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

Tara Read

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Biochemistry)

GRADE / DEGREE

Department of Biochemistry, Microbiology and Immunology

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Elucidating a Novel Gene Associated with 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

Robert Korneluk

Martin Bulman

Gary W. Slater

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

Elucidating a Novel Gene Associated with Myoclonus Dystonia

By

Tara Read

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

Your file Votre référence
ISBN: 978-0-494-59882-5
Our file Notre référence
ISBN: 978-0-494-59882-5

Abstract: Myoclonus Dystonia (MD) is an autosomal dominant disease with high but incomplete penetrance and is characterized by both involuntary myoclonic jerks and dystonic posturing. Our group has found that mutations within the epsilon sarcoglycan (*SGCE*) gene on chromosome 7q21 are associated with MD in 30-40% of affected individuals in 31 families studied, supporting the basis for genetic heterogeneity. Novel mutations have been found in *SGCE* by screening these families for point mutations and large deletions and duplications through the use of sequencing, high performance liquid chromatography (HPLC) and multi-ligation probe amplification (MLPA) analysis. A 10cM genome wide linkage analysis of a large Canadian family provided significant LOD scores for microsatellite markers within the 18p11 region, now designated as the DYT 15 locus. Further haplotype analysis has narrowed a non-recombinant region associated with the disease phenotype to a 3.18 Mb region in this locus. Since the current understanding of Myoclonus Dystonia is poor, it is difficult to predict genes that could be responsible for MD. Sarcoglycans are essential constituents of the dystrophin-glycoprotein complex and are involved in linking the extracellular lamanin matrix to the actin filaments within the cytoplasm; therefore focus is given to the examination of potentially related structural genes that are expressed in the brain. By analyzing such candidate genes in a panel of affected individuals, we believe that a novel gene will be elucidated and provide insight into the mechanism of Myoclonus Dystonia.

Acknowledgements

First and foremost I would like to thank my supervisor Dr. Dennis Bulman for allowing me to study and work on the Myoclonus Dystonia project. He has been an excellent mentor and has always offered insightful guidance to my research questions. The Bulman lab members have played an equal part in my positive experience within the lab and have always been there for scientific discussions as well as out of lab adventures. Thanks are in order to Lemuel Racacho who has offered both technical help with much of the equipment I had used as well as sound advice regarding genetics and genomics resources. Roubing Zao has also been an excellent source of training and technical assistance regarding genotyping. I would also like to thank Dr. Fabin Han who has previously worked on this project and supplied candidate gene data and technical assistance for multiplex PCR in deletion screening. Dr. David Grimes has been essential to this project, as he has collaborated with movement disorder clinics and has allowed for the recruitment of the Myoclonus Dystonia families. The members of my thesis advisory committee, Dr. Martin Holcik and Dr. John Wolfe, have gratefully provided insightful advice as well as constructive criticism for my research. Special thanks to Dr. Wolfe for reading my thesis prior to submission. I would like to thank Dr. Nancy Carson for her training on the MLPA data analysis software and for allowing me to use her facilities to do so. The Toronto Centre for Applied Genomics (TCAG) has performed the FISH experiments and specifically, Ann George and her lab has helped modify the BAC clones when needed. As well, Nancy Cracknell's Biobanking Facility in TCAG created the Myoclonus Dystonia cell lymphoblast cell line. The graduate organizations GSAED, FGPS, BMI as well as Dr. Bulman are graciously thanked for their travel funding that had allowed me to attend the annual

American Society for Human Genetics meeting. The BMI Graduate Student Association is also thanked for allowing my participation and organization of various student events to be an enjoyable experience. Last but not least I would like to give many thanks to all of my many family members and friends who have given moral support throughout my masters, notably my parents and Nicholas, who without their support this endeavor would not have been possible.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

MD	Myoclonus Dystonia
cM	centimorgans
chr	chromosome
LOD	logarithm of the odds ratio
θ	(theta) recombination fraction
DNA	deoxyribonucleic acid
RNA	ribonucleic acid
cDNA	complementary deoxyribonucleic acid
SNP	single nucleotide polymorphism
PCR	polymerase chain reaction
HPLC	high performance liquid chromatography
SG	sarcoglycan
SGCE	epsilon sarcoglycan
DGC	dystrophin glycoprotein complex
BHC	benign hereditary chorea
LAMA1	laminin, alpha 1
PEG10	paternally expressed gene 10
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	milliliter
µl	microlitre
mg	milligram
ng	nanogram
nm	nanometer
WT	wild type
SP1	sporadic (family 1)
NCBI	National Centre for Biotechnology Information (USA)
UCSC	University of California Santa Cruz
OHRI	Ottawa Health Research Institute
TCAG	Toronto Centre for Applied Genomics

Chapter 1.

Introduction: Myoclonus Dystonia

Chapter 1 Introduction

1.1 Myoclonus Dystonia

Myoclonus Dystonia (MD) is an autosomal dominant movement disorder with reduced penetrance characterized by both myoclonic jerks and dystonia. The myoclonic jerks are often described as "lightening fast" and are usually seen within the axial (head or trunk) skeletal muscles as well as the arms and legs. Dystonia is a slow twisting and/or posturing movement which can present itself within one muscle (focal) or many (generalized)[1]. The age of onset is typically within the first two decades of life and the spectrum of disease severity can range from very mild to constant myoclonic jerking, resulting in a significant interference of daily life. Interestingly, the myoclonic symptoms of some patients are alcohol responsive causing an alleviation of the myoclonic jerks after alcohol intake. The mechanism for this condition remains unknown. Mutations in the epsilon sarcoglycan (*SGCE*) gene were found to cause MD[2]. There are now 39 mutations identified in *SGCE*, the majority of which are nonsense mutations; however missense, large deletions and splice site mutations have also been characterized. Myoclonus dystonia is genetically heterogeneous, since a significant portion of MD families do not have *SGCE* mutations. To date, there is no effective treatment, although benzodiazapines, deep brain stimulation and botulinum injections have been shown to have minimal and unreliable effects[3].

1.2 Related Dystonic Movement Disorders

There are a total of 13 dystonia type movement disorders, each of which map to a different locus and are traditionally classified based on age of onset, distribution of symptoms, and the site of onset[4]. Those that are most phenotypically related to Myoclonus Dystonia and are more common include Dopamine Responsive Dystonia (.5-1.0/million), Primary Torsion Dystonia (1/160,000) and Chorea. These are described below.

Primary Torsion Dystonia (PTD) is mapped to the DYT1 locus on chromosome 9q34 and is the most severe form of hereditary dystonia. The age of onset is early, under 28 years of age, and the site of onset is usually the legs but often becomes generalized over time. It was found that a heterozygous GAG deletion in exon 5 in the gene *TOR1A* accounts for the phenotype in 50-70% of early-limb-onset PTD cases[5]. The Torsin A protein is expressed in neural cells and is seen in a punctuate pattern throughout the cell body and neurites[6]. It is thought to act as a molecular chaperone and modulate synaptic vesicle recycling[6,7]. Augood *et al* (1998) found its expression is highest in dopaminergic neurons in the substantia nigra compacta and suggested there may be dysfunction in dopamine (DA) transmission in this disease[8]. Indeed, the release of DA from pre-synaptic neurons was inhibited likely contributing to the symptoms of the human disorder[9]. Pallidal deep brain stimulation has been found to be an effective treatment for patients with the GAG deletion[10].

Dopamine Responsive Dystonia (DRD) is characterized by childhood onset, with the initial symptom usually being leg dystonia that gradually progresses to other parts of the body. Most cases of DRD are autosomal dominant and caused by heterozygous mutations in the GTP cyclohydrolase 1 (*GCH1*) gene on chromosome 14q11[11]. GCH1 is the rate limiting enzyme in the synthesis of tetrahydobiopterin (BH4), which is in turn necessary for

dopamine synthesis. DRD is less commonly inherited as an autosomal recessive disorder with mutations in the tyrosine hydroxylase (TH) gene on chr11p15. TH is an enzyme which is also involved in dopamine biosynthesis pathway. Reduced activity of this pathway is thought to be the underlying mechanism leading to the depletion of dopamine and consequently the DRD phenotype. This is consistent with the dramatic therapeutic response in most patients to levadopa ($_L$-dopa), a synthetic form of dopamine that increases dopamine levels[12]. There is generally no association between DRD and MD. Myoclonus dystonia patients are not responsive to L-dopa treatment, in fact, long-term therapy of L-dopa has been shown to induce dyskinesias such as chorea, as well as dystonia or myoclonus, due to abnormal stimulation of dopamine receptors[13]. Klein *et al* (1999) had screened a MD family that was found to have mutations in both the *SGCE* and the D2 dopmanine receptor (*DRD2*) genes[14]. No such cases have been identified since this report.

Benign Hereditary Chorea (BHC) is not classified as a dystonic syndrome, but is a dyskinesia-type neurological disorder. It is a rare autosomal dominant movement disorder characterized by "jerky" hyperkinetic movements that can present with choreoathetosis, thyroid dysfunction and pulmonary infections in infancy. Some affected families have been shown to also present with dystonia of the neck, axial muscles and limbs[15]. BHC has been questioned in up to 25% of cases of MD and could be a major differential diagnosis, thereby illustrating the importance for proper genetic screening and phenotype assessment[15]. To date, 19 heterozygous mutations of the *NKX2-1* gene on chr14q21 have been found to cause BHC. This gene encodes for the thyroid transcription factor 1 protein (TITF-1), a member of the *Nk* homeobox gene family, which binds and activates the promoter of thyroid specific genes. It is predicted that these mutations cause decreased binding of the transcription factor

to the homeobox domain thereby causing a decrease in transcription of the target genes[16]. Given the similarities between BHC and MD, careful attention should be made to such overlapping phenotypes and should be reflected in the consideration of genetic screening of either *NKX2-1* or *SGCE*.

1.3 The *SGCE* gene is associated with Myoclonus Dystonia

In 1999, a locus for Myoclonus Dystonia was mapped in a large family to a 28cM region of chromosome 7q21-q31[17]. After further narrowing of the region and the testing of potential candidate genes by other groups[18,19], Zimprich *et al* (2001) had found 5 different heterozygous loss of function mutations in the epsilon sarcoglycan (*SGCE*) gene[2]. These were found in 6 families and ranged from nonsense changes, deletions and a splice-site mutation. Pedigrees showed an autosomal dominant mode of inheritance and some patients had psychiatric abnormalities such as panic attacks and obsessive-compulsive behavior. This has been reported by others as well[20,21,22]. A marked difference in penetrance of the disease was observed depending on the parental origin of the disease allele; this was subsequently found to be due maternal imprinting of the *SGCE* locus. There is no known correlation or difference between missense and premature stop mutations when compared to the phenotype or sex of the patient[23]. As well, no genotype-phenotype association exists, except in cases of large deletions encompassing *SGCE* and its neighboring genes[24,25]. The majority of *SGCE* mutations reported in MD families are nonsense mutations that abolish the synthesis of the full length protein. Missense mutations are less common. Three of the disease-associated *SGCE* missense mutations (p.His36Pro, p.His36Arg and p.Leu172Arg) were studied for their effects on the biosynthesis and trafficking of the ε-SG protein[26]. These missense

mutations should result in the synthesis of full length proteins carrying a single altered amino acid[27]. It was found that the mutant sarcoglycan was mislocalized as it did not reach the plasma membrane of neuronal cells. This protein was targeted for degradation by the proteasome via polyubiquitination, likely due to the misfolding of the missense mutated form of ε-SG.

1.4 The *SGCE* gene is imprinted

It is now known that *SGCE* expression is under the control of parent-of-origin specific differential gene expression due to the imprinting of a CpG island located within its promoter and exon 1. This CpG stretch also spans the 5' end of its neighboring gene, *PEG10* which has also been found to be imprinted in mice[28,29,30]. Through the use of bisulfite genomic sequencing this CpG island has was shown to be differentially methylated on the maternal allele[31,32]. *PEG10* is a retrotransposon-derived imprinted gene that has evolved to its present function as an essential placental gene in mice, but not humans[33]. There is evidence that retrotransposon insertion can drive the evolution of genomic imprinting in mammals, since the expression of inserted foreign genes from viruses would likely be under negative selection[33]. It is likely the repression of *PEG10* has indirectly caused the repression of *SGCE* expression on the maternal allele due to its close proximity to *PEG10*. As a result, *SGCE* mutations that are transmitted to offspring on the paternal allele will present with MD. Since *SGCE* expression generally occurs on the paternal allele, maternal transmission of a mutation will not cause the MD phenotype. Interestingly, a case report recently described an individual who had inherited both copies of the maternal chromosome 7, resulting in maternal uniparental disomy of chromosome 7 (mUPD7)[34]. Since both copies of

the *SGCE* gene are methylated, this patient exhibits the MD phenotype. Silver-Russell syndrome is also observed in the patient and is thought to be caused by altered expression of a different imprinted gene. It is possible that maternal imprinting of *SGCE* is under a leaky mechanism since there are reports that 5-10% of symptomatic MD patients have *SGCE* mutations that have been inherited from the mother[32]. Nonetheless, the majority of cases exhibit paternal expression of the *SGCE* gene and has shown to cause for reduced penetrance seen in MD families, as demonstrated in the moderate frequency of obligate carriers[35].

1.5 Characteristics of the *SGCE* gene

Epsilon sarcoglycan was first identified in mouse by Ettinger *et al* (1997) as a homolog to alpha sarcoglycan (α-SG)[36]. Genomic sequence identity was reported as 95% identical between the human and mouse and ε-SG protein was found to be 44% identical between these species[36]. This is similar to α-SG, which has a sequence identity of 88% between mouse and human. The mouse ε-SG begins with a 22-aa signal sequence, followed by 262-aa extracellular domain, a single 23-aa hydrophobic transmembrane domain, and a 98-aa intracellular domain. It has a predicted molecular mass of 43.5 kDa, with an increase of roughly 2kDa when glycosylated. The extracellular domain has a single consensus sequence for *N*-glycosylation and the intracellular domain has 3 potential sites for phosphorylation.

In human, *SGCE* has two main isoforms that are alternatively spliced and are tissue specific (see Figure 1.1). The primary transcript is thought to lack exon 10 and is widely expressed in embryonic and adult tissue, striated and smooth muscle, lung, liver, kidney,

spleen, testis, and sciatic nerve[36]. It has also been localized to smooth muscle as well as the outer Schwann cell membrane of peripheral nerve and to the brain in general[37,38,39]. The other transcripts have either exons 8 or 11b spliced out, and are found less ubiquitously throughout tissues.

As opposed to human, the *SGCE* transcripts have been well studied in mouse. Nishiyama *et al* (2004) had found 4 splice variants of *Sgce* in mice, two were shown to be major transcripts found throughout brain, one of which was also widely expressed in other tissues[40]. They referred to these as ε-SG1, with exon 8 but lacking exons 10 and 11b (8^+, 10^-, $11b^-$) and ε-SG2 (8^-, 10^-, $11b^+$). When investigating human brain, they also found these two transcripts were present[40]. In another study, Yokoi *et al* (2005) had found 3 transcripts present in mice, two of which were found in brain[30]. They referred to these as: Type1 (ε-$SG8^+11c^+$), and Type 2 (ε-$SG8^-11c^+$). These isoforms had an extended coding sequence in 11b that they refer to as 11c, adding 36bp to the transcript and causing a frameshift throughout the remainder of the protein. Both of these isoforms were found to have motifs at the C-terminal that were similar to PDZ binding motifs, which allow for binding of scaffolding proteins. This could prove to be an important finding since the sarcoglycan (SG) complex which is found within the larger dystrophin-glycoprotein complex (DGC) functions to provide structural integrity to the cell membrane.

1.6 ε-SG within the Dystrophin Glycoprotein Complex

Epsilon sarcoglycan is a single-pass transmembrane glycoprotein that is found within the sarcoglycan complex. This complex is made up of four sarcoglycan proteins and is localized to the plasma membrane. The SG complex is found within the larger DGC

which is comprised of such proteins as dytrophin, dystrobrevin, dystroglycans (α and β), and sarcospan (Figure 1.2). This multi-protein complex is critical for the stability of the striated muscle cell membrane. Dystrophin, a gene product responsible for Duchenne muscular dystrophy, associates with β-dystroglycan, which in turn binds to α-dystroglycan[38]. Alpha-Dystroglycan tightly binds to laminin, a major component of the basal lamina, indicating that dystrophin and the dystroglycan complex form a bridge between the extracellular matrix and the actin-based cytoskeletal network. The SG complex plays a role in reinforcing this linkage through its interaction with membrane-spanning β-dystroglycan[41]. It has also been suggested to play role in scaffolding for signal-transduction cascades[42]. There are 6 known sarcoglycans, α, β, δ, γ, ε, and ζ, with epsilon being a homolog of alpha, and zeta a homolog for gamma. A defect in any one of these glycoproteins (excluding ε-SG and ζ-SG) disrupts the entire sarcoglycan complex in striated muscle and leads to limb-girdle muscular dystrophy (LGMD)[43]. The components of the tetrameric sarcoglycan complex differ depending on the particular tissue in which they are expressed. ε-SG is widely expressed as previously mentioned, particularly within the brain, heart and lung tissue. However α-SG expression is limited to striated muscle[38], therefore it is then likely that ε-SG replaces α-SG in the sarcoglycan complex within non-muscle cells. Indeed, it has been shown that ε-SG replaces α-SG to form a novel DGC complex in peripheral nerve cells[38]. As well, Cai *et al* (2007) studied Schwann cells, the myelin forming glial cells of the peripheral nervous system, and found they express ε-SG in rat sciatic nerve[39]. In this cell type, the SG complex exists as a hetero-tetramer as it does in muscle, except with ε, ζ, β, δ instead of α, γ, β, δ. This pattern of sarcoglycan expression is thought to be similar in human.

1.7 Sarcoglycanopathies

The limb-girdle muscular dystrophies (LGMDs) are a heterogeneous group of neuromuscular disorders involving progressive muscle wasting of the shoulder and girdle muscles. There are a number of proteins with a wide spectrum of function and localization that are associated with these dystrophies, mainly the α-, β-, γ-, δ-SGs, but also other proteins necessary for sarcoglycan processing such as calpain-3. This protein is thought to have chaperone activity and was found to cleave filamin C, thereby affecting its ability to interact with γ and δ-sarcoglycans[44].

Sarcoglycanopathies, unlike MD, involve muscle and heart deficiencies, further suggesting there are differences in the pathophysiological mechanisms of the SG complex in the brain compared to muscle. Functional studies by Chen *et al* (2006) have demonstrated that the assembly of sarcoglycans into their tetrameric complex follows a stepwise pathway in COS-1 cells, with β- and δ-SG forming a core to which α- and γ-SG later bind[45]. While investigating the sarcoglycans in rat skeletal muscle, they found the N-terminal half of the extracellular region contains domains that are required for SG-SG interaction[46]. As well, various point mutations in the sarcoglycans were shown to affect SG complex assembly and/or localization to the cell surface[46]. ε-SG was not included in these studies, rather α-SG, its homolog, was analyzed. The BIO14.6 hamster has also been used to study the SG complex within Schwann cells, the myelin forming glial cells of the peripheral nervous system. This LGMD model has a large genomic deletion of the δ-SG gene that deletes the transcriptional initiation site and exon 1. This loss of δ-SG caused a 98% reduction or undetectable levels of β-, δ-, ε-, and ζ-SG in the sciatic nerve membrane fractions containing Schwann cells[39]. This disruption of the SG complex in turn caused a reduction of steady

state levels of the neuronal form of dystrophin Dp116, as well as α-dystroglycan. This is one of the few studies investigating the components of the SG complex, DGC, and their interacting partners within neuronal cells. They demonstrate the importance of the SG complex and its role within the DGC in Schwann cells in the peripheral nervous system (PNS) and could suggest a role within the CNS as well.

1.8 Investigating a functional role of the SGCE protein

Ettinger *et al* (1997) found that *SGCE* was widely expressed in mouse E14 embryonic and adult tissue using Northern analysis[36]. Levels were found to be high in lung, moderate in brain, heart, and skeletal muscle, and low but detectable in kidney and liver. In contrast, *SGCA* expression was largely restricted to mouse striated muscle (cardiac and skeletal) muscle, with very low levels in lung and none in brain. Nishiyama *et al* (2004) found that *SGCE* transcripts and protein were widely distributed throughout the mouse brain, in both neuronal and non-neuronal cells including capillary endothelial cells and astrocytes[40]. A similar expression profile was found in rats[47]. This diverse expression pattern and variance in isoforms is also seen in human tissue. Both ε-SG1 and ε-SG2 isoforms (mRNA and protein) were found in both the neuronal and non-neuronal areas of human brain tissue[40]. Hjermind *et al* (2008) performed analysis on a muscle (quadricep) biopsy from a MD patient and found normal skeletal muscle morphology[48]. In brain, ε-SG is thought to play a structural role within the sarcoglycan complex and DGC in neuronal cell membranes. Dystroglycans and the cytoplasmic members of the DGC have been shown to be expressed in brain however little is known regarding their functional roles and interacting components.

Perhaps the most insight into the mechanism of MD can be gathered from a mouse model. Yokoi *et al* (2005) created a knockout (KO) mouse lacking exon 4 of the *SGCE* gene using a cre/loxP method[49]. A total of 73 bp was excised, causing a frameshift resulting in a premature stop codon in exon 5. Interestingly, these mice exhibited a MD phenotype, complete with myoclonic jerks and deficits in motor coordination and balance[30]. They found that 30% of the KO mice had myoclonic jerks compared to no jerks observed in the wild-type (WT) littermates when observed for 30 min (n= 19 KO, n= 22 WT). A beam test showed that the KO mice had decreased fine motor coordination when compared to WT (122% more slips). The male KO mice had more "anxiety-like behavior" than male WT mice, whereas no significant difference was observed between KO and WT female mice. This behavior was studied with only one test (open field) and its reliability in quantifying behavior is unknown; no such sex differences in MD phenotypes exist in humans. Levels of dopamine, serotonin and metabolites were measured in the striatum of the KO mice, presumably since this region of the brain is associated with movement and cognitive control and contains a high density of dopamine receptors. These receptors have key roles in many processes such as the control of motivation, learning, and fine motor movement. They found DA, DOPAC and HVA (metabolites of DA) had minimal increases in KO mice compared to WT: 1.4, 1.2 and 1.2 fold increases, respectively. As well, no other areas of the brain were studied. The mechanism by which the loss of *SGCE* causes this minimal increase in dopamine in the striatum of mice remains to be investigated. Whether or not the underlying mechanism of MD is related to DA levels remains to be elucidated.

1.9 Cause for genetic heterogeneity of Myoclonus Dystonia

Zimprich *et al* (2001) had first found that mutations in the *SGCE* gene were associated with MD. Five novel mutations were described in 6 MD families[2]. It is generally agreed upon in the literature that roughly 30-40% of MD cases are due to *SGCE* mutations. Tezenas du Montcel *et al* (2006) had screened 49 index patients with MD and found 11 had SGCE mutations (22%)[22]. They found there was an increased probability of finding *SGCE* mutations in individuals that had an age of onset below 25 and if they had inherited the disease allele from their father. Schule *et al* (2004) had screened 7 families with inherited MD in which the mean age of onset was 12 years (range 3-24) and found 2 of these families to have *SGCE* mutations (29%)[50]. Our group has previously screened 11 families, with a total of 20 affected individuals, and had found that 9/11 families, or 16/20 affected individuals (88%), had *SGCE* mutations[51]. The mean age of onset was 4 years (range 6 months - 20 years) and alcohol responsiveness was reported in 4 of the families, absent in 2 and not known in the remaining 5 families. Although age of onset, alcohol sensitivity and presenting symptoms are strong indicators of *SGCE* mutations, they should not be cause for limiting genetic screening of this gene. It is possible that *SGCE* mutations have been implicated in a relatively low number of affected individuals due to the lack of screening for large *SGCE* deletions. Although not as common, multi-exonic deletions have recently been reported [52]. As well, our group has identified two large deletions in two families[53].

In 2002 our group performed a 25cM linkage analysis of a large pedigree with 15 affected individuals[54]. Two-point linkage analysis demonstrated 7 microsatellite markers had a maximum LOD score of 3.9 across a 16.9cM span in the 18p11 region. All of the symptomatic individuals shared a common haplotype for these markers, providing sufficient

evidence that a novel locus for MD had been found. As well, Schule B *et al* (2004) genotyped these markers in 5 families that were *SGCE* mutation-negative and found the chr18p11 region segregated with the disease in one of these families[50]. The affected individuals in the remaining 4 families did not share a haplotype for markers in 7q21 (for *SGCE*) or 18p11, suggesting there are more loci for this disease.

The refinement of the chr18p11 region was performed by our group using direct sequencing of three recombinant single nucleotide polymorphisms (SNPs)[55]. The region was narrowed to 3.18Mb, and direct sequencing of seven known and four predicted genes with EST support did not identify any mutations.

1.10 Rationale, Hypothesis, Aims

1.10.1 Rationale

To date, our research group has recruited thirty-one families with various numbers of affected individuals, ranging from one to fifteen in our largest family (family 1). In agreement with the literature, we have also found that mutations within *SGCE* are found in only 30-40% of these affected individuals, strongly suggesting there is another gene or genes involved. This genetic heterogeneity was the basis for a 25cM and subsequent 10cM genome-wide linkage analysis that was previously performed on family 1[65]. Using this method as well as haplotype analysis of the microsatellite markers used, a 3.18Mb region on chromosome 18p11 was found to strongly segregate with the disease phenotype. Candidate genes within this region were then selected based on their protein characteristics such as expression location, and their coding regions were then sequenced in a panel of six affected individuals, two of them from family 1. There were no mutations found that segregated with

the disease phenotype within the fifteen genes that were analyzed. It is the goal of this project to elucidate the specific gene in this region causing Myoclonus Dystonia in family 1.

1.10.2 Hypothesis

Based on the evidence that only 30-40% of Mycoclonus Dystonia patients have loss of function mutations in *SGCE*, suggesting cause for heterogeneity, we hypothesize that a gene associated with this disease lies in the 3.18 Mb region on chromosome 18p11.

1.10.3 Aims

1. Recruit and screen additional Myoclonus Dystonia families for mutations in the *SGCE* gene.

Through the help of our collaborating neurologist, we are continually recruiting new families affected with Myoclonus Dystonia. This enables us to gather more information regarding the frequency and inheritance of *SGCE* mutations. As well, if the families recruited are large and have multiple affected individuals, it can provide us additional insight into the heterogeneity of the disease through marker genotyping and linkage analysis. During the course of this project we have recruited 5 familial cases of MD. The phenotypic characteristics and mutation status of these families will be discussed.

2. To further ensure linkage of Myoclonus Dystonia to the chr18p11 locus (DYT15) and to exclude linkage elsewhere in the genome for family 1.

To rule out the possibility that linkage to Myoclonus Dystonia is found elsewhere in the genome in family 1, we first accounted for the markers previously genotyped for the

25cM and 10cM genome-wide linkage analysis and identified any chromosomal regions that were not covered by these analyses. It is important to recognize the possibility that linkage to the disease does in fact occur in additional loci, however, linkage analysis performed on the large kindred (family 1) recruited for our studies has implicated chr18p11, and it is this area we have chosen to examine. To exclude linkage elsewhere, a total of thirty-nine microsatellite markers were genotyped for all forty members of family 1. The alleles that were generated from this analysis for each marker were subject to linkage analysis.

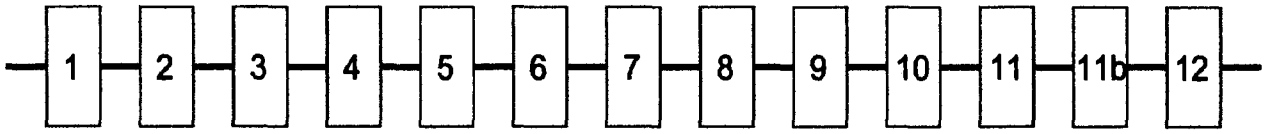
3. Elucidate the causative gene or mutation in the chr18p11 region responsible for Myoclonus Dystonia in family 1.

There are a number of candidate genes within the 3.18Mb region that could potentially be involved with MD, however most of them have been previously analyzed by our group[55]. Remaining candidate genes have therefore been prioritized on their likelihood of being associated with MD. For example, *LAMA1* is an excellent candidate gene for this disease due to its expression of the laminin alpha 1 chain protein which is known to be an extracellular matrix protein. It is similar to the alpha 2 chain form which is known to interact with the DGC at the membrane of muscle cells and has been implicated in congenital muscular dystrophy. However *LAMA1* has been sequenced in a number of MD patients and determined to not have mutations specific to the disease. Candidate genes are accessed based on their tissue expression profile, with brain being a priority, and whether or not they are known to express a protein or if they are a predicted gene. Each candidate gene selected for analysis was sequenced in 2-6 affected individuals where at least 2 were from family 1. Wild-type controls were also used to determine if sequence variants were found in non-

affected individuals and therefore not the cause of MD. Our group has previously confirmed the absence of large deletions or duplications in patients within family 1 through the use of a 500K SNP chip. However, in order to investigate as to whether or not there is an inversion within our 3.18Mb chr18p11 region, we have performed a three color fluorescent *in situ* hybridization (FISH) experiment on this region in an affected individual from family 1.

Figure 1.1. Two common human isoforms of the *SGCE* gene. *SGCE* spans a 71 kb genomic region on 7q21. The human isoforms of *SGCE* share 10 common exons and differ by the three alternatively spliced exons; 8, 10 and 11b. The primary isoform (8^-, 10^-, $11b^-$) encodes a 437-aa protein and is expressed in brain, lung, heart, placenta, testis, kidney and liver. An alternate isoform (8^-, 10^-, $11b^+$) is considered to be the brain-specific and is 462-aa in size. There are two other isoforms that are also present in tissues but not as frequent as the above mentioned isoforms, these are: (8^+, 10^-, $11b^+$) and (8^-, 10^-, $11b^-$). Exon 10 is rarely found in *SGCE* transcripts.

Primary *SGCE* Isoform: $8^{-}10^{-}11b^{-}$



Brain specific *SGCE* Isoform: $8^{-}10^{-}11b^{+}$

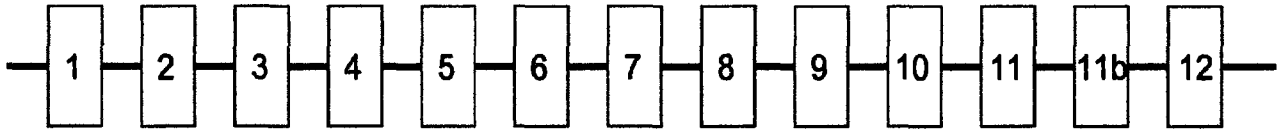
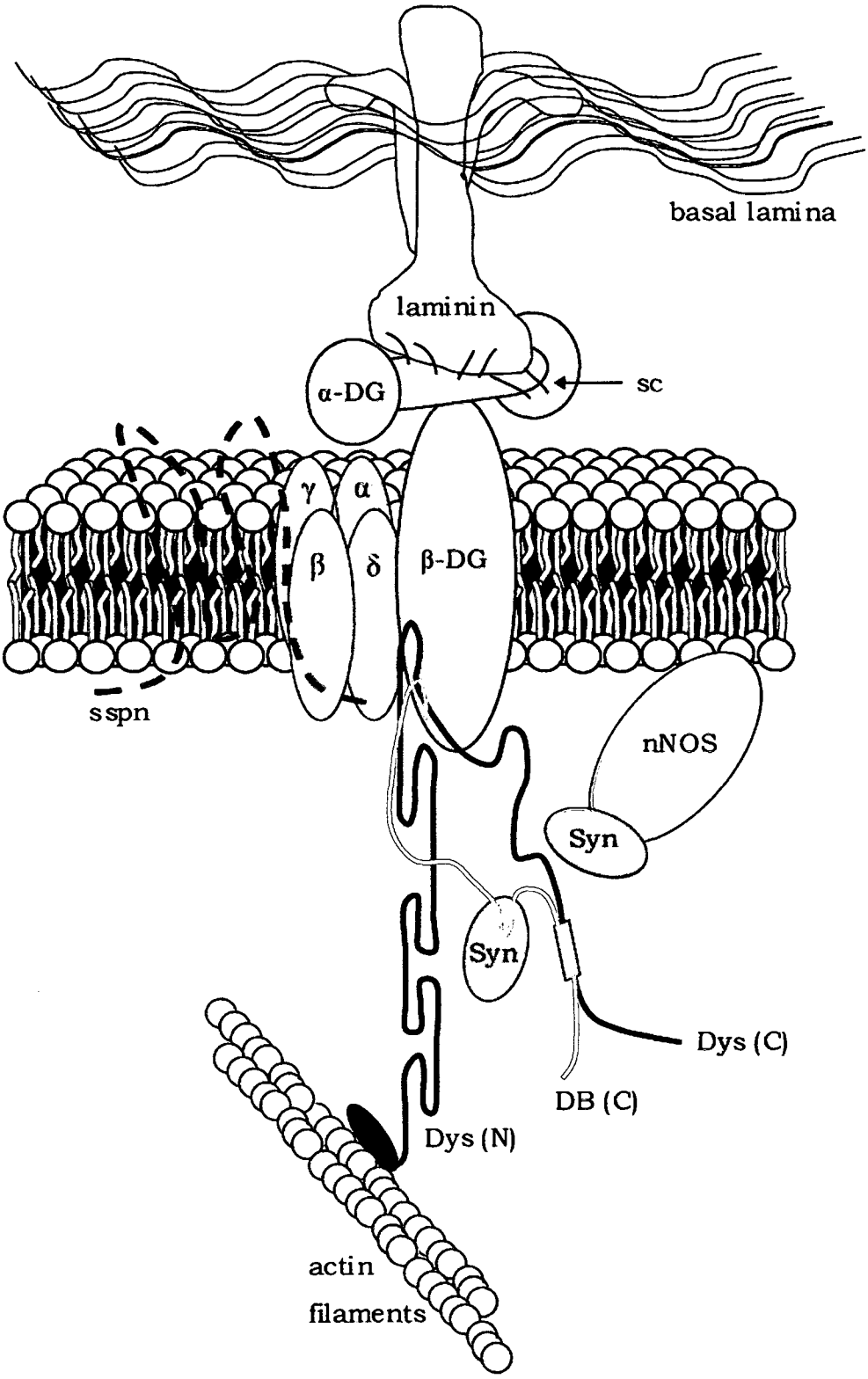


Figure 1.2. The predicted arrangement of the DGC and SG complex in the muscle cell membrane. Dystrophin (black), β-DG (β-dytroglycan) (green), α-DG (blue), form a dystrophin bolt that is thought to be reinforced by the SG complex (red) composed of β-, δ-, α-, and γ-SG in muscle. The dystrophin bolt, through α-DG, binds laminin (light blue) within the extracellular matrix of the basal lamina. β-DG connects to actin filaments in the cytoskeletal network through dystrophin (Dys). (C), C-terminus; (N), N-terminus; sspn, sarcospan; DB, dystrobrevin; Syn, syntrophin; nNOS, neuronal nitric oxide synthase; sc, sugar chain.


basal lamina
laminin
α-DG
sc
γ
α
β
δ
β-DG
sspn
nNOS
Syn
Syn
Dys (C)
DB (C)
Dys (N)
actin
filaments

Chapter 2.

Subjects and General Methodology

Chapter 2 Subjects and General Methodology

2.1 Phenotypic characteristics of recruited patients in families with Myoclonus Dystonia

Consent forms approved by the Ottawa Hospital Research Ethics Board were signed by all individuals donating blood for participation in our studies. If the patient was under the age of 16, parental consent was given. Families were recruited through contact with our collaborating neurologist to physicians within various movement disorder clinics throughout Canada, and some within the United States. Affected individuals were diagnosed by a movement disorder neurologist and if possible, patients were videotaped. Patients also filled out data sheets describing the nature of their signs and symptoms. The phenotypic characteristics of the 13 families studied in this project are listed in Table 2.1.

2.2 DNA and RNA Extraction

Genomic DNA was extracted from whole blood drawn from patients in yellow top ACD Vacutainer tubes. Extraction was performed using the QIAamp DNA Blood Mini Kit. Briefly, 20ul of protease was added to 200ul of whole blood along with 200ul of lysis buffer (AL). Samples were incubated at 56°C for 10 min, after which 200ul of 100% ethanol was added. The mixture was applied to a QIAamp Spin Column and centrifuged at 6000 x g for 1 min to remove cellular debris. Two washes were then performed using buffers AW1 and AW2, and purified genomic DNA was then eluted from the filter using 100ul of either buffer AE (Qiagen) or TE (10mM Tris-HCl, 1mM EDTA, pH 8.2). A yield of ~30ng/ul was

typical with an A_{260}/A_{280} ratio of 1.7-1.9. Remaining patient blood sample was aliquoted and stored at -80°C.

RNA was extracted from whole blood drawn from patients using PAX tubes (PreAnalytiX) containing an RNA stabilizing additive. RNA was extracted roughly one week after blood draw using the PAXgene Blood RNA Kit. Blood tubes were centrifuged for 10 min at 4000 x g, supernatant was removed and replaced with 4ml of RNase-free water. The pellet was resuspended via vortexing and centrifuged again at 4000 x g and supernatant was subsequently removed. The pellet was resuspended in 350ul of Buffer BR1 by vortexing after which 300ul of Buffer BR2 and 40ul of proteinase K was added to the mix. Incubation at 55°C while shaking at 800rpm ensured sufficient protein digestion. The lysate was then transferred to a PAXgene Shredder spin column and centrifuged for 3 min at 20,000 x g. The flow-through fraction containing RNA and DNA was then transferred to a new tube and 350ul of 100% ethanol was added. This was applied to a PAXgene RNA spin column and centrifuged for 1 min at 14,000 x g. The column was washed with 350ul of Buffer BR3 and centrifuged at 14,000 x g. 80ul of DNase I mix (10ul DNase I with 70ul Buffer RDD) was applied to the column and incubated at room temperature for 15 min to ensured digestion of DNA. The column was then washed with 350ul of Buffer BR3 and Buffer BR4. Purified RNA was then eluted off the column using 20ul of Buffer BR5 and centrifugation at 14,000 x g. Eluate containing RNA was incubated for 5 min at 65°C, followed by ice to allow for denaturation for downstream applications. RNA yield was typically between 100-200ng/ul with a A_{260}/A_{280} ratio of 1.7-1.9.

2.3 Primer Design and PCR

Genomic DNA sequence was obtained from the University of California Santa Cruz (UCSC) Genome Browser. Primers were designed using Primer 3 v.0.4.0 (Steve Rozen and Helen J. Skaletsky) and had an optimal length of 20nt, 50% G/C content and a melting temperature (Tm) of 60°C. Optimal amplicon size was dependant on the particular sequence being studied and ranged from 150-3500bp.

Various conditions of the polymerase chain reaction (PCR) were used depending on the characteristics of the product of interest being amplified. Conditions were optimized for each primer set by adjusting the Tm, cycle number and time, altering $MgCl_2$ concentration, or by using additives such as dimethyl sulfoxide (DMSO).

2.4 Sequencing

PCR products were purified using Exonuclease 1 and Shrimp Alkaline Phosphatase (EXO-SAP) (Amersham Biosciences). EXO removed single stranded DNA, including primers and ssDNA produced from PCR and SAP digested unconsumed dNTPs by binding to phosphate groups. A total of 3ul of PCR product was combined with 2ul of EXO-SAP along with 5ul of sterile dH_2O and incubated at 37°C for 15 min, followed by 15 min at 80°C for enzyme inactivation. This product was then sequenced using a single primer (3.2pmol), 5X Sequencing Buffer (Applied Biosystems) and Big Dye 3.1 sequencing pre-mix (Applied Biosystems) was used to fluorescently label the PCR product. The program was as follows: 1 min at 96°C; 25 cycles of 10s at 96°C, 5s at 50°C, 4 min at 60°C; 5°C hold. Labeled sequence was then further purified by either ethanol precipitation or, if using a 96-well plate, an automated magnetic bead system. The 16 capillary ABI 3130xl Genetic

Analyzer was used to perform capillary electrophoresis of the products. Sequence was read using ABI Sequencing Analysis 5.2 software and sequence variants were identified by the program as well as manually.

2.5 High Performance Liquid Chromatography

The dosage of each *SGCE* exon was determined by High Performance Liquid Chromatography (HPLC). A multiplex PCR was used to amplify each *SGCE* exon and exon 7 from BX648242 on chromosome 5 as an internal control. Ten wild type (non-affected) controls were also used for each *SGCE* exon analyzed, and each PCR was done in triplicate. This ensured the average ratio of a normal exon dosage for each *SGCE* exon after HPLC analysis was approximately 1. The non-denatured PCR multiplex products were then run through the DNASep Prep HT column at 50°C. Triethylammonuim Acetate (TEAA) was used as an ion pairing buffer allowing the bridging between the DNA and the polystyrene-divinylbenzene (PS-DVB) copolymer beads within the column. A 0.1M solution of TEAA mixed with 25% acetonitrile was used as the elution buffer. As the concentration of acetonitrile across the column increased the smaller DNA fragments eluted off the column first followed by the larger fragments. The fragments were eluted at a flow rate of 0.9ml/min and were detected by using a UV detector set at 260nm.

Fragment size, determined by elution time, and intensity (in mV) was plotted using Navigator software (Transition Technologies Inc.) in chromatogram form. This allowed for the data retrieval of peak height, area, and size of each fragment. Normalized exon dosage ratios were determined by using the peak area values of the *SGCE* and control fragment exons and inputting them into the following formula:

$$\text{Normalized Ratio} = \frac{\textit{SGCE} \text{ exon / Control exon [patient]}}{\text{Mean } \textit{SGCE} \text{ exon / Control exon [10 W/T controls]}}$$

A total of 3 normalized ratios were generated for each patient per exon since PCR amplifications were done in triplicate. The average of the three was used as the final exon ratio and standard deviation was calculated. If one peak area value deviated outside the range of the other two, it was discounted from normalization calculation. Normal exon dosage threshold was set to 0.75 to 1.25, as suggested by Naimi et al (2005)[56]. Anything above or below this range was considered a potential duplication or deletion, respectively.

2.6 Multi-Ligation Probe Amplification

The dosage of each *SGCE* exon was determined using Multi-Ligation Probe Amplification (MLPA). The SALSA MLPA Kit P099-B2, designed by MRC-Holland, can detect deletions/duplications of one or more exons of the *GCH1*, *TH*, and *SGCE* genes. The probemix contains probes for 5 of the 6 exons of *GCH1*, 5 of the 14 exons for *TH* and all *SGCE* exons excluding alternatively spliced exons 10 and 11b. There are 9 probes that serve as internal controls that are specific for other genes on different chromosomes. For a complete list of probes and sizes please see Table 4.1.

Genomic DNA (100-200ng) was denatured at 98°C for 10 min, after which the probe mix and buffer was added. Ligation of probes to genomic DNA was performed at 60°C for 17 hours. Ligase-65 was then added and incubated at 54°C for 15 min allowing for the ligation of the probe pairs. PCR of the MLPA probes was then performed (33 cycles: 30s at 95°C, 30s at 60°C, 60s at 72°C; 20 min at 72°C) and products were labeled using a 6-

FAM dNTP mix supplied by the kit. The 3130xl Genetic Analyzer was used to perform capillary electrophoresis for separation of the amplified fragments. The instrument protocol FragmentAnalysis36_POP7 was used, with default initial injection voltage and initial injection time changed to 1.2kV and 20s, respectively. The size standard GS600LIZ was used to determine MLPA product fragment sizes.

Raw data from ABI files was entered into the program GeneMarker (SoftGenetics). Fragment peak data including height, area, and size was analyzed for each fragment. There were at least two wild type controls (having both copies of all exons being analyzed) used for determining the normal ratios of each probe in the set. Peak intensity normalization was performed using two methods: the first was based on setting height ratios of the 9 internal control probes to approximately 1, and the second normalized peak intensities based on the most probable median intensities. This second method is unique to the software and corrects for peak variation over size, since the trend is for peak intensities to drop as DNA fragment size increases. After intensity normalization, the data was plotted in two formats: ratio and regression type. In ratio, the intensity ratios of identical probes from the patient sample are compared to the wild type control samples. Outliers deviating outside of the 0.82-1.17 threshold lines (default setting used by CHEO diagnostic lab) were considered to be potential deletions or duplications. The regression plot format was obtained by taking the square root of peak intensities and provided a 99% confidence level of patient ratio compared to wild type ratio for each probe of normal copy number. Those probes that deviate from the regression line are considered suspect (duplications > 1.33 and deletions < 0.75). The combination of the two methods of peak intensity normalization as well as the two ratio plotting formats allowed for few false positives.

2.7 Genotyping and Linkage Analysis

Genotyping was performed using 39 microsatellite markers for analysis of all members of family 1 (see Table 5.1 for a complete list of markers). The di-, tri-, or tetra-nucleotide repeating sequence of these markers was amplified by end-labeled primers listed in the CHLC Human Screening Set (Weber Version 8) designed by Research Genetics Inc.. After amplification, an IR2 Stop Solution containing IRDye 800RS (LI-COR Biosciences) was added to the sample to label the product. These products were then denatured for 10 min at 95°C and run on an 8% polyacrylamide gel in the LI-COR DNA sequencer model 4000. Fragment sizes were separated and analyzed using Saga Generation 2 software. Two alleles were generated based on the sizes of the two PCR products of each individual for each marker.

Linkage analysis was performed with the computer program LINKAGE version 5.1 program using the MS-DOS (Microsoft) operating system. An autosomal dominant mode of inheritance was specified, penetrance was set to 80%, gene frequency (disease frequency) was set to 0.000001 (1/100,000), and theta values (recombination faction; distance from the marker) were set to start at 0 and increase in increments of 0.05. LOD scores were then calculated. A LOD score above 3.0 was considered linkage to the disease.

2.8 Flourescence *In Situ* Hybridization

Three color Fluorescence *In Situ* Hybridization (FISH) was performed on one affected individual from family 1 who has the chromosome 18 disease haplotype. Blood was drawn from the individual and sent to the TCAG Biobanking Facility in Toronto where the lymphocytes were isolated and immortalized with the Epstein Barr Virus (EBV). This

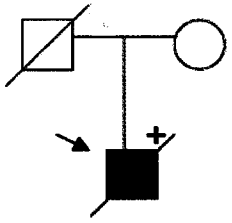
lymphoblastoid cell line was then sent to the Cytogenomics and Genome Resources Facility at the Hospital for Sick Children in Toronto. Eight Bacterial Artificial Chromosome (BAC) clones were chosen from the RPCI-11 Human BAC library based on their genomic location within the chr18 critical region as mapped in UCSC. BAC probe alignment to the genomic region is depicted in Figure 6.4. Clone sizes ranged from 156kb to 231kb and were approximately 500-700kb apart. BACs were fluorescently labeled with red, green or yellow fluorophores and hybridized to the chr18 critical region during metaphase, allowing potential inversions to be visualized.

Table 2.1. Clinical characteristics of the 15 families affected with Myoclonus Dystonia studied. L, left; R. right; UE, upper extremities; LE, lower extremities; N.K., not known.

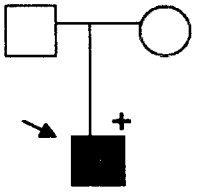
Family/ Pedigree Number	Number of Affected Individuals	Mean age of onset in years (range)	Myoclonus Location	Frequency of Myoclonus	Dystonia Present? (Location)	Alcohol Responsive	Unusual Features	SGCE Mutation Status
2	4	7.5 (1-17)	Face, neck, UE, trunk	N.K.	No	Yes		Exons 2 and 3 deleted
6	3	2 (1-3)	Face, neck, UE, LE	N.K.	Yes (UE, LE)	N.K.		Exons 2-5 deleted
19	1	12	both arms, both legs, face, post-cervical, axial	N.K.	Yes (N.K.)	N.K.		None
20	1	(baby)	R hand, head	N.K.	Yes (N.K.)	Yes	Tremor. Difficulty walking as a child.	None
21	1	5	both arms, left leg, face, neck	Constant	Yes (both arms, neck, voice)	Yes	Postural tremor in head since childhood	c.304C>T
22	1	12	L arm, head, generalized	Constant	Yes (both arms, head)	Yes	Tremor of jaw, lips, tongue	None
23	1	18	axial	N.K.	No	N.K.		None
24	1	6	both arms, face, neck	Constant	No	No		None
25	1	23	generalized, subtle	Constant	No	No	Tremor. Parasthesias in hands/feet.	None
26	1	1.5	both arms, both legs, axial	Constant	Yes (N.K.)	N.K.	Tremor in hand	None
27	1	1.5	both arms, both legs, axial	Constant	Yes (present on streesed gait)	N.K.		None
28	3	N.K.	neck	Constant	N.K.	N.K.	Postural and kinetic tremor in arms	c.300C>T
29	5	5 (2-10)	generalized, both arms, neck, axial	Constant	Yes (mild)	N.K.		c.285delT
30	1	childhood	both arms, face	Constant	Yes (hand)	No		None
31	4	34 (13-55)	both arms, R leg, neck	Constant	Yes (N.K.)	No	Tremor and chorea	None

Figure 2.1. Pedigrees of families 19-24 affected with Myoclonus Dystonia. Pedigrees are numbered based on the chronological order in which families are recruited for these studies. Filled symbols, affected individuals; +, we have received blood for the individual; arrow, the proband of the family; ?, possibly affected.

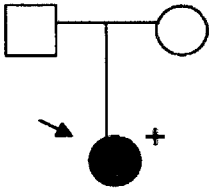
Family 19



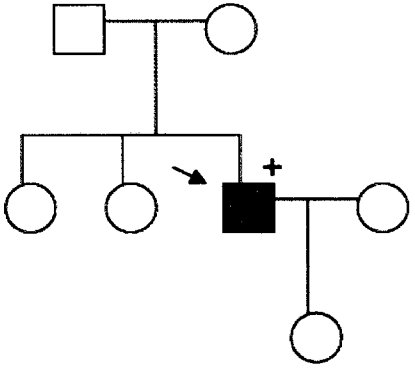
Family 20



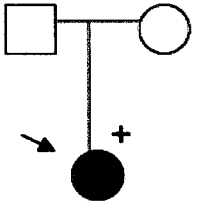
Family 21



Family 22



Family 23



Family 24

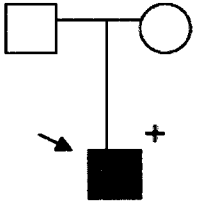
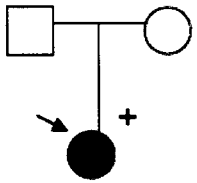
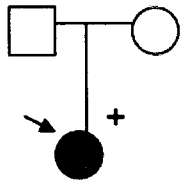


Figure 2.2. Pedigrees of families 25-31 affected with Myoclonus Dystonia. Pedigrees are numbered based on the chronological order in which families are recruited for these studies. Filled symbols, affected individuals; +, we have received blood for the individual; arrow, the proband of the family; ?, possibly affected.

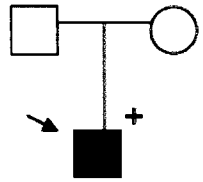
Family 25



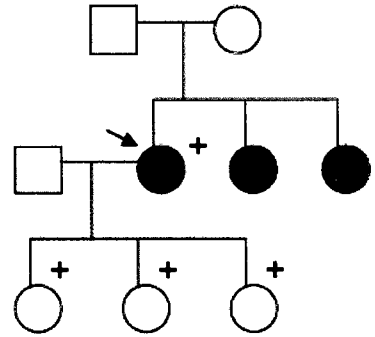
Family 26



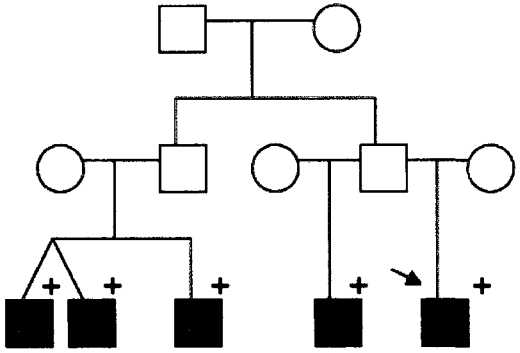
Family 27



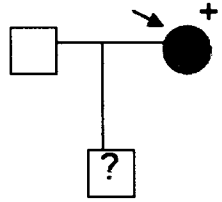
Family 28



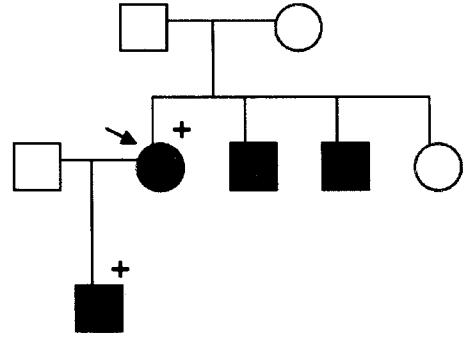
Family 29



Family 30



Family 31



Chapter 3.

Screening for *SGCE* Mutations in Myoclonus Dystonia Families by Direct Sequencing

Chapter 3 Screening for *SGCE* mutations in MD families by direct sequencing

3.1 Introduction

There have been 39 point mutations and 6 large deletions identified that have been found to be responsible for Myoclonus Dystonia[52]. These point mutations range from nonsense, missense, and splice-site mutations. Large deletions can encompass one exon or up to the entire gene and neighboring genes[52,25]. The majority of point mutations are typically found within the first half of the gene, likely due to the fact that the 5' half contains the largest exons. Mutations in *SGCE* typically alter the translational reading frame creating a premature stop codon and in turn a non-functional protein. Studies on three *SGCE* missense mutations have demonstrated that translation of these transcripts result in full length protein products that are quickly targeted for degradation via the proteasome pathway[26]. The *SGCE* gene is under the control of parent-of-origin regulation. This is due to differential methylation of a CpG island within the 5' end of *SGCE* causing transcription to be under the control of maternal imprinting[31,32]. Paternal transmission of an *SGCE* mutation often results in the MD phenotype, whereas maternal transmission is often non-pathogenic because the promoter is methylated.

There have been four families recruited and screened for mutations in the *SGCE* gene during the course of this project. Because the majority of variants are point mutations, individuals are first screened using PCR and direct sequencing. Sequencing will identify point mutations as well as small insertions, duplications and deletions within exons. The *SGCE* exons are sequenced for all individuals participating in the study, regardless of their

affection status. Family pedigrees and details regarding phenotypes can be found in Figure 2.1 and Table 2.1.

Epsilon sarcoglycan mutations account for a portion of Myoclonus Dystonia cases, ranging from 30-80%, depending on the literature[52,57,58]. Some research groups have found that discounting individuals that do not have classic MD phenotypes such as early age of onset, or the presence of both myoclonus and dystonia have increased their chances of finding *SGCE* mutations. However, these groups have overlooked a large subset of patients, some of whom have *SGCE* mutations. Our group has screened all individuals participating in our studies, even those with sporadic MD (without family history). In accordance with the current literature, with a group of four families with 13 affected individuals (from 9 of which we've obtained DNA), we could expect roughly 2-4 out of 9 to have *SGCE* mutations.

3.2 Materials and Methods

3.2.1 Myoclonus Dystonia patients

Families with Myoclonus Dystonia are recruited for our studies through collaboration with a neurologist specializing in movement disorders. Informed consent forms approved by the Ottawa Hospital Research Ethics Board were signed by each patient participating in the study. An assessment was performed by a neurologist and if possible patients were videotaped. Clinical characteristics and family history information was obtained and detailed in an information sheet. Family pedigrees are numbered based on our lab system, with families numbered in the chronological order in which they were recruited.

3.2.2 Sequencing

Blood was received and DNA extracted using the QIAGEN DNA Blood Mini Kit as previously described in chapter 2. Genomic DNA (150ng) was used for PCR of each *SGCE* exon. The PCR program was as follows: 95°C-2:00; 95°C-0:45, 58-62°C-0:45, 72°C-0:45 for 27-30 cycles; 72°C-10:00; 15°C-hold. Primer conditions are listed in Table 3.1. PCR clean-up was then performed using EXO-SAP, and 2ul of this product was then used in the sequencing reaction. BigDye3.1 sequencing pre-mix (ABI) was used to fluorescently label the PCR product. Labeled sequence was then purified using ethanol precipitation, or by an automated magnetic bead capture system. The ABI 3130xl Genetic Analyzer was used to process the product by capillary electrophoresis and Sequencing Analysis 5.2 software was used for analysis.

3.2.3 Mutation Confirmation

The enzyme BsiEI (New England Biosystems) was used to confirm a mutation that created a cut site. PCR product (~1ug) from the patient was digested with 5 units of BsiEI, with 0.5ul of BSA (bovine serum albumin) and 3ul of 10X NEBuffer (final concentration 1X) for a total reaction volume of 30ul. The mix was digested for 2 hours at 60°C, followed by enzyme inactivation for 20 min at 80°C. All 30ul was loaded onto a 1.5% agarose gel, where the expected products were observed.

3.3 Results

3.3.1 A c.300G>A mutation is found in family 28

Genomic sequencing was performed using both the forward and reverse primers for each *SGCE* exon, enabling sequence to be read in both directions. Individual M28-1 is the proband of family 28 and was found to have a heterozygous c.300G>A mutation in exon 3 resulting in a nonsense mutation causing a tryptophan change to a stop codon at amino acid 100 (p.Trp100X) (Figure 3.1). One of her three unaffected daughters, M28-3, was also found to harbor this mutation whereas M28-4 did not share this mutation. DNA was not available for M28-2. This premature stop codon causes a truncated and likely non-functional protein that is targeted for degradation.

Individuals M28-1 and M28-3 also have a heterozygous AC-dinucleotide insertion within the intronic sequence, 64bp downstream of exon 3 (data not shown). During sequence analysis, this created an overlapping read downstream of the insertion and made it difficult to read the forward primer sequence. An alternate primer was designed to bypass this to allow for successful sequence reading (for primer sequence see Table 3.1). This AC-insertion is not within the coding sequence or within a splice site of the gene and has been reported by UCSC Genome Browser to be present within the general population.

3.3.2 A c.285delA mutation is found in family 29

Individual M29-1 is the proband of this family. He is currently 8 years of age and developed myoclonic jerks when he was 2 years old. Through PCR and direct sequencing he

was found to have a heterozygous c.285delA mutation within exon 3 of *SGCE* (Figure 3.2). His paternal half-brother as well as 4 cousins were later recruited for this study. These individuals range in age from 6-16 and have an average age of onset of 5 (range 2-10). All were found to have the same mutation as the proband. We were able to confirm this mutation by restriction enzyme digestion (Figure 3.3). This frameshift mutation creates a BsiEI cut site within the exon 3 amplicon, creating two products: 112bp and 245bp. The full length uncut exon 3 is 357bp. All 5 affected individuals were found to have all three products, hence are heterozygous for the mutation. This single nucleotide deletion results in a frameshift of the downstream coding sequence, thereby changing the remaining amino acid sequence (p.Pro95fs). Since this mutation occurs early in the 5' sequence of the gene, it is likely such a frameshift creates a non-functional protein and is in turn responsible for the disease phenotype seen within the affected members of this family.

3.4 Discussion

There have been 39 *SGCE* mutations previously identified that are responsible for MD and novel mutations continue to be reported. We have found two in our cohort, the c.300G>A mutation which was previously described in a French family[22] and the c.285delA mutation which is novel. Both mutations have been found within the coding sequence of exon 3. Interestingly, 9 out of the now 40 total point mutations identified are found within this exon (23%), while it makes up 12% of the coding sequence. This could indicate that exon 3 is a "hot spot" for point mutations.

The majority of *SGCE* mutations reported in MD families are nonsense mutations that result in a premature stop codon and abolishment of the synthesis of the full length

protein[26]. The nonsense mutation p.Trp100X was found in two individuals in family M28, and is likely to create a truncated protein that is quickly degraded. For the proband of this family, this is a disease-causing mutation. However, her daughter, M28-3, also has the mutation and is asymptomatic. This mutation was transmitted to the daughter on the maternal allele and is likely imprinted and therefore not expressed.

Missense mutations are also commonly found in MD families, and although they do not create a stop codon, they result in a single amino acid change thereby impeding its function. Recently, a group studied 3 *SGCE* missense mutations that were found in MD patients[26]. They found that the mutant *SGCE* proteins were unable to localize properly to the plasma membrane of cortical neurons and neuroblastoma cells. These proteins are instead polyubiquitinated and targeted to the proteasome for degradation. The 1bp deletion mutation (c.285delA) causes a frameshift in the amino acid sequence (p.Pro95fs) and was found in 5 symptomatic individuals in Family 29. This mutation likely causes a downstream stop codon which is responsible for a non-functional ε-SG and subsequently responsible for the disease phenotype in this family. We can predict that this mutation was initially present in the grandmother, who in turn transmitted it to her two sons, who, according to the information we have been given, are asymptomatic. This is likely due to imprinting of the inherited maternal allele of *SGCE*. The affected grand-children in this family have likely inherited the *SGCE* mutation from their fathers, and since the mutated *SGCE* gene is expressed on the paternal allele, they exhibit the MD phenotype. DNA is not available from the parents; therefore we are unable to confirm this mode of transmission.

When possible, *SGCE* mutations found by direct sequencing are confirmed using an alternate method, as it allows for rapid screening of the mutation. Enzyme digestion was

able to confirm the c.285delA mutation in family M29, since this deletion created a BsiEI cut site (Figure 3.3). Unfortunately, we were unable to confirm the c.300G>A mutation by digestion analysis, since this change did not create or impair an existing cut site. Since RNA is also extracted from patient blood samples, reverse-transcription yields cDNA which in turn can be used for mutation confirmation using primers specific to *SGCE* cDNA. However, this proved to be unsuccessful since these primers did not amplify *SGCE* cDNA from lymphocytes. Testing these primer sets on cDNA reverse transcribed from human whole brain RNA allowed for the amplification of *SGCE*. It was therefore determined that *SGCE* is not expressed in lymphocytes. Lymphocyte expression of *SGCE* has not been reported in the literature or tissue expression databases available online.

There is one individual in family 30 that has been diagnosed with MD. Her onset was during childhood and like most patients in this cohort, her myoclonus is constant. She presents with dystonia as well, and her symptoms are not alcohol responsive. The cause of disease in this individual remains unknown since no *SGCE* mutations were found. There are four affected individuals in family 31; we have obtained blood for the mother and a son. Both have chorea and tremor in addition to their MD phenotype. Chorea is typically not seen in conjunction with MD, and has been reported to be a possible differential diagnosis due to its similar "jerky" phenotype as myoclonus[15]. Since the two affected individuals were found to be S*GCE*-mutation negative, the *TITF-1* gene responsible for Benign Hereditary Chorea could be a potential candidate gene to screen for in this family.

We have found that two of the four families have mutations within exon 3 of *SGCE*. In this cohort, 6 of the 9 affected individuals (in which we obtained DNA), or two out of four families, were found to have *SGCE* mutations, equaling 67% of the patient cases. This

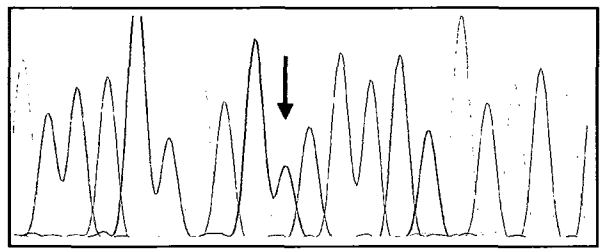
is fairly high in comparison to the 30-40% incidence that was found previously by our group in a cohort of 27 families. Although, it is not an unusually high incidence since reports in the literature range from 30-80%[2,22,51,50]. Factors such as age of onset, presence of both myoclonus and dystonia as well as alcohol responsiveness, are considered classical symptoms of MD and can be modest predictors of *SGCE* mutations. The 4 families within this cohort studied have a mean age of onset of 13 (range 2-55) and alcohol responsiveness is unknown for two families, and not seen in the other two. Half the affected individuals in this cohort are still very young (under 18), consequently alcohol responsiveness is unknown. The classical symptoms of MD are present within the families in this cohort that were *SGCE*-mutation positive, thereby supporting the basis for these as positive indicators of *SGCE* mutations.

Table 3.1 Primer sequences used for amplification of *SGCE* exons. Primer sequences used to amplify each *SGCE* exon, including those that are alternatively spliced. Tm, melting temperature used (°C); [$MgCl_2$], concentration of magnesium chloride (in mM) in the 10X PCR buffer used; 3*, alternate exon 3 forward primer.

Exon	Forward Primer Sequence	Reverse Primer Sequence	Product Size (bp)	Tm	[MgCl]	# of Cycles
1	GTCCGGGACAGAAAGAGAGG	GCTAGAGGGAGTACGGGATT	357	61	1.5	27
2	TCCAGAAATATGCTTTCCTTAACA	GCTGAATTATCAAGGGCGTATC	400	59	2.0	27
3	CTGTGTTTCCATGCACAAAAT	CCAAAGCAACATGTGTGAAAA	357	62	1.5	27
4	TTTCACCATCATAACACACCAT	CTCATTGCCCAGAGAAGGAA	417	62	1.5	27
5	TGCAATAGGCCATCTTCCAT	CTTCCTGCTGCCAGGATTAT	404	62	1.5	27
6	AAACGTTAACTCCAGCCACA	TCCTGCTTTTAAGGTGGATTG	302	60	2.0	30
7	TGAACCAAATGAAACTTTGCT	CAAAGAATGCTTTAGTGTATCCAG	391	61	2.0	27
8	CACATGTATGGAGCATGATGG	TGCAAGATCCCATGAGCATA	185	58	2.0	27
9	TTTCAGGAGAGTAGGGGTGA	TTGATGACCCATCAGGCTAA	411	60	2.0	29
10	TGTTTTGCCTTATTTGGTGAA	GGGGAATTGTGCTGGATTAC	285	58	2.0	27
11	TTGTGGAATGAGAATGAACACAT	TCCATGCCCTGACTAACTCC	499	58	2.0	27
11b	AATAGCGGCATTGTGGTAGG	GCGAGATGGAAGCTACAAGG	439	61	2.0	27
12	TTGTGGAATGAGAATGAACAC	TCCAATAATGCATGAGCTTTTC	302	61	2.0	30
3*	GCTCTTTCTAGGTGTAAATTACCTCA	CCAAAGCAACATGTGTGAAAA	258	62	1.5	27

Figure 3.1 A nonsense mutation in exon 3 of *SGCE* is found in family 28. Sequencing results are seen to the left, followed by the corresponding genomic and amino acid sequence for each individual on the right. Both M28-1 (A) and M28-3 (B) have a c.300G>A, indicated in red, within exon 3 of SGCE. This creates a pTrp100X nonsense mutation. (C) M28-1 does not have this mutation. DNA for M28-4 was not available.

A



M28-1

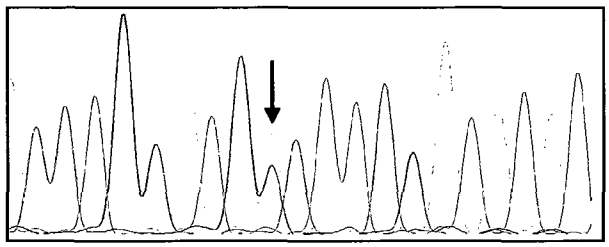
c.300G>A

5' CGA CCT GGA TGG 3'

R P G X

p.Trp100X

B



M28-3

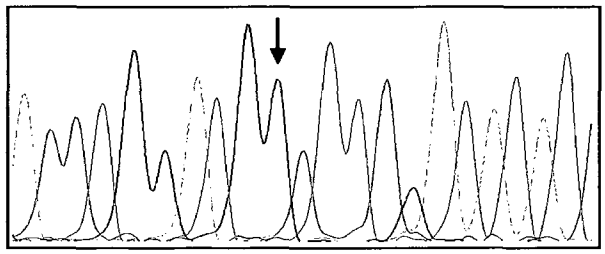
c.300G>A

5' CGA CCT GGA TGG 3'

R P G X

p.Trp100X

C

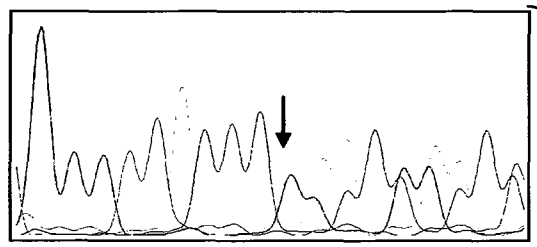


M28-4

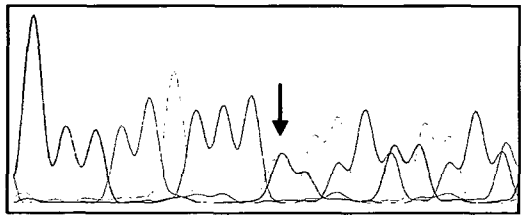
5' CCT GGA TGG CTT CGA 3'

P G W L R

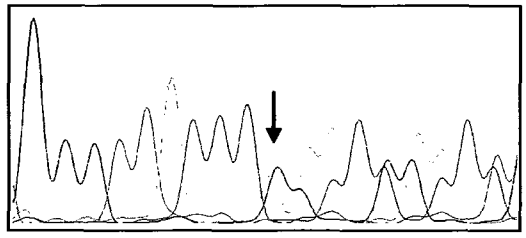
Figure 3.2 A 1bp deletion in exon 3 of *SGCE* is found in family 29. Sequencing results are shown to the left and the corresponding genomic and amino acid sequence for each individual is on the right. All 5 affected individuals were found to have the novel c.285delA mutation. This created a p.Pro95fs, causing a frameshift of the amino acid sequence as shown in red. The WT sequence is shown for reference in brackets below.



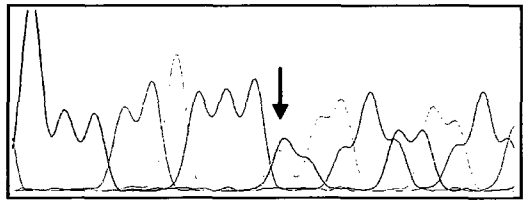
M29-1



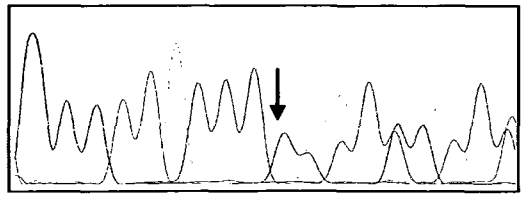
M29-2



M29-3



M29-4



M29-5

c.285delA

5' ATG GGT TAC CCG **ACC GAC CTG** 3'

M G Y **P** **T** **D** **L**

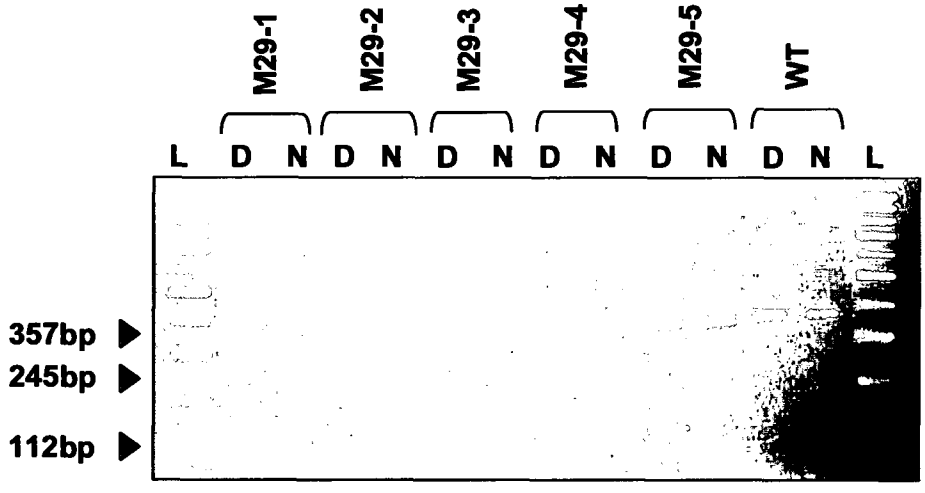
p.Pro95fs

5' ATG GGT TAC CCA GAC CGA CCT 3'

M G Y P D R P

WT sequence

Figure 3.3 Confirmation of the c.285delA mutation found in family 29 by enzyme digestion. Genomic DNA from individuals M29-1 to M29-1 was amplified for exon 3 of *SGCE*. This amplicon was then digested (D) or not digested (N) with the BsiEI enzyme. The c.285delA creates one BsiEI cut site in this amplicon, thereby creating two products: 112bp, 245bp. The uncut product is 357bp in size. This restriction enzyme digestion has confirmed the presence of a heterozygous c.285delA mutation in all five symptomatic MD patients in this family, since all three products are seen. A WT individual (non-affected) was used as a negative control for the digestion. L = ladder (100bp long-range).


M29-1
M29-2
M29-3
M29-4
M29-5
WT
L D N D N D N D N D N D N L
357bp
245bp
112bp

Chapter 4.

Screening for *SGCE* deletions using HPLC and MLPA

Chapter 4 Screening for *SGCE* deletions using HPLC and MLPA

4.1 Introduction

The majority of *SGCE* mutations identified in Myoclonus Dystonia patients involve a single nucleotide change resulting in a nonsense mutation. Since more research groups are now screening for large deletions as well, there is additional evidence to suggest that they are more common than originally thought. Thus far there have been 6 large deletions identified, many of these encompassing more than one exon[52,25]. Direct sequencing will not identify a deletion of an entire exon since primer binding will still occur on the single undeleted allele, enabling amplification and subsequent sequencing of the single exon. It is therefore necessary to perform an exon dosage analysis of the *SGCE* gene when screening MD patients. Missed deletions could account for an unknown portion of reduced penetrance reported. Groups have reported the deletions of exons 1, 5 and 6[57,58]. As well, Asmus *et al* (2007) described large deletions that spanned up to 8.8Mb and encompassed 44 genes on chromosome 7 in one patient, including *SGCE*[25]. This patient presented with MD as well as other symptoms related to the haploinsufficiency of neighboring deleted genes. Our group has also previously identified two large deletions spanning multiple exons in two families[53].

High Performance Liquid Chromatography (HPLC) was shown by Naimi *et al* (2005) as an effective method for exon dosage analysis[56]. They were able to efficiently screen 90 patients with peripheral neuropathies for genomic rearrangements in 17p11.2. By amplifying the exon of interest and a control exon in a multiplex reaction, PCR products were then able to be separated by size on an HPLC column. They obtained the peak area of

each of the two products and then generated a ratio of the test to control exon. When performed in triplicate and normalized to 10 WT unaffected controls, they found their ratios to reflect the dosage of the particular exon of interest. An exon ratio between 0.75-1.25 was considered to have both copies of the exon, whereas anything above or below could be a potential duplication or deletion, respectively. There are other potential methods for exon dosage analysis, such as southern blotting and quantitative-PCR (Q-PCR). However, HPLC was initially chosen to be utilized for exon dosage analysis in families 19-27 due to its cost effectiveness as well as its published success.

Another method of exon dosage has recently been developed by MRC Holland. This company has created over 200 kits that screen numerous genes implicated in a variety of diseases for exon deletions or duplications. They have recently developed a kit that can screen an individual for *SGCE*, *TH1* and *GCH* in a single reaction. As well, as opposed to HPLC which uses one internal control exon, this method allows for 9 controls. The kit contains probes for 5 of the 6 exons of *GCH*1, 5 of the 14 exons for *TH*, which are both are implicated in Dopa Responsive Dystonia. All *SGCE* exons, excluding alternatively spliced exons 10 and 11b, are included in the kit. Multi-Ligation Probe Amplification (MLPA) PCR products are run on capillary electrophoresis where the peak height of each probe is analyzed. This data allows for the generation of a dosage ratio for each exon analyzed. A normal exon dosage would have a ratio between 0.82-1.17, anything above or below would indicate a possible duplication or deletion, respectively. Research groups studying MD have recently employed the use of MLPA for *SGCE* screening with success. Grunewald *et al* (2007) have found two heterozygous deletions of the *SGCE* gene using this kit[52]. As well, this method is considerably more efficient than HPLC, since the *SGCE*, *TH1* and *GCH1*

exons are screened in one reaction. Therefore, we have also employed the use of the SALSA MLPA Kit to screen for deletions in families 19-31.

4.2 Materials and Methods

4.2.1 DNA extraction

Genomic DNA was extracted from whole blood drawn from patients in yellow top ACD Vacutainer tubes. Extraction was performed using the QIAamp DNA Blood Mini Kit. Briefly, 20ul of protease was added to 200ul of whole blood along with 200ul of lysis buffer (AL). Samples were incubated at 56°C for 10 min, after which 200ul of 100% ethanol was added. The mixture was applied to a QIAamp Spin Column and centrifuged at 6000 x g for 1 min to remove cellular debris. Two washes were then performed using buffers AW1 and AW2, and purified genomic DNA was then eluted from the filter with 100ul of TE (10mM Tris-HCl, 1mM EDTA, pH 8.0).

4.2.2 High Performance Liquid Chromatography

Exon dosage of the *SGCE* gene was performed by High Performance Liquid Chromatography (HPLC). A multiplex PCR was used to amplify each *SGCE* exon with exon 7 from BX648242 on chromosome 5 as an internal control. In order to inhibit the over-saturation of product that could impede the quantification by HPLC, 25 cycles was used for all reactions. *SGCE* primer sequences are listed in Table 4.1. Ten WT (non-affected) controls were also used for each *SGCE* exon analyzed, and each PCR was done in triplicate. This ensured the average ratio of a normal exon dosage (2 copies) for each *SGCE* exon to be

approximately 1. The non-denatured PCR multiplex products were then run through the DNASep Prep HT column at 50°C. Triethylammonuim Acetate (TEAA) was used as an ion pairing buffer allowing the bridging between the DNA and the polystyrene-divinylbenzene (PS-DVB) copolymer beads within the column. A 0.1M solution of TEAA mixed with 25% acetonitrile was used as the elution buffer. As the concentration of acetonitrile across the column increased the smaller DNA fragments eluted off the column first followed by the larger fragments. The fragments were eluted at a flow rate of 0.9ml/min and were detected by using a UV detector set at 260nm.

Fragment size, determined by elution time, and intensity (in mV) was plotted using Navigator software (Transition Technologies Inc.) in chromatogram form. This allowed for the data retrieval of peak height, area, and size of each fragment. Normalized ratios were determined by using the peak area values of the *SGCE* and control fragment exons using the following formula:

$$\text{Normalized Ratio} = \frac{\textit{SGCE} \text{ exon / Control exon [patient]}}{\text{Mean } \textit{SGCE} \text{ exon / Control exon [10 WT controls]}}$$

A total of 3 normalized ratios were generated for each patient per exon since amplifications were done in triplicate. The average of the three was used as the final exon ratio and standard deviation was calculated. If one peak area value deviated outside the range of the other two, it was discounted and not used for normalization calculation. Normal exon dosage threshold was set to 0.75 to 1.25, as suggested by Naimi *et al* (2005)[56]. Anything above or below this range was considered a potential duplication or deletion, respectively.

4.2.3 Multi-Ligation Probe Amplification

Exon dosage of the *SGCE* gene was performed using Multi-ligation Probe Amplification (MLPA). The SALSA MLPA Kit P099-B2, designed by MRC-Holland, can detect deletions/duplications of exons. There are a total of 34 different probes with amplification products ranging between 130 and 409 nucleotides (nt) in size, as well as 7 fragments smaller than 130nt that serve as DNA quality and denaturation controls. The probemix contains probes for 5 of the 6 exons of *GCH1*, 5 of the 14 exons for *TH* and 11 *SGCE* exons. There are 9 probes that serve as internal controls that are specific for genes on different chromosomes. For a complete list of probes and sizes please see Table 4.1. It is important to note that the MLPA probes that are used to amplify *SGCE* exon 10 actually correspond to exon 11, and exon 11 probes correspond to exon 12 of the gene. These discrepancies were found through the use of inputting the probe chromosomal start and end points into UCSC Genome Browser.

A volume of 5ul of genomic DNA (100-200ng) was denatured at 98°C for 10 min, after which the MLPA probe mix and ligation buffer was added. Ligation of probes to patient genomic DNA was performed at 60°C for 17 hours. Ligase-65 was then added and incubated at 54°C for 15 min allowing for the ligation of the probe pairs. PCR of the MLPA probes was then performed (33 cycles: 30s at 95°C, 30s at 60°C, 60s at 72°C; 20 min at 72°C) and products were labeled using a 6-FAM dNTP mix supplied by the manufacturer. The 3130xl Genetic Analyzer was used to perform capillary electrophoresis for separation of the amplified fragments. The instrument protocol FragmentAnalysis36_POP7 was used, with default initial injection voltage and initial injection time changed to 1.2kV and 20s,

respectively. The size standard GS600LIZ was used to determine MLPA product fragment sizes.

Raw data from ABI files was inputted into the program GeneMarker (SoftGenetics). Fragment peak data including height, area, and size, was analyzed for each fragment. There were at least two wild type controls (having both copies of all exons being analyzed) used for determining the normal ratios of each probe in the set. Peak intensity normalization was performed using two methods: the first was based on setting height ratios of the 9 internal control probes to approximately 1, and the second normalized peak intensities based on median intensities. This second method is unique to the software and corrects for peak variation over size, since the trend is for peak intensities to drop as DNA fragment size increases. After intensity normalization, the data was plotted in two formats: ratio and regression type. In ratio, the intensity ratios of identical probes from the patient sample are compared to the wild type control samples. Outliers deviating outside of the 0.82-1.17 threshold lines (set by GeneMarker) were considered to be potential deletions or duplications. The MLPA manufacturer claims a threshold of 0.7-1.3 is a suitable normal dosage range; however the range we have used is more stringent and is used by the Children's Hospital of Eastern Ontario (CHEO) diagnostic lab. The regression plot format was obtained by taking the square root of peak intensities and provided a 99% confidence level of patient ratio compared to wild type ratio for each probe of normal copy number. Those probes that deviated from the regression line were considered suspect (duplications > 1.33 and deletions < 0.75).

4.3 Results

4.3.1 HPLC – Dosage Analysis

Myoclonus Dystonia families 19-27 were screened for genomic rearrangements using HPLC. This cohort was previously tested for mutations by direct sequencing, after which it was found that individual M21-1 had a R102X (c.304C>T) nonsense mutation in exon 3 of *SGCE* (data not shown). This individual was therefore not included in the HPLC analysis. Normalized ratios for each *SGCE* exon are given in Figures 4.1 through 4.6. Raw data generated from HPLC as well as the calculated ratios for each triplicate for patients and 10 controls are listed in Appendix 1. A normal exon dosage range was observed for the majority of exons for the patient set. Overall, 8 of the 13 exons tested in the patient set had normalized ratios that were within the normal range of 0.75-1.25. Patients M19-1 and M20-1 showed ratios above 1.25 for exons 4, 6, 7 and M19-1 had a ratio of 3.25 for exon 11b (Figures 4.2, 4.3, 4.4). Results for exon 6 were also unexpected, with ratios above 1.25 for patients M19-1, M20-1, M22-1, M23-1 and M25-1 (Figure 4.2). Exon 8 for individual M22-1 had a ratio of 1.6, whereas all other patients for this exon indicated normal exon dosage (Figure 4.3). This was investigated further by repeating the exon 8 analysis for this patient. Four WT controls were used for normalization and patients M25-1 and M27-1 were also repeated to test the reproducibility of the method. In this second analysis, M22-1 was found to have a ratio within the normal dosage threshold (Figure 4.6, A). Patients M25-1 and M27-1 were found to have normal dosage ratios of 0.95 and 0.8, respectively. These values were consistent with the previous ratios of 0.85 for M25-1 and 0.95 for M27-1 (Figure 4.6, B).

4.3.2 MLPA – Dosage Analysis

Myoclonus Dystonia families 19-31 as well as individual M1-39 were screened for exon deletions or duplications using MLPA. Families 19-27 were previously screened using HPLC however due to contradictory results; MLPA was also used to ensure there were no missed genomic rearrangements. At least two WT controls (having both copies of all exons being analyzed) were used for determining the normal ratios of each probe in the set. As well, probe peak data of WT controls could only be used for normalization of the patient samples of that particular experiment. For this reason, controls are included in each of the three experiments performed, as illustrated in Figures 4.7 through 4.12. Raw data, including probe ratios and peak area data, can be found in Appendix 2. The first experiment included 3 WT controls, families 28-31, and individual M2-1. Our group had previously screened M2-1 using HPLC and had identified a deletion spanning exons 2 and 3 of *SGCE* (data not shown)[53]. This patient was therefore used as a positive control. Indeed the peak ratios of exons 2 and 3 for this individual were 0.57 and 0.78, respectively, which were below the 0.82-1.17 normal threshold (Figure 4.8). Patient M28-1 demonstrated a low peak ratio for exon 3 of *SGCE* (0.79). This individual was previously shown to have a heterozygous c.300G>A mutation within exon 3 and was therefore determined to have both copies of the exon. After repeating MLPA analysis of M28-1 in a subsequent experiment, the peak ratio for this exon was within normal dosage range (0.87) (Figure 4.11). Both individuals from family 31 had a number of probes that deviated from the normal copy number threshold. Individual M31-1 showed *SGCE* exon 6 and *TH* exon 8 ratios of 0.75 and 0.78, respectively (Figure 4.8). Patient M31-2 showed additional deviation, with 5 *SGCE* exons above or below normal range: exon 2 – 0.76; exon 4 – 0.75; exon 5 – 1.28; exon 6 – 0.79; exon 9 –

1.28. The sporadic distribution of these exons as well as their ratios being both above and below normal threshold is indicative of a sample problem resulting in false positives. There were no heterozygous variants found within these exons from direct sequencing to conclude that these family members have both alleles. The next cohort to be tested included a total of 10 individuals from families 19 – 28 (Figures 4.10 and 4.11). Note that M24-1 was unsuccessful and repeated in a subsequent analysis (Figure 4.12). Analysis was successful in 3 controls, however WT-3 had a single probe that was out of range. This control was stricken from statistical analysis to ensure the normalization of this probe was not used for patient analysis. The remaining 2 WT controls were subsequently used (Figure 4.9). Family 6 was previously found to have a large deletion spanning exons 2 through 5, therefore these individuals were used as positive controls[53]. Dosage ratios for these four exons were below the normal range for all three individuals in family 6 thereby confirming this deletion (Figure 4.9). All exon peak ratios were normal for patients in this cohort, with the exception of M27-1 which had 5 exons below normal range (Figure 4.11). Upon re-testing, this patient was found to be within the normal threshold for all exons (Figure 4.12). M1-39 is an individual from family 1 that was not previously tested for large *SGCE* deletions, and was found to have both copies of all exons (Figure 4.12).

4.4 Discussion

Myoclonus Dystonia families 19 through 27 were screened for genomic rearrangements using HPLC. The majority of *SGCE* exons were found to be within the normal dosage range for this cohort of 8 patients. Exons 4, 6, 7, and 11b produced unexpected results with various patients showing ratios above the normal exon dosage

range. There is a pattern of ratios tending to be too high, alluding to the possibility that different PCR efficiencies could exist for the *SGCE* exon versus the control exon. It is possible that the *SGCE* primers for the exons mentioned above have increased binding efficiency compared to the control primers at the particular melting temperature used in the multiplex PCR. This would favor the increase in *SGCE* product over control product and thus create a higher ratio. As well, the control amplicon is larger than the *SGCE* amplicon for all exons tested, with the exception of 11b. It is possible that there is a bias for the smaller *SGCE* products to be amplified over the control, which could also cause an increase in *SGCE* product thereby reflecting high ratios. Prior to HPLC, each multiplex product is run on an agarose gel to ensure the reactions were successful. By gel, it was generally found that the two exon product bands had the same intensity, although small quantity changes could be missed. Two exon 8 analyses via HPLC were performed on patient M22-1 and yielded contradictory results, perhaps illustrating the unreliability of this method. However, HPLC has worked previously for our group and has provided results consistent with MLPA. Unfortunately, HPLC is much more technically demanding and time consuming, hence MLPA is the preferred method of exon dosage analysis.

Families 19 through 27 were examined using MLPA, as well as the 4 families recently recruited during this project. Individual M1-37 was also examined, resulting in a total of 13 families analyzed by MLPA. Families 2 and 6 were previously found by our group to have large deletions[53]. By re-testing these individuals, we were able to both confirm these mutations as well as validate the use of MLPA for exon dosage analysis. The MLPA kit is not specific for finding point mutations, however a variant within the probe binding sequence could possibly decrease probe hybridization to the target genomic DNA.

Individual M28-1 has a c.300G>A mutation within the exon 3 probe binding sequence and was initially found to have a low peak ratio of 0.79 that would indicate decreased probe binding efficiency. When repeated the exon was found to have a normal dosage value of 0.87, but still had the lowest peak ratio of all 34 probes in the kit. Since this mutation is within the probe binding sequence, MLPA probe hybridization is therefore likely to be mildly affected by point mutations within these target sites.

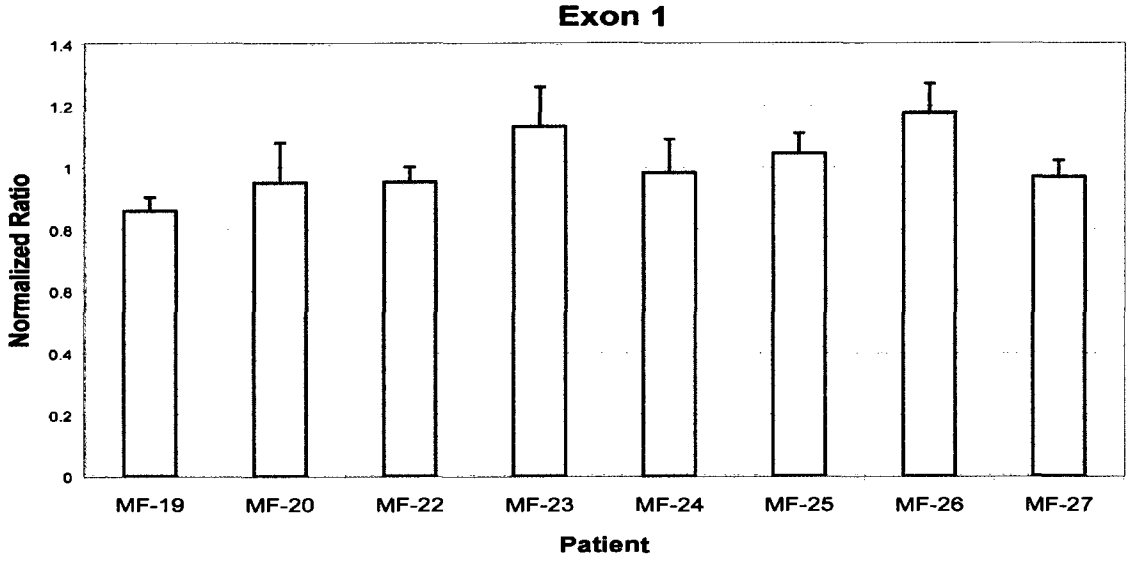
The two individuals in family 31 showed a number of exons with peak ratios above and below the normal threshold. Since a deletion or duplication would be spread through adjacent exons as seen in families 2 and 6, a sporadic distribution of abnormal-dosage exons would indicate a series of false positives. This could be a result of poor sample quality due to impurities such as protein or excess salt within the extracted DNA, or a problem with the MLPA process itself for those particular samples. Unfortunately there were no heterozygous variants found within these exons from direct sequencing to conclude that these family members have both alleles. These samples can be repeated when screening future affected families recruited for our studies.

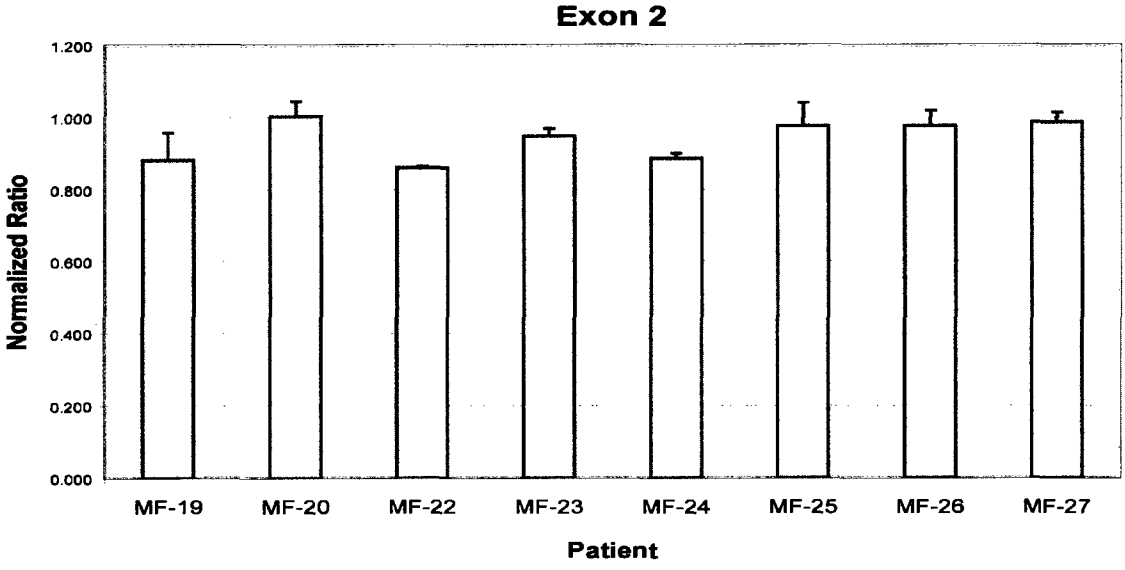
In conclusion, we had previously employed the use of HPLC for exon dosage analysis, but due to contradictory and variable results, we have since moved to MLPA. This method contains many more internal control probes (nine versus one for HPLC) and less starting DNA is required. We confirmed the methodology by detecting deletions which were previously described by our group. In our future studies, this will be an excellent tool for screening MD families for genomic rearrangements.

Table 4.1 SALSA MLPA Kit P099-B2 Probe List. There are a total of 34 different probes with amplification products ranging between 130 and 409nt in size. The probemix contains probes for 5 of the 6 exons of *GCH1*, 5 of the 14 exons for *TH* and 11 *SGCE* exons. There are 9 probes that serve as internal controls that are specific for genes on different chromosomes. *SGCE* exons 11b and 10 which are alternatively spliced are not included in this kit.

Probe #	Fragment Size	Gene	Exon	Chr. Position	Start Point	End Point	Left Probe Sequence	Right Probe Sequence
1	130	IL4	control	05q31.1	132037612	132037671	CTACATTGTCACTGCAAATCGACACCTAT	TAATGGGTCTCACCTCCCAACTGCTTCCCCCT
2	142	GCH1	1A	14q22.2	54439452	54439504	CCTAGGGGCCGGCGCCCTTTTCCT	TCCCTCCCTGCGCTTGCGAACCCCTCCG
3	148	DMR1	control	11p15.5	2121936	2121993	GCGGCCCTTGGGTTTCTTCCT	TACACCCGGACGCCCGCTAAGCTCGGGCTGCCGC
4	154	SGCE	3	07q21.3	94095518	94095590	TTGTCCTTTGGATATATCGAAGCCATCCA	GGTCGGTCTGGGTAACCCATTAAATTTGTATTAAATGTTATGGG
5	160	NSD1	control	05q35.3	176629353	176629426	GGAGAGGGTGACGTGAGCAGCAAGGA	TAAGATGGGCAAAGGAGTGGATGGGACATATAAAAAAGGTAACTTTATCC
6	166	TH	1	11p15.5	2149521	2149573	CCACGCCACAGGCCAAGGGCT	TCCGCAGGGCCGTGTCTGAGCTGGACGCCAA
7	172	SGCE	9	07q21.3	94066006	94066079	TCCTTTACACACAGACAACTATGATAGCACAA	ACATGCCATTGATGCAAACGCAGCAGTAAGTGTCCACTTAG
8	178	GCH1	2	14q22.2	54401796	54401869	GATGAGATGGTGATTGTGAAGGACATAGAC	ATGTTTTCCATGTGTGAGCATCACTTGGTTCCATTTGTTGGAA
9	184	TH	3	11p15.5	2147624	2147678	CGTCCCCGCGGTTCATTGGGCGCA	GGCAGAGCCTCATCGAGGACGCCCGCAAGGA
10	190	SGCE	4	07q21.3	94090319	94090396	TGTGTGTTATGATGGTGAAAACTCAGATAGG	AAGCCAAACAGATAGTCCTCTTATTTCTCTAAGCCTTTTTCATTAC
11	196	G1P2	control	01p36.33	938742	938799	TGGCTGAGAGGCAGCGAACTCA	TCTTTGCCAGTACAGGAGCTTGTGCCGTGGCCCA
12	202	GCH1	3	14q22.2	54396148	54396221	CCGTCTCTTCTCTGAAGGTCCATATTGGT	TATCTTCCTAACAAGCAAGTCCTTGGCCTCAGCAAACTTGCGAG
13	211	TH	4	11p15.5	2146393	2146451	CCATCTAGAGACCCGGCCCGCCCAGA	GGCCGCGAGCTGGGGGCCCCCACCTGGAGTAC
14	219	SGCE	1A	07q21.3	94123265	94123320	CAGGAGCTGGGAGACCCCTGTGCTTGGAC	GGGACAGGGTCGGGGGACACGCAGGATGAGCCCC
15	229	GCH1	5	14q22.2	54382235	54382300	CAAATTGCTGTAGCAATCACGGAAGCCTT	GCGGCCTGCTGGAGTCGGGGTAGTGGTTGAAGCAA
16	238	TH	8	11p15.5	2144737	2144789	GGGGAGCACCTGGAGGCCTTT	GCTTTGCTGGAGCGCTTCAGCGGCTACCGGG
17	247	SGCE	2	07q21.3	94097023	94097093	TCAAAGAGGACACCTGCTGATGGGTATACAT	TCCGATCGGAGTGTACCTTGGAGAAAATACTGTACACTGA
18	256	GCH1	1B	14q22.2	54438895	54438960	CCACGCAGCGAGGAGGATAACGAGCTGA	ACCTCCCTAACCTGGCAGCCGCCTACTCGTCCATCCT
19	265	TH	12	11p15.5	2143503	2143555	GGTGGAGTTCGGGCTGTGTAA	GCAGAACGGGGAGGTGAAGGCCTATGGTGCC
20	274	SGCE	11	07q21.3	94052725	94052798	TCTTCAGTCAGTTTTCTTTCTTCAGGGAT	ACCATTTACCTGCAATGTGTAAGAATGAATATGGGCAGGAGTTA
21	283	GCH1	6	14q22.2	54380457	54380521	GACTCGGGAAGAGTTCCTGACTCTCAT	TAGGAGCTGAGCTTCATTCAGTGTGTGTGCGTTGGTT
22	291	SGCE	5	07q21.3	94086069	94086133	CCAGTGAGGTTCTTGGAGACTTTCTTGGCGCAGT	GAAAAATGTGTGGCAGCCAGAGCGCCTGAACGCCATAA
23	298	TH	14	11p15.5	2142146	2142204	CACCAGGAGCTATGCCTCACGCATCCA	GCGCCCCTTCTCCGTGAAGTTCGACCCGTAC
24	309	NF1	control	17q11.2	26574636	26574712	CTTAGTAGCACAGAAATTCTCAAGTGGTTGCG	GGAAATATTGATCTGCAGGAATAAATTTCTTCTTAAAAATAAGG
25	319	SGCE	7	07q21.3	94067952	94068022	CCCTTCTGATTCTTTGAAAAGCAGAGACT	ATTACACGGATTTCCTAATTACACTGGCTGTGCCCTCGGCA
26	328	APC	control	05q22	112191530	112191603	CCTGCTGTAAGTCTTCACTTTCAGATTTTAGTT	GGGCCACAAGTGCTCTCATGCAGCCTTTCATAGAGCATAG
27	337	SGCE	1B	07q21.3	94122630	94122700	CATTTCTCGGTGACGGAAATCTCGGTT	TTCTTCGGTTTTACGGAGCAGGGAAGCGGAATCATGAAATGCT
28	346	MGP	control	12p13.1	14929956	14930014	GACTGACCTGCAGGACGAAACCA	TGAAGAGCCTGATCCTTCTTGCCATCCTGGCCGCC
29	355	SGCE	10	07q21.3	94055929	94055996	CACTCACCTGTAGTCTGCTGTTGG	GGAATCTGAGTCTGATGTGGCAAGTTCCTAGAAATGATGAACA
30	365	SGCE	6	07q21.3	94070633	94070700	TCGTTTATGTCATGGTTGGTGCAGA	TGTCCCGTTTTCTTCTTGTTTACGAGAAGTTGAAAATCCACAG
31	382	SGCE	8	07q21.3	94066896	94066975	CTAGGGAAAAGAGAAACATGCAAACACCAG	AGTAAGTGTCTTCTTTCCTGTTATTTTCTTCATCATTTTTTCCCACCTC
32	391	RB1	control	13q14.2	47937195	47937265	CGGTCTTCATGCAGAGACTGAAAACAAAT	ATTTTGCAGTATGCTTCCACCAGGGTAGGTCAAAAGTATCC
33	400	SGCE	8	07q21.3	94067023	94067096	CCTGGTTGCATTTAATAGTGACAATGTCA	GCATTTCCACATCATGCAAGATCCCATGAGCATATAGTCTTAAT
34	409	ATM	control	11q23	107626715	107626790	CGCAACTTCCGTAAGGCATCGTAACACATATGG	TGTACGTTCCCCATGTCGCTGTTGGGGTAGAAGCTGAGATAGT

Figure 4.1 HPLC dosage analysis of *SGCE* for exons 1-3. Normalized exons dosage ratios are shown for *SGCE* exons 1-3. A range of 0.75-1.25 is considered normal dosage range for the exon, whereas a ratio above or below this threshold is considered a potential duplication or deletion, respectively.


Exon 1
Normalized Ratio
1.4
1.2
1
0.8
0.6
0.4
0.2
0
MF-19
MF-20
MF-22
MF-23
MF-24
MF-25
MF-26
MF-27
Patient


Exon 2
Normalized Ratio
1.200
1.000
0.800
0.600
0.400
0.200
0.000
MF-19
MF-20
MF-22
MF-23
MF-24
MF-25
MF-26
MF-27
Patient

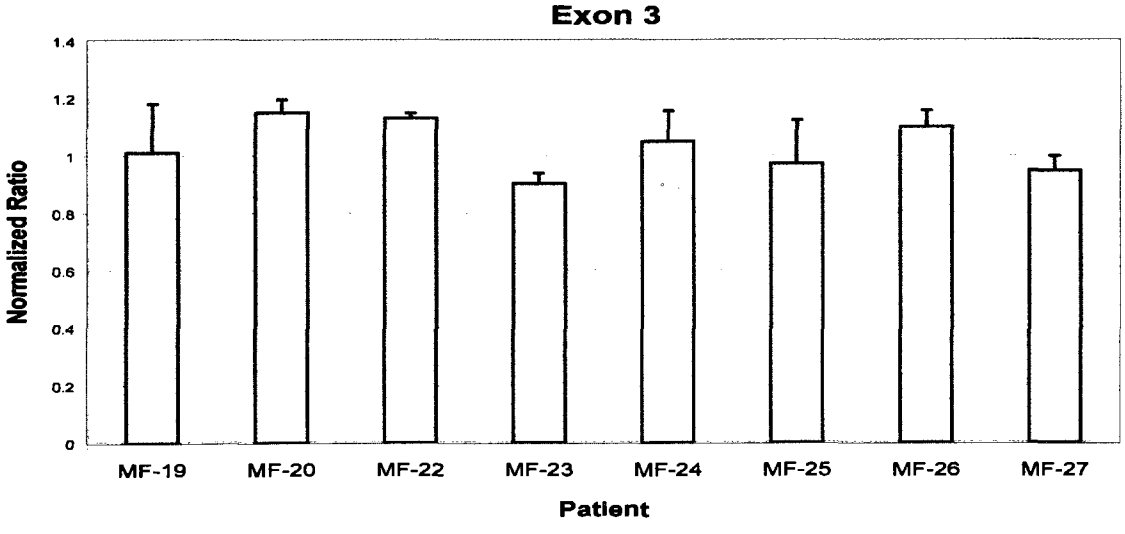
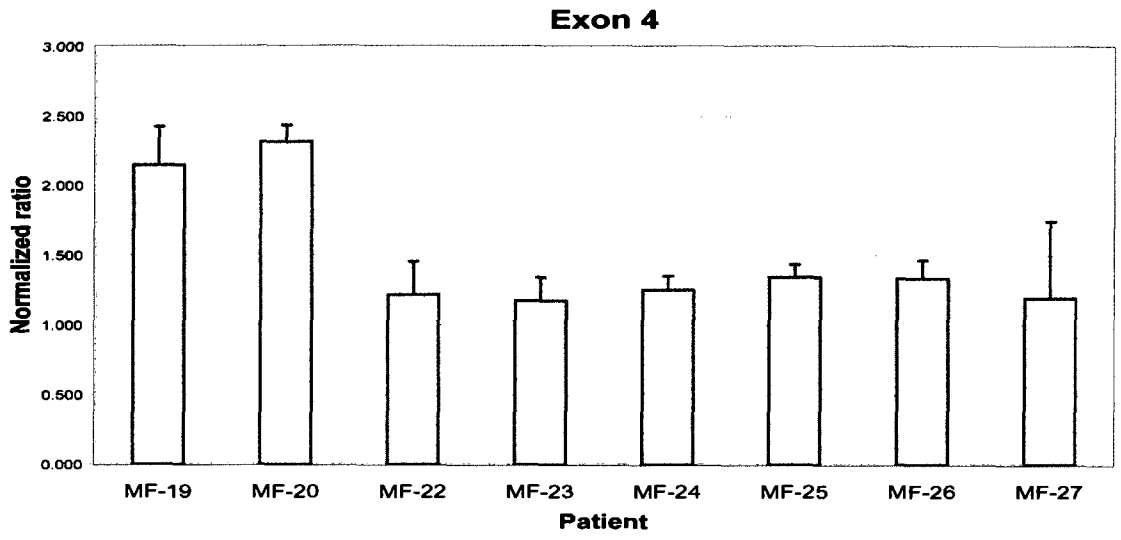
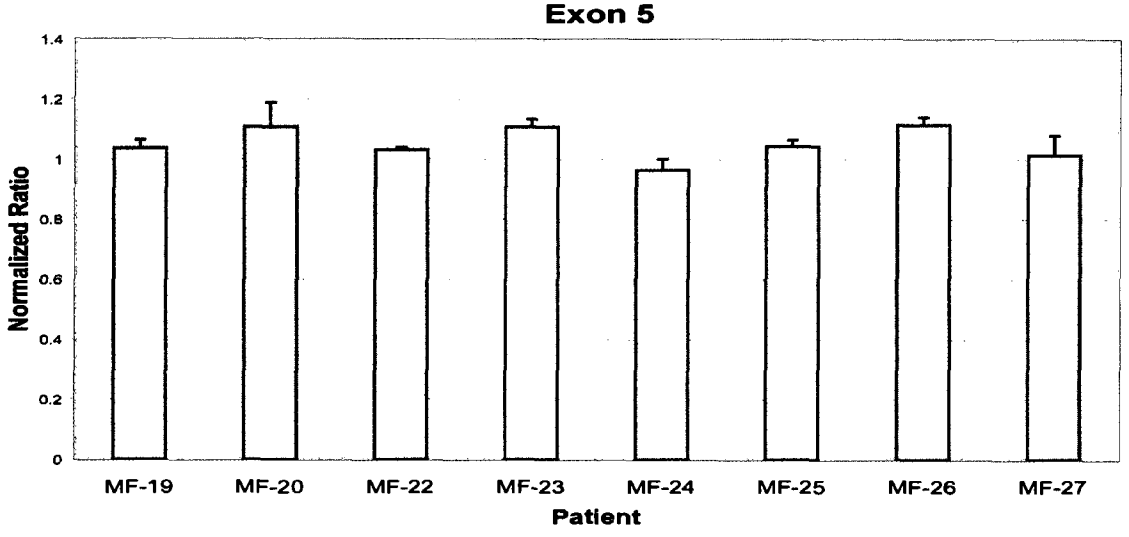

Exon 3
Normalized Ratio
1.4
1.2
1
0.8
0.6
0.4
0.2
0
MF-19
MF-20
MF-22
MF-23
MF-24
MF-25
MF-26
MF-27
Patient

Figure 4.2 HPLC dosage analysis of *SGCE* for exons 4-6. Normalized exons dosage ratios are shown for *SGCE* exons 4-6. A range of 0.75-1.25 is considered normal dosage range for the exon, whereas a ratio above or below this threshold is considered a potential duplication or deletion, respectively.


Exon 4
3.000
2.500
2.000
1.500
1.000
0.500
0.000
Normalized ratio
MF-19
MF-20
MF-22
MF-23
MF-24
MF-25
MF-26
MF-27
Patient


Exon 5
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Normalized Ratio
MF-19
MF-20
MF-22
MF-23
MF-24
MF-25
MF-26
MF-27
Patient

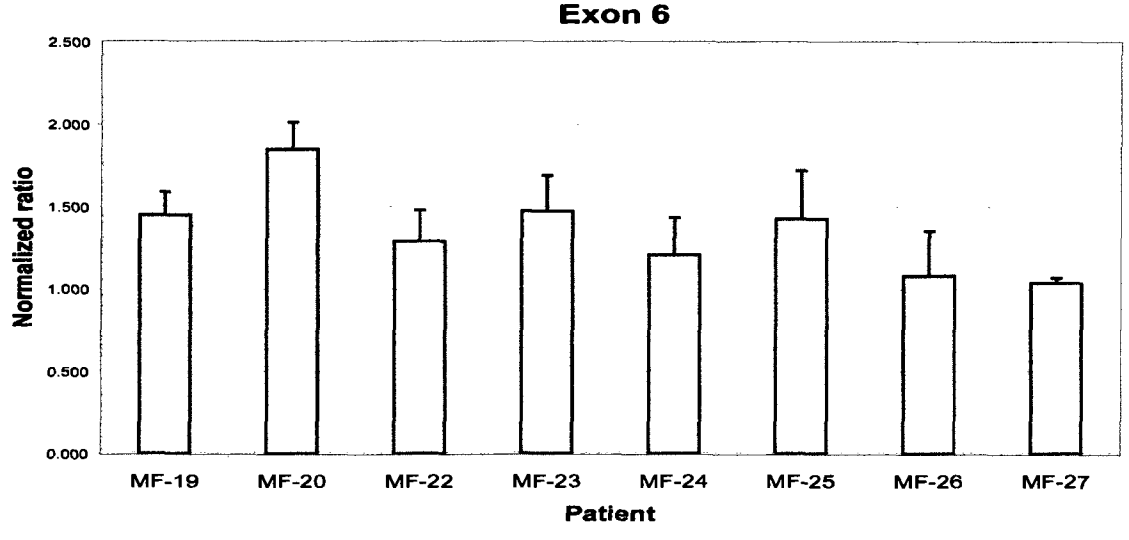
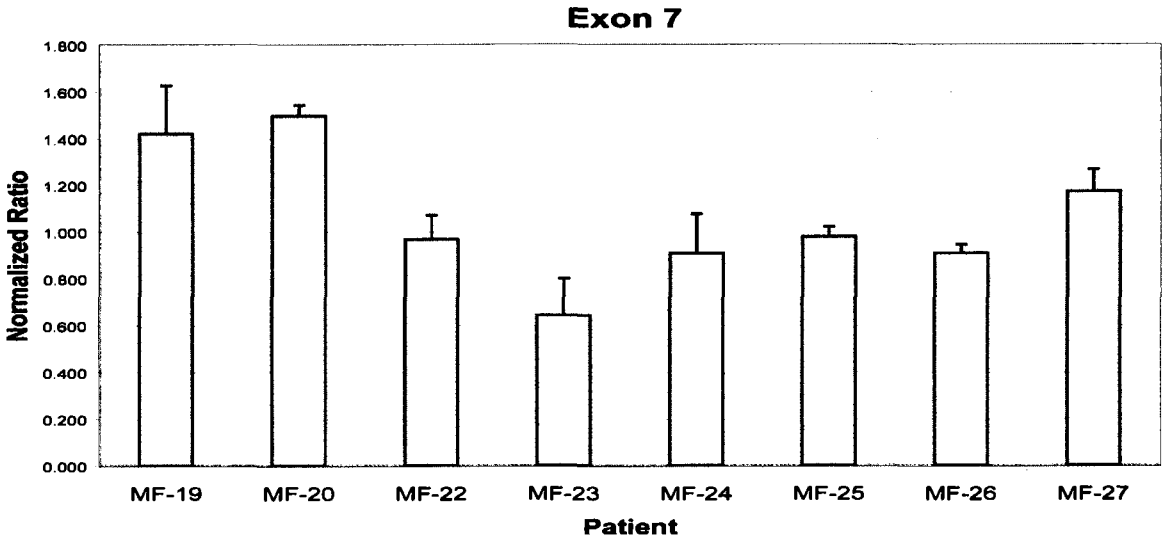
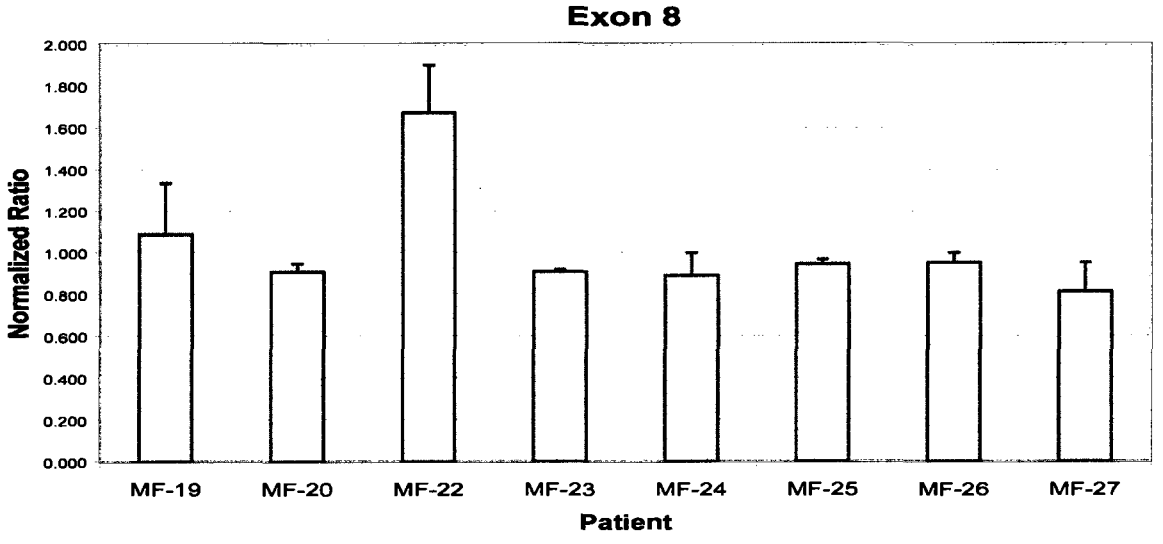

Exon 6
2.500
2.000
1.500
1.000
0.500
0.000
Normalized ratio
MF-19
MF-20
MF-22
MF-23
MF-24
MF-25
MF-26
MF-27
Patient

Figure 4.3 HPLC dosage analysis of *SGCE* for exons 7-9. Normalized exons dosage ratios are shown for *SGCE* exons 7-9. A range of 0.75-1.25 is considered normal dosage range for the exon, whereas a ratio above or below this threshold is considered a potential duplication or deletion, respectively.


Exon 7
Normalized Ratio
1.800
1.600
1.400
1.200
1.000
0.800
0.600
0.400
0.200
0.000
MF-19
MF-20
MF-22
MF-23
MF-24
MF-25
MF-26
MF-27
Patient


Exon 8
Normalized Ratio
2.000
1.800
1.600
1.400
1.200
1.000
0.800
0.600
0.400
0.200
0.000
MF-19
MF-20
MF-22
MF-23
MF-24
MF-25
MF-26
MF-27
Patient

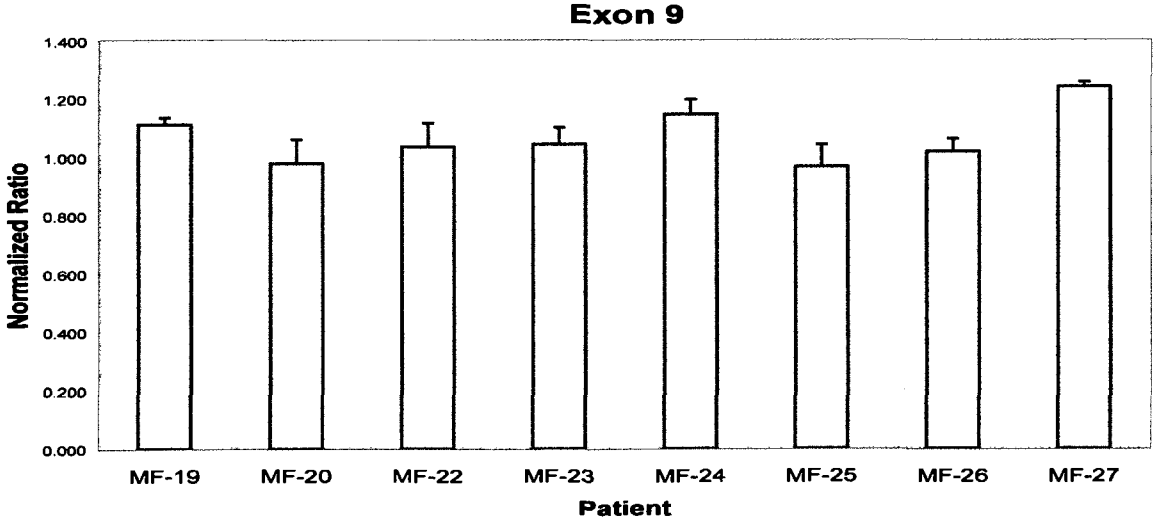
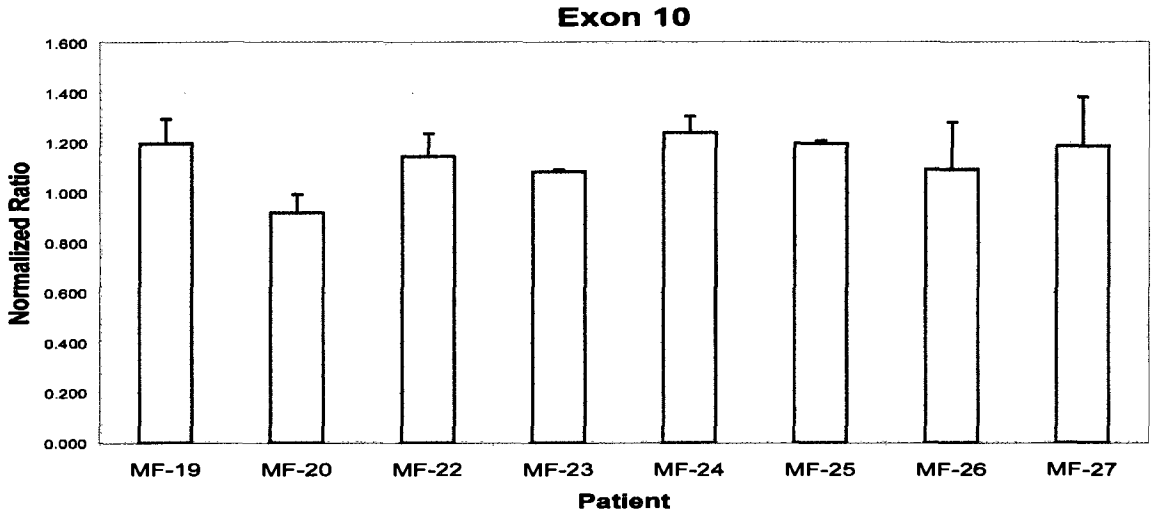
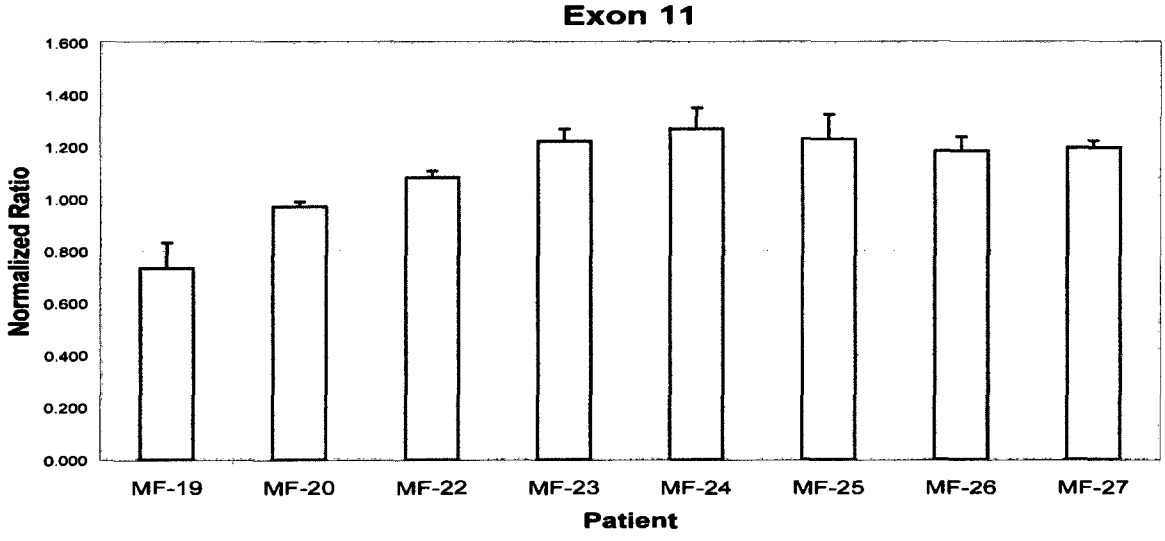

Exon 9
Normalized Ratio
1.400
1.200
1.000
0.800
0.600
0.400
0.200
0.000
MF-19
MF-20
MF-22
MF-23
MF-24
MF-25
MF-26
MF-27
Patient

Figure 4.4 HPLC dosage analysis of *SGCE* for exons 10-11b. Normalized exons dosage ratios are shown for *SGCE* exons 10, 11, and 11b. A range of 0.75-1.25 is considered normal dosage range for the exon, whereas a ratio above or below this threshold is considered a potential duplication or deletion, respectively.


Exon 10
Normalized Ratio
1.600
1.400
1.200
1.000
0.800
0.600
0.400
0.200
0.000
MF-19
MF-20
MF-22
MF-23
MF-24
MF-25
MF-26
MF-27
Patient


Exon 11
Normalized Ratio
1.600
1.400
1.200
1.000
0.800
0.600
0.400
0.200
0.000
MF-19
MF-20
MF-22
MF-23
MF-24
MF-25
MF-26
MF-27
Patient

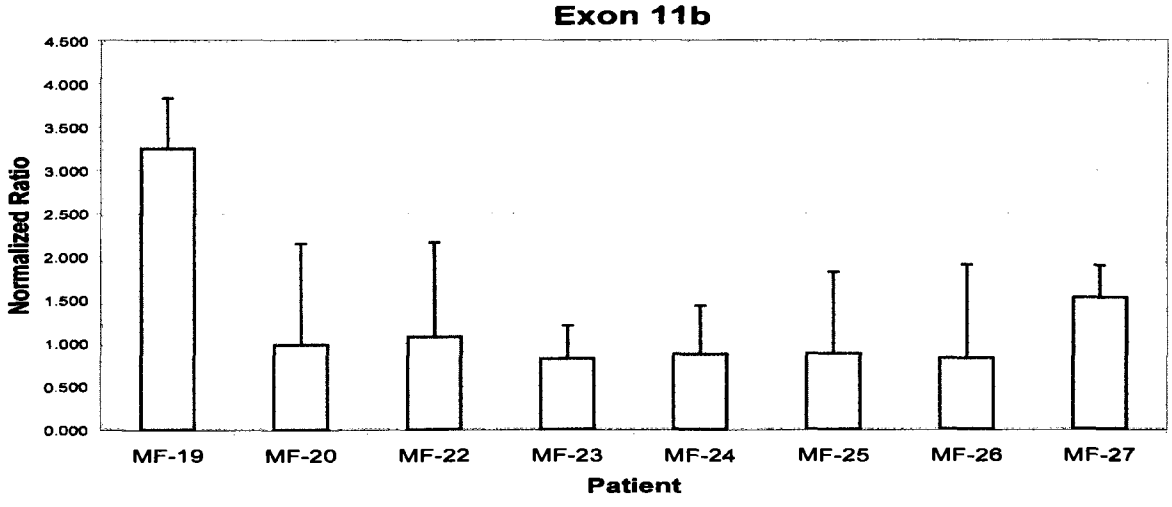

Exon 11b
Normalized Ratio
4.500
4.000
3.500
3.000
2.500
2.000
1.500
1.000
0.500
0.000
MF-19
MF-20
MF-22
MF-23
MF-24
MF-25
MF-26
MF-27
Patient

Figure 4.5 HPLC dosage analysis of *SGCE* for exon 12. Normalized exons dosage ratios are shown for *SGCE* exon 12. A range of 0.75-1.25 is considered normal dosage range for the exon, whereas a ratio above or below this threshold is considered a potential duplication or deletion, respectively.

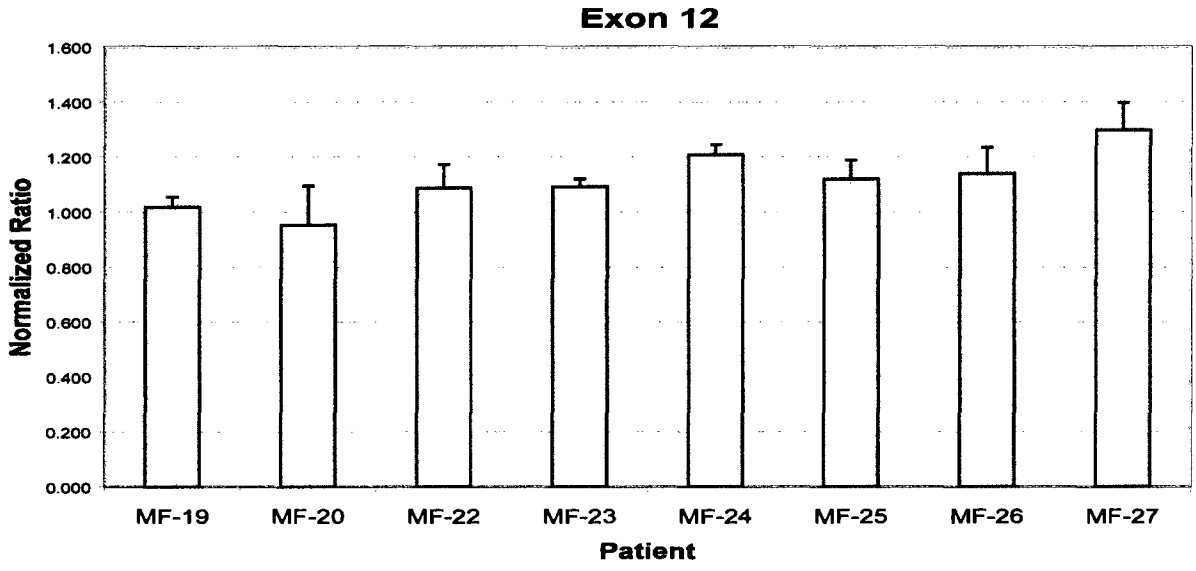
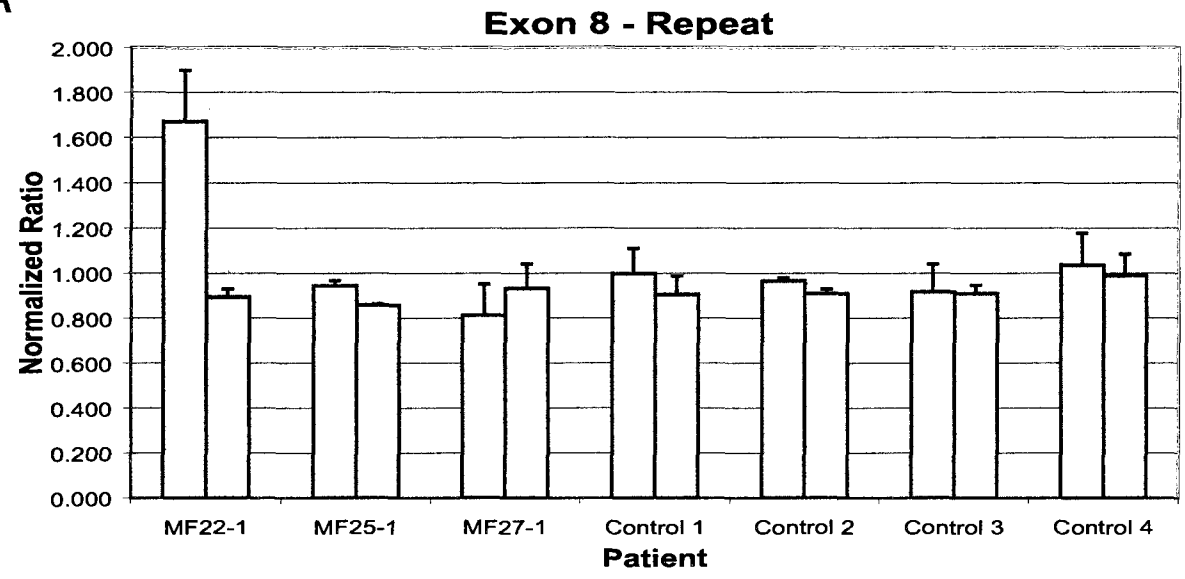

Exon 12
1.600
1.400
1.200
1.000
0.800
0.600
0.400
0.200
0.000
Normalized Ratio
MF-19
MF-20
MF-22
MF-23
MF-24
MF-25
MF-26
MF-27
Patient

Figure 4.6 HPLC dosage analysis of *SGCE* exon 8. Exon dosage analysis was repeated for exon 8 in individuals 22-1, 25-1 and 27-1. Four WT controls were used for normalization of the patient exon ratios. These normalized ratios and their standard deviations (SD) for each patient and experiment (B) is plotted in a graph to show the contradictory results for patient M22-1 (A).

A

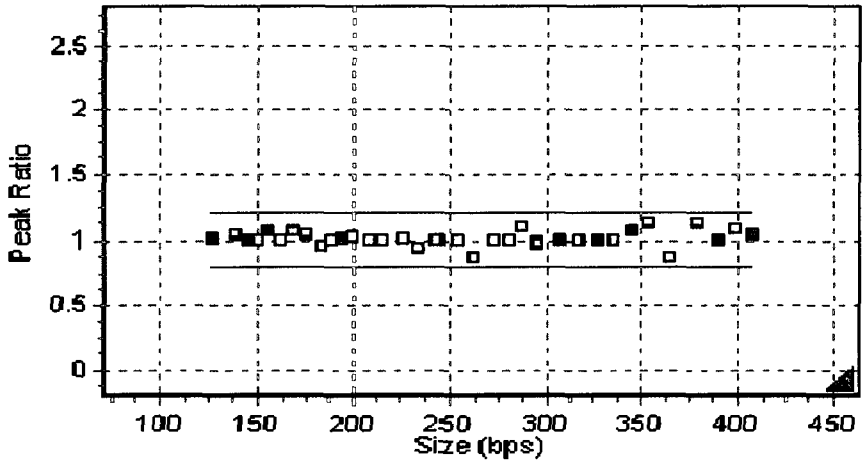


B

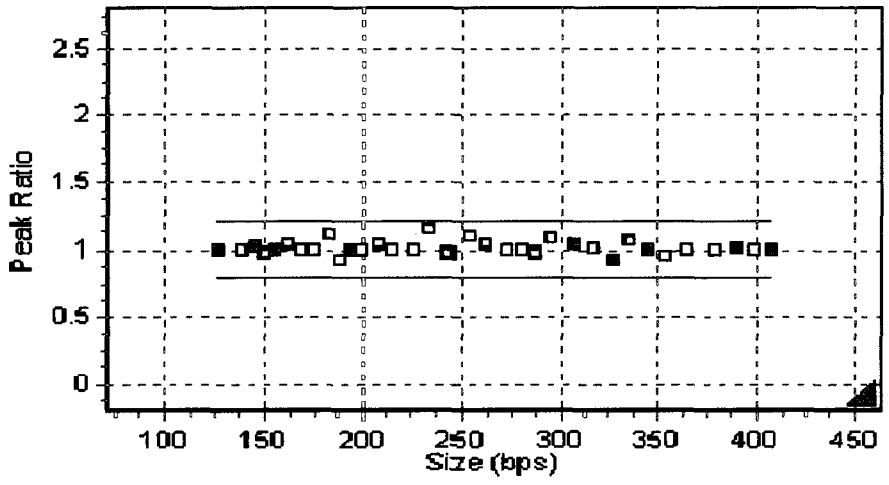
	Normalized Ratios			
Patient	1st Experiment	SD	2nd Experiment	SD
MF22-1	1.671	0.227	0.894	0.036
MF25-1	0.945	0.022	0.859	0.003
MF27-1	0.814	0.138	0.933	0.107
Control 1	0.998	0.112	0.906	0.082
Control 2	0.965	0.014	0.91	0.018
Control 3	0.917	0.124	0.91	0.037
Control 4	1.036	0.142	0.99	0.094

Figure 4.7 MLPA exon dosage analysis of *SGCE* in 3 WT controls. WT non-affected controls were used in the first MLPA analysis as well as 6 affected MD patients. All probes are shown to have peak ratios within the normal exon dosage range. Blue dots represent control probes (a total of 9), green dots represent exons of genes being tested (*SGCE*, *GCH1* and *TH*), and red dots indicate a probe outside the normal exon dosage threshold.

WT - 1


Peak Ratio
2.5
2
1.5
1
0.5
0
100
150
200
250
300
350
400
450
Size (bps)

WT - 2


Peak Ratio
2.5
2
1.5
1
0.5
0
100
150
200
250
300
350
400
450
Size (bps)

WT - 3

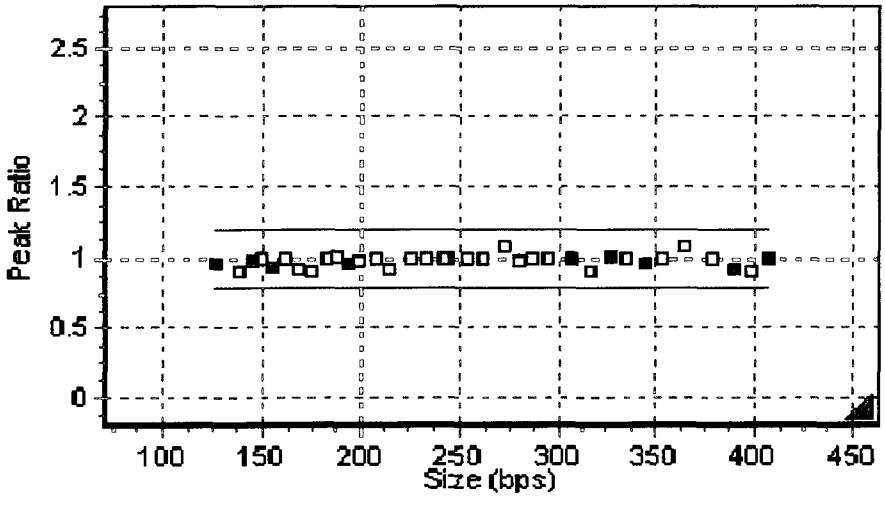
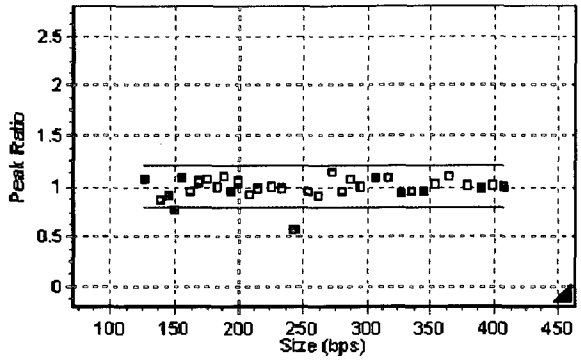
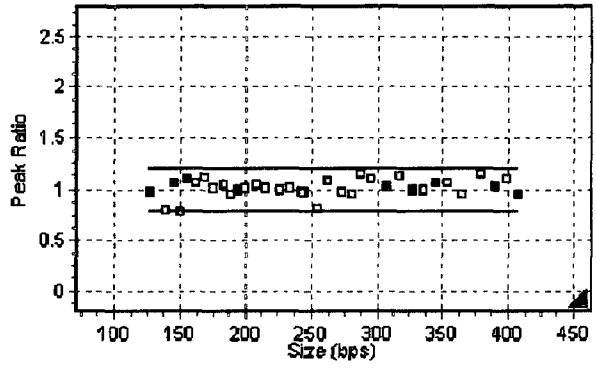

Peak Ratio
2.5
2
1.5
1
0.5
0
100
150
200
250
300
350
400
450
Size (bps)

Figure 4.8 MLPA exon dosage analysis of *SGCE* in 6 Myoclonus Dystonia families. Patients M2-1, M28-1, M29-1, M30-1, M31-1 and M31-2 were analyzed for exon dosage in this first MLPA analysis. Individual M2-1 was previously found to have a large *SGCE* deletion of exons 2-3, shown by the two probes that are below the normal dosage threshold. Blue dots represent control probes (a total of 9), green dots represent exons of genes being tested (*SGCE*, *GCH1* and *TH*), and red dots indicate a probe outside the normal exon dosage threshold.

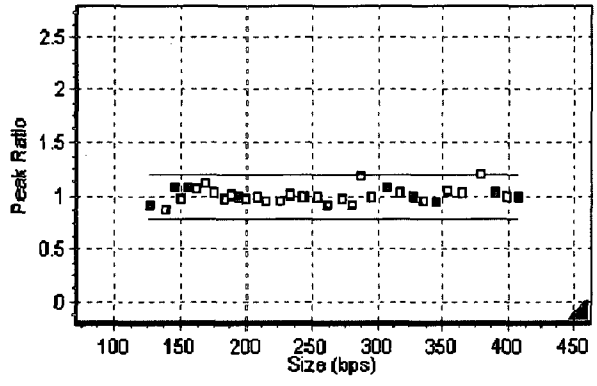
M2-1



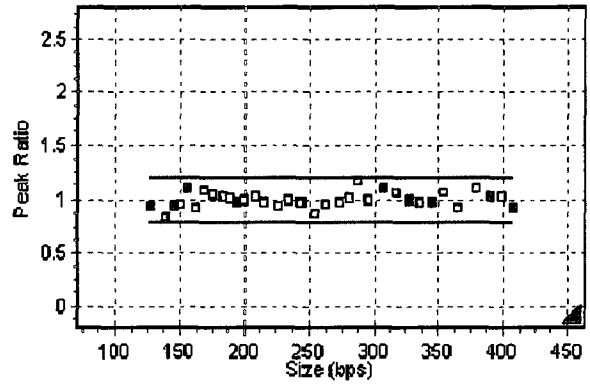
M28-1



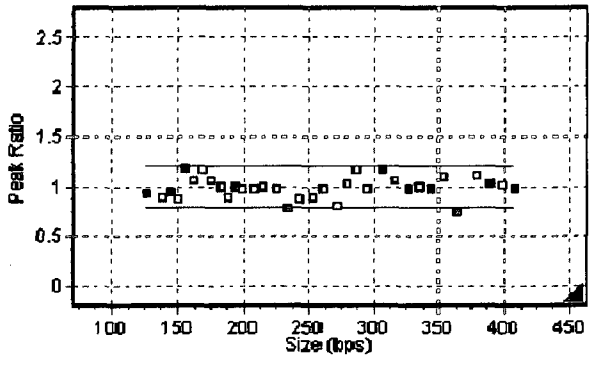
M29-1



M30-1



M31-1



M31-2

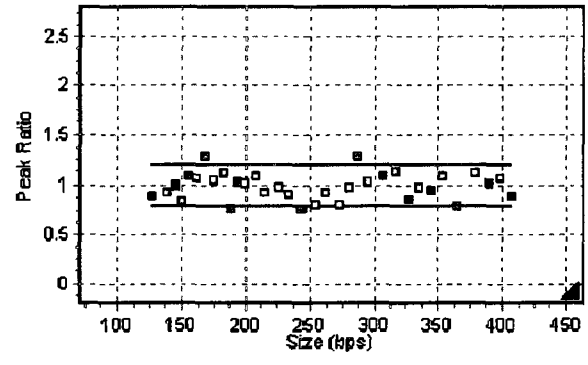
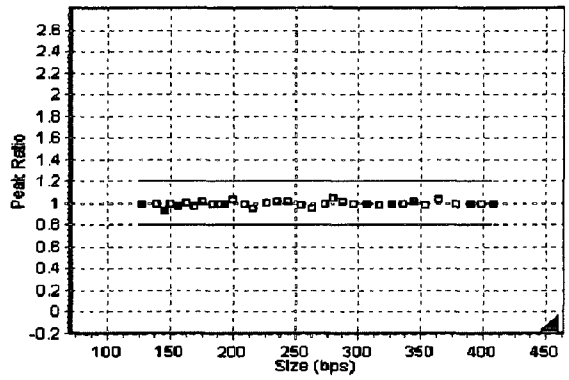
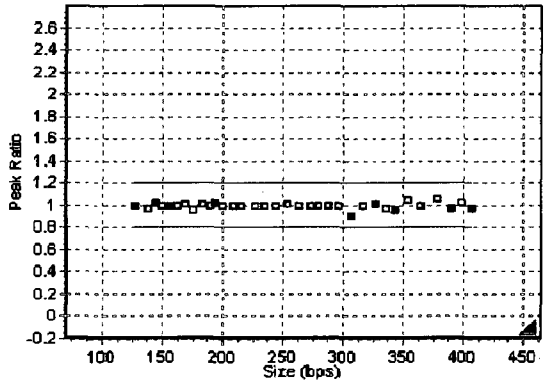


Figure 4.9 MLPA exon dosage analysis of *SGCE* in three WT controls and three Myoclonus Dystonia patients. WT-3 was not used for normalization of patient probes since one probe was out of range; the remaining two WT controls were used. Three individuals in family 6 (M6-1, M6-3, M6-4) were previously found to have a large deletion spanning exons 2-5 in *SGCE* and were found to be below the normal dosage threshold. Blue dots represent control probes (a total of 9), green dots represent exons of genes being tested (*SGCE*, *GCH1* and *TH*), and red dots indicate a probe outside the normal exon dosage threshold.

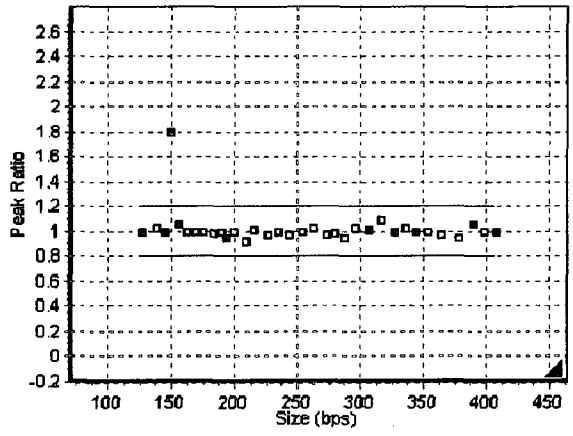
WT - 1



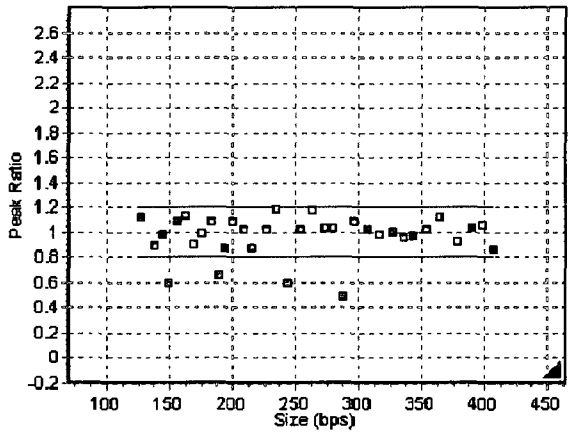
WT - 2



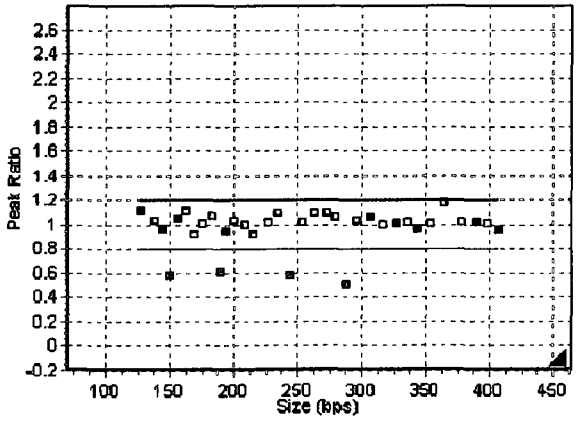
WT - 3



M6-1



M6-3



M6-4

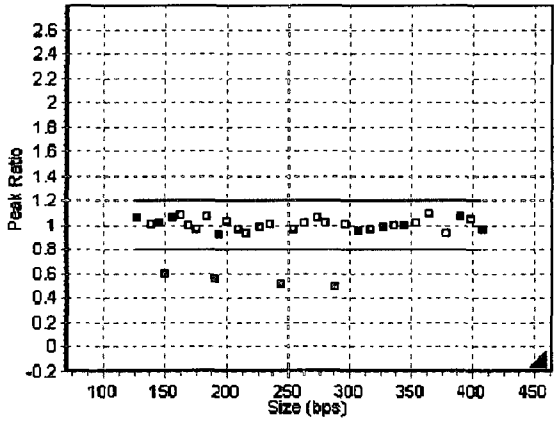
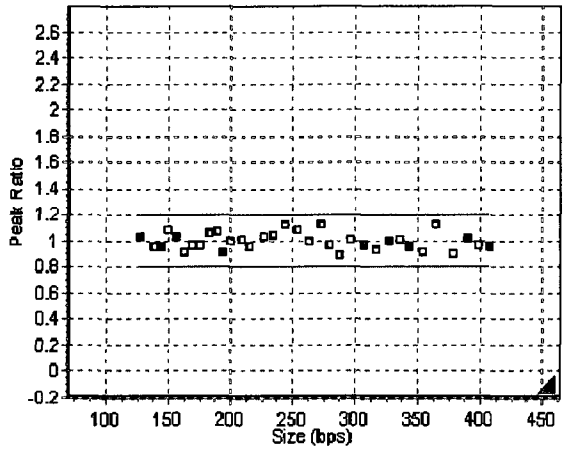
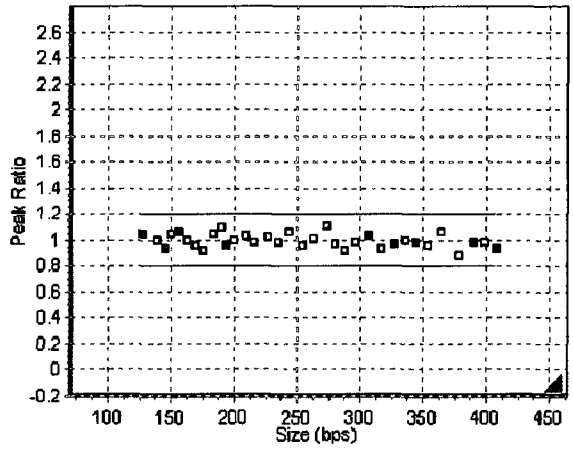


Figure 4.10 MLPA exon dosage analysis of *SGCE* in 5 Myoclonus Dystonia patients. Results from patients M19-1 through to M23-1 are shown. Blue dots represent control probes (a total of 9), green dots represent exons of genes being tested (*SGCE*, *GCH1* and *TH*), and red dots indicate a probe outside the normal exon dosage threshold.

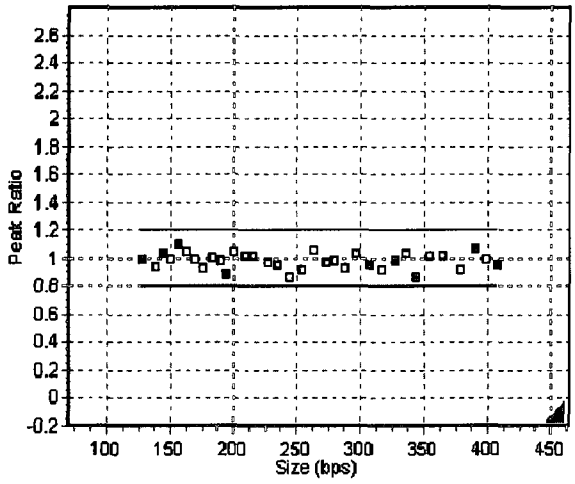
M19-1



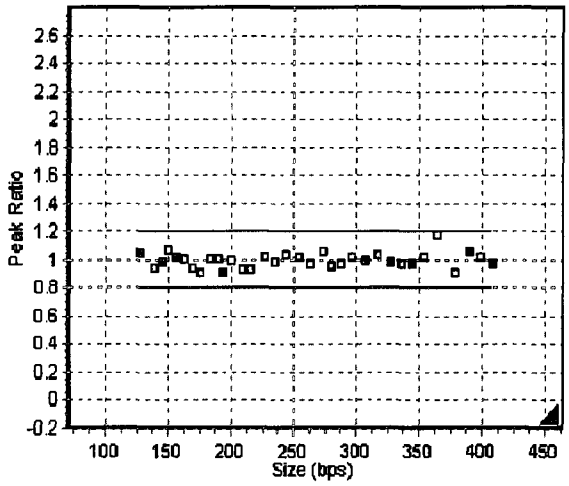
M20-1



M21-1



M22-1



M23-1

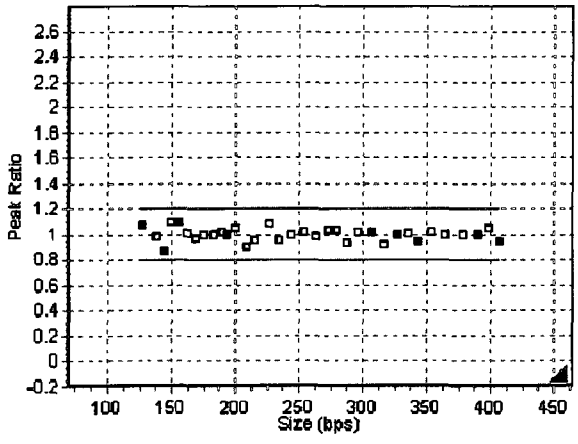
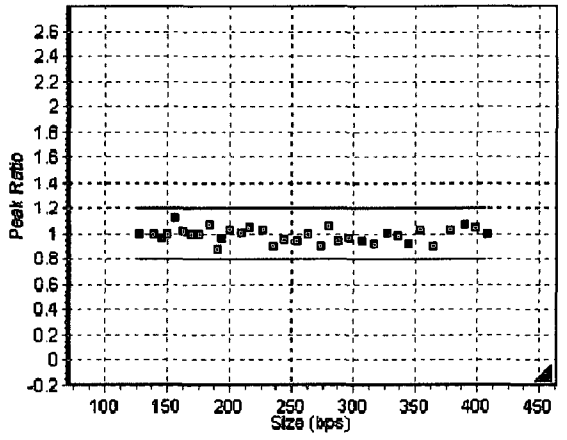
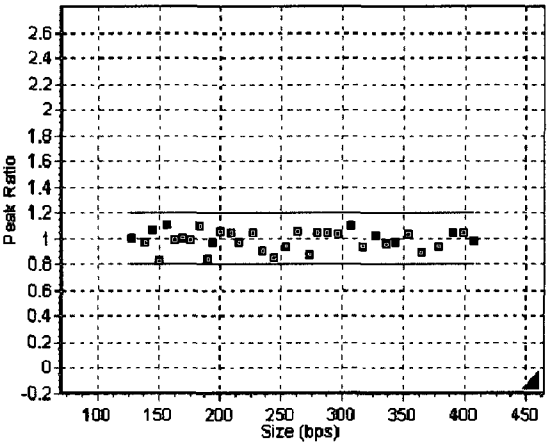


Figure 4.11 MLPA Exon Dosage Analysis of *SGCE*. Results from patients M25-1 through to the repeat of M28-1 are shown. Analysis for individual M24-1 failed and was subsequently repeated. Blue dots represent control probes (a total of 9), green dots represent exons of genes being tested (*SGCE*, *GCH1* and *TH*), and red dots indicate a probe outside the normal exon dosage threshold.

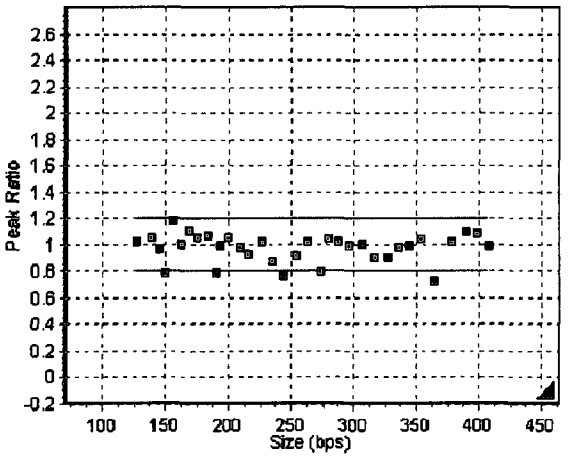
M25-1



M26-1



M27-1



M28-1 (repeat)

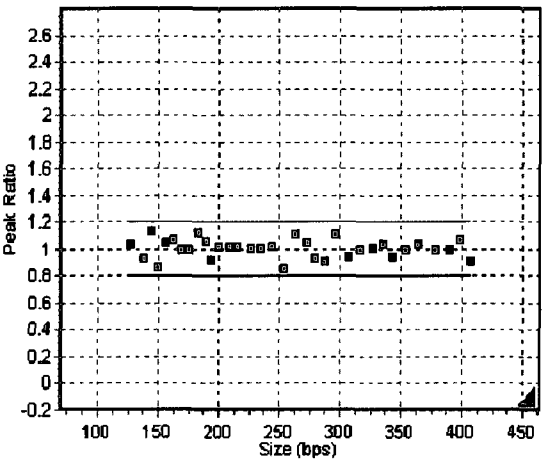
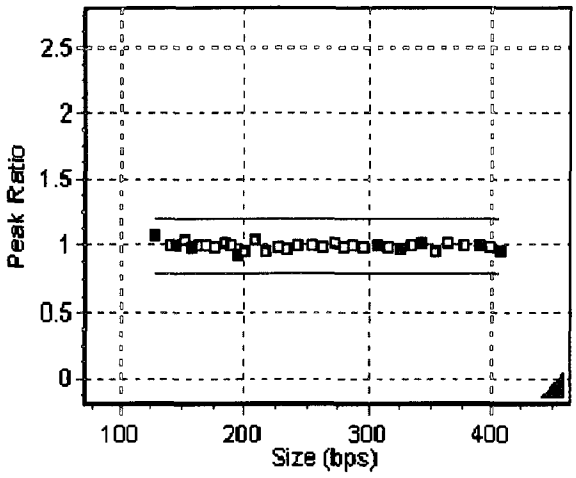
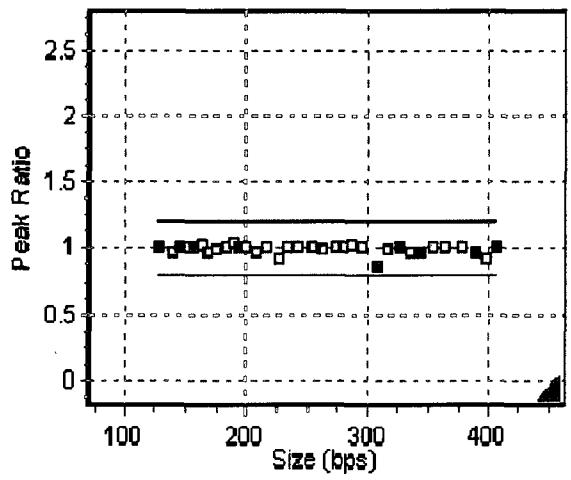


Figure 4.12 MLPA exon dosage analysis of *SGCE* in three WT controls and three patients. Results for M1-39, M24-1 (repeat) and M27-1 (repeat) are shown. Patient M1-39 is from family 1 and was not previously tested for genomic rearrangements via dosage analysis. Blue dots represent control probes (a total of 9), green dots represent exons of genes being tested (*SGCE*, *GCH1* and *TH*), and red dots indicate a probe outside the normal exon dosage threshold.

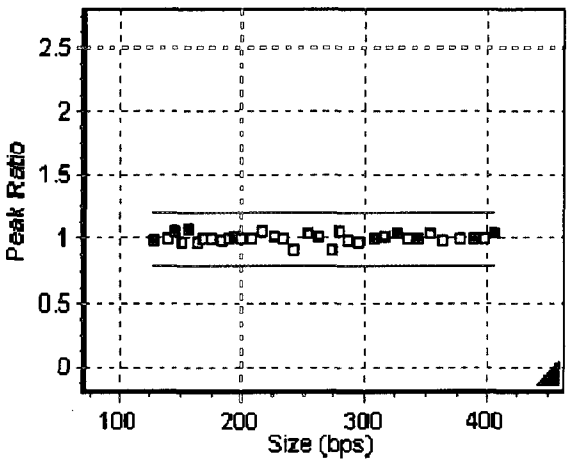
WT - 1



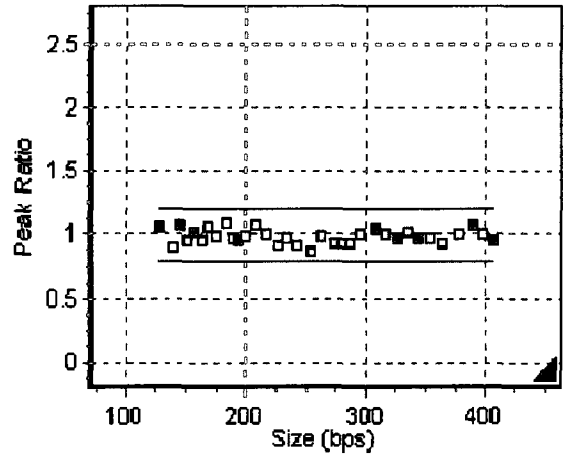
WT - 2



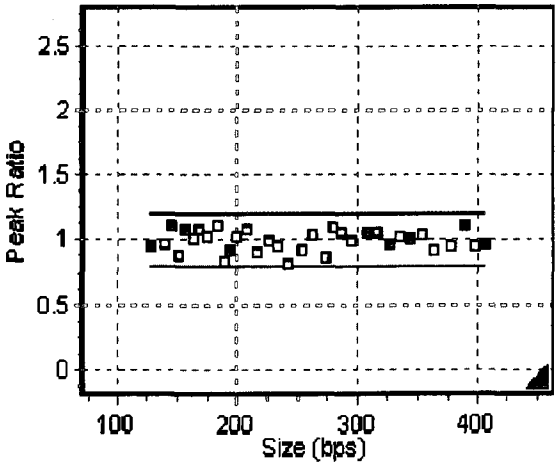
WT - 3



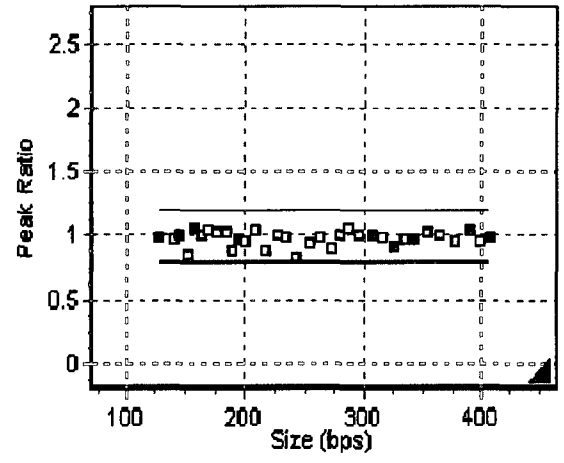
M1-39



M24-1 (repeat)



M27-1 (repeat)



Chapter 5.

Confirmation of chr18p11 as the single disease locus for Myoclonus Dystonia in family 1

Chapter 5 Confirmation of chr18p11 as the single disease locus for Myoclonus Dystonia in family 1

5.1 Introduction

Screening of four families affected with Myoclonus Dystonia by direct sequencing has resulted in the identification of two mutations within exon 3 of *SGCE*. In this cohort 67% of the affected individuals (for which we obtained DNA) were found to have the *SGCE* mutations. This is a fairly high proportion of *SGCE* mutations compared to the 30-40% we found by studying the previous 27 recruited families. The incidence of *SGCE* mutations in MD families varies greatly on the particular cohort being studied. Since many MD families do not have *SGCE* mutations as demonstrated by both us and others, we can conclude that MD is genetically heterogeneous. Other genes causing and accounting for its clinical heterogeneity remain to be elucidated.

Family 1 was the first family recruited by our group and is an excellent pedigree for linkage analysis due its large number of individuals, 15 of which are symptomatic (Figure 5.1). There were no *SGCE* mutations found using both direct sequencing and HPLC. A 25cM whole genome linkage analysis was first performed by genotyping microsatellite markers roughly every 25cM along the genome. This was performed for all family members that had given consent to participate in these studies. Genotypes for each marker were then entered into a linkage program and essentially cross referenced to the affected status of each individual, allowing for a two-point linkage analysis approach to elucidating a haplotype associated with the MD phenotype. LOD scores, indicating linkage to a region, were above 3.0 for a locus mapped to chr18p11. As well, there were two individuals, #38 and #48, who were affected and had recombination events which delineated this region to 3.18Mb. All

affected family members shared alleles at 5 loci within the 3.18Mb region. These shared alleles are referred to as the disease haplotype. There was one individual (#67) who was designated as affected but did not share the disease haplotype. To test the possibility that we did not identify the true locus, our group increased the resolution of our linkage analysis to 10cM. Like the original analysis, chr18p11 was again found to be the disease "critical" region. Upon retrospection of the markers used in the 25 and 10cM linkage analysis, it was found that gaps existed in which some genomic areas (greater than 20cM between markers) were not covered in the analysis. Therefore, we genotyped additional markers to fill these gaps to ensure linkage to the disease was not occurring elsewhere in the genome. We also wanted to confirm that individual #67 did not have the disease haplotype, therefore he and his first degree relatives were genotyped for the chr18p11 critical region markers. This group was also re-bled to ensure that we did not mistakenly mix or mislabel blood samples.

5.2 Materials and Methods

5.2.1 Genotyping and Linkage Analysis

Family 1 was genotyped for 39 microsatellite markers (see Figure 5.1 for a complete list of markers). The di-, tri-, and tetra- nucleotide repeat sequences of these markers were amplified by end-labeled primers listed in the CHLC Human Screening Set (Weber Version 8) designed by Research Genetics Inc.. After amplification, an IR2 Stop Solution containing IRDye 800RS (LI-COR Biosciences) was added to the sample. These products were then denatured for 10 min at 95°C and run on an 8% polyacrylamide gel in the LI-COR DNA sequencer model 4000. Fragment sizes were separated by gel electrophoresis and the PCR products were analyzed using Saga Generation 2 software.

Allele calls were entered into the LINKAGE version 5.1 program using the MS-DOS (Microsoft) operating system. An autosomal dominant mode of inheritance was specified, penetrance was set to 80%, gene frequency (disease frequency) was set to 0.000001 (1/100,000), and theta values (recombination faction; distance from the marker) were set to start at 0 and increased in increments of 0.05. A LOD score above 3.0 was considered linkage to the disease.

5.3 Results

Of the 39 microsatellite markers that were genotyped (Figures 5.2 – 5.7), none showed LOD scores above 3.0 (Table 5.1). However, there were two markers that had LOD scores above 1.0, D4S408 with a score of 1.19 and D17S2196 with 1.89. These markers were further examined by constructing a haplotype surrounding the marker using the previous allele data from the markers used in the 10cM and 25cM analysis. If a haplotype was shared between the affected individuals, then this would be a potential disease locus. The haplotype surrounding D17S2196 is not shared between the symptomatic individuals in family 1, and is therefore ruled out as a possible disease locus (Figure 5.8A). The D4S408 surrounding haplotype has one marker (D4S2417) on its flank that shares a common allele between the symptomatic individuals (Figure 5.8B). However, because this allele is also shared between most of the unaffected individuals within the family (30/46 alleles, 65%), it is likely uninformative. As well, the "1" allele for D4S408 is also common in the asymptomatic family members (29/44 alleles, 66%) making this uninformative as well. To investigate whether or not there is true linkage to this locus, additional informative markers in this locus would need to be genotyped.

These genotyped markers showed an unusually high number of allele calls that were incompatible between individual #67 and his father. In total, there were 16/39 incompatible alleles for #67, and as well, 10/39 were incompatible between #69 and her father. These siblings both had compatible allele calls with their mother. This incompatibility was not observed in any other individuals within the pedigree.

Individual #67 and his first degree relatives were re-genotyped for six chr18p11 critical region microsatellite markers, allowing us to confirm that the original haplotypes for these individuals were correct. We observed the same haplotypes as previously reported, and confirmed that individual #67 does not have the disease haplotype common among all other affected individuals in family 1 (Figure 5.9A). These first degree relatives were also re-bled and this new batch of genomic DNA was genotyped once more, again revealing the same results and thereby ruling out the possibility of a tube or labeling mix-up.

This group within family 1 was also sequenced for the GAG deletion in exon 5 of the Torsin A gene that is known to cause Primary Torsion Dystonia. This was screened because there are two individuals within this group (#35 and #36) that have unexplained severe mental retardation, which is a common phenotype seen in patients with a homozygous GAG deletion in this gene[12]. As well, other groups studying Myoclonus Dystonia often screen patients for this mutation, although it almost always results in WT sequence. Likewise, our sequencing analysis showed no Torsin A mutations in this group (data not shown), thereby ruling out the involvement of Primary Torsion Dystonia within family 1.

5.4 Discussion

A locus for Myoclonus Dystonia has been mapped to a 3.18Mb region on chr18p11 in the large family 1. Because individual #67 does not share the disease haplotype for this region, we thought it necessary to thoroughly complete the 10cM linkage analysis previously performed. This enabled us to ensure that chr18p11 was the only locus segregating with this disease. There were two microsatellite markers that had generated LOC scores above 1.0. One (D17S2196) was ruled out as a possible disease loci by generating haplotypes of the affected individuals using previously analyzed markers directly up and downstream of the marker. The marker D4S408 has a flanking marker (D4S2417) that shares a common allele with the symptomatic individuals in the family, however both of these alleles are uninformative. To properly investigate if there is true linkage to this region, additional informative markers are required to be genotyped.

Individual #67 was confirmed to not have the ch18p11 disease haplotype. It is possible that individual #67 is a phenocopy for the disease. This would occur if there is another locus other than chr18p11 and *SGCE* that causes the same MD phenotype in this patient. Another possibility for this anomaly would be if this patient has been misdiagnosed. However, this individual has been accessed by two neurologists, both of whom agree on the diagnosis. The only counter indication of this patient is the presence of tics, which is not commonly seen with this disease, however all other symptoms point to MD. To rule out *SGCE* as a causative gene, individual #67 was screened for *SGCE* rearrangements by MLPA (Figure 4.1, see M1-39) and direct sequencing (data not shown), both demonstrating WT sequence for this gene.

Strangely, there were obvious incompatibilities seen between the alleles of individual #67, as well as his sister #69, and their father. In comparison, the allele calls match up with their mother's alleles. Questionable paternity was initially suspected, however we first ruled out the possibility of a tube or labeling error by re-genotyping six chr18p11 markers both before and after re-bleeding the 9 individuals within this group in family 1. We then sought out a potential alternate father within the pedigree by haplotype comparison; this resulted in no consistent allele matches to the males in the family. It is possible that there are more individuals in the pedigree that have not been identified for our studies, however we are limited to the information we are provided by the family. Indeed, if the pedigree is incorrect and the father is unknown, it is highly coincidental that individual #67 would have the MD phenotype but not be related to the family. Since these markers clearly indicate that individual #65 is not the father of this individual (#67) and likely his sibling (#69) as well, we are forced to conclude that they are not related to this family. Since the mother is not affiliated with the family by blood, her son #67 and likely her daughter #69 can be removed from the pedigree. The markers for the daughter #68 match both the mother and father, therefore this individual can remain in the pedigree. In conclusion, family 1 demonstrates linkage to the DYT15 locus of chr18p11 and members #67 and #69 are deemed to be unrelated to this family.

Table 5.1 Microsatellite marker LOD scores for the genome-wide 10cM completion. A total of 39 markers were genotyped to fill in gaps that were found in the 10cM analysis. A gap was considered a distance of 20cM or more between two neighboring markers. Marker names are listed along with their corresponsing genetic distance in both Marshfield and DeCode. LOD scored are given at theta (θ) intervals of 0.05, beginning at 0.00 and ending at 0.4. Theta represents an increase in genetic distance (in cM) both up and downstream from the point of the marker (recombination interval), where 0.05 = 5cM and 0.1 = 10cM, etc. A LOD score above 3.0 is considered linkage to the disease.

Marker	cM Marshfield	cM DeCode	LOD Score at θ= 0.00	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4
Chr 1											
D1S1653	164	152	-8.64	-2.48	-1.45	-0.89	-0.53	-0.30	-0.16	-0.08	-0.03
D1S1677	175	162	-19.05	-2.00	-0.83	-0.26	0.06	0.23	0.31	0.31	0.26
D1S1589	192	176	-19.69	-4.06	-2.44	-1.55	-0.97	-0.58	-0.31	-0.13	-0.03
D1S1678	218	203	-17.05	-2.84	-1.66	-1.02	-0.60	-0.32	-0.13	-0.01	0.04
D1S547	267	259	-10.04	-0.48	0.12	0.39	0.50	0.52	0.49	0.41	0.30
Chr 2											
D2S1788	56	NK	-12.19	-2.80	-1.74	-1.15	-0.76	-0.49	-0.29	-0.16	-0.08
D2S441	87	91	-18.73	-4.61	-2.91	-1.97	-1.37	-0.95	-0.64	-0.42	-0.24
D2S410	125	127	-18.92	-4.00	-2.64	-1.83	-1.27	-0.87	-0.56	-0.34	-0.17
D2S125	260	258	-24.47	-0.64	0.24	0.60	0.73	0.74	0.66	0.52	0.35
Chr3											
D3S2409	71	71	-19.11	-5.94	-3.76	-2.59	-1.83	-1.30	-0.92	-0.63	-0.40
D3S2406	102	NK	-10.04	-0.93	-0.38	-0.10	0.07	0.16	0.20	0.20	0.17
D3S3045	124	117	-26.24	-4.94	-3.04	-1.98	-1.29	-0.81	-0.48	-0.26	-0.11
Chr4											
D4S2632	NK	54	-11.91	-2.29	-1.37	-0.88	-0.58	-0.38	-0.26	-0.17	-0.11
D4S3243	NK	88	-13.83	-2.16	-1.21	-0.67	-0.33	-0.12	0.01	0.07	0.09
D4S408	195	189	1.19	1.10	1.01	0.89	0.77	0.65	0.51	0.38	0.25
Chr5											
D5S817	23	30	-11.42	-4.16	-2.65	-1.81	-1.26	-0.88	-0.60	-0.40	-0.24
Chr6											
D6S1031	89	89	-5.52	-1.31	-0.86	-0.59	-0.40	-0.26	-0.16	-0.09	-0.04
D6S1021	112	107	-21.67	-3.67	-1.92	-0.99	-0.41	-0.05	0.16	0.26	0.26
D6S1040	129	NK	-22.84	-2.61	-1.47	-0.86	-0.49	-0.25	-0.10	-0.01	-0.03
Chr7											
D7S531	18	23	-21.26	-2.91	-1.47	-0.69	-0.22	0.08	0.24	0.30	0.27
D7S2195	162	NK	-10.16	0.42	0.83	0.96	0.96	0.89	0.77	0.61	0.42
D7S559	182	183	-21.67	-1.16	-0.23	0.19	0.37	0.43	0.40	0.33	0.23
Chr8											
D8S1130	23	18	-10.80	-5.10	-3.47	-2.51	-1.85	-1.36	-0.98	-0.67	-0.42
D8S2324	94	NK	-13.77	-2.66	-1.67	-1.11	-0.75	-0.49	-0.32	-0.20	-0.11
D8S1179	NK	129	-18.72	-4.20	-2.57	-1.63	-1.01	-0.58	-0.28	-0.08	0.02
Chr10											
GATA121A08	88	NK	-24.54	-4.46	-2.73	-1.79	-1.18	-0.77	-0.48	-0.28	-0.15
Chr11											
D11S1986	105	NK	-28.96	-6.08	-3.80	-2.55	-1.73	-1.17	-0.77	-0.49	-0.29
Chr12											
D12S1301	56	58	-19.48	-3.59	-2.27	-1.49	-0.96	-0.58	-0.31	-0.13	-0.02
D12S1064	95	NK	-23.29	-4.88	-2.99	-1.92	-1.23	-0.75	-0.42	-0.19	-0.06
D12S392	166	168	-13.23	-2.93	-1.92	-1.30	-0.87	-0.56	-0.33	-0.17	-0.07
Chr14											
D14S588	66	68	-18.89	-3.98	-2.57	-1.78	-1.25	-0.87	-0.59	-0.38	-0.22
D14S53	86	76	-11.82	-2.07	-1.23	-0.74	-0.44	-0.24	-0.12	-0.05	-0.01
Chr15											
ACTC	31	NK	-31.31	-7.69	-4.88	-3.30	-2.25	-1.50	-0.96	-0.56	-0.28
Chr16											
D16S621	130	NK	-16.23	-4.21	-2.53	-1.63	-1.05	-0.67	-0.40	-0.23	-0.11
Chr17											
D17S2196	44	47	-3.17	1.89	1.87	1.71	1.49	1.22	0.93	0.63	0.36
D17S784	117	129	-24.46	-3.02	-1.64	-0.92	-0.48	-0.20	-0.03	0.05	0.08
Chr20											
D20S478	54	61	-16.50	-1.76	-0.71	-0.19	0.10	0.25	0.30	0.30	0.24
Chr21											
D21S1440	37	45	-18.40	-3.68	-2.26	-1.48	-0.98	-0.63	-0.39	-0.22	-0.10
Chr22											
D22S685	32	39	-18.93	-2.56	-1.43	-0.81	-0.44	-0.21	-0.08	-0.01	0.01

Figure 5.1 Pedigree and chr18p11 haplotypes of Family 1. All 51 known members of family 1 are illustrated in this pedigree. Of these members, 15 are symptomatic, and 7 have the chr18p11 disease haplotype and are non-symptomatic, indicating reduced penetrance. Haplotypes are listed for the chr18p11 microsatellite markers. Those markers in blue fall within the 3.18Mb critical region. The disease haplotype is identified by a shaded rectangle. Circles, female; square, male; slash, individual deceased; shaded, affected; +, blood collected for individual; arrows, recombination event within the region allow us to narrow the disease haplotype.

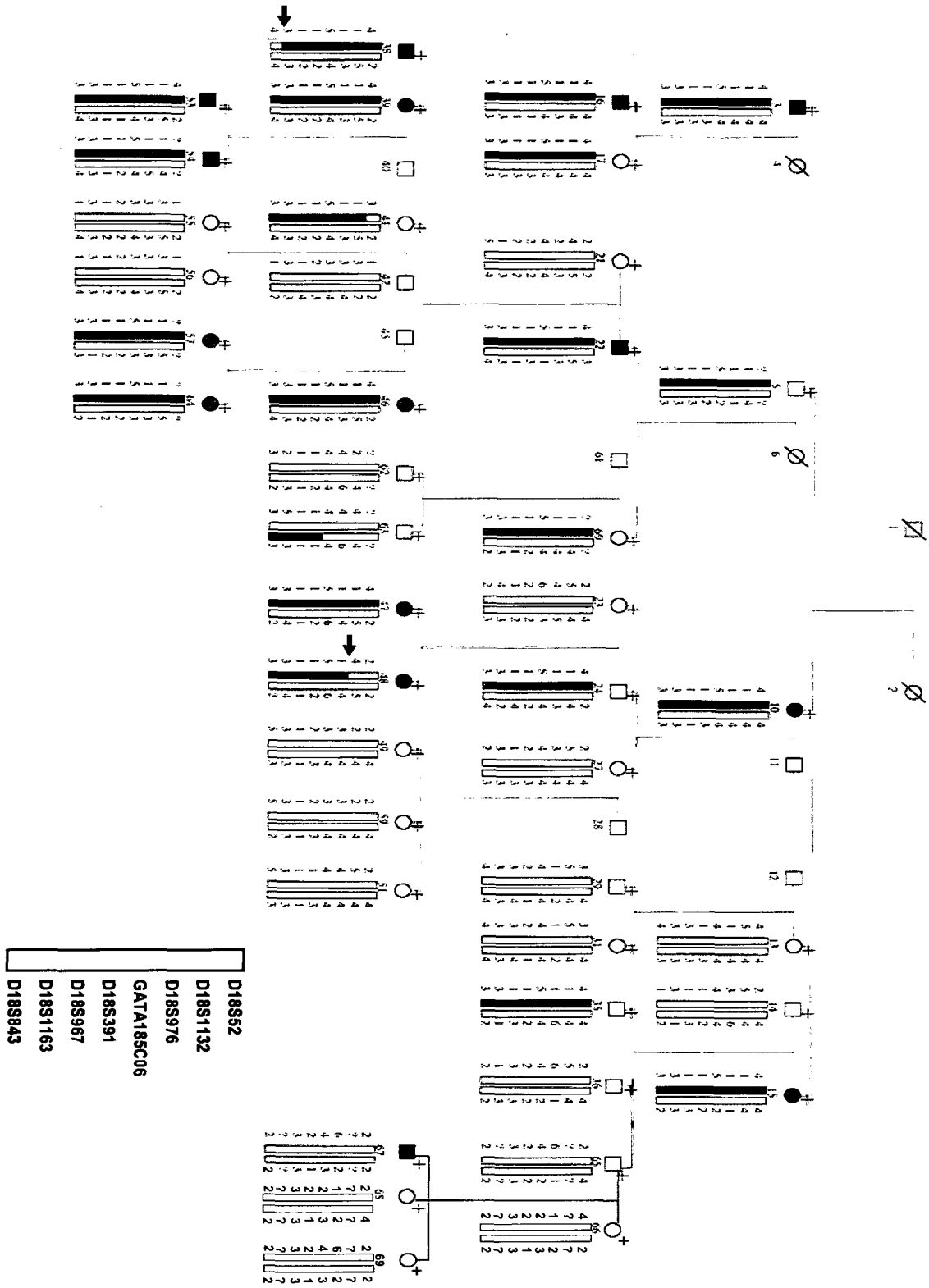
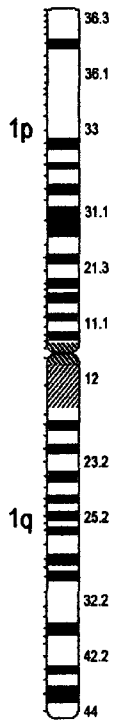
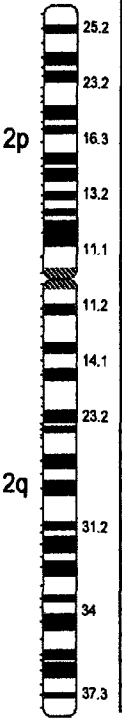

D18S52
D18S1132
D18S976
GATA185C06
D18S391
D18S967
D18S1163
D18S843

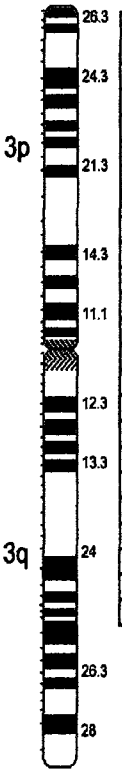
Figure 5.2 Chromosomes 1-4: Markers genotyped for linkage analysis of family 1. Complete lists of markers genotyped on these chromsomes are illustrated. The cytogenic location of the marker as well as its genetic distance in Marshfield and DeCode are indicated. If the marker was previously genotyped, the analysis in which it was used (ie. 25 or 10cM analysis) is given. Those highlighted in yellow were genotyped for completion of the 10cM analysis.



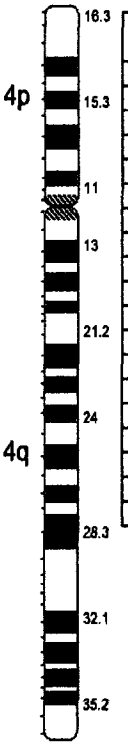
Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D1S468	1p36.32	10cM	4.2	4
D1S1612	1p36.23	25cM	16	14
D1S3669	1p36.13	10cM	37	NK
D1S552	1p36.13	25cM	45	37
D1S1622	1p35.2	10cM	57	49
D1S3721	1p34.2	10cM	73	66
D1S3728	1p32.1	10cM	89	NK
D1S1665	1p31.1	25cM	102	100
d1s1728	1p31.1	10cM	109	NK
D1S551	1p31.1	10cM	113	107
D1S1588	1p22.2	25cM	125	116
D1S1631	1p21.1	10cM	137	126
D1S3723	1p13.3	10cM	140	NK
D1S534	1p12	25cM	152	141
D1S1653	1q23.1		164	152
D1S1677	1q23.3		175	162
D1S1589	1q25.1		192	176
D1S518	1q31.1	25cM	202	188
GATA124F08	NK	10cM	237	NK
D1S1678	1q32.1		218	203
D1S2141	1q41	10cM	233	219
D1S549	1q41	25cM	240	NK
D1S3462	1q42.2	10cM	247	235
D1S235	1q42.3	10cM	254	244
D1S547	1q43		267	259
D1S1609	1q44	25cM	275	NK



Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D2S2976	2p25.3	25cM	NK	3
D2S2952	2p25.1	10cM	18	19
D2S1400	2p25.1	25cM	27	29
D2S1360	2p24.2	10cM	38	40
D2S405	2p23.2	10cM	48	51
D2S1788	2p22.3		56	NK
D2S1356	2p21	25cM	64	NK
D2S441	2p14		87	91
D2S1777	2p12	10cM	99	103
D2S1790	2p11.2	10cM	113	NK
D2S2972	2q11.2	25cM	114	116
D2S410	2q14.1		125	127
D2S1328	2q14.3	25cM	132	127
D2S1334	2q21.3	10cM	145	NK
D2S442	2q22.1	10cM	147	150
D2S1399	2q22.3	25cM	152	NK
D2S1353	2q24.1	10cM	164	168
D2S1776	2q24.3	10cM	173	NK
D2S1391	2q32.1	25cM	186	189
D2S1384	2q33.2	25cM	200	201
GATA4G12	2q34	10cM	210	NK
D2S434	2q35	10cM	216	216
D2S1363	2q36.3	25cM	227	229
D2S427	2q37.1	10cM	237	238
D2S2968	2q37.3	25cM	252	250
D2S125	2q37.3		260	258

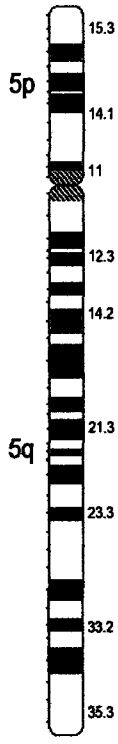


Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D3S2387	3p26.3	25cM	6	2
D3S1304	3p26.1	10cM	22	20
D3S1259	3p25.2	10cM	31	37
D3S3038	3p24.3	10cM	45	43
D3S2432	3p22.3	25cM	58	58
D3S2409	3p21.31		71	71
D3S1766	3p14.2	25cM	79	79
D3S4542	3p14.1	10cM	90	NK
D3S3023	3q13.33	10cM	NK	127
D3S3045	3q13.12	10cM	124	117
D3S2406	3p13		102	NK
D3S4529	3p12.1	25cM	112	NK
D3S3045	3q13.12		124	117
D3S1764	3q23	10cM	153	146
D3S1744	3q24	25cM	161	153
D3S1763	3q26.1	10cM	177	169
D3S3053	3q26.31	10cM	182	NK
D3S2427	3q26.31	25cM	188	NK
D3S1262	3q27.3	10cM	201	194
D3S2398	3q28	10cM	209	203
D3S2418	3q28	25cM	216	NK
D3S1311	3q29	10cM	225	220

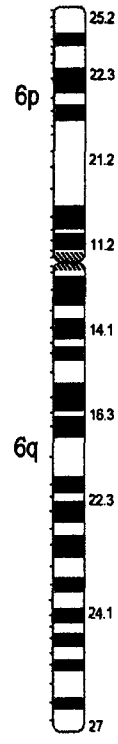


Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D4S2366	4p16.1	25cM	13	12
D4S2639	4p15.32	10cM	21	NK
D4S2397	4p15.2	10cM	43	NK
D4S2632	4p15.1			54
D4S1627	4p13	25cM	60	65
D4S3248	4q13.1	10cM	73	75
D4S2367	4q13.2	10cM	78	NK
D4S3243	4q21.21		NK	88
D4S2361	4q21.23	25cM	93	92
D4S1647	4q23	10cM	105	104
D4S2623	4q25	10cM	114	114
D4S2394	4q28.2	25cM	NK	128
D4S1644	4q31.21	10cM	143	138
D4S1625	4q31.21	25cM	146	140
D4S1629	4q32.1	10cM	158	152
D4S2368	4q32.3	10cM	168	162
D4S2417	4q34.3	10cM	182	176
D4S408	4q35.1		195	189
D4S1652	4q35.2	25cM	208	204

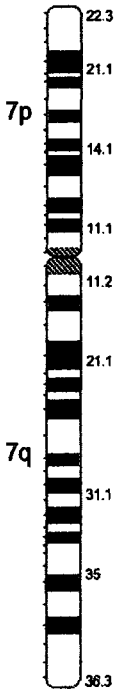
Figure 5.3 Chromosomes 5-8: Markers genotyped for linkage analysis of family 1. Complete lists of markers genotyped on these chromsomes are illustrated. The cytogenic location of the marker as well as its genetic distance in Marshfield and DeCode are indicated. If the marker was previously genotyped, the analysis in which it was used (ie. 25 or 10cM analysis) is given. Those highlighted in yellow were genotyped for completion of the 10cM analysis.



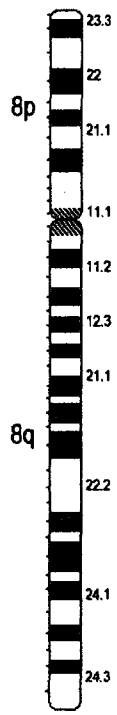
Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D5S2849	5p15.33	10cM	8	NK
D5S2505	5p15.32	10cM	14	16
D5S817	5p15.2		23	30
D5S2845	5p14.3	25cM	36	NK
D5S2848	5p14.1	10cM	40	46
D5S1457	5p13.1	10cM	59	NK
D5S2500	5q11.2	25cM	69	75
D5S1501	5q14.1	10cM	85	94
D5S1462	5q15	10cM	105	NK
D5S1453	5q21.3	10cM	123	NK
D5S1505	5q23.1	25cM	130	125
D5S816	5q31.1	10cM	139	138
D5S1480	5q32	10cM	147	146
D5S820	5q33.3	10cM	159	161
D5S1471	5q34	10cM	172	174
D5S211	5q35.2	10cM	183	191
D5S408	5q35.3	10cM	195	205



Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D6S2434	6p23	25cM	25	31
D6S1959	6p22.3	10cM	34	NK
GATA163B10	6p22.2	25cM	42	48
D6S2427	6p21.2	10cM	54	59
D6S1017	6p21.1	25cM	63	63
D6S1053	6q12	25cM	80	
D6S1031	6q14.1		89	89
D6S1056	6q16.1	25cM	103	100
D6S1021	6q16.3		112	107
D6S474	6q21	25cM	119	NK
D6S1040	6q23.1		129	NK
D6S1009	6q23.3	10cM	138	139
GATA184A08	6q24.3	10cM	146	NK
D6S2436	6q25.2	10cM	155	161
D6S305	6q26	10cM	166	166
D6S1277	6q26	25cM	173	176
D6S1027	6q27	25cM	187	NK



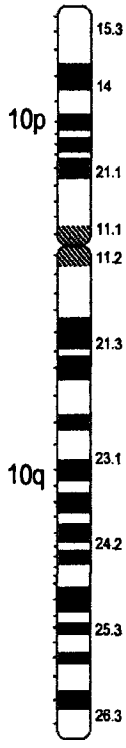
Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D7S1819	7p22.2	25cM	7	8
D7S513	7p21.3		18	23
D7S1808	7p15.1	25cM	42	NK
D7S2846	7p14.1	10cM	58	59
D7S3046	7q11.22	10cM	78	NK
D7S3046(B)	7q11.22	10cM	78	NK
D7S639	7q11.22	10cM	79	NK
D7S645	7q11.22	10cM	80	82
D7S2415	7q11.22	10cM	84	84
D7S653	7q11.22	10cM	85	85
D7S1870	7q11.23	10cM	86	87
D7S2470	7q11.23	10cM	90	90
D7S669	7q21.11	10cM	90	89
D7S660	7q21.11	10cM	94	93
D7S2212	7q21.11	10cM	95	96
D7S806	7q21.11	10cM	95	NK
D7S524	7q21.11	10cM	97	98
D7S820	7q21.11	10cM	98	98
D7S644	7q21.11	10cM	97	99
D7S627	7q21.13	10cM	97	101
D7S492	7q21.13	10cM	100	101
D7S2410	7q21.13	10cM	102	101
D7S657	7q21.3	10cM	105	103
D7S821	7q21.3	10cM	NK	106
D7S2539	7q21.3	10cM	109	109
D7S515	7q22.1	10cM	112	111
D7S796	7q22.1	10cM	113	NK
D7S1799	7q22.1	10cM	114	113
D7S2453	7q22.2	10cM	116	116
D7S2459	7q22.3	10cM	120	118
D7S1842	7q31.32	25cM	128	NK
D7S1824	7q34	10cM	150	NK
D7S2195	7q35		162	NK
D7S1823	7q36.2	25cM	173	NK
D7S559	7q36.3		182	183



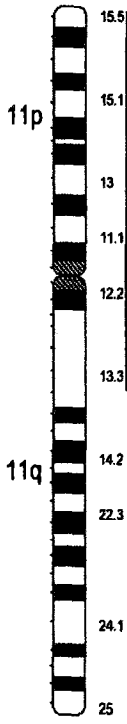
Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D8S264	8p23.3	25cM	1	3
D8S277	8p23.1	10cM	8	15
D8S1130	8p23.1		23	18
D8S1145	8p22	10cM	37	31
D8S136	8p21.3	10cM	44	39
D8S1477	8p12	25cM	60	NK
D8S1110	8q11.23	10cM	67	65
D8S1136	8q13.1	10cM	82	NK
D8S2324	8q21.11		94	NK
GATA1A4	NK	10cM	111	NK
D8S592	8q24.11	10cM	125	119
D8S1179	8q24.13		NK	129
D8S1128	8q24.21	25cM	139	135
D8S256	8q24.22	10cM	148	145
D8S373	8q24.3	25cM	164	164

Figure 5.4 Chromosomes 9-12: Markers genotyped for linkage analysis of family 1. Complete lists of markers genotyped on these chromsomes are illustrated. The cytogenic location of the marker as well as its genetic distance in Marshfield and DeCode are indicated. If the marker was previously genotyped, the analysis in which it was used (ie. 25 or 10cM analysis) is given. Those highlighted in yellow were genotyped for completion of the 10cM analysis.

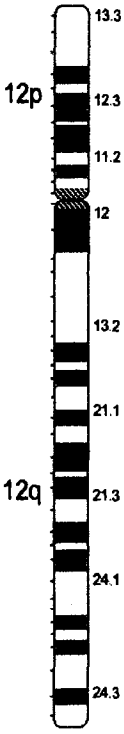
Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D9S2169	9p24.1	25cM	14	NK
D9S921	9p23	10cM	22	NK
D9S925	9p22.2	25cM	32	37
D9S1121	9p21.3	10cM	44	NK
D9S1118	9p21.1	25cM	58	54
D9S301	9q21.12	10cM	66	67
D9S1122	9q21.13	10cM	76	74
D9S922	9q21.31	25cM	80	78
D9S257	9q21.33	10cM	91	92
D9S938	9q31.1	10cM	111	106
D9S930	9q32	10cM	120	117
D9S934	9q33.1	25cM	128	NK
D9S282	9q33.3	10cM	136	136
D9S2157	9q34.2	10cM	147	NK
D9S158	9q34.3	25cM	163	162



Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D10S1435	10p15.3	25cM	NK	NK
D10S189	10p14	10cM	19	20
D10S2325	10p13	25cM	33	31
D10S1423	10p12.33	10cM	NK	44
D10S1426	10p11.23	10cM	59	58
D10S1208	10p11.21	10cM	63	NK
D10S1221	10q21.1	10cM	NK	NK
GATA121A08	10q21.3		88	NK
D10S1432	10q22.1	25cM	94	94
D10S2327	10q22.3	10cM	101	99
D10S1419	10q23.31	10cM	113	111
D10S1239	10q24.32	25cM	NK	122
D10S1237	10q25.3	10cM	135	NK
D10S1230	10q26.12	10cM	143	145
D10S1213	10q26.13	25cM	148	NK
D10S1248	10q26.3	25cM	165	NK
D10S212	10q26.3	10cM	171	177

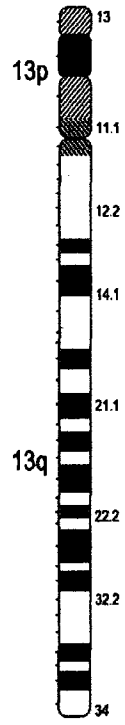


Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D11S1984	11p15.5	25cM	2	0
D11S2362	11p15.4	10cM	9	7
D11S1999	11p15.4	25cM	17	17
D11S1981	11p15.1	10cM	21	26
GATA34E08	NK	10cM	34	NK
D11S1392	11p13	25cM	43	51
D11S1985	11q12.1	10cM	58	NK
D11S2371	11q13.4	25cM	76	NK
D11S2002	11q14.1	10cM	85	87
D11S2000	11q22.3	25cM	100	NK
D11S1986	11q23.1		105	NK
D11S4464	11q24.1	25cM	123	130
D11S912	11q24.3	10cM	131	138
D11S2359	11q25	25cM	148	NK



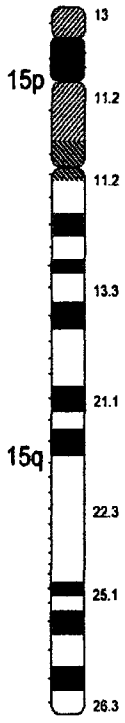
Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D12S372	12p13.32	25cM	6	NK
GATA49D12	12p13.31	10cM	18	NK
D12S373	12p12.3	10cM	36	35
D12S1042	12p11.23	25cM	49	51
D12S1301	12q12		56	58
D12S398	12q13.13	10cM	68	67
D12S1294	12q15	10cM	NK	82
D12S1052	12q21.1	10cM	83	87
D12S1064	12q21.33		95	NK
D12S1300	12q23.1	10cM	NK	99
PAH	12q23.2	25cM	115	NK
D12S2070	12q24.21	10cM	125	NK
D12S395	12q24.23	10cM	136	140
D12S2078	12q24.32	10cM	NK	152
D12S1045	12q24.33	25cM	160	NK
D12S392	12q24.33		166	168

Figure 5.5 Chromosomes 13-16: Markers genotyped for linkage analysis of family 1. Complete lists of markers genotyped on these chromsomes are illustrated. The cytogenic location of the marker as well as its genetic distance in Marshfield and DeCode are indicated. If the marker was previously genotyped, the analysis in which it was used (ie. 25 or 10cM analysis) is given. Those highlighted in yellow were genotyped for completion of the 10cM analysis.

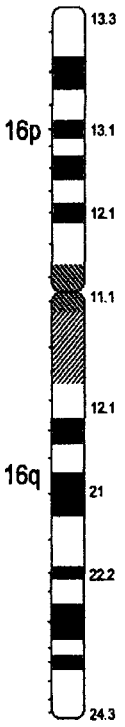


Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D13S787	13q12.12	25cM	9	9
D12S1493	13q13.2	10cM	25	31
D12S894	13q13.3	10cM	NK	39
D12S788	13q14.3	10cM	45	54
D13S317	13q31.1	25cM	64	NK
D13S800	13q22.1	10cM	55	68
D13S325	12q13.13	10cM	69	69
D13S793	13q32.1	25cM	NK	90
D13S779	13q32.3	10cM	83	95
D13S796	13q33.3	25cM	93	109
D13S285	13q34	25cM	110	123

Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D14S742	14q11.2	10cM	12	13
D14S1280	14q12	25cM	25	22
D14S599	14q13.1	10cM	41	NK
D14S587	14q22.2	10cM	56	NK
D14S592	14q23.1	25cM	67	61
D14S588	14q24.1		66	68
D14S53	14q24.3		86	76
D14S606	14q31.1	25cM	92	81
D14S1279	14q31.3	10cM	96	NK
D14S617	14q32.12	10cM	105	94
D14S1434	14q32.13	10cM	113	100
D14S1426	14q32.2	25cM	126	115

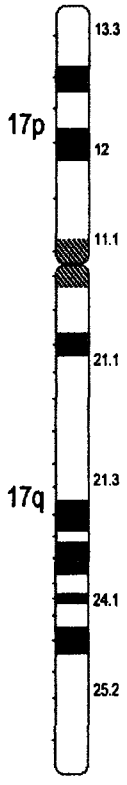


Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D15S165	15q13.3	10cM	20	22
ACTC	15q14		31	NK
D15S659	15q21.1	25cM	43	44
D15S643	15q22.2	25cM	52	59
D15S1507	15q22.31	10cM	60	66
D15S211	15q25.1	10cM	75	85
D15S655	15q25.3	25cM	82	92
D15S652	15q26.1	10cM	90	99
D15S816	15q26.2	10cM	100	110
D15S657	15q26.2	25cM	104	116
D15S642	15q26.3	25cM	122	133

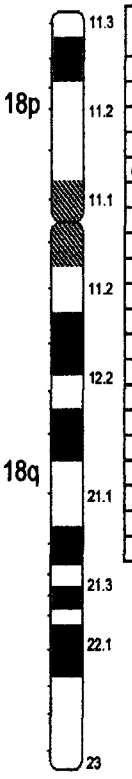


Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D16S2616	16p13.3	10cM	11	NK
D16S748	16p13.13	10cM	22	29
D16S764	16p13.11	25cM	30	36
D16S403	16p12.1	10cM	43	45
D16S769	16p12.1	10cM	50	NK
D16S3396	16q12.1	10cM	63	61
D16S2620	16q21	10cM	81	NK
D16S2624	16q22.3	25cM	87	87
D16S516	16q23.1	10cM	98	100
D16S402	16q23.3	10cM	110	113
D16S539	16q24.1	25cM	124	NK
D16S621	16q24.2		130	NK

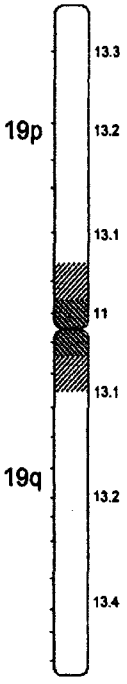
Figure 5.6 Chromosomes 17-20: Markers genotyped for linkage analysis of family 1. Complete lists of markers genotyped on these chromsomes are illustrated. The cytogenic location of the marker as well as its genetic distance in Marshfield and DeCode are indicated. If the marker was previously genotyped, the analysis in which it was used (ie. 25 or 10cM analysis) is given. Those highlighted in yellow were genotyped for completion of the 10cM analysis.



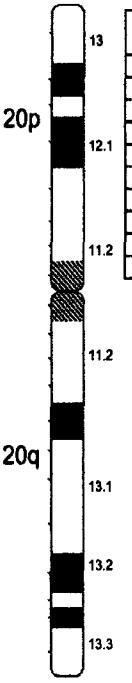
Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D17S1308	17p13.3	25cM	0.6	NK
D17S1298	17p13.2	10cM	11	10
D17S974	17p13.1	10cM	22	NK
D17S947	17p12	10cM	33	31
D17S2196	17p11.2		44	47
D17S740	17p11.2	10cM	44	NK
D17S1925	17p11.2	10cM	45	NK
D17S2196	17p11.2	10cM	44	47
D17S1294	17q11.2	10cM	51	52
D17S1293	17q12	25cM	56	60
D17S1299	17q21.2	10cM	62	NK
D17S2180	17q21.32	10cM	66	74
D17S2193	17q24.2	10cM	89	99
D17S1301	17q25.1	25cM	113	100
D17S784	17q25.3		117	129
D17S928	17q25.3	25cM	126	135



Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D18S481	18p11.31	25cM	7	9
D18S52	18p11.31	10cM	9	13
D18S1132	18p11.31	10cM	11	14
D18S976	18p11.31	10cM	13	16
GATA185C06	18p11.31	10cM	16	17
D18S873	18p11.31	10cM	NK	NK
D18S391	18p11.31	10cM	18	18
D18S452	18p11.31	10cM	19	18
D18S458	18p11.31	10cM	18	19
D18S62	18p11.31	25cM	19	17
D18S344	18p11.31	10cM	19	17
D18S471	18p11.31	10cM	19	18
D18S78	18p11.31	10cM	20	18
D18S967	18p11.31	10cM	20	21
D18S1163	18p11.23	10cM	24	25
D18S843	18p11.22	25cM	28	28
D18S877	18q12.1	10cM	54	52
D18S858	18q21.31	10cM	80	NK
D18S1371	18q22.3	25cM	115	110
D18S844	18q23	10cM	116	NK

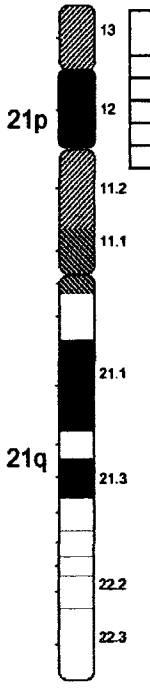


Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D19S591	19p13.3	25cM	10	10
D19S1034	19p13.3	10cM	21	19
D19S586	19p13.2	25cM	33	29
D19S714	19p13.12	10cM	36	NK
D19S433	19q12	25cM	52	51
D19S178	19q13.31	10cM	68	70
D19S246	19q13.33	25cM	78	83
D19S589	19q13.41	10cM	87	NK
D19S254	19q13.43	25cM	101	110

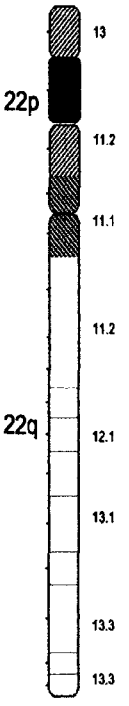


Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D20S103	20p13	10cM	2	2
D20S482	20p13	25cM	NK	13
D20S851	20p12.3	10cM	24	25
D20S604	20p12.1	10cM	33	38
D20S470	20p12.1	25cM	39	NK
D20S477	20p11.21	10cM	NK	52
D20S478	20q11.23		54	61
D20S481	20q13.12	25cM	62	69
D20S480	20q13.2	25cM	80	83
D20S171	20q13.32	25cM	96	98

Figure 5.7 Chromosomes 21-22: Markers genotyped for linkage analysis of family 1. Complete lists of markers genotyped on these chromsomes are illustrated. The cytogenic location of the marker as well as its genetic distance in Marshfield and DeCode are indicated. If the marker was previously genotyped, the analysis in which it was used (ie. 25 or 10cM analysis) is given. Those highlighted in yellow were genotyped for completion of the 10cM analysis.



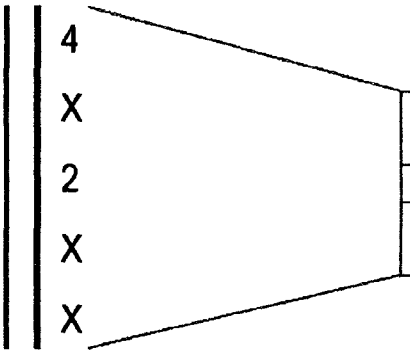
Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D21S1432	21q21.1	25cM	3	9
D21S1437	21q21.1	10cM	13	NK
D21S1440	21q22.13		37	45
D21S2055	21q22.2	25cM	40	NK
D21S1446	21q22.3	25cM	58	NK



Marker	Cytogenic	Origin of data	cM Marshfield	cM DeCode
D22S420	22q11.1	25cM	4	3
D20S608	22q11.23	10cM	18	NK
D22S689	22q12.1	25cM	28	33
D22S685	22q12.3		32	39
D22S683	22q12.3	25cM	36	NK
D22S445	22q12.3	10cM	46	45

Figure 5.8 Common affected haplotypes for markers surrounding D17S2196 (A) and D4S408 (B). The closest markers both up and downstream analyzed previously were used to generate a common haplotype surrounding two markers that had LOD scores above 1.0. "X" indicates the allele is not shared between the symptomatic individuals in family 1, however an allele number indicates this is shared among this group. The blue box indicates the markers that were used to generate the haplotype. Note that in Figure B there are no markers available beyond D4S1652, since this is the telomeric end of chromosome 4.

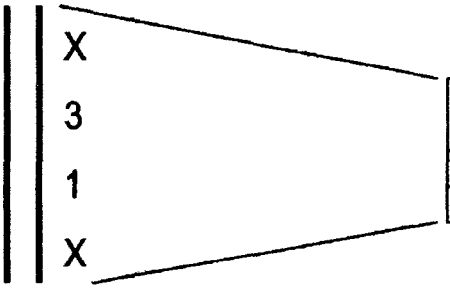
A



	Marker	Genetic Distance	Origin of Data
1	D17S974	22.24	10cM
2	D17S1303	31.25	25cM
3	D17S740	44.62	10cM
4	D17S2196	47.32	"gap fill-in"
5	S17S1294	52.78	10cM
6	D17S1293	60.14	25cM
7	D17S1299	62.01	10cM
8	D17S2180	74.32	10cM

Common affected haplotype

B

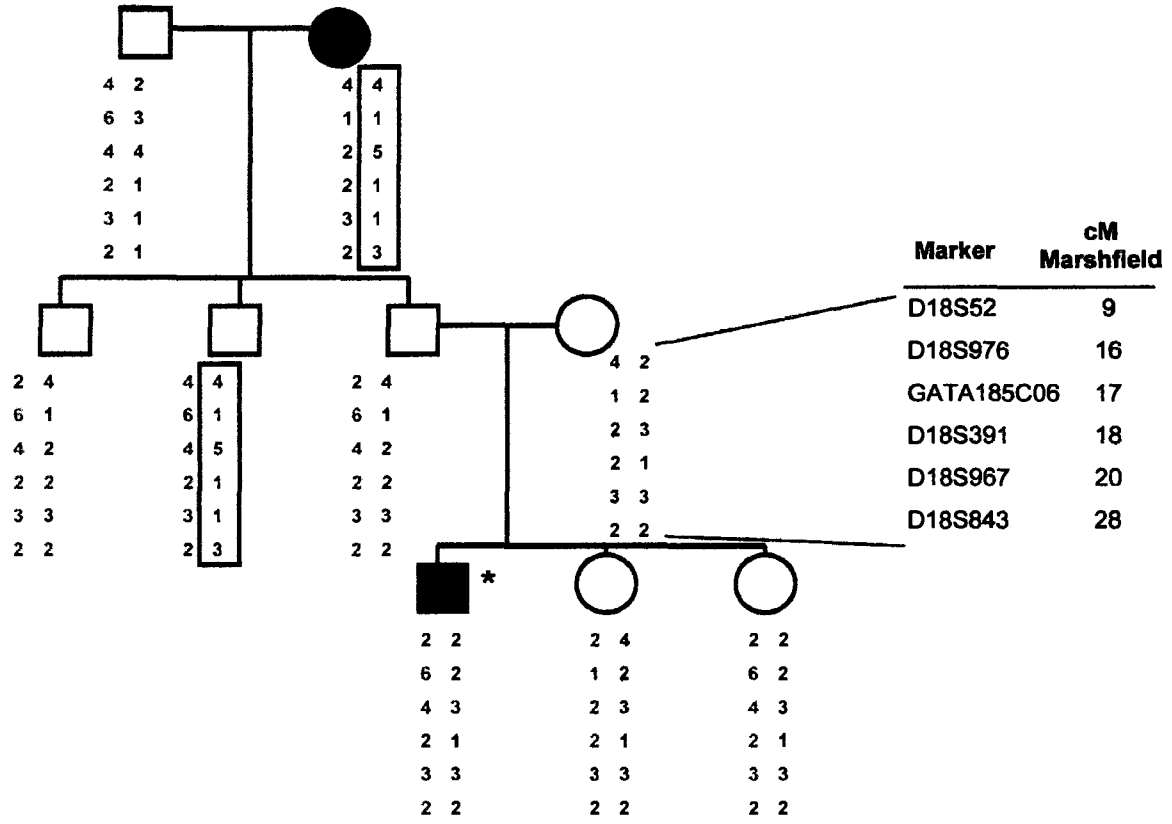


	Marker	Genetic Distance	Origin of Data
1	D4S1625	146	25cM
2	D4S1629	158	10cM
3	D4S2417	182	10cM
4	D4S408	189	"gap fill-in"
5	D4S1652	208	25cM

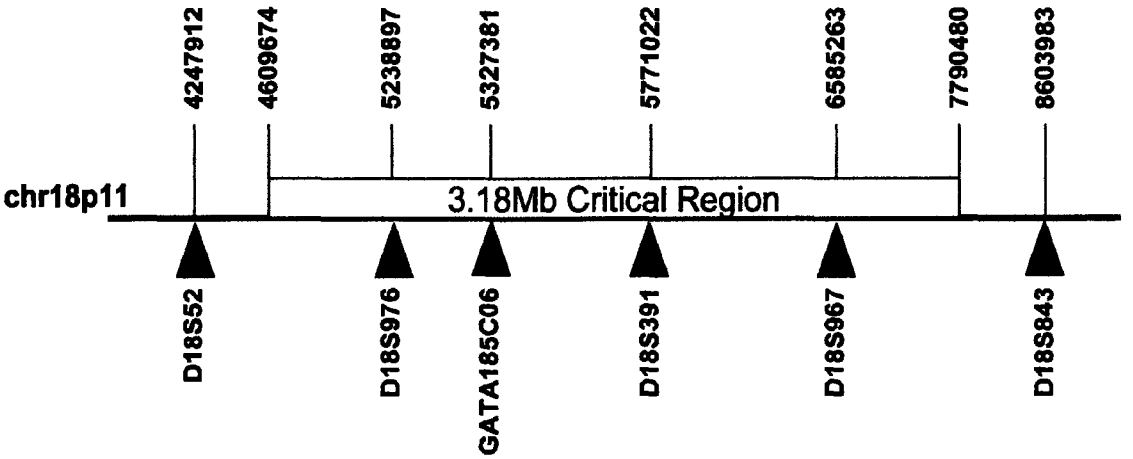
Common affected haplotype

Figure 5.9 Confirmation of the chr18p11 haplotype in individual #67 and his first degree relatives. Haplotypes are given for this group within family 1 generated from 6 markers within and flanking the 3.18Mb critical region (A). Haplotypes are in phase, and the colors allow for tracking the inheritance of alleles, ie. the disease haplotype (indicated by the box) is in red in the grandmother and is transmitted to the unaffected son. Shaded, affected; asterisk, indicates individual #67 who is confirmed to not have the common disease haplotype. B, the physical location of the 6 markers genotyped within the chr18p11 region.

A


Marker
cM Marshfield
D18S52 9
D18S976 16
GATA185C06 17
D18S391 18
D18S967 20
D18S843 28
*

B


4247912
4609674
5238897
5327381
5771022
6585263
7790480
8603983
chr18p11
3.18Mb Critical Region
D18S52
D18S976
GATA185C06
D18S391
D18S967
D18S843

Chapter 6.

Analysis of the chr18p11 critical region: Candidate gene and genomic rearrangement investigation

Chapter 6 Analysis of the chr18p11 critical region: Candidate gene and genomic rearrangement investigation

6.1 Introduction

Completion of the 10cM linkage analysis provided further evidence for chr18p11 as the single disease locus associated with Myoclonus Dystonia within family 1. Recombination events within the disease locus in two affected individuals narrowed this region to 3.18Mb. This region contains a number of potential candidate genes that could be responsible for MD. When candidate gene prioritization of this region began in 2003, there were 17 genes within this region according to the NCBI database. There are now a total of 24 genes, both confirmed and predicted, within the region. Initial prioritization of candidate genes was performed using what was known about their expression patterns, molecular function, map position, and conservation with known genes. As well, genes that could potentially be involved in the MD phenotype, such as those that were expressed but not limited to brain and fetal brain, were given priority. Genes expressing proteins that were known to interact with sarcoglycans, as well as any of the components of the dystrophin glycoprotein complex (DGC) were given the highest priority. One such example is *LAMA1*, which encodes for laminin alpha 1. This protein is an extracellular matrix protein within the same family as the laminin alpha 2 chain, which binds to the DGC in the sarcolemma of muscle[59]. It is also expressed in brain, making it an excellent candidate gene for this disease. This gene and 11 others were previously sequenced in a panel of roughly 6-8 affected individuals, two of which were from family 1 (Table 6.1). There were no common novel variants found that could be responsible for MD. As well, there were no cases of all affected individuals having different novel variants within the same gene, which would have also

indicated a potential gene for MD. Many polymorphisms were found, however they were found to be present in the general population through the HapMap database of single nucleotide polymorphisms (SNPs).

For our studies we catalogued the remaining genes that had not yet been sequenced and prioritized them based on the same criteria described above. Most of the remaining candidate genes are cDNAs that have been sequenced from tissue expression libraries and subsequently mapped to the chr18p11 genomic region by sequence homology. Therefore, it is unknown if these predicted genes encode protein products, or if they are non-coding RNAs. Genes that were strongly predicted to express a protein product (i.e. containing transcriptional start and stop sites) were also prioritized to be sequenced in a panel of affected individuals.

There have been a number of newly identified isoforms of previously sequenced genes that are now present in the UCSC Genome database. Four candidate genes (*LAMA1*, *ZFP161*, *ARHGAP28*, *L3MBTL4*) now have additional coding exons that have not yet been examined. We have sequenced a total of 12 of these exons in a panel of affected individuals in search of novel variants that may be responsible for MD.

We have also investigated the possibility of an inversion within in the chr18p11 critical region. It is possible that an inversion breakpoint within the region could potentially fall within the coding region of a gene, or within regulatory regions that are necessary for proper transcriptional control of a gene cluster. There has indeed been a large pericentric inversion previously identified in a Scottish family with schizophrenia with a breakpoint roughly 100bp upstream of the *TTMA* gene within the critical region[60]. A series of three-color Fluorescence *In Situ* Hybridization (FISH) experiments were designed to investigate a

potential inversion. Specific bacterial artificial chromosome (BAC) clones were selected based on their genomic position within the critical region and were fluorescently labeled (Figure 6.4). Two clones were selected for hybridization outside the critical regions flanks and were used as anchors to orient the position of the other clones in the probe set. A correct order of hybridization of the three BAC clones along the region would indicate the absence of an inversion, whereas an altered order would be a potential inversion. By using this method on an affected individual (#16) from family 1 in comparison to a WT control, we can determine the presence or absence of a pathogenic inversion within the mapped chr18p11 critical region.

6.2 Materials and Methods

6.2.1 DNA and RNA Extraction

The following Myoclonus Dystonia patients were used for sequencing of candidate genes: from family 1 – M1-39, M1-15, M1-3, M1-13 and M1-7. From other families – SP1, M8-1, M12-1, M16-1, M13-1. These individuals were chosen to be sequenced based on their phenotype (early onset of MD and alcohol responsive) and they did not have *SGCE* mutations. Genomic DNA was extracted from whole blood drawn from patients in yellow top ACD Vacutainer tubes. Extraction was performed using the QIAamp DNA Blood Mini Kit. Briefly, 20ul of protease was added to 200ul of whole blood along with 200ul of lysis buffer (AL). Samples were incubated at 56°C for 10min, after which 200ul of 100% ethanol was added. The mixture was applied to a QIAamp Spin Column and centrifuged at 6000 x g for 1 min to remove cellular debris. Two washes were then performed using buffers AW1 and AW2, and purified genomic DNA was then eluted from the filter with 100ul of either

buffer AE (Qiagen) or TE (10mM Tris-HCl, 1mM EDTA, pH 8.0). A yield of ~30ng/ul was typical with an A_{260}/A_{280} ratio of 1.7-1.9. Remaining patient blood sample was aliquoted and stored at -80°C.

6.2.2 Primer Design, PCR, and Bioinformatics Databases

Genomic DNA sequence was obtained from the University of California Santa Cruz (UCSC) Genome Browser. Primers were designed using Primer 3 v.0.4.0 (Steve Rozen and Helen J. Skaletsky) and had an optimal length of 20nt, 50% G/C content and a melting temperature (Tm) of 60°C. Optimal amplicon size was dependant on the particular sequence being studied and ranged from 150-3500bp.

Various conditions of the polymerase chain reaction (PCR) were used depending on the characteristics of the product of interest being amplified. Conditions were optimized for each primer set by adjusting the Tm, cycle number and time, altering $MgCl_2$ concentration, or by using additives such as dimethyl sulfoxide (DMSO).

The UCSC database (http://genome.ucsc.edu/) was used to extract information regarding previously identified polymorphisms within coding regions. NCBI (http://www.ncbi.nlm.nih.gov/), Gepis Tissue (http://www.cgl.ucsf.edu/Research/genentech/genehub-gepis/index.html), and published microarray data available on UCSC were used to pull tissue expression data on candidate genes. The Coriell Institute for Medical Research (http://www.coriell.org/) and Database for Genomic Variants (http://projects.tcag.ca/variation/) was used to obtain information on previously identified structural variants within the chr18p11 region.

6.2.3 Sequencing

PCR products were purified using Exonuclease 1 and Shrimp Alkaline Phosphatase (EXO-SAP). EXO removed single stranded DNA, including primers and ssDNA produced from PCR and SAP digested unconsumed dNTPs by binding to their phosphate groups. A total of 3ul of PCR product was combined with 2ul of EXO-SAP along with 5ul of sterile dH_2O and incubated at 37°C for 15 min, followed by 15 min at 80°C for enzyme inactivation. This product was then amplified using a single primer (3.2pmol), 5X Sequencing Buffer (Applied Biosystems) and Big Dye 3.1 sequencing pre-mix (Applied Biosystems) was used to fluorescently label the PCR product. The program was as follows: 1 min at 96°C; 25 cycles of 10s at 96°C, 5s at 50°C, 4 min at 60°C; 5°C hold. Labeled sequence was then further purified by either ethanol precipitation or, if using a 96-well plate, an automated magnetic bead system. The 16 capillary ABI 3130xl Genetic Analyzer was used to perform capillary electrophoresis of the products. Sequence was read using ABI Sequencing Analysis 5.2 software and sequence variants were identified by the program as well as manually.

6.2.4 Flourescence *In Situ* Hybridization

Three color Fluorescence *In Situ* Hybridization (FISH) was performed on one affected individual from family 1 with the chromosome 18 disease haplotype (#16). Blood was drawn from the individual and sent to the TCAG Biobanking Facility in Toronto where the lymphocytes were isolated and immortalized with the Epstein Barr Virus (EBV). The lymphoblastoid cell line was then sent to the Cytogenomics and Genome Resources Facility at the Hospital for Sick Children in Toronto where the FISH experiments were performed.

Eight Bacterial Artificial Chromosome (BAC) clones were chosen from the RPCI-11 Human BAC library based on their location within the chr18 critical region as mapped in the human Genome Browser (UCSC). Clone sizes ranged from 156kb to 231kb and were approximately 500-700kb apart. BACs were fluorescently labeled with red, green or yellow fluorophores and hybridized to the chr18 critical region during metaphase, allowing potential inversions to be visualized.

6.3 Results

A total of 24 candidate genes fall within the chr18p11 critical region, 12 of these have been previously sequenced in a panel of symptomatic Myoclonus Dystonia patients (Table 6.1). The remaining 12 genes are classified as predicted genes and are potential candidates for this disease. There is limited tissue expression data available for most of these genes besides the tissue library in which the cDNA originated. Two genes were prioritized to have their predicted coding regions sequenced, first *BC041875* due to its expression in fetal brain. It potentially has three coding exons and has over 60 known polymorphisms (Figure 6.2, A). *TTMA* is a single exon predicted gene that was sequenced next due to its wide expression and the presence of transcriptional start and stop sites (Figure 6.2, B). Fifty-one of the known polymorphisms within the *BC041875* gene were catalogued in the panel of 9 affected individuals (Table 6.2). Four individuals in this panel are from family 1, since this locus was mapped in this family. These 51 polymorphisms ranged from single nucleotide polymorphisms (SNPs), to small insertions and deletions. There were also 11 novel variants found, in which case WT unaffected controls were subsequently sequenced to determine if these were also present in individuals not affected with MD. Variant "G" is novel and was

found in one affected individual (M12-1). One WT control was screened and this variant was not seen. Since this variant is not seen in the four affected family 1 individuals or any other individuals in the panel, it is unlikely this variant is causing the disease. In addition, the remaining 10 novel variants were all seen in WT controls and therefore ruled out as causative mutations. If this gene was associated with Myoclonus Dystonia, we would expect to find many of the affected individuals in the panel to carry mutations within the gene, particularly family 1 since it was used to map this region. We have only observed one novel variant in one individual (not in family 1) that was not present in a WT control; therefore we are confident that this gene is not associated with MD. However, future screening of additional controls will be performed for variant "G" to determine if it is seen in non-affected individuals.

The candidate gene *TTMA* was sequenced in a panel of 6 affected individuals and one WT control. This analysis yielded 6 polymorphisms, two of which were novel variants. These were both seen in patient SP1 only, making this unlikely to be associated with the disease for reasons similar to those mentioned above for variant "G". An intronic A-insertion designated as rs34293594 was found in individual M13-1, however given its intronic location as well as its absence in the rest of the panel, it is also unlikely to be a causative mutation.

Four candidate genes, *LAMA1*, *ZFP161*, *ARHGAP28* and *L3MBTL4* are considered to be strong candidate genes based on their reported function (Figure 6.3). *LAMA1* encodes laminin alpha 1, which has been found to mediate the attachment, migration and organization of cells into tissues by interacting with other extracellular matrix components and is closely related to the alpha 2 chain form of the DGC, making *LAMA1* a prime

candidate gene[61]. *ZFP161* expresses a zinc finger protein known to be a transcriptional activator of the dopamine transporter (*DAT*) gene by binding to its promoter[62]. *ARHGAP28* is thought to express a RhoGTPase activating protein involved in intracellular signal transduction[63]. *L3MBTL4* encodes a protein with similar homology to a Drosophilia lethal (3) malignant brain tumor-like 4 (I(3)mbt) protein. It has three mbt repeats and one sterile alpha motif (SAM) domain[64]. These four candidate genes have recently had additional isoforms identified, in turn revealing new coding exons. These 12 new exons were sequenced in a panel of two affected individuals from family 1, and two WT controls. This resulted in the identification of SNPs that have been previously found in the general population (Table 6.4). In total, 36 SNPS were found, however none are likely candidates for a potential disease-causing mutation, since these have previously been identified in the HapMap population. Exon 1 of *ZFP161* isoform 2 was unable to be sequenced due to insufficient amplification using various primer sets. The remaining 11 newly identified exons were successfully sequenced and ruled out as being potential sites for pathogenic mutations within family 1.

A total of 7 FISH experiments were performed on the family 1 symptomatic individual #16 for inversion testing. Two to three metaphase spreads were checked for each experiment to confirm chromosomal localization of the probes to chr18p11. Varying number of nuclei were scored for each experiment, ranging from 122-349. The number of nuclei in which the middle probe of the probe set was seen between its two flanking probes were counted. If there is no inversion present within the probe set, the middle probe would theoretically be present in the center 100% of the time, however a range of 60-80% is

usually observed. A heterozygous inversion would be indicated by the middle probe being centered in roughly 50% of the nuclei counts.

Experiment 1 initially showed that centre probe 2 (yellow) was seen between its flanking probes 1 and 4 in 50% of the 217 nuclei scored (Table 6.7). As well, probe 1 was found to be in the centre of this probe set 28% and 37% of the time, in the original and repeat experiments, respectively. These results are indicative of a possible inversion involving probes 1 and 2. However, the data from this experiment has been deemed not entirely conclusive by the TCAG Cytogenetics facility due to the variability observed when experiment 1 was repeated. Experiments 2 through 5 show percentages that are not indicative of an inversion in these probe sets since these middle probes are most often seen in the centre of these sets. Experiment 6 initially showed the centre probe 5 to be in the middle of the set in 49% of the nuclei counts, and its flanking probe 4 in the centre 42% of the time, thereby identifying a possible inversion involving probes 4 and 5. When repeated, very similar results were found, further supporting this theory. A WT control was also used for this experiment and showed results indicative of the same inversion, suggesting this is potentially a non-pathogenic inversion. In addition to experiment 6, 7 also shows a potential inversion involving probes 4 and 5 which was also seen in a WT control. Both experiments indicate that probes 4 and 5 are flipped heterozygously (50% of the time).

6.4 Discussion

Direct sequencing of the predicted coding regions of candidate genes *BC041875* and *TTMA* showed a number of novel variants. However, since the majority of these were also seen in WT controls, they are not responsible for the disease pathology of Myoclonus

Dystonia. Three novel variants were seen in MD patients and not in one WT control. To properly determine if these are also found in non-affected individuals, additional WT controls will be screened. Sequencing of the newly identified coding exons within new isoforms of four the strong candidate genes *LAMA1*, *ZFP161*, *ARHGAP28* and *L3MBTL4* showed the absence of novel variants in a panel of two affected individuals from family 1. In conclusion, sequencing various coding regions of these 6 genes showed the absence of a mutation responsible for MD.

We have used a series of BAC probes which are flanking and distributed throughout the critical region, allowing us to access the possibility of an inversion as the mutational event in family 1. The FISH experiments performed on the region in individual #16 provided inconclusive evidence of a heterozygous inversion. Data from experiment 1 is suggestive of an inversion of probes 1 and 2 (Figure 6.4; Table 6.5). However, additional interrogation of this area is required to determine a specific inversion breakpoint. Experiments 2, 3, 4 and 5 yielded results consistent with no inversion. Both experiments 6 and 7 indicate probes 4 and 5 are heterozygously inverted, since they are found to be in the middle of their probe sets in 50% of the nuclei counted. However, the results from experiment 5 (involving probes 2, 4 and 5) are compelling as probe 4 was found between probes 3 and 5 approximately 75% of the time. This is contradictory to the results of experiments 6 and 7 which suggest an inversion involving 4 and 5. In summary, based on these results, there is potential for an inversion involving probes 1 and 2. A key experiment to perform would involve a probe set with probes 1, 2 and 3 since it would allow us to narrow the site of a potential breakpoint in this area. The potential inversion involving

probes 4 and 5 is presently inconclusive, given that the results from experiment 5 are contradictory with those from experiments 6 and 7.

There have been previously identified inversion breakpoints within the chr18p11 critical region, however given the results from our FISH experiments; it is unlikely these are responsible for the potential inversion found within the family 1 Myoclonus Dystonia patient. The breakpoint previously identified in a schizophrenia patient falls within the sequence between probes 4 and 5, which according to experiment 5, has been ruled out as a potential breakpoint region in patient #16. There has also been a small inversion (51,241bp) identified and listed in the Database for Genomic Variants in one Chinese individual in the HapMap Han Chinese in Beijing (CHD) population that spans the 5' region of the *LAMA1* gene (indicated by the green bar in Figure 6.5). This inversion is too small to be a candidate for the potential inversion we are seeing in our FISH experiments. However, it certainly should not be ignored as a potential candidate structural variant for Myoclonus Dystonia, since the breakpoint would not have been identified through direct sequencing as it falls within the intronic sequence between exons 42 and 43 of the gene. Searches in the HapMap, Coriell and NCBI databases for clinical phenotypic information for this Chinese individual were unsuccessful.

The genomic area within probes 1 and 2 contains very few genes, therefore the potential effect of an inversion within this region is unknown. Conversely, the potential inversion involving probes 4 and 5 could disrupt a number of genes, including *TTMA*, *L3MBTL4* and 2 predicted genes. However, since the potential inversion was also seen in one WT control, it is possible this is a common, non-pathogenic inversion. Additional controls could be screened to further investigate this possibility. In summary, to fully

investigate the potential inversions within this region and the specific locations of their breakpoints, additional FISH experiments and subsequent sequencing is required.

Table 6.1 List of candidate genes within the 3.18Mb critical region mapped to chr18p11. Candidate genes are listed based on the order in which they were prioritized to be sequenced. Genes within the grey box have previously been sequenced. The genomic start and stop positions of each gene (as determined by UCSC) are listed, as well as their size and a brief description. The status of the gene is either confirmed (C), predicted (P), or discontinued (D). Asterisks indicate the gene is not found within the UCSC database, only NCBI.

Priority	Gene Name	Genomic Start	Genomic End	Size	Status	Gene Description
1	LAMA1	6,931,886	7,107,813	175,927	C	Laminin, aplha 1. Extracellular matrix protein
2	EPB41L3	5,382,388	5,533,986	151,598	C	Erythrocyte membrane protein band 4.1-like 3
3	PTPRM	7557314	8396859	839546	C	Protein tyrosine phosphatase, receptor type, M
4	ARHGAP28	6,840,790	6,888,720	47,930	C	Rho GTPase activating protein 28 isoform b
5	L3MBTL4	6,019,387	6,404,901	385,514	C	l(3)mbt-like 4 (Drosophilia)
6	ZFP161	5,279,379	5,286,039	6,660	C	Zinc finger protein 161 homolog (mouse)
7	LRRC30	7,221,137	7,222,042	906	C	Leucine rich repeat containing 30
8	MGC17515	5,228,042	5,226,717	1,305	P	Homo sapiens chromosome 18 open reading frame 18, mRNA
9	LOC400643*	6915477	6918551	3074	P	Hypothetical gene supported by EST AK095347
10	LOC645423*	7,124,929	7,125,851	32,100	P	Similar to mitochondrial carrier triple repeat 1
11	LOC388460	6,452,073	6,453,036	963	D	Similar to 60S ribosomal protein L6
12	LOC390826*	5122722	5123790	1,068	D	Similar to peptidyl-Pro cis trans isomerase
13	BC041875	5,228,083	5,236,504	8422	P	expression)
14	TTMA	5,880,184	5,882,103	1,920	P	Hypothetical protein LOC645369 (widely expressed)
15	DA799915	4765053	4994523	229,470	P	EST isolated from human fetal brain
16	AK095347	6,915,477	6,918,551	3,075	P	Homo sapiens cDNA clone CTONG2013222
17	AK097613	6,109,495	6,919,804	10,310	P	Homo sapiens cDNA clone CTONG2013222
18	DQ596733	6,917,186	6,917,208	23	P	NK
19	BC040631	6,246,746	6,250,933	4,187	P	Homo sapiens, clone, mRNA
20	LOC339290	5,222,875	5,228,525	5,651	P	Homo sapiens hypothetical protein, mRNA (cDNA clone)
21	LOC642597*	5187507	5135284	52,223	P	Similar to hCG1795978
22	BC035413	6,915,477	6,919,868	4,392	P	Homo sapiens cDNA, clone CTONG2013222
23	LOC645355*	5740577	5783968	43,391	P	Hypothetical gene, cDNA clone
24	LOC729309*	5888221	5885038	3183	P	Hypothetical gene, cDNA clone

Table 6.2 Novel variants and polymorphisms identified in the candidate gene *BC041875.* All novel variants and previously characterized polymorphisms found within this gene are listed. A panel of 9 affected individuals were sequenced for the 3 potential coding exons of this predicted gene. Four of these individuals originate from family 1 (indicated by M1-n). Variants that have been previously identified in the HapMap populations are given an "rs" designation, and novel variants that have been found are named A through K. The specific change is listed with its genomic location. WT unaffected controls were screened for the novel variants found, the number of controls screened is stated in brackets.

					Affected Individuals									
Exon	Designation	Change	Base Position	Status	M1-39	M1-15	M1-3	M1-13	SP1	M8-1	M12-1	M16-1	M13-1	Seen in controls?
1	rs35793050	G/A	chr18: 5228119	known	NK	NK	NK	NK	NK	NK	NK	NK	NK	none tested
1	A	C/T	chr18: 5228207	novel	C/T	C/T	NK	NK	C/T	C/T	NK	C/C	NK	yes (6/8)
1	rs1061666	A/G	chr18: 5228208	known	A/G	A/G	NK	NK	A/A	A/G	A/A	A/A	A/A	no (1)
1	rs11363931	G insertion	chr18: 5228230	known	NK	NK	NK	NK	NK	NK	NK	NK	NK	none tested
1	rs9627	C/A	chr18: 5228412	known	C/C	C/C	NK	NK	C/A	C/C	NK	C/A	NK	no (1)
1	rs11795	G/C	chr18: 5228442	known	G/G	G/G	NK	NK	G/G	G/G	NK	G/C	NK	no (1)
1	rs12163	C/T	chr18: 5228499	known	C/C	C/C	NK	NK	C/C	C/C	NK	C/C	NK	yes (1/1)
1	B	A/C	chr18: 5228731	novel	A/C	A/C	NK	NK	A/C	A/C	NK	A/C	NK	yes (1/1)
1	rs3087772	C/T	chr18: 5228828	known	C/T	C/T	NK	NK	C/C	C/T	C/C	C/C	C/C	yes (3/5)
1	rs11541047	T/C	chr18: 5228946	known	NK	NK	NK	NK	NK	NK	NK	NK	NK	none tested
1	rs1047413	C/T	chr18: 5229264	known	NK	NK	NK	NK	NK	NK	NK	NK	NK	none tested
2	rs1016374(3)	G/A	chr18: 5231517	known	G/G	G/G	NK	NK	G/A	G/G	G/G	G/G	G/A	none tested
2	rs4621073	A/G	chr18: 5231755	known	A/G	A/G	NK	NK	A/A	A/G	A/A	A/A	A/A	none tested
3	C	T/C	chr18: 5233459	novel	T/C	T/T	T/C	T/C	T/T	T/T	T/T	T/T	T/T	yes (2/6)
3	E	T/C	chr18: 5233497	novel	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/C	yes (2/5)
3	F	G/A	chr18: 5233527	novel	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/A	yes (1/11)
3	rs8084872	C/T	chr18: 5233567	known	C/T	NK	C/T	C/T	C/C	C/T (ho)	C/C	C/C	C/T	none tested
3	rs11278848	8bp deletion	chr18: 5233614-622	known	no del	NK	no del	no del	no del	no del	no del	no del	no del	no (1)
3	rs11278595	8bp deletion	chr18: 5233617-625	known	no del	NK	no del	no del	no del	no del	no del	no del	no del	no (1)
3	D	9bp deletion	chr18: 5233618-627	novel	no del	no del	no del	no del	no del	no del	9bp del	no del	9bp del	yes (2/5)
3	rs1212303	A/C	chr18: 5233780	known	A/A	NK	A/A	A/A	A/A	A/A	A/A	A/A	A/A	no (1)
3	G	G/A	chr18: 5233960	novel	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	no (1)
3	rs11267425	10bp deletion	chr18: 5234169-179	known	no del	NK	no del	no del	no del	no del	no del	no del	no del	no (1)
3	rs59474841	11bp deletion	chr18: 5234172	known	no del	no del	no del	no del	no del	no del	11bp del	no del	11bp del	yes (1/2)
3	rs1062837	C/T	chr18: 5234173	known	C/T	NK	C/T	C/T	C/T	C/C	C/C	C/C	C/T	no (1)
3	rs3184460	A/G	chr18: 5234261	known	A/A	NK	A/A	A/G	A/G	A/A	A/A	A/A	A/A	no (1)
3	rs4798325	G/A	chr18: 5234325	known	G/G	NK	G/G	G/G	G/G	G/G	G/G	G/G	G/G	no (1)
3	rs1127990	T/C	chr18: 5234343	known	T/T	NK	T/T	T/C	T/C	T/T	T/T	T/T	T/T	no (1)
3	rs1127991	C/T	chr18: 5234487	known	C/C	NK	C/C	C/T	C/C	C/C	C/C	C/C	C/C	no (1)
3	rs1127992	C/T	chr18: 5234541	known	C/C	NK	C/C	C/C	C/C	C/C	C/C	C/C	C/C	no (1)
3	rs9962322	C/T	chr18: 5234751	known	C/C	NK	C/C	C/T	C/T	C/C	C/C	C/C	C/C	no (1)

Table 6.3 Novel variants and polymorphisms identified in the candidate gene *BC041875.* All novel variants and previously characterized polymorphisms found within this gene are listed. A panel of 9 affected individuals were sequenced for the 3 potential coding exons of this predicted gene. Four of these individuals originate from family 1 (indicated by M1-n). Variants that have been previously identified in the HapMap populations are given an "rs" designation, and novel variants that have been found are named A through K. The specific change is listed with its genomic location. WT unaffected controls were screened for the novel variants found, the number of controls screened is stated in brackets.

					Affected Individuals									
Exon	Designation	Change	Base Position	Status	M1-39	M1-15	M1-3	M1-13	SP1	M8-1	M12-1	M16-1	M13-1	Seen in controls?
3	rs1044656	T/C	chr18: 5234771	known	T/C	NK	T/C	T/C	T/C	T/C	T/T	T/T	T/C	no (1)
3	rs1044658	G/C	chr18: 5234776	known	G/C	NK	G/C	G/C	G/C	G/C	G/G	G/G	G/C	no (1)
3	rs34514668	3bp insertion	chr18: 5234934-933	known	no ins	NK	no ins	no ins	no ins	no ins	no ins	no ins	3bp ins	yes (1/4)
3	H	TAG insertion	chr18: 5234935-934	novel	no ins	no ins	no ins	3bp ins (ho)	no ins	no ins	no ins	no ins	no ins	yes (2/3)
3	rs10689927	ACT insertion	chr18: 5234935-934	known	3bp ins	NK	3bp ins	no ins	3bp ins	3bp ins (ho)	3bp ins	no ins	no ins	yes (2/4)
3	rs34495101	2bp insertion	chr18: 5234963-962	known	no ins	NK	2bp ins	2 bp ins (ho)	no ins	2bp ins (ho)	no ins	no ins	2bp ins	yes (2/4)
3	rs34232073	2bp insertion	chr18: 5234964-963	known	2bp ins	NK	2bp ins	2 bp ins	no ins	no ins	no ins	no ins	no ins	no (1)
3	I	TG insertion	chr18: 5234966-967	novel	no ins	no ins	no ins	2bp ins (ho)	2bp ins	no ins	no ins	no ins	no ins	yes (1/2)
3	rs15920	T/C	chr18: 5234995	known	T/T	NK	T/T	T/C	T/T	T/C	T/T	T/T	T/C	no (1)
3	rs15764	C/T	chr18: 5235020	known	C/C	NK	C/C	C/C	C/C	C/C	C/C	C/C	C/C	no (1)
3	rs35684186	T insertion	chr18: 5235074	known	no del	NK	no del	no del	no del	no del	no del	no del	no del	no (1)
3	rs11874542	A/G	chr18: 5235176	known	A/A	NK	A/A	A/A	A/A	A/A	A/A	A/A	A/A	no (1)
3	rs11877479	C/T	chr18: 5235228	known	C/C	NK	C/C	C/T	C/C	C/C	C/C	C/C	C/C	no (1)
3	rs11877315	G/A	chr18: 5235247	known	G/G	NK	G/G	G/A	G/G	G/G	G/G	G/G	G/G	no (1)
3	rs34934500	TGTG insertion	chr18: 5235372-371	known	no ins	NK	no ins	no ins	no ins	no ins	no ins	no ins	no ins	no (1)
3	rs34149046	6bp insertion	chr18: 5235383-382	known	no ins	NK	no ins	no ins	no ins	no ins	no ins	no ins	no ins	no (1)
3	J	TATA insertion	chr18: 5235401-405	novel	no ins	no ins	no ins	4bp ins	no ins	no ins	no ins	no ins	no ins	yes (1/2)
3	rs34497956	ATAT insertion	chr18: 5235409-408	known	no ins	NK	no ins	no ins	no ins	no ins	no ins	no ins	no ins	no (1)
3	rs1542948	G/T	chr18: 5235409	known	G/G	NK	G/G	G/G	G/G	G/G	G/G	G/G	G/G	no (1)
3	rs4548972	T/G	chr18: 5235411	known	T/T	NK	T/T	T/G	T/T	T/T	T/T	T/T	T/T	no (1)
3	rs10689221	GAGA insertion	chr18: 5235415-414	known	no ins	NK	no ins	no ins	no ins	no ins	no ins	no ins	no ins	no (1)
3	rs16947948	A/G	chr18: 5235578	known	A/A	NK	A/A	A/A	A/A	A/A	A/A	A/A	A/A	no (1)
3	rs34274740	C insertion	chr18: 5235638-637	known	no ins	NK	no ins	no ins	no ins	no ins	no ins	no ins	no ins	no (1)
3	rs16947949	T/G	chr18: 5235658	known	T/T	NK	T/T	T/T	T/T	T/T	T/T	T/T	T/T	no (1)
3	rs17379984	G/A	chr18: 5235824	known	G/G	NK	G/A	G/A	G/G	G/G	G/G	G/G	G/G	no (1)
3	K	G/C	chr18: 5235860	novel	NK	NK	G/C	G/C	G/G	G/C	G/G	G/G	G/G	yes (1/2)
3	rs755619	C/T	chr18: 5235882	known	C/C	NK	C/C	C/C	C/C	C/C	C/C	C/C	C/C	no (1)
3	rs36058156	2bp insertion	chr18: 5236007-006	known	no ins	NK	2 bp ins	2 bp ins (ho)	2bp ins	2bp ins (ho)	no ins	no ins	no ins	no (1)
3	rs33924005	AC insertion	chr18: 5236014-013	known	no ins	NK	no ins	no ins	no ins	no ins	no ins	no ins	no ins	no (1)
3	rs8576	C/T	chr18: 5236398	known	C/C	NK	C/T (ho)	C/T	C/T	C/T (ho)	C/C	C/C	C/T	no (1)

Table 6.4 Novel variants and polymorphisms identified in the candidate gene *TTMA*. All novel variants and previously characterized polymorphisms found within this gene are listed. A panel of 6 affected individuals and 1 WT control were sequenced for the 1 predicted coding exon of this gene. Two novel variants were found, and were designated as "A" and "B".

				Affected Individuals						
Exon	Designation	Change	Base Position	M1-3	M1-13	SP1	M8-1	M16-1	M13-1	Seen in Control?
1	rs34293594	A insertion	chr18:5880016-5880015	no ins	no ins	no ins	no ins	no ins	A ins	no (1)
1	rs11751	C/A	chr18:5880385	C/C	C/C	C/C	C/C	C/C	C/C	no (1)
1	rs7506026	C/T	chr18:5880571	C/C	C/C	C/C	C/C	C/C	NK	yes (1/1)
1	rs28656885	A/C	chr18:5881637	A/A	A/A	A/A	A/A	A/A	A/A	no (1)
1	A	C/A	chr18:5880589	C/C	C/C	C/A	C/C	C/C	C/C	no (1)
1	B	C/A	chr18:5881636	C/C	C/C	C/A (ho)	C/C	C/C	C/C	no (1)

Table 6.5 Sequencing newly identified exons in new isoforms of four strong candidate genes. Newly identified coding exons from genes *LAMA1*, *ZFP161*, *ARHGAP28* and *L3MBTL4* were sequenced in a panel of 2 affected individuals from family 1, as well as 2 controls. One WT control was a blood relative within family 1 (M1-7), and the other non-related (Q63). The polymorphism designation, their genomic location and heterozygosity (population diversity) are listed. A deletion is indicated by "-/" followed by the nucleotide(s) which is deleted, and an insertion is indicated by "+/" followed by its inserted nucleotide(s).

Gene	Isoform	Exon	Designation	Change	Heterozygosity	Function	Location	Affected Patients		W/T controls	
								M1-13	M1-3	M1-7	Q63
LAMA1	1	1	rs1940970	C/A	0.496 +/- 0.046	intronic	chr18:6946194	C/A	C/A	C/A	C/G
LAMA1	1	1	rs949215	G/T	0.360 +/- 0.224	intronic	chr18:6945676	G/T	G/T	G/G	G/G
LAMA1*	2	1									
ZFP161*	1	1									
ZFP161	1	2	rs11327444	-/A	NK	intronic	chr18:5282218	-/A (ho)	-/A	-/A	-/A
ZFP161	1	2	rs620652	A/G	0.233 +/- 0.249	coding-synon	chr18:5282030	A/G (ho)	A/G	A/G (ho)	A/G (ho)
ZFP161**	2	1									
ZFP161	2	2	rs11665417	G/A	0.432 +/- 0.171	intronic	chr18:5283943	G/G	G/A	G/A	G/A
ARHGAP28	1	17	rs12185460	C/T	0.488 +/- 0.077	intronic	chr18:6888415	C/T	C/T	C/T (ho)	C/T (ho)
ARHGAP28*	2	1									
ARHGAP28*	3	16									
ARHGAP28	3	17	rs1056408	A/C	0.371 +/- 0.219	coding-missense	chr18:6902143	A/C (ho)	A/C (ho)	A/C (ho)	A/C (ho)
ARHGAP28	3	17	rs2163467	C/T	0.322 +/- 0.239	3' UTR	chr18:6902317	C/T (ho)	C/T (ho)	C/T (ho)	C/T (ho)
ARHGAP28	3	17	rs1431390	A/G	0.398 +/- 0.202	3' UTR	chr18:6902332	A/G (ho)	A/G (ho)	A/G (ho)	A/G (ho)
ARHGAP28	3	17	rs4398174	A/G	0.382 +/- 0.212	3' UTR	chr18:6903378	A/G (ho)	A/G (ho)	A/G (ho)	A/G (ho)
ARHGAP28	3	17	rs4486988	T/C	0.359 +/- 0.225	3' UTR	chr18:6903420	T/C (ho)	T/C (ho)	T/C (ho)	T/C (ho)
ARHGAP28	3	17	rs4588059	T/C	NK	3' UTR	chr18:6903447	T/C (ho)	T/C (ho)	T/C (ho)	T/C (ho)
ARHGAP28	3	17	rs7236448	C/T	0.178 +/- 0.239	3' UTR	chr18:6903534	C/T	C/T	C/C	C/C
ARHGAP28	3	17	rs35586134	-/T	NK	3' UTR	chr18:6903627	-/T (ho)	-/T (ho)	-/T (ho)	-/T (ho)
ARHGAP28	3	17	rs1052890	T/C	0.476 +/- 0.107	3' UTR	chr18:6903671	T/C (ho)	T/C (ho)	T/C (ho)	T/C (ho)
ARHGAP28	3	17	rs1052891	A/G	NK	3' UTR	chr18:6903674	A/G (ho)	A/G (ho)	A/G (ho)	A/G (ho)
ARHGAP28	3	17	rs1052892	A/G	0.368 +/- 0.220	3' UTR	chr18:6903716	A/G (ho)	A/G (ho)	A/G (ho)	A/G (ho)
ARHGAP28	3	17	rs1431388	A/G	0.293 +/- 0.246	3' UTR	chr18:6904033	A/G (ho)	A/G (ho)	A/G (ho)	A/G (ho)
ARHGAP28	3	17	rs1431387	G/C	0.380 +/- 0.214	3' UTR	chr18:6904054	G/C (ho)	G/C (ho)	G/C (ho)	G/C (ho)
ARHGAP28	3	17	rs1060012	T/C	0.293 +/- 0.246	3' UTR	chr18:6904458	T/C (ho)	T/C (ho)	T/C (ho)	T/C (ho)
ARHGAP28	3	17	rs34628151	-/A	NK	3' UTR	chr18:6904589	-/A (ho)	-/A (ho)	-/A (ho)	-/A (ho)
ARHGAP28	3	17	rs6506454	A/G	0.290 +/- 0.247	3' UTR	chr18:6905219	A/G (ho)	A/G (ho)	A/G (ho)	A/G (ho)
ARHGAP28	3	17	rs56059909	+/T	NK	3' UTR	chr18:6905526	+/T (ho)	+/T (ho)	+/T (ho)	+/T (ho)
ARHGAP28	3	17	rs16735	-/ACTA	0.401 +/- 0.199	3' UTR	chr18:6905543-546	-/ACTA (ho)	-/ACTA (ho)	-/ACTA (ho)	-/ACTA (ho)
ARHGAP28	3	17	rs1060027	T/C	0.349 +/- 0.230	3' UTR	chr18:6905639	T/C (ho)	T/C (ho)	T/C (ho)	T/C (ho)
L3MBTL4	1	12	rs10670178	+/CAA	NK	3' UTR	chr18:5944818-817	+/CAA	+/CAA	+/CAA	NK
L3MBTL4	1	12	rs3833206	+/AAC	NK	3' UTR	chr18:5944813-812	+/AAC	+/AAC	no ins	NK
L3MBTL4	1	12	rs3811	T/A	0.262 +/- 0.250	3' UTR	chr18:5944871	T/A	T/A	T/T	NK
L3MBTL4	1	12	rs1940621	T/C	0.474 +/- 0.111	3' UTR	chr18:5944987	T/C	T/C	C/C	T/C
L3MBTL4	1	12	rs1940622	G/A	0.474 +/- 0.110	3' UTR	chr18:5945181	G/A	G/A	G/G	G/A
L3MBTL4	1	12	rs719153	A/G	0.701 +/- 0.099	3' UTR	chr18:5945554	A/G	A/G	G/G	A/G
L3MBTL4	1	12	rs2298535	C/T	0.479 +/- 0.101	3' UTR	chr18:5945911	C/T	C/T	C/C	C/T
L3MBTL4	1	12	rs2298534	G/A	0.468 +/- 0.122	3' UTR	chr18:5945932	G/A	G/A	G/G	G/A
L3MBTL4	1	12	rs3737351	G/A	0.477 +/- 0.104	3' UTR	chr18:5946116	G/A	G/A	G/G	G/A
L3MBTL4	1	12	rs3737352	T/C	0.478 +/- 0.103	3' UTR	chr18:5946145	T/C	T/C	T/T	T/C
L3MBTL4	1	12	rs3737353	C/T	0.478 +/- 0.103	coding-fn NK	chr18:5946238	C/T	C/T	C/C	C/T
L3MBTL4*	3	1									

Table 6.6 Specifications of BAC clones used. (A) List of BAC clones used and their genomic start and stop sites. (B) The distances between probes used in each experiment are listed.

A

Clone #	BAC clone	Genomic start	Genomic end
1	RP11-1106L23	3,824,819	4,003,843
2	RP11-398H11	4,680,166	4,884,468
3	RP11-835E18	5,127,620	5,309,190
4	RP11-241F3	5,852,452	6,049,963
5	RP11-431J12	6,286,196	6,458,429
6	RP11-76K13	7,039,214	7,195,350
7	RP11-653O10	7,580,286	7,737,010
8	RP11-434M7	8,382,056	8,542,655

B

Expt #	Clone	Distance (Kb)	Clone	Distance (Kb)	Clone
1	1	678	2	614	4
2	8	627	7	965	5
3	1	919	3	756	5
4	8	1169	6	996	4
5	2	614	4	415	5
6	7	965	5	415	4
7	4	415	5	581	6

Table 6.7 Nuclei counts for 3-color metaphase FISH for each test probe set. Results for a total of 7 FISH experiments are listed. Probes used for each experiment are given in their fluor-labeled color, as well as the probe number in brackets. Two to three metaphase spreads were used for each experiment, and the number of times a specific probe was seen in the centre of its respective probe set was counted. A percentage is determined using the number of nuclei counts in which a probe was seen in the middle, in relation to the total number of nuclei counts.

Expt #		Inversion Test Probe Set (color indicates fluor)			Total nuclei
		RP11-1106L23 (1)	(2)	(4)	
#1	1	30	44	27	
	2	17	30	13	
	3	13	34	9	
		28%	**50%**	**22%**	217
#1 repeat		RP11-1106L23 (1)	(2)	(4)	
	1	35	48	21	
	2	38	48	14	
	3	49	55	24	
		37%	**46%**	**17%**	332
#2		(5)	RP11-653O10 (7)	(8)	
	1	17	76	20	
	2	16	91	30	
		13%	**67%**	**20%**	250
#2 repeat		(5)	RP11-653O10 (7)	(8)	
	1	15	36	20	
	2	7	70	30	
		12%	60%	28%	250
#3		RP11-1106L23 (1)	(3)	(5)	
	1	23	85	26	
	2	8	53	16	
		15%	**65%**	**20%**	211
#4		(4)	RP11-76K13 (6)	(8)	
	1	19	82	15	
	2	15	55	5	
		18%	**72%**	**28%**	191
#5		RP11-398H11 (2)	(4)	(5)	
	1	13	78	16	
	2	3	55	25	
		9%	**70%**	**21%**	190
#5 repeat		RP11-398H11 (2)	(4)	(5)	
	1	6	60	9	
	2	3	59	10	
	3	4	55	12	
		6%	**80%**	**14%**	218
#6		(4)	(5)	RP11-653O10 (7)	
	1	42	47	9	
	2	61	59	11	
	3	38	61	11	
		42%	**49%**	**9%**	339
#6 repeat		(4)	(5)	RP11-653O10 (7)	
	1	61	59	10	
	2	39	62	7	
	3	50	53	8	
		43%	**50%**	**7%**	349
#6 on WT		(4)	(5)	RP11-653O10 (7)	
control	1	26	29	4	
	2	15	26	1	
	3	11	11	0	
		42%	**54%**	**4%**	123
#7		(4)	(5)	RP11-76K13 (6)	
	1	37	46	9	
	2	21	67	10	
	3	38	62	10	
		32%	**58%**	**10%**	299
#7 on WT		(4)	(5)	RP11-76K13 (6)	
control	1	21	23	4	
	2	18	22	6	
	3	11	15	2	
		41%	**49%**	**10%**	122

Figure 6.1 Genes found within the chr18p11 critical region, in relation to the microsatellite markers used to generate the disease haplotype. Schematic representation of the 3.18Mb reduced region for DYT15 on chr18p11. The previous flanking recombinant (R), non-recombinant (NR) microsatellite markers, and the recombinant (red) SNPs are indicated as nucleotide positions of these markers. Black arrows indicate genes that have previously been sequenced; blue indicates those that have not yet been sequenced. Asterisks indicate genes in which the NCBI record has been discontinued. Physical distances are determined by UCSC Genome Browser.

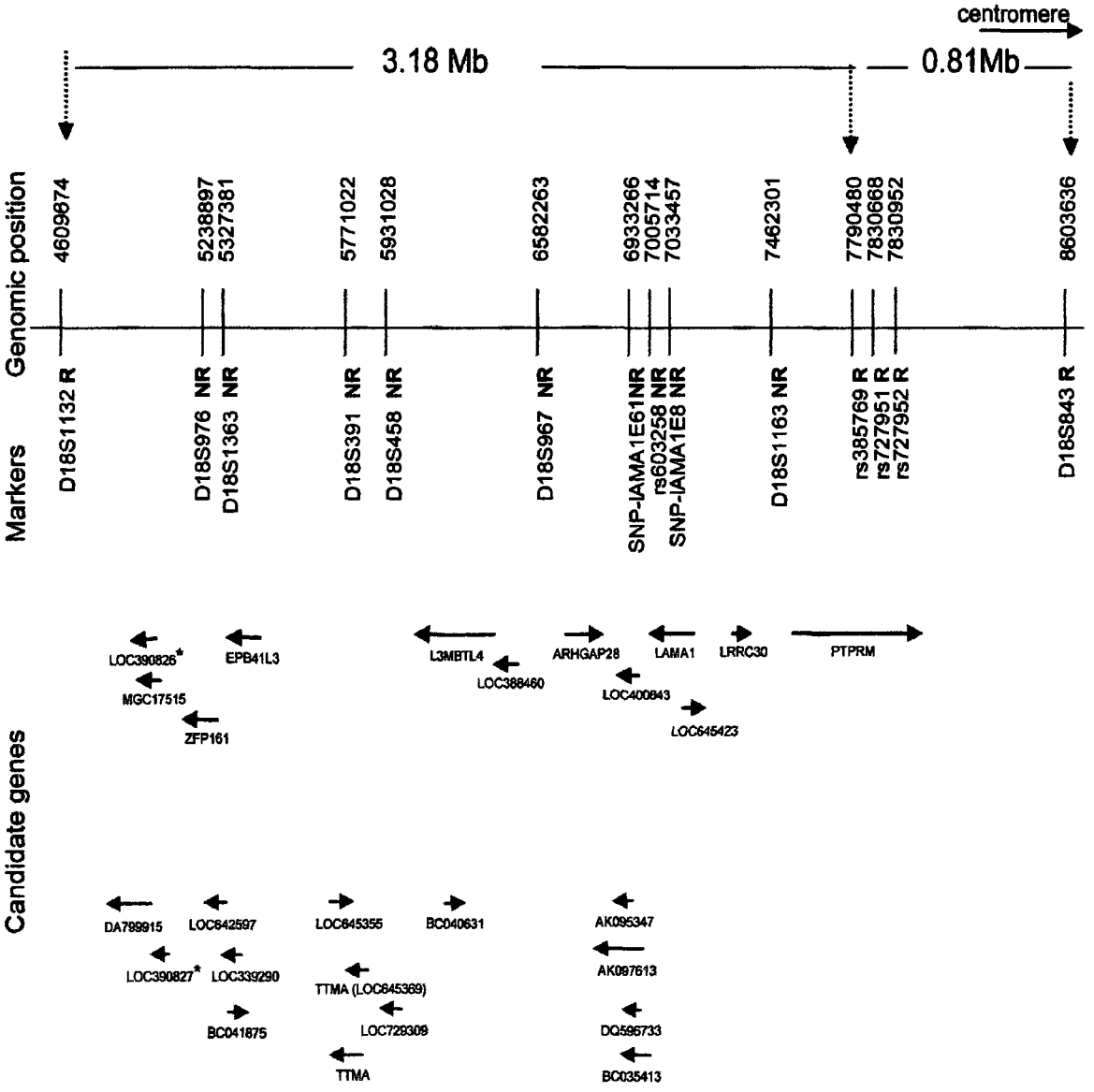
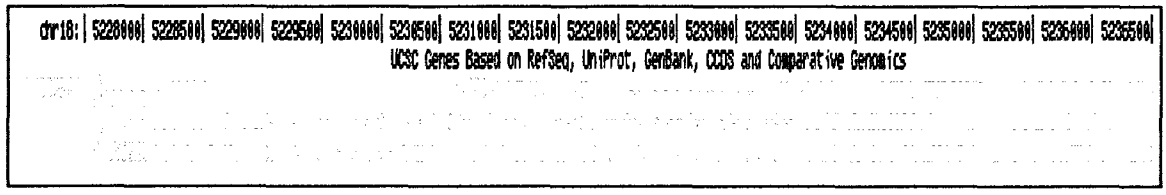

centromere
3.18 Mb
0.81Mb
Genomic position
4609674
5238897
5327381
5771022
5931028
6582263
6933266
7005714
7033457
7462301
7790480
7830668
7830952
8603636
Markers
D18S1132 R
D18S976 NR
D18S1363 NR
D18S391 NR
D18S458 NR
D18S967 NR
SNP-JAMA1E61NR
rs603258 NR
SNP-JAMA1E8 NR
D18S1163 NR
rs385769 R
rs727951 R
rs727952 R
D18S843 R
Candidate genes
LOC390826*
EPB41L3
MGC17515
ZFP161
L3MBTL4
ARHGAP28
LAMA1
LRRC30
PTPRM
LOC388460
LOC400843
LOC645423
DA789915
LOC642597
LOC645355
BC040631
AK095347
LOC390827*
LOC339290
TTMA (LOC645369)
AK097613
BC041875
LOC729309
DQ596733
TTMA
BC035413

Figure 6.2 UCSC representation of the two candidate genes, BC041875 and TTMA, sequenced in MD affected patients. (A) BC041875 is represented and is predicted to have 3 coding exons. It is expressed in fetal brain and a good candidate for sequencing. It is overlapped by three additional predicted genes. (B) TTMA is represented and has one predicted coding exon. It is widely expressed and contains a transcriptional start and stop site. It was also considered to be a good candidate gene for sequencing.

A


chr18: 5228000 5228500 5229000 5229500 5230000 5230500 5231000 5231500 5232000 5232500 5233000 5233500 5234000 5234500 5235000 5235500 5236000 5236500
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics

B

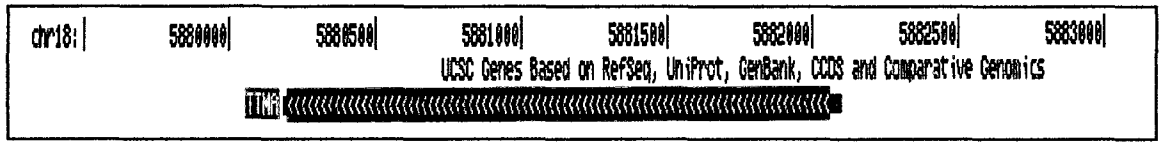
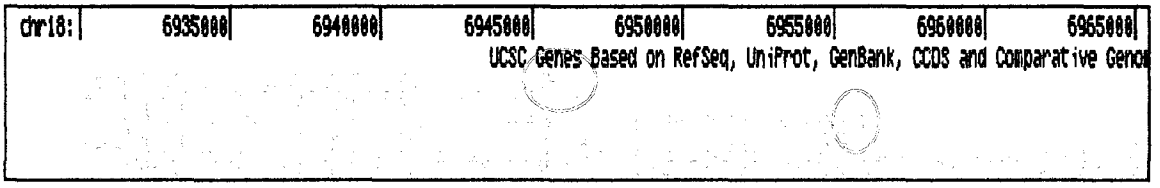
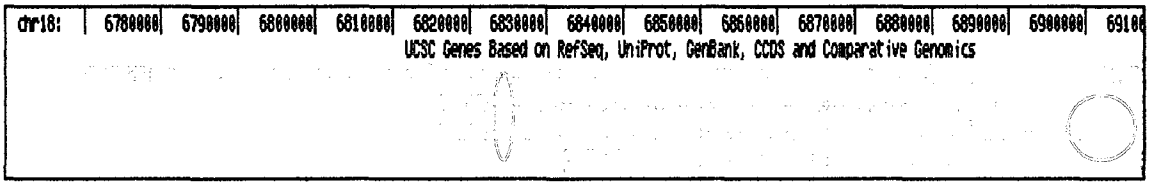

chr18: 5880000 5880500 5881000 5881500 5882000 5882500 5883000
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
TTMP

Figure 6.3 UCSC representation of four candidate genes which have recently had newly identified isoforms reveal additional coding exons. All four strong candidate genes (A – *LAMA1*, B – *ARHGAP28,* C – *L3MBTL4*, D – *ZFP161*) that were found to have newly identified coding exons (circled in red) were sequenced for these exons in a panel of two affected individuals. These representations are taken from UCSC Genome Browser.

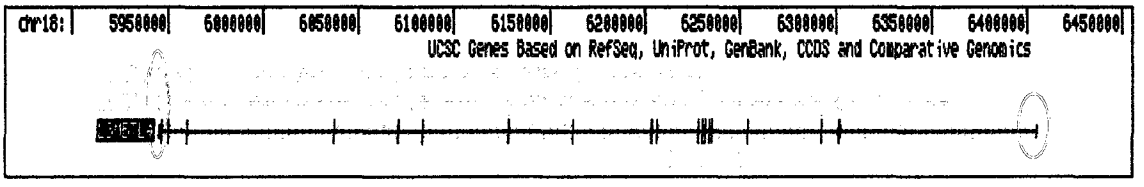
A


chr18:
6935000
6940000
6945000
6950000
6955000
6960000
6965000
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genom

B


chr18:
6780000
6790000
6800000
6810000
6820000
6830000
6840000
6850000
6860000
6870000
6880000
6890000
6900000
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics

C


chr18:
5950000
6000000
6050000
6100000
6150000
6200000
6250000
6300000
6350000
6400000
6450000
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics

D

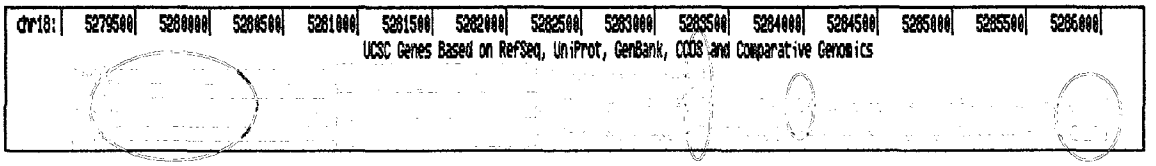
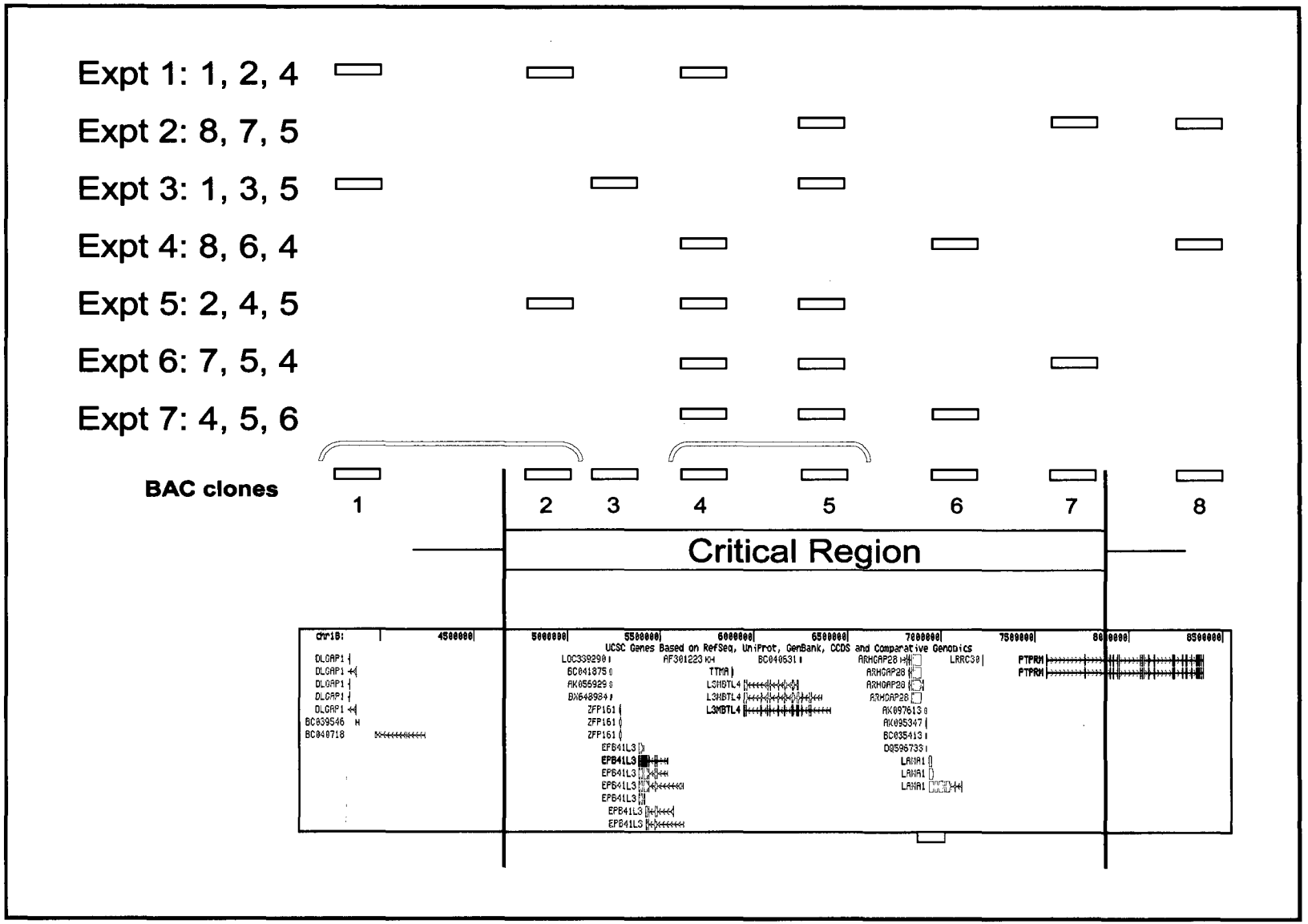

chr18:
5279500
5280000
5280500
5281000
5281500
5282000
5282500
5283000
5283500
5284000
5284500
5285000
5285500
5286000
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics

Figure 6.4 FISH experiment design. The design of 7 FISH experiments is illustrated. Each experiment consists of three fluorescently labeled BAC clones, the particular color of each probe in each experiment is indicated and their location corresponding to the 3.18Mb critical region is shown. Green brackets above the probe locations indicate the probes that are involved in potential inversions. The small green bar under the *LAMA1* gene corresponds to a previously identified inversion.


Expt 1: 1, 2, 4
Expt 2: 8, 7, 5
Expt 3: 1, 3, 5
Expt 4: 8, 6, 4
Expt 5: 2, 4, 5
Expt 6: 7, 5, 4
Expt 7: 4, 5, 6
BAC clones
1
2
3
4
5
6
7
8
Critical Region
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics

Chapter 7.

Future Directions

Chapter 7 Future Directions

7.1 Recruitment and Screening of Myoclonus Dystonia families

Recruitment of additional Mycoclonus Dystonia families is ongoing and on average results in the gain of one family every 1-2 months for our studies. Our current collaborating neurologist will continue to be in contact with movement disorder clinics across the country for this purpose. Newly recruited families will be screened for *SGCE* mutations using PCR and direct sequencing, as well as MLPA for identifying large deletions. The recruitment of large families (and of additional family members of *SGCE*-negative families) could allow us to perform linkage analysis which could either confirm the chr18p11 as a disease locus for MD, or provide additional loci. A study involving the linkage analysis of 5 MD families revealed linkage to the chr18p11 region in one family, but not in the remaining four[50]. These four families did not have linkage to chr7q21 (*SGCE* locus) either, suggesting additional loci could be responsible for MD. Recruiting additional families for our studies is therefore an integral part of identifying the genetic components of Myoclonus Dystonia.

7.2 Elucidating a novel gene in the DYT 15 locus responsible for Myoclonus Dystonia

A total of 39 microsatellite markers were genotyped to complete the 10cM linkage analysis on family 1. Additional informative markers that map to the region surrounding D4S408 will need to be genotyped to confirm the chr18p11 (DYT15) locus as the single disease locus in family 1. Analysis of the various candidate genes using PCR and direct sequencing of their coding regions has not resulted in the identification of a gene responsible

for MD in this family. Since only coding regions are studied, it is possible that a mutation is located within a regulatory region and has been missed. We have sequenced the intron-exon boundaries of the coding exons, therefore we would expect to identify potential donor and acceptor splice-site mutations. It is also possible that a large deletion in one of these candidate genes is responsible for MD. This would not be identified by direct sequencing unless the deletion breakpoint was within the exon coding sequence. However, since we have not observed a large stretch of homozygosity in the genes we have studied, which would indicate a deletion, we can predict that none are present.

Currently, there are a total of 24 candidate genes, both confirmed and predicted, within the chr18p11 region. Ten are yet to be sequenced and are classified as predicted genes. The continuation of candidate gene analysis by exon sequencing has its flaws, as previously mentioned, and is time consuming. Since whole sequencing technologies have advanced and are now increasingly cost efficient, we have decided to employ the use of a sequence capture method followed by high throughput sequencing to analyze the entire 3.18Mb chr18p11 region. This will be performed using extracted genomic DNA from two affected individuals in family 1. A custom derived chip (designed by NimbelGen) containing oligonucleotides specific to the region will be used to hybridize the genomic DNA, thereby allowing for the capturing of our region of interest. A coverage of 60-70% of the region of interest is typical of this chip based capture method. The chip is washed after hybridization and the region-specific DNA is all that remains. This is then denatured and labeled for downstream high-throughput Roche 454 pyrosequencing (performed by TCAG in Toronto). A sequencing fold coverage of 25-30X is expected, allowing for sufficient accuracy of the sequence read. By using a WT reference sequence for comparison, we will

collect a database of all novel variants. These will then be cross referenced to the SNP data we have accumulated through our sequencing of candidate genes. We can then disregard these as disease-causing SNPs, since they have been found in WT patient screening. As well, these SNPs can be used as positive controls to validate this method. The novel variants remaining and shared between the two affected individuals will then be considered candidate variants and will be screened in WT controls.

There are many advantages to using this sequence capture method. Firstly, it is cost and time efficient, with the entire process taking roughly 1 month. In comparison with direct sequencing, this is quite significant. Secondly, this will identify all polymorphisms and novel variants, including point mutations, insertions, deletions, etc. It will also provide coverage of all non-coding regions which could potentially harbor regulatory regions, non-coding RNAs, or even genes that have not yet been identified. This method will also investigate inversions, as a breakpoint would be identified by a sudden change in sequence that maps to a different region of the chromosome. This is advantageous, since we have identified a potential inversion within the chr18p11 region in individual #16 from family 1. One could argue the results from the FISH experiments performed on the region are contradictory, given that experiment 5 does not agree with experiments 6 and 7. As well, the potential inversion involving probes 4 and 5 are also seen in one WT control, and the potential inversion involving probes in experiment 1 is unknown. These lingering uncertainties lend additional support to the use of the proposed sequence capture method. This whole region sequencing would confirm or refute an inversion within the chr18p11 region, and its breakpoints would be specifically mapped allowing us to screen WT controls and subsequently determine if this is a non-pathogenic polymorphic inversion. In

conclusion, we are confident that this method will identify a mutation, and subsequently a gene responsible for Myoclonus Dystonia.

References

1. Vercueil L. Myoclonus and movement disorders. Neurophysiologie Clinique. 2006; 36: 327-331.
2. Zimprich A, Grabowski M, Asmus F, Naumann M, Berg D, Bertram M, Scheidtmann K, Kern P, Winkelmann J, Müller-Myhsok B, Riedel L, Bauer M, Müller T, Castro M, Meitinger T, Strom TM, Gasser T. Mutations in the gene encoding ε-sarcoglycan cause myoclonus-dystonia syndrome. Nature Genetics. 2001; 29: 66-69.
3. Magariños-Ascone C.M., Regidor I., J C Martínez-Castrillo, M Gómez-Galán, Figueiras-Méndez F. Pallidal stimulation relieves myoclonus–dystonia syndrome.
4. De Carvalho Aguiar PM, Ozelius LJ. Classification and genetics of dystonia. Lancet Neurol. 2002; 1(5): 316-25.
5. Klein C, Ozelius LJ, Hagenah J, Breakefield XO, Risch NJ, Vieregge P. Search for a founder mutation in idiopathic focal dystonia from Northern Germany. Am. J. Hum. Genet. 1998; 63(6): 1777-82.
6. Nemeth AH. The genetics of primary dystonias and related disorders. Brain. 2002; 125:695-721.
7. Granata A, Watson R, Collinson L.M., Schiavo G., Warner T.T. The Dystonia-associated Protein TorsinA Modulates Synaptic Vesicle Recycling. J Biol Chem. 2008; 283(12): 7568-79.
8. Augood S.J., Penney J.B., Friberg I. K., Breakefield X. O., Young A.B., Ozelius L.J., Standaert D.G. Expression of the early-onset torsion dystonia gene (DYT1) in human brain. Ann. Neurol. 1998; 43(5): 669-673.
9. Balcioglu A., Kim M.O., Sharma N, Cha J.H., Breakefield X.O., Standaert D.G. Dopamine release is impaired in a mouse model of DYT1 dystonia. J. Neurochem. 2007; 102(3): 783-8.
10. Alterman R.L., Snyder B.J. Deep brain stimulation for torsion dystonia. Acta Neurochir Suppl. 2007; 97:191-199.
11. Ichinose H, Ohye T, Takahashi E, Seki N, Hori T, Segawa M, Nomura Y, Endo K, Tanaka H, Tsuji S. Hereditary progressive dystonia with marked diurnal fluctuation caused by mutations in the GTP cyclohydrolase I gene. Nat. Genet. 1994; 8(3): 236-242.
12. Nemeth A.H. The genetics of primary dystonias and related disorders. Brain. 2002; 125: 695-721.
13. Zesiewicz TA, Sullivan KL, Hauser RA. Levodopa-induced dyskinesia in Parkinson's disease: epidemiology, etiology, and treatment. Curr. Neurol. Neurosci. Rep. 2007; 7(4): 302-310.
14. Klein C, Brin M., Kramer P., Sena-Esteves M., de Leon D., Doheny D, Bressman S, Fahn S, Breakefield X.O., Ozelius L.J. Association of a missense change in the D2 dopamine receptor with myoclonus dystonia. Proc. Natl. Acad. Sci. USA 1999; 96: 5173-5176.
15. Asmus F., Devlin A., Munz M., Zimprich A., Gasser T., Chinnery P.F. Clinical differentiation of genetically proven benign hereditary chorea and Myoclonus-Dystonia. Movement Disorders. 2007; 22(14): 2104-2109.

16. do Carmo Costa M, Costa C, Silva AP, Evangelista P, Santos L, Ferro A, Sequeiros J, Maciel P. Nonsense mutation in TITF1 in a Portuguese family with benign hereditary chorea. Neurogenetics. 2005; 58: 792-797.
17. Nygaard TG, Raymond D, Chen C, Nishino I, Greene PE, Jennings D, Heiman GA, Klein C, Saunders-Pullman RJ, Kramer P, Ozelius LJ, Bressman SB. Localization of a gene for myoclonus-dystonia to chromosome 7q21-q31. Ann Neurol. 1999; 46(5):794-8.
18. Asmus F, Zimprich A, Naumann M, Berg D, Bertram M, Ceballos-Baumann A, Pruszak-Seel R, Kabus C, Dichgans M, Fuchs S, Müller-Myhsok B, Gasser T. Inherited myoclonus-dystonia syndrome: Narrowing the 7q21-q31 locus in German families. Ann Neurol. 2001; 49(1); 121-4.
19. Klein C, Schilling K, Saunders-Pullman RJ, Garrels J, Breakefield XO, Brin MF, deLeon D, Doheny D, Fahn S, Fink JS, Forsgren L, Friedman J, Frucht S, Harris J, Holmgren G, Kis B, Kurlan R, Kyllerman M, Lang AE, Leung J, Raymond D, Robishaw JD, Sanner G, Schwinger E, Tabamo RE, Tagliati M. A major locus for myoclonus-dystonia maps to chromosome 7q in eight families. Am. J. Hum. Genet. 2000; 67: 1314-1319.
20. Hess CW, Raymond D, Aguiar Pde C, Frucht S, Shriberg J, Heiman GA, Kurlan R, Klein C, Bressman SB, Ozelius LJ, Saunders-Pullman R. Myoclonus-dystonia, obsessive-compulsive disorder, and alcohol dependence in SGCE mutation carriers. Neurology. 2007; 68(7): 522-4.
21. Saunders–Pullman R, Shriberg J, Heiman G, Raymond D, Wendt K, Kramer P, Schilling K, Kurlan R, Klein C, Ozelius LJ, Risch NJ, Bressman S.B. Myoclonus dystonia: Possible association with obsessive–compulsive disorder and alcohol dependence. Neurology. 2002; 58(2): 242-245.
22. Tezenas du Montcel S, Clot F, Vidailhet M, Roze E, Damier P, Jedynak CP, Camuzat A, Lagueny A, Vercueil L, Doummar D, Guyant-Maréchal L, Houeto JL, Ponsot G, Thobois S, Cournelle MA, Durr A, Durif F, Echenne B, Hannequin D, Tranchant C, Brice A; French Dystonia Network. Epsilon sarcoglycan mutations and phenotype in French patients with myoclonic syndromes. J. Med. Genet. 2006; 43(5): 394-400.
23. Nardo Nardocci, Giovanna Zorzi, Chiara Barzaghi, Federica Zibordi, Claudia Ciano, Daniele Ghezzi, Barbara Garavaglia. Myoclonus-dystonia syndrome: Clinical presentation, disease course, and genetic features in 11 families. Movement Disorders. 2008; 23(1): 28-34.
24. DeBerardinis RJ, Conforto D, Russell K, Kaplan J, Kollros PR, Zackai EH, Emanuel BS. Myoclonus in a Patient with a deletion of the epsilon-sarcoglycan locus on chromosome 7q21. American Journal of Medical Genetics. 2003; 121A: 31-36.
25. Asmus F, Hjermind LE, Dupont E, Wagenstaller J, Haberlandt E, Munz M, Strom TM, Gasser T. Genomic deletion size at the epsilon-sarcoglycan locus determines the clinical phenotype. Brain. 2007; 130: 2736-2745.
26. Esapa CT, Waite A, Locke M, Benson MA, Kraus M, McIlhinney RA, Sillitoe RV, Beesley PW, Blake DJ. SGCE missense mutations that cause myoclonu-dystonia syndrome impair e-sarcoglycan trafficking to the plasma membrane: modulation by ubiquitination and torsionA. Human Molecular Genetics 2007; 16(3): 327-342.

27. Hedrich K, Meyer EM, Schüle B, Kock N, de Carvalho Aguiar P, Wiegers K, Koelman JH, Garrels J, Dürr R, Liu L, Schwinger E, Ozelius LJ, Landwehrmeyer B, Stoessl AJ, Tijssen MA, Klein C. Myoclonus Dystonia: Detection of novel, recurrent, and *de novo* SGCE mutations. Neurology. 2004; 62: 1229-1231.
28. Gerrits MC, Foncke EM, Koelman JH, Tijssen MA. Pediatric writer's cramp in myoclonus-dystonia: Maternal imprinting hides positive family history. Eur J Paediatr Neurol. 2008; Jun 19 [Epub ahead of print].
29. Ryuichi Ono, Hirosuke Shiura, Hiroyuki Aburatani, Takashi Kohda, Tomoko Kaneko-Ishino, Fumitoshi Ishino. Identification of a Large Novel Imprinted Gene Cluster on Mouse Proximal Chromosome 6. Genome Res. 2003; 13: 1696-1705.
30. Fumiaki Yokoi, Mai Tu Dang, Jianyong Li, and Yuqing Li. Myoclonus, Motor Deficits, Alterations in Emotional Responses and Monoamine Metabolism in epsilon-Sarcoglycan Deficient Mice. J. Biochem. 2006; 140: 141-146.
31. Grabowski M., Zimprich A., Lorenz-Depiereux B., Kalscheuer V., Asmus F., Gasser T., Meitinger T. and Strom, T.M. The epsilon-sarcoglycan gene (SGCE), mutated in myoclonusdystonia syndrome, is maternally imprinted. Eur. J. Hum. Genet. 2003; 11: 138–144.
32. Grabowski M, Zimprich A, Lorenz-Depiereux B, Kalscheuer V, Asmus F, Gasser T, Meitinger T, Strom TM. The ε-sarcoglycan gene (SGCE), mutated in Myoclonus-Dystonia syndrome, is maternally imprinted. Eur. J. Hum. Genet. 2003; 11: 138-144.
33. Suzuki S, Ono R, Narita T, Pask AJ, Shaw G, Wang C, Kohda T, Alsop AE, Marshall Graves JA, Kohara Y, Ishino F, Renfree MB, Kaneko-Ishino T. Retrotransposon silencing by DNA methylation can drive mammalian genomic imprinting. PLOS Genetics. 2007; 3(4): 1-7.
34. Guettard E, Portnoi MF, Lohmann-Hedrich K, Keren B, Rossignol S, Winkler S, El Kamel I, Leu S, Apartis E, Vidailhet M, Klein C, Roze E. Myoclonus-Dystonia due to maternal uniparental disomy. Arch. Neurol. 2008; 65(10): 1380-1385.
35. Muller B., Hedrich K., Kock N., Dragasevic N., Svetel M., Garrels J., Landt O., Nitschke M., Pramstaller P.P., Reik W., Schwinger E., Sperner J, Ozelius L., Kostic V., Klein C. Evidence that paternal expression of the epsilon-sarcoglycan gene accounts for reduced penetrance in Myoclonus-Dystonia. 2002; 71: 1303-1311.
36. Ettinger AJ, Feng G, Sanes JR. epsilon-Sarcoglycan, a broadly expressed homologue of the gene mutated in limb-girdle muscular dystrophy 2D. J. Biol. Chem. 1997; 272(51): 32534-8.
37. Volker Straub, Audrey J. Ettinger, Madeleine Durbeej, David P. Venzke, Susan Cutshall, Joshua R. Sanes, and Kevin P. Campbell. Epsilon-sarcoglycan replaces alpha-sarcoglycan in smooth muscle to form a unique dystrophin-glycoprotein complex. J. Biol. Chem., 1999; 274: 27989-96.
38. Michihiro Imamura, Kenji Araishi, Satoru Noguchi, and Eijiro Ozawa. A sarcoglycan–dystroglycan complex anchors Dp116 and utrophin in the peripheral nervous system. Hum. Mol. Genet. 2000; 9: 3091-3100.
39. Cai H, Erdman RA, Zweier L, Chen J, Shaw JH 4th, Baylor KA, Stecker MM, Carey DJ, Chan YM. The sarcoglycan complex in Schwann cells and its role in myelin stability. Exp Neurol. 2007; 205(1): 257-69.

40. Nishiyama A, Endo T, Takeda S, Imamura M. Identification and characterization of epsilon-sarcoglycans in the central nervous system. Brain Res. Mol. Brain Res. 2004; 125(1-2): 1-12.
41. Araishi K, Sasaoka T, Imamura M, Noguchi S, Hama H, Wakabayashi E, Yoshida M, Hori T, Ozawa E. Loss of the sarcoglycan complex and sarcospan leads to muscular dystrophy in beta-sarcoglycan-deficient mice. Hum. Mol. Genet. 1999; 8: 1589-1598.
42. Yoshida M, Hama H, Ishikawa-Sakurai M, Imamura M, Mizuno Y, Araishi K, Wakabayashi-Takai E, Noguchi S, Sasaoka T, Ozawa E. Biochemical evidence for association of dystrobrevin with the sarcoglycan-sarcospan complex as a basis for understanding sarcoglycanopathy. Hum. Mol. Genet. 2000; 9: 1033-1040.
43. Michael J. Allikian, Elizabeth M. McNally. Processing and Assembly of the Dystrophin Glycoprotein Complex. Traffic. 2007; 8(3): 177-183.
44. Guyon JR, Kudryashova E, Potts A, Dalkilic I, Brosius MA, Thompson TG, Beckmann JS, Kunkel LM, Spencer MJ. Calpain 3 cleaves filamin C and regulates its ability to interact with gamma- and delta-sarcoglycans. Muscle Nerve. 2003; 28(4): 472-83.
45. Shi W., Chen Z., Schottenfeld J., Stahl R.C, Kunkel L.M., Chan Y.M. Specific assembly pathway of sarcoglycans is dependent on beta- and delta-sarcoglycan. Muscle Nerve 2004; 29: 409–419.
46. Chen J., Shi W., Zhang Y., Sokol R., Cai H., Hun M., Moore B.F., Farber M.J., Stepanchick J.S., Bonnemann C.G., Chan, Y.M. Identification of functional domains in sarcoglycans essential for their interaction and plasma membrane targeting. Experimental Cell Research. 2006; 312: 1610-1625.
47. Jianfeng Xiao, Mark S. LeDoux. Cloning, developmental regulation and neural localization of rat epsilon-sarcoglycan. Molecular Brain Research. 2003; 119(2): 132-143.
48. Hjermind LE, Vissing J, Asmus F, Krag T, Lochmüller H, Walter MC, Erdal J, Blake DJ, Nielsen JE. No muscle involvement in myoclonus-dystonia caused by epsilon-sarcoglycan gene mutations. Eur J Neurol. 2008; 15(5): 525-9.
49. Yokoi F, Dang MT, Mitsui S, Li Y. Exclusive paternal expression and novel alternatively spliced variants of epsilon-sarcoglycan mRNA in mouse brain. FEBS Lett. 2005; 29; 579(21): 4822-8.
50. Schule B., Kock N., Svetel M., Dragasevic N., Hedrich K., de Carvalho Aguiar P., Liu L., Kabakci K., Garrels J., Meyer E.M., Berisavac I., Schwinger E., Kramer P.L., Ozelius L.J., Klein C., Kostic V. Genetic heterogeneity in ten families with Myoclonus Dystonia. J Neurol Psychiatry. 2004; 75: 1181-1185.
51. Han F., Lang A.E., Racacho L., Bulman D.E., Grimes, D.A. Mutations in the epsilon-sarcoglycan gene found to be uncommon in seven Myoclonus-Dystonia families. Neurology 2003; 61: 244-246.
52. Grünewald A, Djarmati A, Lohmann-Hedrich K, Farrell K, Zeller JA, Allert N, Papengut F, Petersen B, Fung V, Sue CM, O'Sullivan D, Mahant N, Kupsch A, Chuang RS, Wiegers K, Pawlack H, Hagenah J, Ozelius LJ, Stephani U, Schuit R, Lang AE, Volkmann J, Münchau A, Klein C. Myoclonus-Dystonia: Significance of Large *SGCE* Deletions. Hum. Mutat. 2008; 29(2): 331-2.

53. Han F, Racacho L, Yang H, Read T, Suchowersky O, Lang AE, Grimes DA, Bulman DE. Large deletions account for an increasing number of mutations in SGCE. Mov Disord. 2008; 15; 23(3):456-60.
54. Grimes D.A., Han F., Lang A.E., St. George-Hyssop P., Racacho, L., Bulman D.A. A novel locus for inherited Myoclonus-Dystonia on 18p11. Neurology. 2002; 59: 1183-1186.
55. Han F., Racacho L., Lang A.E., Bulman D.E., Grimes D.A. Refinement of the DYT15 locus in Myoclonus Dystonia. Mov. Disord. 2007; 22(6): 888-892.
56. Naïmi M, Tardieu S, Depienne C, Ruberg M, Brice A, Dubourg O, Leguern E. Detection of genomic rearrangements by DHPLC: a prospective study of 90 patients with inherited peripheral neuropathies associated with 17p11.2 rearrangements. Am. J. of Med. Genet. 2005; 136: 136-139.
57. Asmus F, Zimprich A, Tezenas Du Montcel S, Kabus C, Deuschl G, Kupsch A, Ziemann U, Castro M, Kühn AA, Strom TM, Vidailhet M, Bhatia KP, Dürr A, Wood NW, Brice A, Gasser T. Myoclonus-Dystonia syndrome: epsilon-sarcoglycan mutations and phenotype. Ann. Neurol. 2002; 52: 489-492.
58. Asmus F, Salih F, Hjermind LE, Ostergaard K, Munz M, Kühn AA, Dupont E, Kupsch A, Gasser T. Myoclonus-Dystonia due to genomic deletions in the epsilon-sarcoglycan gene. Ann. Neurol. 2005; 58: 792-797.
59. Häger M, Bigotti MG, Meszaros R, Carmignac V, Holmberg J, Allamand V, Akerlund M, Kalamajski S, Brancaccio A, Mayer U, Durbeej M. Cib2 binds integrin alpha7Bbeta1D and is reduced in laminin alpha2 chain-deficient muscular dystrophy. J. Biol. Chem. 2008; 283(36): 24760-24769.
60. Pickard B.S., Malloy M.P., Clark L., Lehellard S., Ewald H.L., Mors O., Porteous D.J., Blackwood D.H., Muir W.J. Candidate psychiatric illness genes identified in patients with pericentric inversions of chromosome 18. Psychiatric Genetics 2005; 15(1): 37-44.
61. Gu YC, Kortesmaa J, Tryggvason K, Persson J, Ekblom P, Jacobsen SE, Ekblom M. Laminin isoform-specific promotion of adhesion and migration of human bone marrow progenitor cells. Blood. 2003; 101(3): 877-885.
62. Lee K.H., Kwak Y.D., Kim D.H., Chang M.Y., Lee Y.S. Human zinc finger protein 161, a novel transcriptional activator of the dopamine transporter. Biochem. Biophys. Res. Commun. 2004; 313(4): 969-976.
63. Nagase T., Kikuno R., Ishikawa K.I., Hirosawa M., Ohara O. Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro. DNA Res. 2000; 7(1): 65-73.
64. Ota T, Suzuki Y, Nishikawa T, Otsuki T, Sugiyama T, Irie R, Wakamatsu A, Hayashi K, Sato H, Nagai K, Kimura K, Makita H, Sekine M, Obayashi M, Nishi T, Shibahara T, Tanaka T, Ishii S, Yamamoto J, Saito K, Kawai Y, *et al.* Complete sequencing and characterization of 21,243 full-length human cDNAs. Nat. Genet. 2004; 36(1): 40-45.
65. Grimes DA, Bulman D, George-Hyslop PS, Lang AE. Inherited Myoclonus-Dystonia: evidence supporting genetic heterogeneity. Mov. Disord. 2001; 16(1): 106-110.

Appendix 1. Raw peak data from High Performance Liquid Chromatography (HPLC) triplicate experiments of *SGCE* exons. Exon dosage analysis of all *SGCE* exons was performed on 8 individuals affected with Myoclonus Dystonia. The peak area under the HPLC curve is given for the *SGCE* exon and a control exon (BX648242 – exon 7), which are then used calculate the ratio of each triplicate. These ratios are then normalized using the average of 10 WT control samples of each triplicate. An average of these three ratios gives the normalized ratio for the *SGCE* exon for the MD patient. A ratio between 0.75 and 1.25 is considered to be the normal dosage range, whereas a ratio above or below this is indicative of a potential duplication or deletion, respectively. Box with slash, indicates the value is not consistent with the other two values in the triplicate for this individual, and was removed from normalization calculation.

SGCE - EXON 1

	Triplicate									Average	Normalized ratios			
	1			2			3							
Sample	Exon 1	Int Control	Ratio	Exon 1	Int Control	Ratio	Exon 1	Int Control	Ratio	Ratio	1	2	3	Average
MF-19	45992	35303	1.303	39947	28540	1.400	38900	28055	1.387	1.363	0.842	0.829	0.91	0.860
MF-20	41236	24340	1.694	40130	28179	1.424	42309	30301	1.396	1.505	1.094	0.844	0.916	0.951
MF-22	32167	20851	1.543	30958	20321	1.523	31033	21244	1.461	1.509	0.997	0.902	0.959	0.953
MF-23	45791	25533	1.793	44311	26413	1.678	40821	21481	1.900	1.790	1.158	0.994	1.247	1.133
MF-24	21574	13841	1.559	25443	17442	1.459	21349	12987	1.644	1.554	1.007	0.864	1.079	0.983
MF-25	14311	ND	ND	23816	14130	1.685	19352	11643	1.662	1.674	ND	0.998	1.091	1.045
MF-26	48015	27595	1.740	47260	21768	2.171	49021	28705	1.708	1.873	1.124	1.286	1.121	1.177
MF-27	22335	14151	1.578	22870	14850	1.540	22705	15283	1.486	1.535	1.019	0.912	0.975	0.969
control 1	23528	15777	1.491	18362	2745	4.803	24012	15606	1.539	1.515				
control 2	25862	18484	1.399	26691	19830	1.346	22742	17090	1.331	1.359				
control 3	33822	25630	1.320	35558	20622	1.724	35337	25096	1.408	1.484				
control 4	50616	29299	1.728	40007	21306	1.878	39047	21342	1.830	1.812				
control 5	29698	19537	1.520	19413	16232	1.196	37279	26310	1.417	1.378				
control 6	39251	23353	1.681	40975	21198	1.933	35778	23259	1.538	1.717				
control 7	27567	18305	1.506	29419	18355	1.603	26340	17276	1.525	1.544				
control 8	23127	15150	1.527	22681	14049	1.614	23414	15883	1.474	1.538				
control 9	29976	19960	1.502	25618	14851	1.725	28511	19551	1.458	1.562				
control 10	59392	32880	1.806	39101	17963	2.177	41734	24239	1.722	1.902				
SGCE Average			1.525			1.528			1.575	1.566				
Control Average			1.549			1.971			1.525	1.575				

SGCE - EXON 2

	Triplicate									Average	Normalized ratios			
	1			2			3							
Sample	Exon 2	Int Control	Ratio	Exon 2	Int Control	Ratio	Exon 2	Int Control	Ratio	Ratio	1	2	3	Average
MF-19	24154	30869	0.782	25691	31296	0.821	26848	31798	0.844	0.816	0.804	0.886	0.955	0.882
MF-20	33966	36126	0.940	33658	36478	0.923	27575	29781	0.926	0.930	0.966	0.996	1.048	1.003
MF-22	21037	25182	0.835	14479	24060	0.602	18520	24242	0.764	0.800	0.858	ND	0.864	0.861
MF-23	30498	33557	0.909	16517	26731	0.618	27187	31961	0.851	0.880	0.934	ND	0.963	0.949
MF-24	13818	15810	0.874	14988	18569	0.807	13062	16591	0.787	0.823	0.898	0.871	0.89	0.886
MF-25	16222	16632	0.975	14797	15604	0.948	11468	14354	0.799	0.908	1.002	1.023	0.904	0.976
MF-26	33427	33597	0.995	31289	34924	0.896	26169	31452	0.832	0.908	1.022	0.967	0.941	0.977
MF-27	17234	17539	0.983	16089	17494	0.920	13820	16334	0.846	0.916	1.01	0.992	0.957	0.986
control 1	17414	18316	0.951	16262	18450	0.881	16301	19508	0.836	0.889				
control 2	18272	12915	1.415	19672	20585	0.956	17588	17910	0.982	0.969				
control 3	31088	31552	0.985	31295	29922	1.046	28195	28320	0.996	1.009				
control 4	43385	34935	1.242	28638	26304	1.089	17636	22888	0.771	1.034				
control 5	16506	17984	0.918	20433	22065	0.926	21261	22282	0.954	0.933				
control 6	23991	26475	0.906	26793	30516	0.878	24351	28456	0.856	0.880				
control 7	18790	20127	0.934	17718	20939	0.846	16253	21099	0.770	0.850				
control 8	17533	17593	0.997	16276	18306	0.889	16780	16636	1.009	0.965				
control 9	18151	21254	0.854	19695	24613	0.800	17049	22031	0.774	0.809				
control 10	28938	29912	0.967	29350	30514	0.962	23408	26258	0.891	0.940				
SGCE Average			0.868			0.754			0.834	0.850				
Control Average			1.034			0.939			0.864	0.929				

SGCE - EXON 3

	Triplicate									Average	Normalized ratios			
	1			2			3							
Sample	Exon 3	Int Control	Ratio	Exon 3	Int Control	Ratio	Exon 3	Int Control	Ratio	Ratio	1	2	3	Average
MF-19	35741	31687	1.128	31194	35560	0.877	29184	33197	0.879	0.961	1.205	0.919	0.906	1.010
MF-20	36335	32776	1.109	34548	31174	1.108	35938	33765	1.064	1.094	1.185	1.161	1.097	1.148
MF-22	24568	23127	1.062	30058	27607	1.089	24928	23163	1.076	1.076	1.135	1.142	1.109	1.129
MF-23	28267	32756	0.863	29148	33085	0.881	25759	30916	0.833	0.859	0.922	0.923	0.859	0.901
MF-24	19296	17608	1.096	16770	17828	0.941	15880	16585	0.957	0.998	1.171	0.986	0.987	1.048
MF-25	15238	14810	1.029	10327	13453	0.768	16318	16685	0.978	0.925	1.099	0.805	1.008	0.971
MF-26	37831	34885	1.084	37324	37530	0.995	36019	34021	1.059	1.046	1.158	1.043	1.092	1.098
MF-27	15524	16773	0.926	18256	19965	0.914	16220	18863	0.860	0.900	0.989	0.958	0.887	0.945
control 1	15342	27179	0.564	15633	29097	0.537	16244	29683	0.547	0.550				
control 2	14540	23286	0.624	12674	20149	0.629	16045	26825	0.598	0.617				
control 3	40158	43710	0.919	28162	28460	0.990	31333	32104	0.976	0.961				
control 4	29176	25949	1.124	27925	26670	1.047	32833	31945	1.028	1.066				
control 5	21605	20507	1.054	26690	22351	1.194	24402	23429	1.042	1.096				
control 6	23403	25510	0.917	26127	29306	0.892	37079	31681	1.170	0.993				
control 7	21336	19896	1.072	23000	22508	1.022	14020	14946	0.938	1.011				
control 8	19145	19061	1.004	19314	17979	1.074	22160	20663	1.072	1.050				
control 9	21237	22604	0.940	23835	25389	0.939	25297	22907	1.104	0.994				
control 10	37931	33187	1.143	36705	30065	1.221	39115	31883	1.227	1.197				
SGCE Average			1.052			0.979			0.962	0.998				
Control Average			0.931			0.899			0.920	0.917				

SGCE - EXON 4

	Triplicate									Average	Normalized ratios			
	1			2			3							
Sample	Exon 4	Int Control	Ratio	Exon 4	Int Control	Ratio	Exon 4	Int Control	Ratio	Ratio	1	2	3	Average
MF-19	15832	9675	1.636	13268	7299	1.818	26265	17582	1.494	1.649	1.999	2.470	1.974	2.148
MF-20	16332	8625	1.894	16001	8919	1.794	14997	9003	1.666	1.784	2.313	2.437	2.201	2.317
MF-22	10052	12938	0.777	9560	9986	0.957	20634	19416	1.063	0.932	0.949	1.301	1.404	1.218
MF-23	ND	ND	ND	7687	9860	0.780	11963	12214	0.979	0.880	ND	1.059	1.294	1.177
MF-24	8702	9061	0.960	8058	8939	0.901	12099	11696	1.034	0.965	1.173	1.225	1.367	1.255
MF-25	8684	8510	1.020	9029	8991	1.004	11602	9535	1.217	1.080	1.247	1.364	1.428	1.346
MF-26	9373	9450	0.992	10456	10717	0.976	14685	13164	1.116	1.028	1.212	1.326	1.474	1.337
MF-27	5932	11102	0.534	12914	9986	1.293	25950	29171	0.890	0.906	0.653	1.757	1.175	1.195
control 1	16532	24747	0.668	14789	28585	0.517	30019	37763	0.795	0.660				
control 2	17992	24343	0.739	19597	27334	0.717	27581	33482	0.824	0.760				
control 3	16033	20325	0.789	20593	26965	0.764	23480	27762	0.846	0.799				
control 4	16528	19989	0.827	16475	22106	0.745	22221	29170	0.762	0.778				
control 5	18069	22259	0.812	19028	23654	0.804	21155	26631	0.794	0.804				
control 6	26835	31100	0.863	19566	27249	0.718	24000	30834	0.778	0.786				
control 7	20612	25083	0.822	22615	27429	0.824	23979	33327	0.720	0.789				
control 8	19583	31885	0.614	24094	34316	0.702	26447	42427	0.623	0.647				
control 9	11196	10992	1.019	27374	31834	0.860	14393	21504	0.669	0.849				
control 10	10785	10441	1.033	16009	22610	0.708	ND	ND	ND	0.870				
SGCE Average			1.116			1.190			1.182	1.153				
Control Average			0.818			0.736			0.757	0.774				

SGCE - EXON 5

	Triplicate									Average	Normalized Ratios			
	1			2			3							
Sample	Exon 5	Int Control	Ratio	Exon 5	Int Control	Ratio	Exon 5	Int Control	Ratio	Ratio	1	2	3	Average
MF-19	38824	24661	1.574	33023	19917	1.658	33824	21924	1.543	1.592	1.037	1.065	1.008	1.037
MF-20	40089	22116	1.813	42423	26364	1.609	36244	21397	1.694	1.705	1.194	1.033	1.101	1.109
MF-22	39575	25453	1.555	31826	19660	1.619	23928	15027	1.592	1.589	1.024	1.04	1.035	1.033
MF-23	46485	26939	1.726	46602	27186	1.714	35087	20977	1.673	1.704	1.137	1.101	1.088	1.109
MF-24	22939	15136	1.516	21280	14751	1.443	20017	13369	1.497	1.485	0.999	0.927	0.973	0.966
MF-25	22630	14083	1.607	18571	11284	1.646	18401	11735	1.568	1.607	1.059	1.057	1.02	1.045
MF-26	53959	31909	1.691	46706	27446	1.702	25947	14733	1.761	1.718	1.114	1.093	1.145	1.117
MF-27	22085	13700	1.612	20574	13615	1.511	21143	~~9795~~	~~2.158~~	1.562	1.062	0.97	ND	1.016
control 1	20549	23548	0.873	19900	22587	0.881	15558	18208	0.854	0.869				
control 2	16630	18936	0.878	18584	20474	0.908	17129	18542	0.924	0.903				
control 3	20530	12359	1.661	18997	10491	1.811	19632	10316	1.903	1.792				
control 4	24716	15304	1.615	23358	15499	1.507	24492	16023	1.529	1.550				
control 5	5711	ND	ND	4105	ND	ND	4359	ND	ND	ND				
control 6	23222	11906	1.950	21341	10565	2.020	22673	11653	1.946	1.972				
control 7	15215	9213	1.651	11660	5697	2.047	13331	7356	1.812	1.837				
control 8	10268	5357	1.917	8921	~~2436~~	~~3.663~~	8085	4602	1.757	1.837				
control 9	22030	16404	1.343	16779	11412	1.470	18045	13281	1.359	1.391				
control 10	52040	29397	1.770	42509	23456	1.812	39920	22699	1.759	1.780				
SGCE Average			1.637			1.609			1.600	1.615				
Control Average			1.583			1.641			1.691	1.613				

SGCE - EXON 6

	Triplicate									Average	Normalized Ratios			
	1			2			3							
Sample	Exon 6	Int Control	Ratio	Exon 6	Int Control	Ratio	Exon 6	Int Control	Ratio	Ratio	1	2	3	Average
MF-19	26974	9529	2.831	25444	7952	3.200	21665	9241	2.344	2.792	1.312	1.593	1.440	1.448
MF-20	27181	7284	3.732	24809	ND	ND	26033	8142	3.197	3.464	1.730	ND	1.964	1.847
MF-22	31919	13687	2.332	25678	9514	2.699	22186	9412	2.357	2.463	1.081	1.344	1.448	1.291
MF-23	27291	10342	2.639	23256	7319	3.177	23421	8914	2.627	2.815	1.223	1.582	1.614	1.473
MF-24	29172	12888	2.264	22784	8296	2.746	7254	ND	ND	2.505	1.049	1.368	ND	1.208
MF-25	19147	8085	2.368	17886	5825	3.071	18089	6712	2.695	2.711	1.098	1.529	1.655	1.427
MF-26	12414	6984	1.777	23390	8540	2.739	12468	7230	1.724	2.080	0.824	1.364	1.059	1.082
MF-27	22257	10215	2.179	26725	12942	2.065	17777	10158	1.750	1.998	1.010	1.028	1.075	1.038
control 1	24126	12649	1.907	24313	14604	1.665	30340	19483	1.557	1.710				
control 2	29088	13773	2.112	18146	9486	1.913	26241	16708	1.571	1.865				
control 3	28987	13380	2.166	22752	9953	2.286	20598	11091	1.857	2.103				
control 4	25820	11572	2.231	24610	11622	2.118	14497	7498	1.933	2.094				
control 5	27166	12458	2.181	24118	11268	2.140	26580	15379	1.728	2.016				
control 6	31386	15826	1.983	24359	12625	1.929	17296	10985	1.575	1.829				
control 7	26938	13762	1.957	22350	12277	1.820	24732	15592	1.586	1.788				
control 8	24602	14961	1.644	ND	ND	ND	20630	15993	1.290	1.467				
control 9	19704	6764	2.913	20612	9452	2.181	15677	8954	1.751	2.282				
control 10	21608	8723	2.477	20160	9982	2.020	18010	12549	1.435	1.977				
SGCE Average			2.515			2.814			2.385	2.603				
Control Average			2.157			2.008			1.628	1.913				

SGCE - EXON 7

	Triplicate									Average	Normalized Ratios			
	1			2			3							
Sample	Exon 7	Int Control	Ratio	Exon 7	Int Control	Ratio	Exon 7	Int Control	Ratio	Ratio	1	2	3	Average
MF-19	15036	8369	1.797	4923	ND	ND	9220	5141	1.794	1.795	1.366	ND	1.477	1.422
MF-20	ND	ND	ND	15867	7814	2.031	7145	ND	ND	2.031	ND	1.524	ND	1.524
MF-22	11385	8319	1.369	25444	21033	1.210	26137	21888	1.194	1.257	1.04	0.907	0.983	0.977
MF-23	ND	ND	ND	15025	11185	1.343	5414	ND	ND	1.343	ND	1.008	ND	1.008
MF-24	ND	ND	ND	15356	10672	1.439	5414	5979	0.906	1.172	ND	1.08	0.746	0.913
MF-25	6772	ND	ND	11672	8862	1.317	4599	ND	ND	1.317	ND	0.988	ND	0.988
MF-26	17219	14002	1.230	28238	23939	1.180	11595	9197	ND	1.205	0.935	0.885	ND	0.910
MF-27	24017	14537	1.652	48581	28893	1.681	6152	ND	ND	1.667	1.255	1.261	ND	1.258
control 1	10238	8187	1.251	12335	8835	1.396	ND	ND	ND	1.323				
control 2	11766	8933	1.317	3792	ND	ND	9411	6773	1.390	1.353				
control 3	6032	6199	0.973	29687	27831	1.067	18524	15053	1.231	1.090				
control 4	19196	12826	1.497	37029	30924	1.197	17218	11779	1.462	1.385				
control 5	18484	13161	1.404	13643	9192	1.484	21081	14974	1.408	1.432				
control 6	26524	21510	1.233	27449	24616	1.115	ND	9827	ND	1.174				
control 7	15222	10966	1.388	14259	10243	1.392	6659	11327	0.588	1.123				
control 8	27111	19544	1.387	36163	26371	1.371	ND	ND	ND	1.379				
control 9	17391	12932	1.345	11787	7775	1.516	14646	14032	1.044	1.301				
control 10	20797	14458	1.438	19326	13278	1.455	20594	14900	1.382	1.425				
SGCE Average			1.797			1.468			1.494	1.486				
Control Average			1.325			1.320			1.154	1.313				

SGCE - EXON 8

	Triplicate									Average	Normalized Ratios			
	1			2			3							
Sample	Exon 8	Int Control	Ratio	Exon 8	Int Control	Ratio	Exon 8	Int Control	Ratio	Ratio	1	2	3	Average
MF-19	7810	5719	1.366	32727	29041	1.127	45475	46265	0.983	1.159	1.339	1.076	0.846	1.087
MF-20	16703	18293	0.913	34631	37466	0.924	43328	39179	1.106	0.981	0.895	0.883	0.952	0.910
MF-22	17114	8894	1.924	22064	14702	1.501	34815	17691	1.968	1.798	1.886	1.434	1.694	1.671
MF-23	21383	23221	0.921	25539	26533	0.963	22517	21462	1.049	0.978	0.903	0.92	0.903	0.909
MF-24	21995	27999	0.786	24450	23855	1.025	15554	14498	1.073	0.961	0.771	0.979	0.923	0.891
MF-25	15904	16087	0.989	19905	20408	0.975	12597	11610	1.085	1.016	0.97	0.931	0.934	0.945
MF-26	21305	23287	0.915	21693	21128	1.027	17047	14996	1.137	1.026	0.897	0.981	0.978	0.952
MF-27	42427	57814	0.734	28221	35946	0.785	36354	32197	1.129	0.883	0.72	0.75	0.972	0.814
control 1	23470	25928	0.905	9250	8136	1.137	13186	11293	1.168	1.070				
control 2	28813	27469	1.049	27495	26635	1.032	24592	24083	1.021	1.034				
control 3	24590	26054	0.944	24270	27501	0.883	23726	21174	1.121	0.983				
control 4	40955	43251	0.947	33439	28135	1.189	21832	18226	1.198	1.111				
control 5	15100	14993	1.007	17810	17520	1.017	477	ND	ND	1.012				
control 6	13192	11735	1.124	12448	12747	0.977	5810	4922	1.18	1.094				
control 7	36146	38988	0.927	31191	33755	0.924	35560	33152	1.073	0.975				
control 8	11248	10695	1.052	8112	7223	1.123	16706	14616	1.143	1.106				
control 9	18702	15860	1.180	17168	16032	1.071	10247	8077	1.269	1.173				
control 10	18848	17637	1.069	18681	16666	1.121	10466	8160	1.283	1.158				
SGCE Average			1.069			1.041			1.191	1.100				
Control Average			1.020			1.047			1.162	1.072				

SGCE - EXON 8 (Repeat)

	Triplicate									Average	Normalized Ratios			
	1			2			3							
Sample	Exon 8	Int Control	Ratio	Exon 8	Int Control	Ratio	Exon 8	Int Control	Ratio	Ratio	1	2	3	Average
MF-22	25274	30406	0.831	21642	24177	0.895	26129	33855	0.772	0.833	0.9	0.927	0.856	0.894
MF-19	18094	16217	1.116	16888	13183	1.281	20419	17878	1.142	1.180	1.2	1.327	1.269	1.265
MF-25	18172	22867	0.795	15680	19015	0.825	18942	24391	0.777	0.799	0.861	0.855	0.861	0.859
MF-27	37177	42574	0.873	33271	40528	0.821	41195	51321	0.803	0.832	1.054	0.851	0.895	0.933
control 1	28319	33231	0.852	28548	28554	1.000	29978	34644	0.865	0.906				
control 2	25271	27347	0.924	28932	32538	0.889	30571	33369	0.916	0.910				
control 3	25202	28832	0.874	28721	31011	0.926	28339	29962	0.946	0.915				
control 4	26410	25363	1.041	30294	28946	1.047	32135	36487	0.881	0.990				
SGCE Average			0.973			1.088			0.957	1.006				
Control Average			0.861			0.884			0.840	0.862				

SGCE - EXON 9

	Triplicate									Average	Normalized Ratios			
	1			2			3							
Sample	Exon 9	Int Control	Ratio	Exon 9	Int Control	Ratio	Exon 9	Int Control	Ratio	Ratio	1	2	3	Average
MF-19	20932	16413	1.275	29978	25313	1.184	24943	21760	1.146	1.202	1.138	1.097	1.099	1.111
MF-20	32712	27273	1.199	29956	30428	0.985	33316	33517	0.994	1.059	1.071	0.913	0.953	0.979
MF-22	20130	17783	1.132	20886	17787	1.174	24566	21902	1.122	1.143	0.943	1.088	1.076	1.036
MF-23	18724	15155	1.236	24060	21483	1.120	25737	24940	1.032	1.129	1.104	1.038	0.989	1.044
MF-24	15260	11649	1.310	21337	16743	1.274	23946	21125	1.134	1.239	1.17	1.181	1.087	1.146
MF-25	~~3102~~	~~1133~~	~~2.737~~	15613	15859	0.984	22212	20871	1.064	1.595	ND	0.912	1.02	0.966
MF-26	24140	21623	1.116	22870	19841	1.153	11947	11550	1.034	1.101	0.996	1.069	0.991	1.019
MF-27	38674	28048	1.379	41487	30555	1.358	20755	16142	1.286	1.341	1.231	1.259	1.233	1.241
control 1	29241	31905	0.916	30445	33671	0.904	37172	43211	0.860	0.894				
control 2	27530	26769	1.028	32067	32139	0.998	31735	32710	0.970	0.999				
control 3	25792	26273	0.982	31968	33843	0.945	26064	28225	0.923	0.950				
control 4	37141	27401	1.355	45225	34728	1.302	32361	27225	1.189	1.282				
control 5	17831	17044	1.046	22109	21529	1.027	10646	9688	1.099	1.057				
control 6	18822	16111	1.168	18858	16889	1.117	12549	11207	1.120	1.135				
control 7	44345	34180	1.297	55539	42490	1.307	~~1871~~	~~1577~~	~~1.186~~	1.264				
control 8	17367	15532	1.118	26746	25311	1.057	13656	12522	1.091	1.088				
control 9	22284	19702	1.131	25128	23469	1.071	28733	27196	1.057	1.086				
control 10	23193	20001	1.160	21003	19808	1.060	16615	17782	0.934	1.051				
SGCE Average			1.531			1.120			1.082	1.228				
Control Average			1.141			1.117			1.073	1.111				

SGCE - EXON 10

	Triplicate									Average	Normalized Ratios			
	1			2			3							
Sample	Exon 10	Int Control	Ratio	Exon 10	Int Control	Ratio	Exon 10	Int Control	Ratio	Ratio	1	2	3	Average
MF-19	24680	17315	1.425	30817	24100	1.279	26865	17872	1.503	1.402	1.191	1.1	1.298	1.196
MF-20	25964	21620	1.201	22793	22233	1.025	25691	25344	1.014	1.080	1.004	0.881	0.876	0.920
MF-22	17827	12154	1.467	18639	13810	1.350	20022	16519	1.212	1.343	1.227	1.161	1.047	1.145
MF-23	20203	15559	1.298	22612	18095	1.250	25608	20246	1.265	1.271	1.085	1.075	1.092	1.084
MF-24	16871	10908	1.547	21281	14505	1.467	24981	18463	1.353	1.456	1.293	1.261	1.168	1.241
MF-25	16006	11339	1.412	19001	13576	1.400	22511	16207	1.389	1.400	1.181	1.204	1.2	1.195
MF-26	26255	18723	1.402	24345	17044	1.428	16976	16713	1.016	1.282	1.172	1.228	0.877	1.092
MF-27	19300	16844	1.146	37516	25318	1.482	37916	24744	1.532	1.387	0.958	1.274	1.323	1.185
control 1	19192	13326	1.440	22124	17235	1.284	27262	19771	1.379	1.368				
control 2	23436	16640	1.408	25546	18806	1.358	27082	21077	1.285	1.351				
control 3	22475	15587	1.442	23366	16269	1.436	27094	18559	1.460	1.446				
control 4	33153	31990	1.036	35983	39566	0.909	40045	42592	0.940	0.962				
control 5	28706	27656	1.038	35761	30992	1.154	26100	27552	0.947	1.046				
control 6	37183	31763	1.171	33821	30171	1.121	41729	37206	1.122	1.138				
control 7	35461	40859	0.868	32021	36780	0.871	45793	46455	0.986	0.908				
control 8	32221	26860	1.200	34018	29088	1.169	~~9495~~	14888	~~0.638~~	1.002				
control 9	34379	29904	1.150	38243	32955	1.160	33573	30044	1.117	1.143				
control 10	35984	29879	1.204	34861	29987	1.163	56477	47518	1.189	1.185				
SGCE Average			1.388			1.274			1.269	1.310				
Control Average			1.236			1.244			1.234	1.229				

SGCE - EXON 11

	Triplicate									Average	Normalized Ratios			
	1			2			3							
Sample	Exon 11	Int Control	Ratio	Exon 11	Int Control	Ratio	Exon 11	Int Control	Ratio	Ratio	1	2	3	Average
MF-19	~~6468~~	~~4911~~	~~1.317~~	9802	12293	0.797	8151	9734	0.837	0.984	ND	0.665	0.803	0.734
MF-20	36859	24559	~~1.501~~	36472	31821	1.146	42527	41498	1.025	1.224	ND	0.957	0.984	0.971
MF-22	21291	13043	~~1.632~~	19559	15355	1.274	18964	16550	1.146	1.351	ND	1.063	1.1	1.082
MF-23	26904	17807	1.511	27753	19845	1.398	27536	21043	1.309	1.406	1.24	1.167	1.256	1.221
MF-24	23254	15307	1.519	26257	18236	1.440	29685	21001	1.414	1.458	1.246	1.202	1.357	1.268
MF-25	22752	15553	1.463	21491	15519	1.385	23954	17227	1.391	1.413	1.2	1.156	1.335	1.230
MF-26	26913	19430	1.385	25513	18170	1.404	26530	20479	1.295	1.362	1.136	1.172	1.242	1.183
MF-27	39129	26723	1.464	35373	25344	1.396	36424	28698	1.269	1.376	1.201	1.165	1.218	1.195
control 1	31799	22302	1.426	36091	26670	1.353	19067	17293	1.103	1.294				
control 2	25453	19183	1.327	24633	18990	1.297	21613	20233	1.068	1.231				
control 3	36910	29795	1.239	35407	30841	1.148	35688	34628	1.031	1.139				
control 4	37010	29980	1.234	36619	30260	1.210	39558	35714	1.108	1.184				
control 5	35965	28572	1.259	34960	27978	1.250	41941	35660	1.176	1.228				
control 6	56718	49791	1.139	33257	28212	1.179	37594	34562	1.088	1.135				
control 7	47362	40570	1.167	30240	27259	1.109	35804	39178	0.914	1.064				
control 8	32998	30801	1.071	37009	34801	1.063	40280	42418	0.950	1.028				
control 9	38657	33056	1.169	8688	16203	~~0.536~~	31906	30252	1.055	0.920				
control 10	4918	4251	1.157	2662	2266	1.175	36246	38833	0.933	1.088				
SGCE Average			1.496			1.211			1.146	1.284				
Control Average			1.310			1.284			1.144	1.243				

SGCE - EXON 11b

	Triplicate									Average	Normalized Ratios			
	1			2			3							
Sample	Exon 11b	Int Control	Ratio	Exon 11b	Int Control	Ratio	Exon 11b	Int Control	Ratio	Ratio	1	2	3	Average
MF-19	21286	4293	~~4.959~~	31719	15961	1.987	33682	15003	2.245	2.116	ND	2.839	3.668	3.254
MF-20	22475	11551	1.946	16249	29580	0.549	21946	30451	0.721	1.072	2.985	0.784	1.178	1.649
MF-22	18513	9664	1.916	14166	16214	0.874	10534	19265	0.547	1.112	2.939	1.249	0.894	1.694
MF-23	11822	12344	0.958	7817	15312	0.511	9120	16436	0.555	0.674	1.469	0.73	0.907	1.035
MF-24	12854	10775	1.193	7686	14091	0.545	9737	16520	0.589	0.776	1.83	0.779	0.962	1.190
MF-25	13307	8112	1.640	7111	11881	0.599	6838	12447	0.549	0.929	2.515	0.856	0.897	1.423
MF-26	14181	8065	1.758	8388	16421	0.511	11840	21185	0.559	0.943	2.696	0.73	0.913	1.446
MF-27	12040	17717	0.680	20056	16207	1.238	20808	26723	0.779	0.899	1.043	1.768	1.273	1.361
control 1	14401	16685	0.863	19506	13911	1.402	23000	20199	1.139	1.135				
control 2	7298	14994	0.487	7490	16750	0.447	9671	21894	0.442	0.459				
control 3	12749	25783	0.494	13138	28290	0.464	17105	35258	0.485	0.481				
control 4	16653	25995	0.641	15890	27390	0.580	18387	34500	0.533	0.585				
control 5	16972	24893	0.682	14885	26438	0.563	19163	34239	0.560	0.602				
control 6	33617	37956	0.886	14566	25569	0.570	18131	35701	0.508	0.654				
control 7	19739	23839	0.828	14374	27941	0.514	17608	33533	0.525	0.623				
control 8	14304	25971	0.551	19020	26188	0.726	21015	39594	0.531	0.603				
control 9	16625	29665	0.560	26639	21742	1.225	32038	36176	0.886	0.890				
control 10	16052	30327	0.529	14315	28374	0.505	19259	37787	0.510	0.515				
SGCE Average			2.194			0.893			0.931	1.150				
Control Average			0.696			0.689			0.608	0.731				

SGCE - EXON 12

	Triplicate									Average	Normalized Ratios			
	1			2			3							
Sample	Exon 12	Int Control	Ratio	Exon 12	Int Control	Ratio	Exon 12	Int Control	Ratio	Ratio	1	2	3	Average
MF-19	~~7964~~	12939	~~0.616~~	9060	10107	0.896	21657	22681	0.955	0.926	ND	0.994	1.046	1.020
MF-20	25403	25154	1.010	19439	24986	0.778	18633	23220	0.802	0.863	1.117	0.863	0.883	0.954
MF-22	23528	21987	1.070	24179	25253	0.957	23410	25164	0.930	0.986	1.183	1.062	1.019	1.088
MF-23	21074	21987	0.958	18190	18226	0.998	17519	17293	1.013	0.990	1.06	1.108	1.11	1.093
MF-24	15495	13888	1.116	15958	14988	1.065	~~4149~~	~~3269~~	~~1.269~~	1.090	1.235	1.182	ND	1.209
MF-25	13744	12680	1.084	13750	14001	0.982	12338	12616	0.978	1.015	1.199	1.09	1.071	1.120
MF-26	23190	24163	0.960	19271	17166	1.123	20226	19932	1.015	1.032	1.062	1.246	1.112	1.140
MF-27	25976	20424	1.272	28481	24755	1.151	30904	27935	1.106	1.176	1.407	1.277	1.211	1.298
control 1	16028	17865	0.897	16286	17677	0.921	17324	19073	1.009	0.942				
control 2	14757	15326	0.963	16386	15764	1.040	16950	16868	1.005	1.002				
control 3	22518	25990	0.866	22242	24663	0.902	23671	26697	0.887	0.885				
control 4	20733	23816	0.871	24130	27244	0.886	24894	26265	0.948	0.901				
control 5	20068	21398	0.938	~~7258~~	17371	~~0.418~~	ND	5376	ND	0.938				
control 6	17489	21638	0.808	19445	22568	0.862	23338	30916	0.755	0.808				
control 7	21674	23785	0.911	22163	24772	0.895	25798	25629	1.007	0.938				
control 8	22761	22937	0.992	23267	25702	0.905	23427	27838	0.842	0.913				
control 9	23277	24743	0.941	22155	25251	0.877	27803	31038	0.896	0.905				
control 10	26523	31055	0.854	28081	34248	0.820	32385	37119	0.872	0.849				
SGCE Average			0.954			0.939			1.042	0.971				
Control Average			0.957			0.904			0.963	0.964				

Appendix 2.1. Raw data derived from fragment analysis of Multi-Ligation Probe Amplification (MLPA) products. Myoclonus Dystonia patients were tested for exon dosage alterations using MLPA. Each of the 34 probes in the SALSA MLPA Kit P099-B2 are listed according to their size (nt). Raw data values for height and area are determined for each fragment and subsequently entered into the GeneMarker (SoftGenetics) program. Normalized ratios for each exon are subsequently calculated based on WT control values as well as the 9 internal control fragments within the probe set.

				Control - 01		Control - Q60		Control - Q63		M2-1		M28-1		M29-1		M30-1		M31-1		M31-2	
Probe Name	Location	Size (nt)	Gene / Exon	Height	Area	Height	Area	Height	Area	Height	Area	Height	Area	Height	Area	Height	Area	Height	Area	Height	Area
Control probe 0797-L00463	05q31	130	control	1324	7626	1075	6604	1231	8527	673	4061	1116	6507	1603	8529	1598	8642	2492	14320	1397	8441
GCH1 probe 4619-L03894	14q22	142	14-054.4 GCH1 Exon 01A	1539	9223	1281	8414	1380	9640	636	4063	1034	6790	1687	9481	1623	9219	2677	15922	1718	10596
Control probe 2218-L01712	11p15	148	control	796	4997	679	4593	785	5391	362	2427	746	4580	1117	6141	1003	5489	1548	9122	1004	6628
SGCE probe 7004-L07150	07q21	154	07-094.1 SGCE Exon 03	1598	9581	1307	9061	1582	11043	614	3831	1108	6702	2039	11017	1950	11014	2811	16756	1652	10421
Control probe 2597-L02068	05q35	160	control	994	5919	801	5562	905	6252	513	3395	890	5564	1357	7399	1401	7575	2268	13455	1278	7928
TH probe 3143-L02611	11p15	166	11-002.1 TH Exon 01	1127	7716	1028	7519	1147	8778	550	4070	1054	7295	1548	9389	1427	8526	2524	15869	1532	10982
SGCE probe 7892-L07706	07q21	172	07-094.1 SGCE Exon 09	983	6100	805	5470	878	6330	499	3459	913	5848	1309	7586	1320	7724	2254	13811	1523	9933
GCH1 probe 3139-L02607	14q22	178	14-054.4 GCH1 Exon 02	1326	8540	1131	8118	1234	9232	707	5028	1179	7654	1768	10377	1786	10793	2884	18431	1738	12582
TH probe 3144-L03587	11p15	184	11-002.1 TH Exon 03	951	6173	934	6800	1010	7357	499	3598	908	6072	1197	7246	1352	8046	2004	12698	1363	9783
SGCE probe 7005-L07151	07q21	190	07-094.1 SGCE Exon 04	1707	11124	1394	10708	1880	14849	1001	7449	1536	10613	2314	13825	2399	14608	3331	21372	1747	12208
Control probe 4689-L04452	01p36	195	control	1350	8896	1201	8807	1365	10668	675	4937	1216	8818	1752	10802	1803	11174	2826	20105	1801	13391
GCH1 probe 3140-L02608	14q22	202	14-054.4 GCH1 Exon 03	1031	6934	897	6684	1078	8370	554	4212	941	6620	1297	7824	1380	8618	2093	14053	1343	9288
TH probe 3145-L02613	11p15	211	11-002.1 TH Exon 04	784	5611	757	6099	856	6953	395	2997	749	5831	1017	6543	1122	7470	1652	11672	1140	8817
SGCE probe 3382-L05789	07q21	219	07-094.1 SGCE Exon 01A	928	6699	859	6781	953	7992	498	3895	877	6794	1179	7692	1290	8396	2018	15053	1187	9399
GCH1 probe 3141-L02609	14q22	229	14-054.4 GCH1 Exon 05	1129	8100	1038	8314	1251	10194	618	4796	1045	7714	1410	9537	1525	10243	2426	17076	1512	11909
TH probe 3146-L02614	11p15	238	11-002.1 TH Exon 08	1295	9902	1466	11892	1548	12988	728	5801	1336	10324	1818	12416	1931	13296	2376	17997	1697	13491
SGCE probe 7003-L07149	07q21	246	07-094.1 SGCE Exon 02	1144	8923	1069	9244	1345	11816	364	3182	1081	8941	1513	11254	1674	12157	2247	17436	1247	10553
GCH1 probe 3682-L03586	14q22	256	14-054.4 GCH1 Exon 01B	662	5469	698	5988	771	6904	358	3015	532	4344	855	5836	872	6409	1339	11059	764	6465
TH probe 3147-L02615	11p15	265	11-002.1 TH Exon 12	928	7369	1081	9251	1235	10722	538	4527	1097	8999	1273	9105	1456	10663	2321	18503	1392	11442
SGCE probe 7009-L07153	07q21	274	07-094.1 SGCE Exon 11	1376	11427	1347	11730	1775	15988	910	8110	1362	11332	1771	13529	2053	15164	2511	21172	1601	13938
GCH1 probe 3686-L03101	14q22	283	14-054.4 GCH1 Exon06	792	6580	784	6815	928	8199	439	3842	760	6366	968	7137	1189	8764	1859	15036	1151	9707
SGCE probe 3384-L05790	07q21	291	07-094.1 SGCE Exon 05	469	3798	424	3756	522	4678	264	2408	482	4225	654	5065	746	5866	1126	9401	796	6952
TH probe 3148-L05945	11p15	298	11-002.1 TH Exon 14	837	7067	931	8198	1067	9732	494	4330	932	8029	1120	8518	1279	9746	1970	16529	1307	11159
Control probe 2506-L02619	17q11	309	control	560	4786	612	5438	712	6584	365	3319	603	5283	802	6204	960	7574	1549	12887	911	8073
SGCE probe 3385-L02778	07q21	319	07-094.1 SGCE Exon 07	471	4011	509	4486	545	5111	310	2796	558	4795	654	5170	791	6164	1199	10120	819	7245
Control probe 1774-L01340	05q22	330	control	826	7240	808	7345	1073	10273	477	4397	848	7527	1066	8866	1263	10441	1884	16785	1102	9988
SGCE probe 7002-L06608	07q21	337	07-094.1 SGCE Exon 01B	1117	9786	1282	11782	1453	13876	664	6198	1178	10499	1399	11634	1734	14113	2696	23152	1721	15620
Control probe 3288-L02645	12p13	346	control	507	4553	512	4822	598	5789	287	2740	534	4769	591	5055	737	6144	1124	10032	707	6579
SGCE probe 3386-L02779	07q21	355	07-094.1 SGCE Exon 10	749	6745	692	6375	882	8566	425	4029	740	6854	921	8014	1146	9593	1772	15998	1138	10675
SGCE probe 7006-L08347	07q21	366	07-094.1 SGCE Exon 06	1129	10142	1383	13033	1837	18202	887	8352	1303	12019	1708	14579	1897	15990	2329	20904	1625	15112
SGCE probe 7891-L07705	07q21	382	07-094.1 SGCE Exon 08B	669	6194	656	6225	801	7993	384	3655	728	6893	911	7968	1050	9269	1642	15253	1077	10286
Control probe 1795-L01358	13q14	391	control	648	6118	741	7399	842	8555	414	4138	718	6850	879	7677	1091	9680	1645	15708	1081	10470
SGCE probe 7931-L06613	07q21	400	07-094.1 SGCE Exon 08A	722	6935	776	7739	867	9007	450	4551	807	7760	878	7854	1136	10247	1698	16313	1208	11875
Control probe 2669-L02136	11q	409	control	454	4218	499	5016	625	6414	288	2928	455	4410	557	5035	679	6007	1057	10172	648	6554

Probe Name	Location	Size (nt)	Gene / Exon	Control - 01		Control - 02		Control - 03		SP3		SP3-3		SP3-4		M19-1	
				Height	Area	Height	Area	Height	Area	Height	Area	Height	Area	Height	Area	Height	Area
Control probe 0797-L00463	05q31	130	control	3298	17935	1975	11872	1401	8133	1657	10924	1255	7632	908	5424	2162	11585
GCH1 probe 4619-L03894	14q22	142	14-054.4 GCH1 Exon 01A	3305	18889	2030	12434	1479	8679	1323	9281	1222	7648	917	5971	2094	11532
Control probe 2218-L01712	11p15	148	control	1825	10043	1183	7429	833	4781	853	5821	673	4432	536	3364	1236	6742
SGCE probe 7004-L07150	07q21	154	07-094.1 SGCE Exon 03	4365	25275	2711	16415	3416	20112	1191	7782	882	5702	688	4335	3170	17308
Control probe 2597-L02068	05q35	160	control	2339	13077	1416	8798	1083	6330	1199	8271	905	5700	713	4440	1606	8884
TH probe 3143-L02611	11p15	166	11-002.1 TH Exon 01	2842	17269	1636	11923	1177	7504	1453	10865	1095	8011	823	5902	1717	10099
SGCE probe 7892-L07706	07q21	172	07-094.1 SGCE Exon 09	2520	14766	1536	9961	1095	6705	1082	8209	865	5561	721	4813	1675	9764
GCH1 probe 3139-L02607	14q22	178	14-054.4 GCH1 Exon 02	3670	22138	2026	13396	1545	9655	1689	12375	1323	9052	988	6658	2397	14171
TH probe 3144-L03587	11p15	184	11-002.1 TH Exon 03	2231	13650	1348	9302	949	5994	1143	8610	880	6371	684	4648	1616	9669
SGCE probe 7005-L07151	07q21	190	07-094.1 SGCE Exon 04	4921	30387	2928	20931	2098	13342	1525	11599	1101	8106	765	5193	3631	22274
Control probe 4689-L04452	01p36	195	control	3661	22587	2120	15753	1439	9614	1485	11453	1242	9192	955	7051	2323	14089
GCH1 probe 3140-L02608	14q22	202	14-054.4 GCH1 Exon 03	2620	16396	1467	10483	1080	7255	1331	10303	960	7305	754	5377	1796	10892
TH probe 3145-L02613	11p15	211	11-002.1 TH Exon 04	1899	12397	1109	8065	721	5039	947	7514	696	5358	530	4098	1342	8422
SGCE probe 3382-L05789	07q21	219	07-094.1 SGCE Exon 01A	2142	14206	1300	9707	964	6762	959	7763	763	5962	612	4731	1518	9767
GCH1 probe 3141-L02609	14q22	229	14-054.4 GCH1 Exon 05	2904	19620	1648	12741	1216	8646	1517	12517	1121	9162	855	6290	2233	14366
TH probe 3146-L02614	11p15	238	11-002.1 TH Exon 08	3767	26258	2122	16955	1576	11651	2252	18807	1559	12509	1124	8768	2880	18864
SGCE probe 7003-L07149	07q21	246	07-094.1 SGCE Exon 02	3456	25769	1968	16194	1404	11035	1037	9049	757	6578	522	4288	2947	19759
GCH1 probe 3682-L03586	14q22	256	14-054.4 GCH1 Exon 01B	1640	12463	975	8169	709	5497	918	8131	672	5640	499	4262	1401	9414
TH probe 3147-L02615	11p15	265	11-002.1 TH Exon 12	2450	18751	1476	12154	1108	8923	1617	13981	1105	9081	810	6844	1920	14571
SGCE probe 7009-L07153	07q21	274	07-094.1 SGCE Exon 11	4126	32976	2261	19692	1647	13756	2275	20545	1763	15383	1309	11188	3497	26145
GCH1 probe 3686-L03101	14q22	283	14-054.4 GCH1 Exon06	2127	15985	1136	9481	846	6891	1126	9868	864	7289	650	5424	1619	11691
SGCE probe 3384-L05790	07q21	291	07-094.1 SGCE Exon 05	1276	10341	704	6184	500	4210	325	2940	256	2327	193	1603	887	6891
TH probe 3148-L05945	11p15	298	11-002.1 TH Exon 14	2014	16430	1119	9963	872	7517	1202	10998	832	7341	648	5715	1613	12249
Control probe 2506-L02619	17q11	309	control	1644	13756	828	7503	701	6049	937	8677	716	6375	511	4473	1285	10427
SGCE probe 3385-L02778	07q21	319	07-094.1 SGCE Exon 07	1224	10064	688	6148	575	4954	693	6489	504	4567	394	3431	938	7530
Control probe 1774-L01340	05q22	330	control	2106	18115	1154	10605	889	8037	1211	11676	873	8193	674	6225	1730	14397
SGCE probe 7002-L06608	07q21	337	07-094.1 SGCE Exon 01B	2873	24593	1499	13609	1229	10883	1582	15407	1210	11366	941	8538	2351	19461
Control probe 3288-L02645	12p13	346	control	1335	11707	669	6314	554	5126	741	7103	533	5186	430	4086	1070	9102
SGCE probe 3386-L02779	07q21	355	07-094.1 SGCE Exon 10	1706	15186	960	9221	717	6661	1055	10254	721	6923	600	5588	1402	12116
SGCE probe 7006-L08347	07q21	366	07-094.1 SGCE Exon 06	3744	33581	1943	18237	1493	13781	2476	24423	1819	17644	1360	12781	3619	30887
SGCE probe 7891-L07705	07q21	382	07-094.1 SGCE Exon 08B	1683	15266	936	9043	658	6392	966	9865	746	7285	538	5110	1419	12701
Control probe 1795-L01358	13q14	391	control	1549	14278	781	7818	693	6544	1012	10347	700	7014	579	5574	1400	12443
SGCE probe 7931-L06613	07q21	400	07-094.1 SGCE Exon 08A	1626	15221	860	8803	677	6750	1098	11527	729	7518	597	5785	1472	13169
Control probe 2669-L02136	11q	409	control	1179	11179	590	5874	485	4735	647	6847	497	5249	405	4098	1025	9482

Probe Name	Size (nt)	Gene / Exon	M20-1		M21-1		M22-1		M23-1		M25-1		M26-1		M27-1		M28-1	
			Height	Area	Height	Area	Height	Area	Height	Area	Height	Area	Height	Area	Height	Area	Height	Area
Control probe 0797-L00463	130	control	2038	11571	1945	11102	1041	6159	637	4319	1288	2096	1149	6814	3699	21786	1167	6908
GCH1 probe 4619-L03894	142	14-054.4 GCH1 Exon 01A	2115	11919	1917	11246	1009	6171	626	4417	1364	2213	1148	7165	4013	24361	1064	6540
Control probe 2218-L01712	148	control	1170	6669	1242	7135	597	3669	324	2260	759	2266	729	4734	2168	13174	778	4870
SGCE probe 7004-L07150	154	07-094.1 SGCE Exon 03	3015	16791	2670	15537	1474	8969	916	6514	1777	2325	1310	8063	3838	23669	1327	8408
Control probe 2597-L02068	160	control	1621	9008	1587	9277	777	4729	501	3435	1104	2375	929	5843	3246	20089	882	5496
TH probe 3143-L02611	166	11-002.1 TH Exon 01	1798	10985	1776	11241	897	5998	540	4170	1169	2438	1003	7272	3202	21487	1053	7558
SGCE probe 7892-L07706	172	07-094.1 SGCE Exon 09	1640	9606	1590	9556	781	4895	496	3529	1080	2501	945	5965	3390	21634	934	6168
GCH1 probe 3139-L02607	178	14-054.4 GCH1 Exon 02	2224	13535	2097	13175	1061	6782	716	5243	1520	2570	1287	8724	4428	29238	1277	8852
TH probe 3144-L03587	184	11-002.1 TH Exon 03	1610	9780	1426	9084	737	5038	442	3346	1001	2642	898	6260	2857	19200	926	6404
SGCE probe 7005-L07151	190	07-094.1 SGCE Exon 04	3615	22463	2958	19068	1601	10952	1013	7779	1861	2697	1524	10638	4645	30395	1873	13221
Control probe 4689-L04452	195	control	2381	14756	1976	12924	1057	7660	728	5581	1484	2742	1287	9369	4225	29332	1196	8337
GCH1 probe 3140-L02608	202	14-054.4 GCH1 Exon 03	1793	11269	1679	11291	842	5749	546	4095	1093	2804	986	7015	3248	22256	933	6686
TH probe 3145-L02613	211	11-002.1 TH Exon 04	1354	9249	1182	8877	569	4074	357	2852	812	2884	725	5360	2231	16136	721	5622
SGCE probe 3382-L05789	219	07-094.1 SGCE Exon 01A	1497	10018	1406	10011	680	5299	452	3774	1000	2958	802	6003	2475	18372	844	6300
GCH1 probe 3141-L02609	229	14-054.4 GCH1 Exon 05	2147	15086	1812	12849	991	7719	683	5560	1313	3065	1152	8864	3698	27194	1093	8268
TH probe 3146-L02614	238	11-002.1 TH Exon 08	2689	19773	2306	17238	1256	9661	761	6269	1475	3139	1301	10240	4048	29852	1422	10816
SGCE probe 7003-L07149	246	07-094.1 SGCE Exon 02	2721	20256	1875	14344	1213	9930	740	6339	1430	3227	1127	9058	3242	25933	1315	10848
GCH1 probe 3682-L03586	256	14-054.4 GCH1 Exon 01B	1219	9283	989	8038	594	4876	391	3393	709	3338	600	5055	1938	16671	558	4853
TH probe 3147-L02615	265	11-002.1 TH Exon 12	1996	15318	1722	13538	844	6920	559	4829	1164	3418	1051	8673	3395	28194	1116	9355
SGCE probe 7009-L07153	274	07-094.1 SGCE Exon 11	3463	28266	2501	20713	1478	12495	965	8439	1669	3524	1396	12188	4202	35786	1668	14224
GCH1 probe 3686-L03101	283	14-054.4 GCH1 Exon06	1627	12693	1306	10637	713	5950	480	4231	994	3597	840	7128	2791	23482	764	6467
SGCE probe 3384-L05790	291	07-094.1 SGCE Exon 05	929	7536	772	6415	436	3607	276	2341	556	3669	520	4593	1680	14500	450	4002
TH probe 3148-L05945	298	11-002.1 TH Exon 14	1593	12900	1355	11398	727	6183	490	4385	915	3755	837	7404	2694	23265	890	7653
Control probe 2506-L02619	309	control	1372	11340	1025	8722	577	4958	403	3589	732	3868	720	6689	2235	19586	617	5368
SGCE probe 3385-L02778	319	07-094.1 SGCE Exon 07	951	7733	731	6234	454	3949	278	2589	537	3961	472	4296	1539	13286	503	4371
Control probe 1774-L01340	330	control	1731	15142	1338	12069	744	6770	522	4847	1005	4069	877	8460	2612	23846	884	7943
SGCE probe 7002-L06608	337	07-094.1 SGCE Exon 01B	2405	20435	1919	16827	1003	8841	717	6633	1313	4150	1122	10400	3858	35035	1221	11098
Control probe 3288-L02645	346	control	1130	9763	755	6910	467	4316	308	2875	566	4235	522	5055	1822	16433	512	4844
SGCE probe 3386-L02779	355	07-094.1 SGCE Exon 10	1506	13587	1192	11035	670	6215	447	4336	847	4330	736	7213	2544	23611	717	6840
SGCE probe 7006-L08347	366	07-094.1 SGCE Exon 06	3535	31391	2476	22962	1635	15272	949	9012	1620	4434	1355	13549	3802	35509	1606	14969
SGCE probe 7891-L07705	382	07-094.1 SGCE Exon 08B	1438	12931	1081	10073	599	5796	449	4423	852	4578	670	7066	2508	23740	725	6786
Control probe 1795-L01358	391	control	1416	13120	1088	10428	619	6128	421	4176	833	4687	695	7545	2552	24313	671	6361
SGCE probe 7931-L06613	400	07-094.1 SGCE Exon 08A	1536	14595	1097	10715	652	6618	462	4677	867	4782	719	7817	2620	25488	763	7494
Control probe 2669-L02136	409	control	1035	9865	747	7430	432	4361	310	3104	584	4866	492	5291	1774	17191	486	4692

Probe Name	Location	Size (nt)	Gene / Exon	Control - 01		Control - 02		Control - 04		M24-1		M27-1		M1-39	
				Height	Area	Height	Area	Height	Area	Height	Area	Height	Area	Height	Area
Control probe 0797-L00463	05q31	130	control	1919	10647	2193	12291	1901	11250	1191	7420	2092	13234	1419	8781
GCH1 probe 4619-L03894	14q22	142	14-054.4 GCH1 Exon 01A	2362	13711	2784	15702	2511	15337	1620	11322	2765	19334	1615	10870
Control probe 2218-L01712	11p15	148	control	1294	7424	1521	9006	1420	8717	1018	6997	1526	10625	1043	6848
SGCE probe 7004-L07150	07q21	154	07-094.1 SGCE Exon 03	2314	13490	2654	15449	2263	13861	1389	9637	2255	14941	1579	10705
Control probe 2597-L02068	05q35	160	control	1588	9404	1883	11252	1749	10987	1233	8504	2009	14326	1223	8107
TH probe 3143-L02611	11p15	166	11-002.1 TH Exon 01	1857	11870	2241	14105	1882	13220	1326	10033	2199	16585	1368	10062
SGCE probe 7892-L07706	07q21	172	07-094.1 SGCE Exon 09	1560	9651	1782	10960	1615	10279	1217	8775	1935	14011	1257	8935
GCH1 probe 3139-L02607	14q22	178	14-054.4 GCH1 Exon 02	2072	13256	2345	14937	2096	14056	1532	11357	2546	18587	1546	11372
TH probe 3144-L03587	11p15	184	11-002.1 TH Exon 03	1517	9916	1688	10901	1480	10099	1218	9088	1830	13482	1240	9036
SGCE probe 7005-L07151	07q21	190	07-094.1 SGCE Exon 04	2279	15164	2710	17284	2238	15617	1359	10705	2377	18178	1694	12740
Control probe 4689-L04452	01p36	195	control	2244	15256	2651	18112	2350	17097	1601	12225	2744	20787	1733	13381
GCH1 probe 3140-L02608	14q22	202	14-054.4 GCH1 Exon 03	1559	10525	1874	12839	1594	11592	1224	9670	1911	14490	1239	9362
TH probe 3145-L02613	11p15	211	11-002.1 TH Exon 04	1299	9393	1373	9796	1190	9128	965	8010	1569	12516	1035	8120
SGCE probe 3382-L05789	07q21	219	07-094.1 SGCE Exon 01A	1474	11082	1730	12627	1555	12292	1033	8428	1647	13235	1230	10156
GCH1 probe 3141-L02609	14q22	229	14-054.4 GCH1 Exon 05	1777	13359	1809	13308	1697	13398	1292	11080	2133	17675	1311	10864
TH probe 3146-L02614	11p15	238	11-002.1 TH Exon 08	1942	15241	2197	17016	1867	14866	1444	12312	2330	19420	1524	12887
SGCE probe 7003-L07149	07q21	246	07-094.1 SGCE Exon 02	1627	13310	1746	14122	1353	11588	985	8939	1596	13938	1177	10720
GCH1 probe 3682-L03586	14q22	256	14-054.4 GCH1 Exon 01B	1046	8651	1120	9141	986	8301	718	6529	1206	10558	729	6443
TH probe 3147-L02615	11p15	265	11-002.1 TH Exon 12	1591	13266	1710	14024	1494	12531	1294	11607	1955	16858	1309	11697
SGCE probe 7009-L07153	07q21	274	07-094.1 SGCE Exon 11	1903	16660	1948	16554	1497	13291	1239	11423	2034	18031	1444	13274
GCH1 probe 3686-L03101	14q22	283	14-054.4 GCH1 Exon06	1239	10690	1287	10691	1140	9991	1021	9374	1493	13414	948	8711
SGCE probe 3384-L05790	07q21	291	07-094.1 SGCE Exon 05	796	7115	832	7315	679	6061	649	6090	1037	9579	606	5653
TH probe 3148-L05945	11p15	298	11-002.1 TH Exon 14	1390	12281	1387	12201	1146	10324	1067	9938	1709	15501	1155	11005
Control probe 2506-L02619	17q11	309	control	935	8540	808	7158	774	7050	766	7449	1147	10732	815	7866
SGCE probe 3385-L02778	07q21	319	07-094.1 SGCE Exon 07	879	7914	868	7712	745	6836	727	6991	1085	10125	738	7017
Control probe 1774-L01340	05q22	330	control	1284	12285	1280	11801	1097	10349	1014	10095	1497	14366	1090	10786
SGCE probe 7002-L06608	07q21	337	07-094.1 SGCE Exon 01B	1700	16046	1565	14553	1344	12741	1363	13534	1976	19054	1443	14369
Control probe 3288-L02645	12p13	346	control	833	7928	752	7046	622	6105	648	6538	972	9526	656	6597
SGCE probe 3386-L02779	07q21	355	07-094.1 SGCE Exon 10	1121	11019	1108	10445	951	9369	975	9889	1468	14667	992	10005
SGCE probe 7006-L08347	07q21	366	07-094.1 SGCE Exon 06	1533	14729	1380	13290	1121	11065	1091	11165	1828	18145	1200	12252
SGCE probe 7891-L07705	07q21	382	07-094.1 SGCE Exon 08B	1096	10624	1008	9735	807	8025	829	8816	1284	13021	935	9708
Control probe 1795-L01358	13q14	391	control	1014	10148	896	8818	744	7497	911	9745	1286	13475	939	9904
SGCE probe 7931-L06613	07q21	400	07-094.1 SGCE Exon 08A	1209	12549	1003	10168	880	9225	938	10168	1430	15000	1051	11680
Control probe 2669-L02136	11q	409	control	732	7523	669	6744	568	5895	608	6630	927	9912	627	6940

Appendix 2.2 Normalized ratios of each MLPA fragment for eighteen Myoclonus Dystonia patients. Normalized ratios are listed for each WT control and MD patient included in the MLPA exon dosage analysis. Note that these results were obtained in groups, and that WT controls run within one group can only be used to normalize the MD patient values in that same group. Controls and MD patients within the same MLPA run are therefore listed in the same table. Shaded boxes indicate ratios that are not within the normal exon dosage range of 0.82-1.17, the default threshold of the GeneMarker program set by the Children's Hospital of Eastern Ontario (CHEO) diagnostic lab.

			Ratios								
Probe Name	Exon / location	Bin Size	Control O1	Control Q60	Control Q63	M2-1	M28-1	M29-1	M30-1	M31-1	M31-2
Control	11p15	144	1	1.034	0.978	0.916	1.064	1.085	0.942	0.941	0.991
Control	11q	407	1.048	0.994	1	0.996	0.968	0.994	0.932	0.975	0.891
Control	12p13	343	1.06	1	0.955	0.969	1.068	0.94	0.972	0.977	0.94
Control	13q14	389	1	1.011	0.933	0.981	1.023	1.023	1.023	1.023	1.009
Control	17q11	306	0.992	1.051	1	1.084	1.034	1.083	1.103	1.165	1.079
Control	1p36	193	1.013	1	0.966	0.96	0.99	1.005	0.974	0.997	1.038
Control	5q22	326	1	0.928	1.011	0.95	1	1	1	0.979	0.865
Control	5q31	126	1.013	1	0.971	1.07	0.981	0.925	0.95	0.925	0.89
Control	5q35	155	1.065	1	0.95	1.083	1.094	1.081	1.091	1.181	1.076
GCH1	Ex1	138	1.05	1	0.907	0.977	0.815	0.877	0.841	0.888	0.922
GCH1	Ex1b	254	0.992	1.094	1	0.964	0.834	0.995	0.885	0.904	0.809
GCH1	Ex2	175	1.044	1	0.911	1.065	1.016	1.037	1.04	1.072	1.045
GCH1	Ex3	199	1.029	1	0.984	1.045	1.02	0.984	0.996	0.973	1.008
GCH1	Ex5	226	1.018	1	0.995	1.003	0.991	0.958	0.953	0.98	0.973
GCH1	Ex6	279	1	1.001	0.976	0.965	0.968	0.935	1.013	1.029	0.985
SGCE	Ex1	215	1.004	1	0.935	0.987	1.013	0.955	0.975	0.994	0.937
SGCE	Ex10	352	1.132	0.955	1	1.022	1.06	1.052	1.061	1.091	1.082
SGCE	Ex11	272	1	0.995	1.079	1.135	0.987	0.975	0.982	0.81	0.814
SGCE	Ex1b	335	0.995	1.072	1	0.961	1	0.961	0.987	0.99	0.988
SGCE	Ex2	242	1	0.983	1.005	0.574	0.977	1	0.987	0.88	0.764
SGCE	Ex3	150	1.004	0.988	1	0.781	0.796	0.988	0.959	0.881	0.84
SGCE	Ex4	188	1	0.922	1.018	1.102	0.962	1.008	1.008	0.896	0.755
SGCE	Ex5	287	1.099	0.981	1	1.064	1.154	1.187	1.17	1.164	1.281
SGCE	Ex6	363	0.884	1	1.075	1.098	0.968	1.036	0.921	0.751	0.799
SGCE	Ex7	316	1	1.019	0.915	1.081	1.123	1.038	1.059	1.063	1.134
SGCE	Ex8	377	1.135	1	0.989	1.019	1.151	1.194	1.106	1.119	1.118
SGCE	Ex8b	398	1.087	1	0.906	1.015	1.093	0.999	1.022	1.008	1.059
SGCE	Ex9	168	1.064	1	0.923	1.054	1.106	1.106	1.08	1.169	1.282
TH	Ex1	162	1	1.043	0.997	0.967	1.059	1.068	0.936	1.072	1.068
TH	Ex12	262	0.884	1.053	1	0.915	1.081	0.926	0.964	0.975	0.927
TH	Ex14	295	0.985	1.076	1	0.993	1.091	1	0.995	0.979	1.025
TH	Ex3	183	0.97	1.112	1	1.001	1.054	0.975	1.032	0.992	1.108
TH	Ex4	208	1	1.047	0.991	0.937	1.044	0.992	1.024	0.973	1.082
TH	Ex8	233	0.944	1.152	1	0.985	1.01	1.021	0.99	0.786	0.914

			Ratios													
Probe Name	Exon / location	Bin Size	Control O1	Control O2	SP3	SP3-3	SP3-4	M19-1	M20-1	M21-1	M22-1	M23-1	M25-1	M26-1	M27-1	M28-1
Control	11p15	144	0.922	1	0.963	0.959	1.09	0.959	0.933	1.041	0.988	0.854	0.959	1.053	0.959	1.116
Control	11q	407	1	0.977	0.863	0.959	0.971	0.964	0.931	0.961	0.975	0.944	1	0.989	1	0.911
Control	12p13	343	1	0.939	0.951	0.952	0.977	0.965	0.978	0.871	0.965	0.921	0.896	0.953	0.974	0.923
Control	13q14	389	1	0.969	1.031	1.024	1.074	1.019	0.979	1.072	1.064	0.996	1.072	1.055	1.099	0.996
Control	17q11	306	1	0.907	1.023	1.06	0.957	0.973	1.032	0.963	0.995	1.022	0.949	1.095	1.005	0.944
Control	1p36	193	0.981	1	0.865	0.924	0.908	0.924	0.958	0.897	0.913	0.981	0.958	0.947	0.977	0.9
Control	5q22	326	0.992	1	1	1	0.979	1	0.976	0.988	0.988	0.997	1	1.015	0.902	1
Control	5q31	126	0.994	1	1.132	1.108	1.063	1.041	1.054	1	1.051	1.075	1	1.016	1.023	1.038
Control	5q35	155	0.968	1	1.093	1.046	1.059	1.033	1.062	1.097	1.024	1.103	1.129	1.109	1.194	1.048
GCH1	Ex1	138	1	0.975	0.892	1.036	1.017	0.963	0.999	0.949	0.949	0.983	1	0.97	1.064	0.928
GCH1	Ex1b	254	0.982	1	1.018	1.01	0.966	1.093	0.963	0.926	1.017	1.014	0.936	0.932	0.908	0.849
GCH1	Ex2	175	1	0.936	0.975	0.983	0.952	0.971	0.917	0.937	0.911	0.975	0.975	0.975	1.043	0.975
GCH1	Ex3	199	1	0.965	1.052	0.994	0.998	0.993	0.998	1.045	0.992	1.011	0.998	1.027	1.028	0.984
GCH1	Ex5	226	1	0.987	1.005	1.013	0.971	1.039	1.019	0.968	1.028	1.075	1.018	1.041	1.01	1.002
GCH1	Ex6	279	1	0.954	0.984	1.012	0.977	0.976	0.968	0.986	0.962	0.991	1.011	1.002	0.998	0.895
SGCE	Ex1	215	0.959	1	0.878	0.915	0.937	0.963	0.981	1.019	0.937	0.963	1.048	0.972	0.929	1.019
SGCE	Ex10	352	0.938	1	0.977	0.954	0.978	0.919	0.953	1.019	1.019	0.975	0.981	0.984	0.997	0.943
SGCE	Ex11	272	1	0.998	1.032	1.103	1.054	1.128	1.116	0.978	1.063	1.035	0.909	0.881	0.798	1.047
SGCE	Ex1b	335	1	0.968	0.957	1.02	0.996	1.007	0.995	1.035	0.977	1.01	0.989	0.965	0.988	1.036
SGCE	Ex2	242	1	0.974	0.58	0.566	0.506	1.127	1.059	0.865	1.041	0.976	0.93	0.835	0.748	1.002
SGCE	Ex3	150	0.995	1	0.597	0.579	0.593	1.093	1.054	0.994	1.075	1.099	0.994	0.829	0.791	0.87
SGCE	Ex4	188	0.999	1	0.659	0.616	0.555	1.069	1.095	0.982	1.012	1.026	0.883	0.838	0.791	1.059
SGCE	Ex5	287	1	0.974	0.478	0.49	0.477	0.894	0.926	0.939	0.966	0.904	0.926	1.025	0.997	0.888
SGCE	Ex6	363	1	0.966	1.091	1.139	1.065	1.123	1.061	1.022	1.18	0.963	0.88	0.863	0.707	0.998
SGCE	Ex7	316	0.982	1	0.986	0.993	0.972	0.935	0.929	0.922	1.04	0.925	0.915	0.942	0.904	0.992
SGCE	Ex8	377	0.947	1	0.878	0.966	0.889	0.906	0.883	0.924	0.911	0.946	0.982	0.902	0.97	0.946
SGCE	Ex8b	398	0.976	1	1.038	0.988	1.021	0.977	0.985	1.003	1.021	1.021	1.021	1.025	1.057	1.045
SGCE	Ex9	168	0.962	1	0.901	0.92	0.991	0.972	0.958	1	0.949	0.968	0.997	1.006	1.105	0.988
TH	Ex1	162	1	0.978	1.124	1.1	1.071	0.925	1.003	1.045	1.015	1.001	1.015	0.982	0.993	1.065
TH	Ex12	262	0.954	1	1.183	1.095	1.027	1	1.007	1.062	0.967	0.981	1	1.055	1.028	1.113
TH	Ex14	295	1	0.998	1.084	1.036	1.014	1.009	0.989	1.033	1.026	1.024	0.977	1.038	0.992	1.114
TH	Ex3	183	0.987	1	1.075	1.066	1.058	1.059	1.053	1.016	1.016	0.988	1.055	1.083	1.066	1.11
TH	Ex4	208	0.997	1	1.024	0.991	0.973	1.005	1.039	1.025	0.939	0.906	1.01	1.04	0.978	1.024
TH	Ex8	233	1	0.974	1.162	1.073	0.986	1.044	0.982	0.961	0.99	0.936	0.883	0.883	0.853	0.987

			Ratios					
Probe Name	Exon / location	Bin Size	Control O1	Control O2	Control O4	M24-1	M27-1	M1-39
Control	11p15	144	0.979	1	1.055	1.119	1.003	1.086
Control	11q	407	0.95	1	1.037	0.973	0.984	0.946
Control	12p13	343	1.014	0.974	1	1.004	0.977	0.964
Control	13q14	389	1	0.973	1	1.118	1.04	1.081
Control	17q11	306	1.004	0.87	1	1.054	1.013	1.047
Control	1p36	193	0.933	1	1.019	0.918	0.968	0.953
Control	5q22	326	0.986	1	1.034	0.98	0.925	0.98
Control	5q31	126	1.078	1	0.994	0.948	0.992	1.069
Control	5q35	155	0.999	1	1.08	1.084	1.065	1.022
GCH1	Ex1	138	1.005	0.964	1	0.979	0.979	0.904
GCH1	Ex1b	254	0.989	1	1.041	0.912	0.943	0.865
GCH1	Ex2	175	1.009	0.983	1	1.02	1.021	0.986
GCH1	Ex3	199	0.957	1.002	1	1.022	0.952	0.984
GCH1	Ex5	226	1	0.925	1.026	0.988	1.002	0.922
GCH1	Ex6	279	0.999	1	1.061	1.09	1.001	0.941
SGCE	Ex1	215	0.953	1	1.059	0.907	0.89	1.01
SGCE	Ex10	352	0.954	1	1.04	1.035	1.017	0.977
SGCE	Ex11	272	1.027	1	0.917	0.866	0.895	0.941
SGCE	Ex1b	335	1	0.976	1.007	1.032	0.964	1.026
SGCE	Ex2	242	1	1.009	0.916	0.819	0.822	0.912
SGCE	Ex3	150	1.055	1	0.964	0.881	0.846	0.962
SGCE	Ex4	188	0.994	1.047	1	0.831	0.877	0.97
SGCE	Ex5	287	1	1.019	0.995	1.06	1.067	0.936
SGCE	Ex6	363	1.028	1	0.995	0.917	1.014	0.938
SGCE	Ex7	316	1	0.992	1.028	1.057	0.998	1
SGCE	Ex8	377	0.986	1.015	1	0.955	0.946	0.999
SGCE	Ex8b	398	1	0.926	1.004	0.953	0.953	1.003
SGCE	Ex9	168	1	0.966	1.003	1.078	1.035	1.054
TH	Ex1	162	1	1.016	0.972	1.003	1.008	0.957
TH	Ex12	262	1	0.996	1.023	1.051	0.981	0.992
TH	Ex14	295	1.001	1	0.977	0.988	1.009	1.005
TH	Ex3	183	1.017	1	0.991	1.113	1.027	1.103
TH	Ex4	208	1.05	0.972	1	1.07	1.049	1.069
TH	Ex8	233	0.983	1	1.009	0.958	0.981	0.963