794	0.0118	1.0522
M17-1	38135	38180	0.9988	31486	32685	0.9633	38638	40004	0.9659	0.9760	0.0198	1.0486
M18-1	23550	25380	0.9279	25422	26294	0.9668	26596	31348	0.8484	0.9144	0.0604	0.9824

SGCE-exon 12

patients#	Exon12-A	control	Ratio-A	Exon12-B	control	Ratio-B	Exon12-C	control	Ratio-C	Average Ratio	Standard Deviation	Normalized Ratio
M1-27	8948	3705	2.4151	10665	6171	1.7282	14487	5567	2.6023	2.2486	0.4602	0.8678
M2-4	9103	3391	2.6845	11647	4501	2.5876	11301	4309	2.6227	2.6316	0.0490	1.0156
M5-1	12944	4453	2.9068	16925	6393	2.6474	13631	5159	2.6422	2.7321	0.1513	1.0544
M6-1				7866	3310	2.3764	8054	3415	2.3584	2.3674	0.0127	0.9137
M8-1	12403	4537	2.7337	10597	4302	2.4633	9334	3813	2.4479	2.5483	0.1608	0.9835
M9-1	16359	5790	2.8254	13266	5111	2.5956	12185	4729	2.5767	2.6659	0.1385	1.0289
M10-11	22756	9320	2.4416	20302	8323	2.4393	16029	9809	1.6341	2.1717	0.4655	0.8381
M11-1	24219	10560	2.2935	23462	9707	2.4170	17688	8490	2.0834	2.2646	0.1687	0.8740
M12-1				11268	5450	2.0675	10322	5141	2.0078	2.0377	0.0422	0.7864
M13-1	15426	5414	2.8493	14081	5063	2.7812	15219	5123	2.9707	2.8671	0.0960	1.1065
M14-1	10759	3884	2.7701	13234	4954	2.6714	10792	4690	2.3011	2.5808	0.2473	0.9960
M16-1	10289	4777	2.1539	9437	4525	2.0855	8345	3644	2.2901	2.1765	0.1041	0.8400
M17-1	16080	5937	2.7084	19048	7985	2.3855	19131	9101	2.1021	2.3987	0.3034	0.9257
M18-1	12623	5119	2.4659	14003	5993	2.3366	14157	6847	2.0676	2.2900	0.2032	0.8838

Appendix 2-1. Detected single nucleotide polymorphisms (SNPs) in the *L3MBTL4* gene for *SGCE*-negative families with M-D. M1-16 and M1-54 are affected individuals from family 1 while M1-21 is a normal control married into family 1. M2-1 and M2-4 are affected individuals from family 2 while M2-5 is normal control married into family 2. M5-1, M6-1, M8-1 and M10-11 are affected individuals of families 5, 6, 8 and 10. N/N indicates unknown nucleotide.

				Affected and unaffected Individuals								
Gene name	Exons	Base Position	SNP#	M1-21	M1-16	M1-54	M2-4	M5-1	M6-1	M8-1	M10-11	Mutation status
AK093255	4	6291903		C/T	C/C	C/C	C/C	C/C	C/C			CAC(T), wobble,
	20	6033576		T/T	C/T	C/T	C/T	C/T	C/T			GT(C)G, missense
	20	6033573		A/A	A/C	A/C	A/C	A/C	A/C			GA(C)G, missense
	20	6033543		T/T	C/T	C/T	C/T	C/T	C/T			GT(C)G, missense
	20	6033528		T/T	C/T	C/T	C/T	C/T	C/T			GT(C)G, missense
	21	6024189		C/T	T/T	T/T	C/T	C/C	C/T			intron, 9 bp upstream
NM_173464	22	6020306		G/T	G/T	G/T	G/T	G/T	G/T			ATG(T), wobble
	22	6020268		T/G	G/G	G/G	T/G	T/T	G/G			G(T)CC, missense
	22	6020267		G/C	G/G	G/G	G/C	G/C	G/G			GC(G)C, missense
	22	6020262	reported	G/A	A/A	A/A	G/A	G/G	A/A			GG(A)C, missense
	22	6020244		T/G	T/T	T/T	T/G	T/G	T/T			TG(T)A, stop codon
	22	6020205		C/C	C/T	C/C	C/C	C/C	C/T			3'-UTR
	22	6020192		C/T	T/T	T/T	C/T	C/T	T/T			3'-UTR
	22	6020169	reported	A/G	G/G	G/G	A/G	A/A	A/G			3'-UTR,
	New-A-21	5946116		T/C	T/T	T/T	C/C	C/C	T/T	C/T	C/T	Intron, 26 bp upstream
	NewA-21	5946145		G/A	G/G	G/G	A/A	A/A	G/G	A/G	A/G	intron, 55 bp upstream
	New-A-21	5946182		C/C	C/T	C/T	C/C	C/C	C/T	C/C	C/C	intron, 92 bp upstream
	New-A-21	5946238		A/G	A/A	A/A	G/G	G/G	A/A	A/G	A/G	intron, 118 bp upstream
	New-A-20	5950165		C/T	T/T	T/T	C/C	C/C	T/T	C/T	C/T	intron, 10 bp upstream
	New-C	6019525		C/C	C/C	C/C	C/C	N/N	N/N	C/T	C/C	intron, 34 bp downstream
	New-F	6162254		G/G	G/G	G/G	A/G	G/G	A/G	G/G	A/G	intron, 77 bp downstream
	New-F	6162236		G/G	G/G	G/G	A/G	G/G	A/G	G/G	A/G	intron, 95 bp downstream
	New-B	6186622		C/C	C/C	C/C	C/T	C/C	C/T	C/T	C/C	intron , 18 bp upstream

Appendix 2-2. Detected single nucleotide polymorphisms (SNPs) in the *ARHGAP28* gene for *SGCE*-negative families with M-D. M1-16, M1-54, M1-67, M2-1, M2-4, M5-1, M6-1, M8-1 and M10-11 are affected individuals; M1-21 and M2-5 are unaffected individuals.

				Affected and unaffected Individuals from families with M-D											
Gene name	Exons	Position	SNP#	M1-21	M1-16	M1-54	M1-67	M2-5	M2-1	M2-4	M5-1	M6-1	M8-1	M10-11	Mutation status
ARHGAP28-Variant a	1	6824599	novel	T/T	T/T	T/T				T/A	T/T	T/T	T/T	T/T	intron, 61 bp downstream of exon 1, 5'-UTR
	12	6879973	reported	A/A	G/A	G/G	G/G	G/A	G/A	G/A	G/A	A/A			CCG(A), wobble, Sense mutation
	13	6880434	reported	A/A	G/A	G/G	G/G	G/A	G/A	G/A	G/A	A/A			CCG(A), wobble, Sense mutation
	16	6898972		G/G	G/G	G/G				G/G	G/G	G/G	C/G	G/G	TCG(C), wobble
	16	6899075		A/G	A/G	A/G				A/G	A/G	A/G	G/G	G/G	intron, 51 bp downstream of exon 16
	17	6902005	novel	A/G	A/A	A/A				A/A	A/G	A/A	A/A	N/N	intron, 55 bp upstream of exon 17
	17	6902143	reported	A/C	C/C	C/C				C/C	A/C	C/C	C/C	N/N	CAA/CCA Glutamine/Proline
	17	6902317	novel	C/T	T/T	T/T				T/T	C/T	T/T	T/T	N/N	3'-UTR
	17	6902332	novel	A/G	G/G	G/G				G/G	A/G	G/G	G/G	N/N	3'-UTR
	17	6902539	novel	C/T	C/C	C/C				C/C	C/C	C/T	C/C	N/N	3'-UTR
	17	6902634	novel	C/G	C/C	C/C				C/C	C/C	C/C	C/C	N/N	3'-UTR
	17	6902670	novel	A/T	A/A	A/A				A/A	A/A	A/A	A/A	N/N	3'-UTR
	17	6902682	novel	A/T	A/A	A/A				A/T	A/A	A/A	A/A	N/N	3'-UTR
	17	6902690	novel	T/C	C/C	C/C				C/C	C/C	C/C	C/C	N/N	3'-UTR
ARHGAP28-Variant b (FLJ10312, AK001174, Loc388462, AB037735)	2	6814937	novel	C/C	C/C	C/C	C/C	C/C	T/C	T/C	C/C	C/C			CCG/CTG missense, mutation Proline to Leucine
	18	6888415	reported	C/C	T/C	T/T	T/T	T/C	T/C	T/C	T/T	T/C			intron, 20 bp upstream of exon18

Appendix 2-3. Detected single nucleotide polymorphisms (SNPs) in the *ZFP161*, *EPB41L3* and *PTPRM* genes for *SGCE*-negative families with M-D. M1-15, M1-16, M1-54, M2-4, M5-1, M6-1 are affected individuals; M1-21, M1-42, and M2-5 are unaffected individuals.

				Affected and unaffected Individuals from families with M-D									
Gene name	Exons	Base Position	SNP#	M1-42	M1-15	M1-21	M1-16	M1-54	M2-5	M2-4	M5-1	M6-1	mutation status
ZFP-161	2	5282030	reported	C/C	C/C	T/C	C/C	C/C					Coding, wobble ACC/T, Threonine
EPB41L3 (KIAA0987)	10	5414370	novel	T/C	T/T	T/C	T/C	T/C					non-coding, intron10 8 bp upstream of exon 10
	13	5400574	reported	C/T	C/C	C/C	C/C	C/C					Coding, wobble, ACC/T, Threonine
	18	5388136	novel			G/G	G/G	G/C		G/G	G/G	G/G	Coding, mis-sense CAG/C, Glutamine/ Histidine
PRPRM	9	7939205				G/A	G/A	G/A		A/G			Coding, wobble CTG/A, Leucine
	24	824335	reported			C/T	T/T	T/T		T/T			Coding, Mis-sense CT(C)C, Leucine/Proline
	28	8333530				A/G	G/G	G/G	A/G	A/G	G/G	A/G	intron 28, 20 bp downstream of exon 28
	35	8374543				C/T	C/T	C/T	C/C	C/C	C/C	C/C	Intron 34, 15 bp upstream of exon 35
	36	8377065				T/T	T/C	T/T	T/T	T/C	T/T	T/T	Intron 35, 5 bp upstream of exon 36

Appendix 2-4. Detected single nucleotide polymorphisms (SNPs) in the *MGC17515* and *LOC400643* genes for *SGCE*-negative families with M-D. M1-16, M1-54, M1-67, M2-4, M5-1, M6-1, M8-1 M9-1, and M10-11 are affected individuals; M1-21 and M2-5 are unaffected individuals.

				Affected and unaffected individuals											
Gene name	Exons	Base Position	SNP#	M1-21	M1-16	M1-54	M1-67	M2-5	M2-4	M5-1	M6-1	M8-1	M9-1	M10-11	Mutation status
MGC17515	1A	5227612	rs11875069	C/C	C/T	C/T		C/C	C/C	C/C	C/C				coding, wobble GCC/T, Alanine
	1B	5227605	rs7242964	C/C	C/C	C/C		N/N	C/G	C/G	C/G				coding, C/GCC Proline/Alanine
	1C	5227575	rs11875052	C/C	C/A	C/A		N/N	C/C	C/C	C/C				coding, C/ACA Proline/threonine
	1D	5227478	rs7244432	A/G	A/A	A/G		N/N	A/G	A/G	A/G				coding, A(A/G)A Lysine/Glutamic acid
	1E	5227476	novel	T/T	T/T	T/T		N/N	C/T	C/T	C/T				coding, T/CTA Leucine/Leucine
LOC400643	1	6918422	novel	C/C	C/T	C/T	C/C	T/T	C/T	C/C	C/T	C/T	C/C	C/T	non-coding, 14 bp downstream of exon 1
	2A	6918291	rs4797266	C/T	C/C	C/T	C/T	C/T	T/T	C/T	C/C	T/T	C/T	C/T	non-coding, 76 bp upstream of exon 2
	2B	6918257	rs4798517	C/T	T/T	C/T	C/T	C/T	C/C	C/T	T/T	C/C	C/T	C/T	non-coding, 41 bp upstream of exon 2

Appendix 2-5. Detected single nucleotide polymorphisms (SNPs) in the *LOC388459* gene for *SGCE*-negative families with M-D. M1-16, M1-54, M2-4, M5-1, M6-1, M8-1 and M10-11 are affected individuals; M1-21 is unaffected individual.

				Affected and unaffected Individuals from M-D families								
Gene name	Exons	Base Position	SNP#	M1-21	M1-16	M1-54	M2-4	M5-1	M6-1	M8-1	M10-11	Mutation status
LOC388459	4	5620531	novel	A/A	A/A	A/A	A/G	A/A	A/A	A/A	A/A	Coding, wobble, CGG(A), Arginine
	4	5620676	novel	G/G	T/T	G/T	G/T	T/T	G/T	G/T	G/G	Coding, missense mutation G(T)CT, Alanine/Threonine DHPLC confirmed SNP
	6	5636700	novel	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	non-coding, intron 6 18 bp downsream of exon 6
	7	5759897	rs3752062	T/T	T/T	T/T	T/T	T/C	T/C	T/T	T/T	non-coding, intron 7 32 bp downstream of exon 7
	9	5829411	novel	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/G	coding, wobble, ACG(A), Threonine
	10	5832789	novel	C/T	C/C	C/C	C/C	C/C	C/C	C/T	C/C	Coding, missense mutation C(T)TC, Leucine/Phenylananine Normal control has C/T
	13	5859371	novel	A/A	A/A	A/A	T/A	A/A	A/A	A/A	A/A	non-coding, intron 13 17 bp downstream exon 13
	14	5862327	novel	C/C	C/T	T/T	T/T	T/T	C/T	C/T	T/T	non-coding, intron 14, 31 bp downstream of exon 14

Appendix 2-6. Detected single nucleotide polymorphisms (SNPs) in the *LOC388461* gene for *SGCE*-negative families with M-D. M1-16, and M1-54, M2-1, M2-4, M5-1, M6-1, M8-1, and M10-11 are affected individuals; M1-21 and M2-5 are unaffected individuals. INS represents insertion.

				Affected and unaffected Individuals from families with M-D												
Gene name	Exons	Base Position	SNP#	M1-21	M1-16	M1-54	M1-67	M2-5	M2-4	M5-1	M6-1	M8-1	M9-1	M10-11	M12-1	Mutation status
LOC388461	1	6741758		T/T	T/T	T/T	T/C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	non-coding, 9 bp of exon 1
	1	6741746		C/C	T/T	C/T	CC	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/T	coding, missense ACG/ATG, Threonine/Methionine
		6730785	reported	A/G	A/A	A/A	G/G	A/G	G/G	G/G	A/G	A/A	A/G	A/G	A/G	non-coding, 93 bp downstream of exon 2
	3	6720016		C/G	G/G	G/G	C/C	G/G	C/C	C/C	C/G	G/G	G/C	C/G	C/G	non-coding, 80 bp upstream of exon 3
	3	6719674	novel						INS 5'-CGCGGCTGTCAGGGGCCCCCGGGCGA-3'							non-coding, 26 bp insertion at 45 bp downstream of exon 3 confirmed in 2 of 50 normal controls
	4	6719046		A/G	G/G	G/G	A/A	A/G	A/A	A/A	A/G	G/G	A/G	A/G	A/G	Coding, wobble CTA(G), Leucine normal control has A/G
	4	6719021		C/T	T/T	T/T	C/C	C/T	C/C	C/C	C/T	T/T	C/T	C/T	C/T	Coding, wobble GA(C/T), Aspartic acid normal control has C/T

CURRICULUM VITAE

Fabin Han
Email: fhan@ohri.ca
Lab Phone: 613-737-8960

EDUCATION

January 2003-Decembe 2006: PhD Student in Microbiology and Immunology-Human Molecular Genetics program, Ottawa Health Research Institute, University of Ottawa, Ottawa, Canada.

Supervisors: Dr. Dennis E Bulman, PhD (Geneticist) and Dr. David A Grimes, MD (Neurologist)

August 1997 – December 1999: World Health Organization Fellowship, Molecular genetics Program, Texas A & M University College Station, USA.

Supervisor: Dr. David L Busbee, PhD (Geneticist)

August 1986-July 1989: MPH, Shandong Medical University, Jinan, Shandong, P.R. China.

Supervisor: Dr. Tian Ding, MD (Genetic Toxicologist)

September 1979-July 1984: MD, Shandong Medical University, Jinan, Shandong, P.R. China.

Awards

Faculty of Medicine Awards of Excellence December 2006 – Graduate Studies Microbiology and Immunology, University of Ottawa, Ottawa, Canada.

October, 2006, University of Ottawa Travel award for the 36th annual meeting of Society of Neuroscience in Atlanta, USA.

May, 2006: Excellence for Sanofi-Aventis Biotech Challenge in East Ontario, Ottawa, Canada.

January, 2003: PhD Admission Scholarship, University of Ottawa, Canada.

November, 2003: University of Ottawa Travel Award for the 53th annual meeting, American Society of Human Genetics, Los Angeles, CA, USA.

Expertises

Molecular genetics and genomics (DNA techniques): Preparation and construction of plasmid DNA, Transfection of plasmid DNA into E.coli and mammalian cells, Isolation of DNA from human blood, tissues and cell cultures, Polymerase Chain reaction (PCR) and real time PCR, DNA cloning, cDNA cloning, Southern blotting, restriction endonuclease analysis of DNA, gel electrophoresis, DNA sequencing, DNA genotyping with

microsatellite markers (STSs) and SNP markers, mutation analysis with restriction enzymes, allele-specific oligonucleotide hybridization, sequencing, DHPLC, and SSCP.

RNA, Protein techniques and gene expression: Isolation of total RNA and mRNA from human blood, tissues and cell cultures, RT-PCR, Northern blotting, Western blotting and immuno-percipitation analysis of proteins, Kinase assay for protein phosphorylation analysis, RNAi to knock down the gene expression in cultured cells.

Cell biology and In vitro cell culture: Human fibroblast, breast cancer cell culture, immunofluerescece and immuno-histochemical analyses of the cultured cells labelled with fluorescent antibodies, flow cytometry technology (FACS).

Genetic linkage and Bioinformatics: Genetic Linkage analysis and association studies, PCR-based markers such as STSs (sequence tagged sites) and ESTs (Expressed sequence tags), Polymorphic markers such as RFLP and SNP, genetic map, physical map, and RH map, using the bioinformatics resources such as NCBI, genome browser (UCSC), GDB to compare the DNA sequences, protein sequences by BLAST and Alignment analysis. Genomic sequence analysis of repetitive DNA such as the Alu sequence of short interspersed nuclear elements (SINEs) and long interspersed nuclear elements (LINEs) using the bioinformatics tools (REPEATMASKER).

Animal models: Pharmacological and toxicological studies of new developed drugs (efficacy, metabolism, acute and chronic toxicity). Anti-virus experiments in mouse influenza and pneumonia models, antitumoric effects of Chinese herbs and fish isolates on implanted tumor cells in mice. Generation of knockout mouse models with mutated *SGCE* gene.

EMPLOYMENT

Research Assistant: Drs. David Grimes and Dennis Bulman Lab, Molecular Neurogenetics,
Ottawa Health Research Institute,
University of Ottawa, Ottawa, Canada
Jan. of 2000-Dec. of 2002

World Health Organization Fellowship:
Dr. David L. Busbee lab, Molecular and genetic Program, Texas A & M
University College Station, USA
Aug. 1997 - December.1999

Assistant Professor: Aug.1991-July, 1997
Department of Genetic Toxicology
Shandong Medical University, P. R. China

Teaching and Research Assistant: Aug. 1989 - July 1991

Department of Genetic Toxicology
Shandong Medical University, Shandong, P. R. China

Teaching assistant and Residency: August 1984 - July 1986
Internal Medicine (Occupational diseases),
University Hospital,
Shandong Medical University, Shandong, P. R. China

Teaching and Supervision Experience:

2001-2006: Helping my supervisors, Drs Bulman and Grimes to supervise one summer undergraduate student per year,
Helping my supervisors to supervise one honour's undergraduate student for thesis project in 2004,
Teaching DNA techniques and chemical reactions to primary school students in LET'S TALK SCIENCE program, Fall of 2005-Spring of 2006, University of Ottawa, Canada
Supervising two high school students for the after-school research project (2006 sanofi-aventis biotech challenge in Eastern Ontario),
Drs. Bulman and Grimes lab, Ottawa Health Research Institute, University of Ottawa, Ottawa, ON, Canada.

1989-1997: Teaching genetic toxicology course, 40 hours/year (100 Medical students),
Supervising one or two undergraduate students for thesis projects yearly,
Shandong Medical University, Jinan, Shandong, P.R.China

PUBLICATION:

1) DA. Grimes, L. Racacho, **F. Han,** M. Panisset, R. Zou, K. Westaff, and Bulman DE., Mutation analysis of LRRK2 gene in 202 cases with Parkinson's disease. Submitted.
2) **F. Han**, L. Racacho, A.E. Lang, D. E. Bulman, D.A. Grimes, Refinement of the DYT15 Locus for Myoclonus-Dystonia and Sequencing Analyses of the Candidate Genes, **Movement Disorder,** in press.
3) DA. Grimes, **F. Han** (co-first authors), M. Panisset, L. Racacho, F. Xiao, R. Zou, K. Westaff, and Bulman DE., A translated mutation in the Nurr1 gene as a cause for Parkinson's disease, **Movement Disorder, 2006 Jul;21(7):906-9.**
4) **Han F.,** Bulman DE, Panisset M, and Grimes DA, Neurofilament M gene in a French-Canadian population with Parkinson's disease, **Can J Neurol Sci. 2005; 32(1):68-70.**
5) **F. Han**, A.E. Lang, L. Racacho, D. E. Bulman, and D.A. Grimes, Mutations in the ε-sarcoglycan gene found to be uncommon in seven myoclonus-dystonia families. **Neurology** 2003; 61:244-246.
6) Grimes DA, **Han F**, Bulman DE, Klimek ML, Suchowersky O. Hereditary chin trembling-A New Family with exclusion of the chromosome 9q13-q21 Locus. **Mov**

Disord, 2002, Nov; 17(6):1390-2.

7) Grimes DA, Grimes JD, Racacho l, Scoggan KA, **Han F**, Schwarz BA, Woulfe J, and Bulman DE.A large French Canadian family with lewy body Parkinsonism: Exclusion of Known Loci, **Mov Disord** 2002Nov;17(6):1205-12.
8) Grimes DA, **Han F**, Lang AE, St George-Hyssop P, Rachacho L and Bulman D, A novel locus for inherited myoclonus-dystonia on 18p11. **Neurology** 2002;59:1183-1186.
9) Miller S, **Han F**, and Busbee D, Estradiol receptors in dolphin kidney endothelial cells: Detection and comparison with estrogen receptors from MCF-7 cells. **In Vitro & Molecular Toxicology,** 2000, 13(3):199-211, USA.
10) Miller S, Kennedy D, **Fabin Han** and Busbee DL. A rapid and sensitive reporter gene that uses green fluorescent protein expression to detect chemicals with estrogenic activity. **Toxicological Sciences,** USA, 2000 May; 55(1):69-77.
11) **Fabin Han**, Junfeng Hu and Houquan Xu, The effects of some Chinese herbal medicine and green tea antagonizing mutagenesisby cigarette tar. **Chung hua Yu Fang I Hsueh Tsa Chih**, 31(2):71-4, 1997.
12) **Fabin Han** and Houquan Xu. The detection of mutagenicity of lake water organic extract, **Environmental Toxicology**, 10 (4); 194, 1995.
13) **Fabin Han** and Houquan Xu, The retrospective cohort analysis on the relationship between the water pollution and the cancer incidence of residents. **Shandong Health and Prevention,** 14 (2): 67, 1994.
14) Houquan Xu, **Fabin Han** and Yureng Han, The analysis of marrow micronucleus and chromosome abberation induced by methyltrichloride in mice, **Shandong Health and Prevention 13(1):1993.**
15) **Fabin Han**, Fulin Wang, and Ding Tian, The detection of mutagenicity of polluted water and drinking water with Salmonella Typhimurium Microsome Test and E Coli. SOS Chromatest. **J. of China Eivironmental science,** 11(4):285, 1991.
16) **Fabin Han**, Fulin Wang, and Ding Tian, Teratogenicity of drinking water extract in drosophila, **J. of Carcinigenesis, Teratogenesis and Mutagenisis.** 2(1):41, 1990.

CONFERENCE ABSTRACTS:

1) **Fabin Han,** Dennis E Bulman, Michael Panisset, David A Grimes. Screening of the neurofilament M Gene Gly336Ser mutation in a French-Canadian population with Parkinson's disease. **The 53rd annual meeting of American Society of Human Genetics.** Los Angeles, USA, November 5-8, 2003.
2) **Fabin Han,** David A Grimes, Lemuel Racacho, and Dennis E Bulman. Refinement of the Gene Locus *DYT15* for Myoclonus-Dystonia. **The 54th annual meeting of American Society of Human Genetics.** Toronto, Ontario, Canada, October 26-31, 2004.
3) **Fabin Han,** Dennis E Bulman, Lemuel Racacho and David A Grimes. Sequencing of the candidate gene *LAMA1* in *DYT15* Locus for myoclonus-dystonia. **The 40th annual meeting of Canadian Congress of Neurological Sciences.** Ottawa, Ontario, Canada, May 14-18, 2005.

4) **Fabin Han,** Lemuel Racacho, Dennis Bulman, and David Grimes. Gene Dosage Alterations of the Epsilon-Sarcoglycan (*SGCE*) Gene in Families with Myoclonus-Dystonia. **The 36 annual meeting of Society of Neuroscience.** Atlanta, GA, USA, October 14-18, 2006.