



LIBRARY  
Michigan State  
University

This is to certify that the  
dissertation entitled

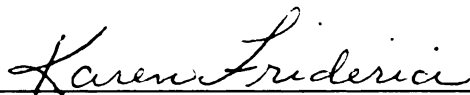
ANALYSIS OF 17p11.2 DELETIONS AND CANDIDATE  
GENE CHARACTERIZATION IN SMITH-MAGENIS  
SYNDROME

presented by

CHRISTOPHER NICHOLAS VLANGOS

has been accepted towards fulfillment  
of the requirements for the

Ph.D. degree in Genetics



Major Professor's Signature

5/11/2005

Date

**PLACE IN RETURN BOX** to remove this checkout from your record.  
**TO AVOID FINES** return on or before date due.  
**MAY BE RECALLED** with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

**ANALYSIS OF 17p11.2 DELETIONS AND CANDIDATE GENE  
CHARACTERIZATION IN SMITH-MAGENIS SYNDROME**

**By**

**Christopher Nicholas Vlangos**

**A DISSERTATION**

**Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of**

**DOCTOR OF PHILOSOPHY**

**Genetics Program**

**2005**



## **ABSTRACT**

### **ANALYSIS OF 17p11.2 DELETIONS AND CANDIDATE GENE CHARACTERIZATION IN SMITH-MAGENIS SYNDROME**

**By**

**Christopher Nicholas Vlangos**

Smith-Magenis syndrome (SMS) is a multiple congenital anomalies and mental retardation disorder usually associated with an interstitial deletion of chromosome 17p11.2. The SMS phenotype includes craniofacial/skeletal anomalies, otolaryngeal abnormalities, and neurological anomalies. The incidence of SMS is estimated at 1:25,000 births. Unfortunately, the syndrome is not widely known and under-diagnosis is likely.

In order to identify the genes responsible for SMS, we constructed a physical and transcription map of the region. Using the small number of markers known to map to SMS critical region (SMCR), we identified large insert DNA clones mapping to 17p11.2. We established the location of the known markers within these clones by Southern analysis and created a genomic contig across the ~1.5 Mb SMCR. The order and orientation of 29 potential candidate genes was also determined within the SMCR.

The candidate genes were initially characterized by bioinformatic analysis in order to determine any known or predicted protein motifs in these sequences. The developmentally regulated GTP-binding protein 2 (*DRG2*) gene was chosen for further investigation by immunohistochemistry on sectioned mouse embryos where expression was found localized within the developing nervous system. *DRG2* was localized to the endoplasmic reticulum and the Golgi apparatus with a GFP fusion

product. *Drg2* may play a role in embryonic neuronal development by regulating protein trafficking.

Concurrent to construction of the physical map and gene characterization we used fluorescent *in situ* hybridization to characterize the deletion sizes carried by a cohort of 22 SMS patients in order to decrease the size of the SMS critical region. Through extensive FISH analysis of these SMS patient samples, we reduced the smallest region of overlap involving 17p11.2 deletion associated with SMS from 1.5 Mb to ~700 kb. This refinement reduced the number of candidate genes from >40 to 11.

Within our cohort of 22 patients there was a group of 7 patients in whom a 17p11.2 deletion could not be detected with FISH analysis. We hypothesized that these patients might carry mutations in gene(s) mapping to the SMCR. Using systematic direct sequencing of candidate genes, we identified 4 SMS patients with dominant frame shift mutations in the retinoic acid induced 1 (*RAI1*) gene. Haploinsufficiency of the *RAI1* gene is the cause of the physical, neurological, and behavioral characteristics of Smith-Magenis syndrome.

## **ACKNOWLEDGEMENTS**

The work presented in this dissertation would not have been possible without the assistance and support from many people. I cannot begin to describe how lucky I was to perform this work under the guidance of Dr. Sarah Elsea. Sarah provided an amazing atmosphere for learning. The dedication and support she provides to her students is unparalleled. The knowledge and skills that I have learned from Sarah are priceless and will serve me well in the future.

My guidance committee included Dr. Karen Friderici, Dr. Kathy Meek, Dr. Helga Torriello, Dr. Vilma Yuzbasian-Gurkan, and Dr. Michael Grotewiel. Their assistance and time is greatly appreciated.

My co-workers and friends in the Elsea Laboratory were also key in the success of this endeavor and include fellow graduate students Rebecca Slager and Santhosh Girirajan, research technicians Dennis Lettau and Barbara Szomju, and undergraduate students Tanisha Jain and Suzanne Shunn.

A special thank you goes out to the Genetics Program at MSU, especially to Jeannine Lee. Without the help of Jeannie no Genetics student would be successful in their graduate career.

Additionally, I would like to thank those in the Departments of Pediatrics and Human Genetics at Virginia Commonwealth University. In my brief stay at VCU as a visiting student they made me feel welcome and treated me as one of their own.

Amazing support also came from my family and friends. Without constant encouragement from my parents Peter Vlangos and Sue Smith, sister Georgette

Vlangos, and grandparents Nick and Carol Smith I would not have come this far. I also have to acknowledge Jim and Sue Vogel for feeding me dinner for months during the more lean times in my graduate career.

It is during the course of my graduate career that my friend Scott Asakevich introduced me to Jim and Sue's daughter Michelle Vogel. Michelle is an amazing and loving woman who provided encouragement and love. As we begin our lives together as husband and wife I know I can always count on Michelle.

## TABLE OF CONTENTS

<b>LIST OF TABLES.....</b>	<b>ix</b>
<b>LIST OF FIGURES.....</b>	<b>x</b>
<b>LIST OF ABBREVIATIONS.....</b>	<b>xii</b>
<b>CHAPTER 1. INTRODUCTION.....</b>	<b>1</b>
The SMS phenotype.....	2
17p11.2 deletions associated with SMS.....	9
Microdeletion syndromes: from critical region to gene identification.....	12
Animal modeling of Smith-Magenis syndrome.....	16
Summary.....	18
<b>CHAPTER 2. PHYSICAL MAPPING OF THE SMITH-MAGENIS SYNDROME CRITICAL REGION.....</b>	<b>19</b>
Background.....	19
The Smith-Magenis syndrome common and critical regions.....	20
Rationale.....	23
Results.....	23
Mapping of new genes and markers using a somatic cell hybrid mapping panel.....	23
Construction of the physical map and contig of the SMS critical region.....	26
Analysis and characterization of ESTs mapped to the SMS critical region.....	29
Summary of known candidate genes mapping within the SMS critical region.....	37
Summary.....	42
Materials and methods.....	43
<b>CHAPTER 3. ANALYSIS OF 17p11.2 DELETIONS USING FLUORESCENT IN SITU HYBRIDIZATION.....</b>	<b>48</b>
Background.....	48
Rationale.....	51
Results.....	52
Commercially available diagnostic FISH probes for SMS will not detect all cases.....	56
The SMS common deletion occurs less frequently than reported.....	63
Summary.....	65

Materials and mehods.....	67
<b>CHAPTER 4. ANALYSIS OF SMITH-MAGENIS SYNDROME PATIENTS WITHOUT A FISH DETECTABLE 17P11.2 DELETION.....</b>	<b>72</b>
Background and rationale.....	72
Results.....	73
Mutation analysis in SMS patients without a FISH detectable deletion.....	73
The retinoic acid induced 1 ( <i>RAI1</i> ) gene.....	84
Summary.....	86
Materials and methods.....	87
<b>CHAPTER 5. DEVELOPMENTALLY REGULATED GTP-BINDING PROTEIN(<i>DRG2</i>).....</b>	<b>94</b>
Background.....	94
Cloning and identification of mammalian DRG genes.....	95
Rationale.....	96
Results.....	97
Identification of the STS WI-13499 as <i>DRG2</i> and mapping to the SMS critical region.....	97
Bioinformatic analysis of <i>DRG2</i> .....	99
Expression levels of <i>DRG2/Drg2</i> using northern analysis.....	101
Spatial and temporal expression of <i>Drg2</i> using immunohistochemistry.....	102
Cellular localiztion of <i>Drg2</i> .....	110
Summary.....	114
Materials and methods.....	115
<b>CHAPTER 6. TARGET OF MYB-1 (CHICKEN) LIKE 2 (<i>TOM1L2</i>).....</b>	<b>120</b>
Background.....	120
Rationale.....	120
Results.....	120
Bioinformatic analysis fo <i>TOM1L2</i> .....	121
Northern analysis of <i>TOM1L2</i> .....	122
<i>Tom1l2</i> gene trapped ES cell line.....	124
XG909 genotyping and creation of XG909 mice.....	125
Analysis of the F2 generation of <i>Tom1l2</i> gene-trapped mice.....	126
Summary.....	130
Materials and methods.....	135
<b>CHAPTER 7. DISCUSSION.....</b>	<b>136</b>
Future studies in the Elsea Laboratory.....	137
<i>TOM1L2/Tom1l2</i> .....	137
<i>DRG2/Drg2</i> .....	140

<i>RAI1/Rai1</i> .....	141
Conclusion.....	145
<b>APPENDIX A</b> .....	<b>147</b>
<b>APPENDIX B</b> .....	<b>156</b>
<b>APPENDIX C</b> .....	<b>157</b>
<b>REFERENCES</b> .....	<b>163</b>

## LIST OF TABLES

Table 1. ESTs mapped by Lucas <i>et al.</i> , within the 2001 ~1.5 Mb critical interval identified as known genes.....	34
Table 2. Novel ESTs mapped by the Elsea Lab within the 2001 ~1.5 Mb critical interval.....	35
Table 3. ESTs recently mapped by other groups within the 2001 ~1.5 Mb critical interval.....	36
Table 4. SMS candidate gene disruption in mouse.....	39
Table 5. Phenotypic comparison of SMS patients with 17p deletions.....	61
Table 6. Frequency and type of SMS deletions.....	66
Table 7. Phenotypic comparison of SMS patients with deletions and <i>RAI1</i> mutations.....	82



## LIST OF FIGURES

**Images in this dissertation are presented in color.**

Figure 1. Chromosome 17p somatic cell hybrid mapping panel.....	21
Figure 2. Somatic cell mapping of the <i>RAI1</i> gene.....	25
Figure 3. Screening of the gridded RP11 BAC library with D17S740.....	27
Figure 4. Mapping of <i>RAI1</i> to BACs/PACs.....	28
Figure 5. Contig and transcription map of the SMS critical region.....	32
Figure 6. Human <i>RAI1</i> northern analysis.....	33
Figure 7. Examples of deleted and not deleted FISH experiments.....	50
Figure 8. Summary of 17p11.2 deletions identified in this study.....	55
Figure 9. SMS patients with unusual deletions used to refine the SMCR.....	58
Figure 10. Location of SMS FISH probes in relation to 17p11.2 deletions.....	60
Figure 11. SMS129 <i>RAI1</i> mutation analysis.....	74
Figure 12. SMS153 <i>RAI1</i> mutation analysis.....	77
Figure 13. SMS156 <i>RAI1</i> mutation analysis.....	79
Figure 14. SMS159 <i>RAI1</i> mutation analysis.....	81
Figure 15. Somatic cell mapping of <i>DRG2</i> .....	98
Figure 16. Assignment of DRG2 amino acid sequences.....	100
Figure 17. Northern analysis of <i>DRG2/Drg2</i> .....	103
Figure 18. Transfection of <i>Drg2</i> into COS-7 cells.....	105
Figure 19. Temporal expression of <i>Drg2</i> during mouse development.....	107
Figure 20. Cellular localiztion of <i>Drg2</i> with GFP.....	113

Figure 21. <i>TOM1L2</i> northern analysis.....	123
Figure 22. Genetic analysis of the XG909 cell line.....	128
Figure 23. Statistical analysis of XG909 mouse measurements.....	132

## LIST OF ABBREVIATIONS

<b>a.a.:</b>	amino acid
<b>AGS:</b>	Alagille syndrome
<b>ARF:</b>	ADP-ribosylation factor
<b>ATPAF2:</b>	ATP synthase mitochondria F1 complex assembly factor 2
<b>aMT6:</b>	6-sulphatoxymelatonin
<b>ANOVA:</b>	analysis of variance
<b>BAC:</b>	bacterial artificial chromosome
<b>BLAST:</b>	basic local alignment search tool
<b>CEPH:</b>	Centre d'Etude du Polymorphisme Humain
<b>CLP:</b>	coactosin like protein
<b>CLYN2:</b>	cytoplasmic linker 2
<b>del:</b>	deletion or deleted
<b>DRG2:</b>	developmentally regulated GTP-binding protein
<b>ELN:</b>	elastin
<b>ES:</b>	embryonic stem
<b>EST:</b>	expressed sequence tag
<b>FISH:</b>	fluorescent <i>in situ</i> hybridization
<b>GFP:</b>	green fluorescent protein
<b>FLII:</b>	<i>Drosophila</i> flightless I homolog
<b>FZD3:</b>	frizzled 3
<b>HBSS:</b>	Hank's balanced salt solution
<b>HGP:</b>	Human Genome Project
<b>HTGS:</b>	high-throughput genome sequencing
<b>IHC:</b>	immunohistochemistry
<b>kb:</b>	kilobase
<b>JAG1:</b>	jagged 1
<b>LCR:</b>	low-copy repeats
<b>LIMK2:</b>	LIM domain kinase 2
<b>LLGL1:</b>	lethal giant larvae homolog 1
<b>Mb:</b>	megabase
<b>MR:</b>	mental retardation
<b>NAHR:</b>	nonallelic homologous recombination
<b>NCBI:</b>	National Center for Biotechnology Information
<b>NHGRI:</b>	National Human Genome Research Institute
<b>NIH:</b>	National Institutes of Health
<b>OMIM:</b>	Online Mendelian Inheritance in Man
<b>PCR:</b>	polymerase chain reaction
<b>PAC:</b>	P1-artificial chromosome
<b>PEMT2:</b>	phosphatidylethanolamine N-methyltransferase
<b>RA:</b>	retinoic acid
<b>RAI1:</b>	retinoic acid induced 1
<b>RASD1:</b>	Ras, dexamethasone-induced 1

<b><i>RFC2:</i></b>	replication factor c subunit 2
<b>REM sleep:</b>	rapid eye movement sleep
<b>RSTS:</b>	Rubinstein-Taybi syndrome
<b>SCN:</b>	suprachiasmatic nucleus
<b>SIB:</b>	self-injurious behavior
<b>SMS:</b>	Smith-Magenis syndrome
<b>SMCR:</b>	Smith-Magenis critical region
<b>SMS-REPD:</b>	distal SMS repetitive element
<b>SMS-REPM:</b>	middle SMS repetitive element
<b>SMS-REPP:</b>	proximal SMS repetitive element
<b>SNP:</b>	single nucleotide polymorphism
<b><i>SRP:</i></b>	signal recognition particle from canine pancreas
<b>SSCP:</b>	signal strand conformation polymorphism
<b><i>STX1A:</i></b>	syntaxin 1
<b>SVAS:</b>	supravalvular aortic stenosis
<b><i>TOM1L2:</i></b>	Target of myb-1 (chicken) like 2
<b><i>TRE:</i></b>	human tumor specific antigen
<b>UTR:</b>	untranslated region
<b>VHS:</b>	Vps27p, Hrs and STAM
<b>WS:</b>	Williams syndrome
<b>YAC:</b>	yeast artificial chromosome

## **Chapter I**

### **Introduction**

Smith-Magenis syndrome (SMS; OMIM#182290) is a multiple congenital anomalies and mental retardation syndrome usually associated with a *de novo* interstitial deletion of chromosome 17p11.2 (Smith et al. 1986; Stratton et al. 1986; Lockwood et al. 1988; de Almeida et al. 1989). The 17p deletions associated with SMS are difficult to detect using standard cytogenetic Giemsa staining (G-banding) techniques and are thus termed microdeletions. In addition to SMS, there are many other microdeletion syndromes including Rubinstein-Taybi syndrome (del 16p13.3), Miller-Dieker syndrome (del 17p13.3), Alagille syndrome (del 20p11.23), DiGeorge/velocardiofacial syndrome (del 22q11.2), and Williams syndrome (del 7q11.2). As cytogenetic techniques improved, additional microdeletion disorders were described. The SMS-associated 17p11.2 deletions usually require high resolution banding to the 550 or 850 band level in order to be detected with G-banding. This technology did not become reliably available until the early 1980s, which is likely why the SMS phenotype was not associated with 17p11.2 deletions until 1982 with the initial description by Ann Smith and Ellen Magenis. With the advent of fluorescent *in situ* hybridization (FISH), diagnostic testing for SMS has become more routine. In just 20 years, the number of known SMS cases now totals in the thousands worldwide. Unfortunately, the disease is not widely known and under-diagnosis is likely. Birth incidence of SMS was estimated in one study at 1:25,000 live births (Greenberg et al. 1991).

The association of SMS with a microdeletion of chromosome 17p11.2 makes the syndrome an excellent candidate as a contiguous gene disorder. Contiguous gene syndromes are a group of identifiable disorders caused by chromosomal abnormalities such as deletions or duplications affecting normal gene dosage (Schmickel 1986). The genes affected by the deletion or duplication are generally related only by their proximity to each other along the chromosome and not by their function. The dosage sensitive genes are thought to contribute independently to the characteristics of the contiguous gene syndrome; although, it is possible that a change in the gene dosage of some genes would have no biochemical effect and thus may not contribute to the overall phenotype. It is also possible that dosage changes in one gene causes the entire phenotype seen in the microdeletion syndrome, making the disease a single gene disorder. Additionally, some aspects of the contiguous gene disorder phenotype may segregate as Mendelian inherited traits (Schmickel 1986). The ultimate goal when studying contiguous gene disorders is to find the gene or genes responsible for the characteristics that make up the syndrome phenotype.

### **The SMS Phenotype**

The SMS newborn is usually born at term of normal weight and length; APGAR scores are generally normal. The baby usually has a facial appearance that is described as cherubic, caused by mid-face hypoplasia, full cheeks, a tented upper lip, upward slanting eyes, an upturned nose, and micrognathia. The facial features can be subtle and are easily overlooked (Allanson et al. 1999). Infantile hypotonia is usually present, and the infant may display poor sucking/swallowing and feeding. Failure to

thrive has been noted in some cases (Smith et al. 1986; Greenberg et al. 1996). Parents describe the SMS infant as the perfect baby who cries infrequently, has a smiling disposition, and sleeps well through the night. The infant is usually described as complacent, lethargic, and generally happy. These features overlap with other genetic disorders, and SMS children are often misdiagnosed based on phenotypic characteristics with Down syndrome or Prader-Willi syndrome (Allanson et al. 1999). Additional congenital anomalies including cleft lip/palate, renal/urinary tract anomalies, seizures, thyroid abnormalities, and heart defects have been noted in ~25% of patients (Chen et al. 1996; Greenberg et al. 1996).

During childhood, the SMS phenotype begins to become more apparent. Though overall length is normal at birth, growth is slow and children are generally short during childhood. Slower growth results in a shifted growth curve, but the patients usually end up within the normal height range after puberty. Though height is within normal limits, SMS patients are usually shorter than their genetically predicted heights. Brachydactyly is almost always present. With age, the facial features begin to become more distinctive to SMS. The eyes remain up slanting. The lower jaw becomes more broad and protrusive causing a distinct square looking face especially during puberty. Mid-face hypoplasia becomes more noticeable, and mild brachycephaly is common. The brow is usually noted as heavy with the presence of synophrys adding to the effect (Allanson et al. 1999). Scoliosis can become problematic during childhood and is reported in over 50% of SMS children over age 4 (Greenberg et al. 1996).

Further pediatric evaluations in SMS patients can reveal additional anomalies.

Ophthalmic abnormalities have been thoroughly studied in SMS patients with a variety of irregularities reported. These include iris abnormalities, microcornea, myopia, and strabismus occurring in ~50 to 60% of patients (Finucane et al. 1993; Chen et al. 1996). Blood workup shows hypercholesterolemia in over 50% of patients (Smith et al. 2002), and decreased immunoglobulins, specifically IgA, in <25% of patients examined (Greenberg et al. 1996; Cassidy and Allanson 2001).

By childhood, the presence of mental retardation (MR) is apparent and is usually in the mild to moderate range (IQ 40-54) (Dykens et al. 1997); however, cases of borderline MR to severe MR have been reported (Greenberg et al. 1996). Parents of SMS patients can expect delay of motor and speech developmental milestones. Signs of peripheral neuropathy including pes planus, pes cavus, depressed deep tendon reflexes, and insensitivity to pain are seen in >75% of SMS patients (Greenberg et al. 1996).

SMS patients also display a spectrum of otolaryngeal anomalies. Otitis media is common leading to placement of ventilation tubes in the ears to alleviate the frequent infections. Sensorineural or conductive hearing loss are frequent and occur in ~68% of patients (Greenberg et al. 1996). Almost all patients have a very characteristic hoarse deep voice. Laryngeal anomalies reported include polyps, nodules, edema and/or partial vocal cord paralysis (Cassidy and Allanson 2001). Additional otolaryngologic problems seen in SMS patients include, velopharyngeal insufficiency, weak bilabial seal, palatal anomalies, limited tongue motion, and frequent drooling (Greenberg et al. 1996; Cassidy and Allanson 2001). SMS children also have an increased incidence of sinusitis, some requiring surgical intervention



(Cassidy and Allanson 2001). Parents also complain that children are sometimes affected with bouts of constipation (Smith et al. 1998b).

The neurobehavioral hallmarks of SMS discussed below are the most recognizable characteristic and give parents and professionals the most difficulty in dealing with SMS patients. Sleep disturbance is common in >75% of SMS patients (Smith et al. 1998b). Though sleep disturbance is generally not a problem at birth, alteration of sleep patterns begins around 9 months, and sleep progressively deteriorates. In an initial sleep study performed in 1996, 50% of SMS participants (12/24) had a reduced amount of REM sleep (Greenberg et al. 1996). Further study of sleep in SMS patients has revealed an increased number of awakenings during the night and early awakening time in the morning leading to an overall decrease in the amount of total sleep time, daytime sleepiness, and increased number of daytime naps (Smith et al. 1998b).

One of the main hormonal components of sleep is melatonin. Melatonin secretion is controlled by the suprachiasmatic nucleus (SCN) of the brain. The SCN is sometimes referred to as the human biologic clock and is an endogenous keeper of circadian rhythm. The SCN controls secretion of melatonin by signaling through the neurotransmitter norepinephrine to the pineal gland. Light/dark cycles have an impact on the SCN through the retinohypothalamic tract that connects the retina to the SCN. Thus, melatonin secretion can be affected by changes in exposure to light and darkness. Exposure of the photoreceptor cells to light causes their hyperpolarization, inhibiting both the release of norepinephrine and melatonin synthesis. When light is reduced, the cells become less polarized and norepinephrine

is released to the pineal gland, which in turn produces and secretes melatonin to the blood stream (Brzezinski 1997). In humans, melatonin secretion peaks in the middle of the night somewhere between 2 and 4 a.m. and is at its lowest around 12 hours later. Melatonin is quickly metabolized in the liver by hydroxylation and conjugation with sulfuric or glucuronic acid to 6-sulphatoxymelatonin (aMT6s) which is excreted in the urine. Levels of aMT6s in urine mirror blood serum levels of melatonin (Lynch et al. 1975). Thus, measurement of aMT6s in urine is a reliable way to measure melatonin levels over a period of time without the need for repetitive blood draws. Two studies have reported the circadian rhythm of melatonin in SMS patients. These studies have shown an inverted rhythm in 37 of 38 SMS cases studied (Potocki et al. 2000b; De Leersnyder et al. 2001a). In the SMS individual, melatonin peaks midday and is at its lowest in the middle of the night. This inverted circadian rhythm of melatonin may contribute to the nighttime restlessness and early awakening experienced by SMS patients.

In order to try to reduce the nighttime awakenings and early rising, parents of SMS patients have tried melatonin supplementation. Improvement in sleep has not been widely seen with this supplementation. This is likely due to the fact that increasing the amount of melatonin alone will only boost the total amount in the system resulting in two peaks rather than reversion of circadian rhythm back to normal. Two studies were conducted in France in an attempt to correct the sleep disturbance in SMS patients (De Leersnyder et al. 2001b; De Leersnyder et al. 2003). In the initial study, 10 SMS children were treated with the selective  $\beta_1$ -adrenergic antagonist acebutolol in the early morning. With this treatment the patient's midday

melatonin peak went away and sleep at night improved (De Leersnyder et al. 2001b). In the follow-up study children were treated with the  $\beta_1$ -adrenergic antagonist in the morning, and a melatonin dose in the evening (De Leersnyder et al. 2003). Following this scheme restored the normal circadian rhythm of melatonin. The sleep in the children involved in the study improved, nighttime awakenings diminished, and wake time was delayed thus allowing for a longer more restful sleep (De Leersnyder et al. 2001b; De Leersnyder et al. 2003). Currently, clinical trials using similar methodologies and further sleep studies to better understand the SMS sleep disturbance are underway at The Oregon Health and Science University and at The National Institutes of Health.

Self-injurious (SIB) and maladaptive behaviors are another neurobehavioral hallmark of SMS patients. In fact, these behaviors are the most recognizable characteristic that usually causes the physician to suspect SMS. The behaviors associated with SMS can begin as early as age two and usually begin with head-banging and wrist biting (Smith et al. 1998a). These behaviors are often sparked when the child becomes upset or frustrated. Uncontrollable tantrums usually develop in combination with the head-banging and hand/wrist biting. As the children age and become stronger, they can become a danger to themselves or others during periods of tantrum. With age, additional self-injurious behaviors begin to manifest including onychotillomania (pulling/picking of finger and toenails), and polyembolokoilomania (insertion of objects into bodily orifices), usually beginning after age 5 (Greenberg et al. 1996). Skin and nail picking and biting can be persistent and quite severe causing permanent scarring. Additional stereotypical behaviors

common in SMS patients include the spasmodic upper-body squeeze (known as “self-hugging”) and “licking and flipping.” The self-hug takes on two forms, one is an upper body hug, while the other includes clapping together the hands at chin level followed by the interlocking of the fingers and pushing of the hands together with the elbows pointed away from the body. These self-hugging behaviors occur usually when the child is very excited and appear involuntary (Dyken et al. 1997; Smith et al. 1998a). The licking and flipping behavior involves licking the finger before turning the pages of a book. The motion is quick and repetitive with little stopping between pages (Dyken et al. 1997). Various drugs have been tried to combat the behaviors associated with SMS including antidepressants, antipsychotics, and anticonvulsants though none has proven extremely helpful over time and effectiveness seems to vary between individuals (Smith et al. 1998a).

After puberty and throughout adulthood the self-injurious and maladaptive behaviors become less prominent. However, SIB can still be an issue during times of anger. The facial features continue to become more coarse, the chin can become very prominent, and the brow very heavy. Scoliosis can become more severe with age. The life expectancy of people with SMS seems to be normal. The oldest known living SMS patient is over 80 years old. There does not seem to be any increase in the incidence of cancer in SMS patients. Though the phenotype in infants and children is well known, less is known about what can be expected in the SMS adult. Studies of the natural history of SMS are being currently being conducted by Ann Smith at NHGRI/NIH in order to refine the characteristics of the phenotype over the lifetime.

### **17p11.2 deletions associated with SMS**

Association between microduplication and microdeletion syndromes and flanking repetitive elements termed low-copy repeats (LCRs) is known (Ji et al. 2000). It is estimated that LCRs cover between 5%-10% of the human genome (Eichler 1998; Ji et al. 2000). The LCRs can range in size from ~10-400 kb, averaging between 150-200kb, and having >97% homology (Eichler 1998; Ji et al. 2000; Stankiewicz et al. 2003). Because of the size similarity to BACs and PACs used in the human genome project and the repetitive nature of the LCRs (independently occurring in between 2-10 copies), they caused considerable difficulty in mapping and sequencing of the human genome (Bailey et al. 2001). LCRs are different than the highly repetitive repeats, i.e. *Alu*, SINE/LINE, and satellite repeats, which are present in numerous copies and were identified based on reassociation kinetics. LCRs are made up of genes, gene fragments, pseudogenes, and retroviral sequences. Data have shown that nonallelic homologous recombination (NAHR) between the LCRs causes the segmental anomalies seen in the microdeletion/microduplication disorders (Ji et al. 2000). NAHR can occur interchromosomally or intrachromosomally resulting in deletion, duplication, or segmental rearrangement depending on the individual characteristics of the LCRs involved (Stankiewicz and Lupski 2002).

The first physical map of the SMS deletion region constructed in 1997 by Chen *et al.* using yeast artificial chromosomes (YACs) identified the LCRs mapping to the SMS region (Chen et al. 1997). Using a combination of PCR, somatic cell

hybrid mapping, and cosmid mapping, the authors were able to locate three LCRs mapping within 17p11 which were termed SMS-REPs. Each SMS-REP was represented by YACs that contained unique flanking sequence solidifying the hypothesis that there were three unique repetitive regions. The authors termed the regions SMS-REPP (proximal), SMS-REPM (middle), and SMS-REPD (distal). The SMS-REPs were determined to be repeated gene clusters of ~200 kb, though it was not clear whether the genes within the REPs are expressed or are pseudogenes. Three genes map to all three REP sequences including type-1 keratin (*KER*), signal recognition particle from canine pancreas (*SRP*), and the human tumor specific antigen TRE (*TRE*). Another gene, coactosin like protein (*CLP*) only maps to the proximal and distal REPs, and not the middle REP. Based on Southern analysis, the authors estimated the size between the proximal and distal REP to be 5 Mb. The existence of the REP sequences led the authors to investigate whether they could detect the junction fragment that would remain after recombination events leading to 17p deletions associated with SMS. Using a cDNA representing the *CLP* gene, the authors were able to demonstrate an SMS specific junction fragment in 90% of patient samples studied. The authors hypothesized that NAHR between SMS-REPP and SMS-REPD facilitates the majority of the deletions seen in SMS, and that 90% of all SMS patients carry the same deletion that was estimated to cover ~5 Mb (Chen et al. 1997). The reciprocal of the SMS deletion at 17p11.2 would be a segmental duplication. The segmental duplication of 17p11.2 was identified in 7 patients with mild mental retardation and dental anomalies (Potocki et al. 2000a).

Genetic proof of NAHR between the distal and proximal SMS-REPs was

provided via genotyping of 24 patient samples harboring 17p11.2 deletions/duplications. The analysis of the 24 cases showed 11 rearrangements resulting from interchromosomal recombination and 13 from intrachromosomal recombination indicating no preference for exchange between or within sister chromatids or homologs (Shaw et al. 2002). Recombination between the proximal and distal SMS-REPs causes the majority of SMS deletions, but these are not the only 17p11.2 deletions associated with SMS. Deletions overlapping the SMS common interval have been reported (Juyal et al. 1996; Trask et al. 1996; Elsea et al. 1997). These unusual sized deletions can be larger or smaller than the common deletion, and are currently reported to occur in ~25% of patients, and are mediated by NAHR between the distal and middle SMS-REPs and/or other newly discovered low copy repeats in the region (Stankiewicz et al. 2003; Vlangos et al. 2003).

Unusual deletions provide an invaluable resource for the study of microdeletion disorders. Overlapping deletions from patients who retain all aspects of the syndrome phenotype delineate the critical region of the disorder. The gene(s) causing the disease phenotype are most likely to map within the disease critical region. The SMS critical region (SMCR) was first described in 1997 (Elsea et al. 1997). The SMCR was represented by an unusual deletion harbored by a single patient. This patient's deletion was located between the distal and middle SMS-REPs and spanned ~1.5 Mb. Haploinsufficiency of a gene, or genes, mapping within this 1.5 Mb deletion was hypothesized to be the cause of the Smith-Magenis syndrome phenotype.

### **Microdeletion syndromes: from critical region to gene identification**

Determining the role that any of the positional candidate genes within the microdeletion syndrome critical region play in producing the characteristics of the disease phenotype is a complicated process. Multiple approaches are usually employed simultaneously in order to successfully determine the role the candidate genes may play in disease etiology. The techniques employed in this research have been successfully used to determine disease etiology in other microdeletion syndromes including Williams syndrome (WS; del 7q11.2; OMIM#194050) and Alagille syndrome (AGS; del 20p12; OMIM#118450).

Williams syndrome is an autosomal dominant disorder associated with deletion of chromosome 7q11.23. Most deletions occur *de novo*, though incidences of parental transmission have been reported. Patients with Williams syndrome have a distinct facial appearance including a wide mouth, flat nasal bridge, small mandible, and prominent cheeks (Morris et al. 1988). Puberty is often precocious, and there can be a loss of subcutaneous tissue giving rise to a look of premature aging (Donnai and Karmiloff-Smith 2000).

Cardiac anomalies in the form of supravalvular aortic stenosis (SVAS) and peripheral pulmonary artery stenosis are common to WS patients and occur in >80% of patients (Osborne 1999). Stenosis of other peripheral vessels including the subclavian, coronary, carotid, mesenteric, and renal have also been reported.

The behavioral/cognitive profile of WS patients is a striking array of extremes. WS patients are usually affected with moderate to severe mental retardation; average IQ ranges between 51 and 70 (Udwin and Yule 1991). Though



MR is present WS patients do not have problems with language or face and emotional processing. This is in stark contrast to the extreme problems with visuospatial cognition (Donnai and Karmiloff-Smith 2000). An amazing propensity for music is displayed, but hypercussis is present and patients are often startled and display extreme anxiety in the presence of loud noise. Anxiety is also present when the WS patient is placed in a new situation. However, their personality is extremely friendly and social, and fear of strangers is not present (Osborne 1999).

The deletion associated with WS is mediated by recombination between LCR sequences as is seen in SMS (Urban et al. 1996). The WS deletion interval was defined in 1998 by assessing 200 WS patients using FISH. The interval spans ~1.5 Mb of DNA. In order to facilitate mapping of candidate genes a complete BAC/PAC physical map was constructed across the region (Meng et al. 1998).

When the physical map of the WS region was constructed the etiology of one aspect of the phenotype was already known. Isolated SVAS, an autosomal dominant trait, was found to map to the same chromosomal region by linkage analysis and by discovery of a translocation involving chromosome 7q co-segregating in a family with SVAS (Curran et al. 1993; Ewart et al. 1993). It was determined that the translocation breakpoints in these patients disrupted the elastin gene (*ELN*). Investigation of Williams syndrome patients determined that >98% were hemizygous for *ELN* (Lowery et al. 1995). In evaluation of DNA samples from patients with isolated SVAS nonsense and splice site mutations were found within the elastin gene (Li et al. 1997a; Tassabehji et al. 1997). Haploinsufficiency of *ELN* in humans causes the hardening of the vessels seen in SVAS and WS.

Elastin is produced in smooth muscles and eventually is formed into concentric rings of elastic lamellae alternating with rings of smooth muscle (Uitto et al. 1991). Haploinsufficiency of *ELN* accounts for the SVAS seen in WS, but not the other characteristics of the syndrome. Thus, Williams syndrome is a true contiguous gene syndrome where other genes in the critical region are responsible for other aspects of the phenotype. Additional genes mapping to the region include LIM domain kinase-2 (*LIMK2*) and cytoplasmic linker 2 (*CLYN2*) thought to be responsible for the impaired visuospatial cognition and the neurologic alterations respectively. Replication factor c subunit 2 (*RFC2*) is also a positional candidate for WS (Osborne 1999). Additional genes that were thought to be excellent candidate genes included syntaxin 1 (*STX1A*) and frizzled (*FZD3*). Study of two classic WS patients with unusual deletions showed that these genes are not included in the WS critical region as these patients are not deleted for these genes (Botta et al. 1999). To date, haploinsufficiency of the elastin gene is the only known cause for characteristics seen in Williams syndrome. The search for additional causative genes in WS continues through attempts to further narrow the critical region and characterization of the positional candidate genes mapping to the region.

Alagille syndrome (AGS) is an autosomal dominant disorder with variable expression that can be associated with a deletion of chromosome 20p12. Phenotypic diagnosis of AGS is given upon confirmation of paucity of the interlobular bile ducts upon liver biopsy, and at least three of the five major clinical features described by Alagille in 1975 including chronic cholestasis, cardiac disease, skeletal anomalies, ocular anomalies, and characteristic facial features (Alagille et al. 1975; Krantz et al.

1999). Hepatic, cardiac, and ophthalmologic manifestations are the most common features seen in AGS. Most patients are of normal intelligence, though mild delay, delay of gross motor skills, and mild MR have been associated with the disease. Additional anomalies associated with the syndrome include delayed puberty, hearing loss, renal defect, and pancreatic insufficiency (Alagille et al. 1975; Li et al. 1997b; Oda et al. 1997b; Krantz et al. 1999).

From the initial description of the disease it was clear that AGS was an autosomal dominant disorder with variable expressivity based on familial inheritance of the disease (Alagille et al. 1975). In fact, clinical manifestations in AGS are highly variable between individuals, ranging from subclinical to life threatening. Birth incidence is reported between 1:70,000 to 1:100,000 births (Oda et al. 1997b; Krantz et al. 1999). Cytogenetic analysis of most AGS patients reveal a normal karyotype, though a small subset (~7%), have interstitial deletions or rearrangement of chromosome 20p12. Patients who harbor large cytogenetically visible deletions of chromosome 20p are usually more affected with the spectrum of phenotypic characteristics. Thus, AGS was an excellent candidate as a contiguous gene syndrome.

The AGS critical region was delineated using overlapping 20p deletions, and was estimated to span between 8 and 11.3 cM based on CEPH mapping and linkage analysis in AGS patients without a 20p deletion (Hol et al. 1995; Pollet et al. 1995). In the mid 1990s the number of mapped genes and available genetic markers were small. A physical map of a continuous path of YAC clones was constructed spanning 3.7 Mb of 20p12 including the AGS critical region in order to facilitate mapping of

new markers and identification of possible candidate genes via elucidation of CpG islands (Pollet et al. 1995). This first physical map containing 15 new anonymous markers was an excellent start toward identification of AGS candidate genes. Further analysis of AGS patients with 20p deletions narrowed the critical region to ~250 kb. A contig of the refined region was created using BAC clones which are less susceptible to recombination than the more unstable YAC clones. At the time only one gene mapped wholly within this refined AGS critical region, human *Jagged1* (*JAG1*), a gene orthologous to the rat *Jagged* gene which is involved in the notch signaling pathway (Oda et al. 1997b).

The *JAG1* gene encodes a ligand for the Notch receptor. The notch receptor through association with its ligands creates a signaling pathway used in directing cell fate (Luo et al. 1997; Oda et al. 1997a). Because of its possible function and map location *JAG1* was an excellent candidate gene for AGS. Analysis of DNA samples from patients without a deletion of chromosome 20p was performed via SSCP and direct sequencing of PCR amplicons of the *JAG1* gene. Two separate groups identified 11 mutations in the *JAG1* gene segregating in affected AGS patients (Li et al. 1997b; Oda et al. 1997b). Haploinsufficiency of the *JAG1* gene via hemizygous deletion or mutation is the cause of Alagille syndrome. In contrast to the theory of the contiguous gene syndrome hemizygous deletion of other genes flanking the *JAG1* gene does not seem to affect the syndrome phenotype.

### **Animal modeling of Smith-Magenis syndrome**

The mouse is an excellent organism for use in animal modeling of human

disease due to the similarity of physiology, anatomy, and genetics. Mice are also relatively inexpensive to house and propagate. The techniques for manipulating the mouse genome have become readily available and are performed routinely. In addition, the completion of sequencing of the mouse and human genomes facilitates comparative analysis aiding experimental design. Genetic mutation of the mouse can include inactivation of an individual gene, or manipulation of entire segments of DNA via chromosomal engineering.

In the mouse chromosome 11 is syntenic to human chromosome 17. Comparative analysis of the SMS critical region between mice and humans shows that order and orientation of 19 genes are conserved (Bi et al. 2002). Chromosomal engineering was undertaken in order to examine the phenotypic effect of deleting the SMS common region in mouse (Walz et al. 2003). Mice lacking a 3 Mb segment of chromosome 11 syntenic to the SMS common region were created and termed *Df(11)17* animals. Heterozygous animals displayed craniofacial anomalies, EEG abnormalities, weight differences, and decreased fertility when compared to normal littermates in a mixed genetic background. Gross organ abnormalities in the heart, urinary tract, and eyes were not noted (Walz et al. 2003). Behaviorally, the male *Df(11)17* mice were reported to be hypoactive, while the female heterozygotes had normal locomotor activity. Both male and female heterozygotes showed shorter circadian periods when placed in constant darkness after light entraining (Walz et al. 2004). This response indicates a reduced precision in the circadian clock. Heterozygous mice had no self-injurious behaviors, decreased sensation to pain, or startle responses (Walz et al. 2004). These results indicate that a dosage sensitive

gene(s) maps within the chromosomal engineered region. Targeting of individual genes in the region will hopefully aid in understanding the genotype:phenotype correlation in humans.

## **SUMMARY**

**The goal of the research project presented here was to determine the gene(s) responsible for Smith-Magenis syndrome. The work explained in the following chapters follows the methods employed in elucidating the etiology of Alagille and Williams syndromes, as well as other microdeletion disorders. Chapter 2 focuses on work performed in creation of a physical and transcriptional map of the SMS critical region. Chapter 3 describes narrowing of the SMS critical region using FISH, as well as determination of a more efficient diagnostic SMS probe. Chapter 4 illustrates direct sequencing efforts of candidate genes in SMS patients where deletion of chromosome 17p11.2 could not be detected. This research led to the identification of mutations in the *RAI1* gene which is likely responsible for the majority of SMS characteristics. Chapters 5 and 6 focus on characterization of two positional candidate genes, *DRG2* and *TOM1L2* respectively, which may play a role in the less penetrant SMS characteristics. Chapter 7 concludes the dissertation by discussing how the work presented could be incorporated into future research projects studying Smith-Magenis syndrome.**

## **Chapter II**

### **Physical mapping of the Smith-Magenis syndrome critical region**

#### **BACKGROUND**

The goal of the research projects in the Elsea Lab is to identify the dosage sensitive gene(s) mapping to chromosome 17p11.2 responsible for the characteristics making up the SMS phenotype. Prior to my joining the project, much work had been performed toward this goal. The SMS critical region (SMCR) had been defined, a somatic cell hybrid mapping panel had been created, and many markers had been mapped to the short arm of chromosome 17 by the lab. Using the markers previously mapped, we embarked on construction of a continuous path of large insert clones (a “contig”) across the SMS critical region. The creation of the contig also allowed for mapping of additional markers, genes, and ESTs to the SMS critical region. The final contig includes the order and orientation of the genes, ESTs, and markers mapped to the critical region. We initially characterized all the genes and ESTs mapping to the region using various bioinformatic methods and northern analysis. The genes were prioritized for further study based on initial characterization as candidate genes for SMS. Our lab published this work in the European Journal of Human Genetics (Lucas et al. 2001).

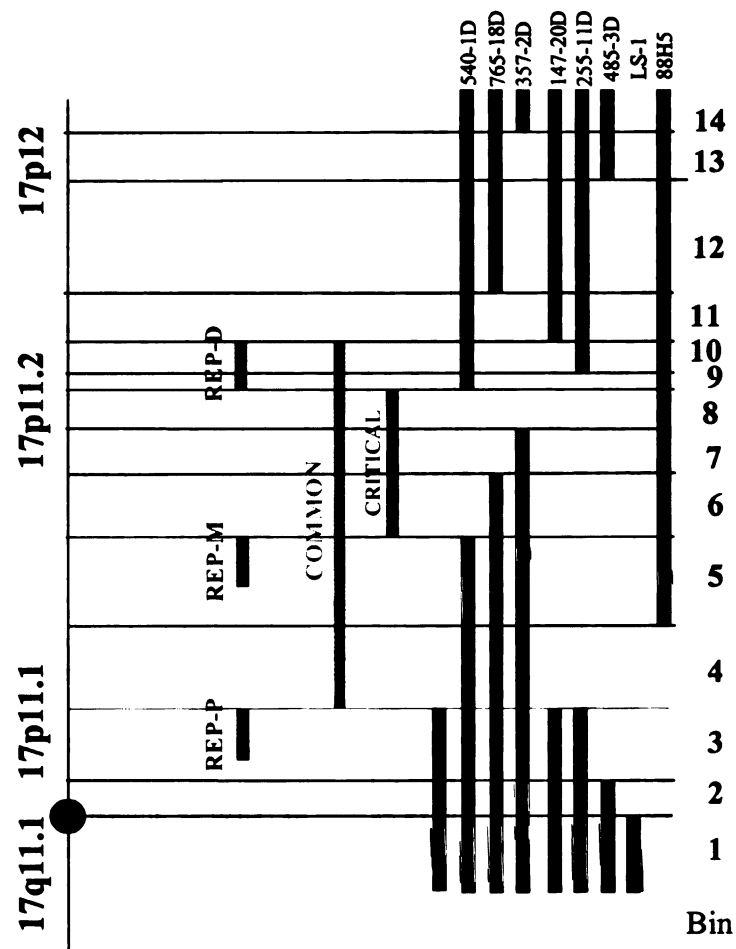
## **The Smith-Magenis syndrome common and critical regions**

As discussed in Chapter I, the deletions associated with SMS can range in size from <2 Mb to >9 Mb (Trask et al. 1996), although it has been reported that >95% of SMS patients carry a common ~3.5 Mb deletion (Chen et al. 1997). Though there are a spectrum of deletions associated with SMS, the phenotype remains consistent in most patients. The dosage-sensitive genes responsible for the syndrome phenotype are theorized to map to the smallest shared region of deletion in patients displaying the full syndrome phenotype. This region is referred to as the critical region. The SMS common deletion and the critical region were delineated in 1996 by studying the deletions carried by 62 SMS patients (Juyal et al. 1996).

At the start of this project, the SMS critical region was defined by the deletion carried by a single patient, HOU142-540 (Figure 1) (Elsea et al. 1997). This patient carries an unusual deletion spanning ~1.5 Mb of DNA roughly between the distal and middle SMS-REPs. Molecularly, the region is bounded by the genomic marker D17S29 proximally and an anonymous cosmid cCI17-638 (Figure 1) (Juyal et al. 1996; Elsea et al. 1997). The deletion size was delineated by FISH using cosmid probes for the small number of known markers mapping to the region (Juyal et al. 1996; Elsea et al. 1997). At the beginning of this project, 7 genes, 13 different ESTs, and 7 markers had been mapped within the SMS critical region (Elsea et al. 1997).

At the same time that the SMS critical region was delineated, the common region boundaries were also identified. The common interval was shown to map between genomic marker D17S58 proximally and the anonymous cosmid cCI17-498





**Figure 1. Chromosome 17p somatic cell hybrid mapping panel.**

Rodent-human somatic cell hybrids retaining a deleted chromosome 17 are illustrated at the right. The breakpoints of the naturally occurring deletions divide the region into bins. Genes and markers mapping to bins were used as anchors in the construction of the critical interval contig. The common deletion seen in SMS is indicated in orange and spans ~3.5 Mb and bins 4-10. The SMS critical interval indicated in red spans ~1.5 Mb and bins 6-8. The approximate location of the SMS-REPs are indicated in blue.

This region encompasses ~3.5 Mb of DNA roughly between the proximal and distal SMS-REPs (Figure 1) (Juyal et al. 1996; Elsea et al. 1997). In molecularly defining the SMS critical region, markers were also mapped to common SMS region along chromosome 17p. The authors mapped 85 additional markers to chromosome 17p surrounding the SMS critical region (Elsea et al. 1997).

Mapping of the markers was facilitated by the creation of a somatic cell hybrid mapping panel of overlapping 17p deletions (Figure 1) (Guzzetta et al. 1992; Patel et al. 1992; Elsea et al. 1997). The rodent:human hybrid cells were grown and maintained to harbor the human chromosome 17 of interest (Guzzetta et al. 1992). The boundaries of the overlapping deletions were used to divide the p-arm of chromosome 17 into 14 bins. The SMS critical interval contained bins six, seven, and eight as delineated by the unusual deletions carried by HOU142-540, and patients with deletions overlapping the SMS critical region carried by HOU261-765 and HOU92-357 (Figure 1) (Elsea et al. 1997). The SMS common deletion was reported to span bins 3 and 10/11 as delineated by numerous patients with overlapping 17p deletions (Figure 1) (Elsea et al. 1997).

A preliminary physical map of the region created using YAC clones was reported in 1997 along with the data showing the presence of the SMS-REPs (Chen et al. 1997). The map and contig reported contained many gaps, and only some of the markers were in the correct order. The orientation of some of the map was inverted based on data reported in the delineation of the critical region. The problems with this initial map likely stem from the lack of markers in the region, and the fact that YAC clones are very prone to recombination during propagation.

## **RATIONALE**

We undertook construction of a physical map and contig of the 1.5 Mb critical interval using BAC, PAC, and cosmid clones. These clones are more stable than the YAC clones used in previous physical maps (Chen et al. 1997). In construction of the physical map, new markers, ESTs, and genes were mapped to the SMS region. Initial characterization of the genes using bioinformatics and northern expression analysis was performed in order to prioritize the candidate genes for further study to determine the role, if any, they play in the SMS phenotype.

## **RESULTS**

### **Mapping of new genes and markers using a somatic cell hybrid mapping panel**

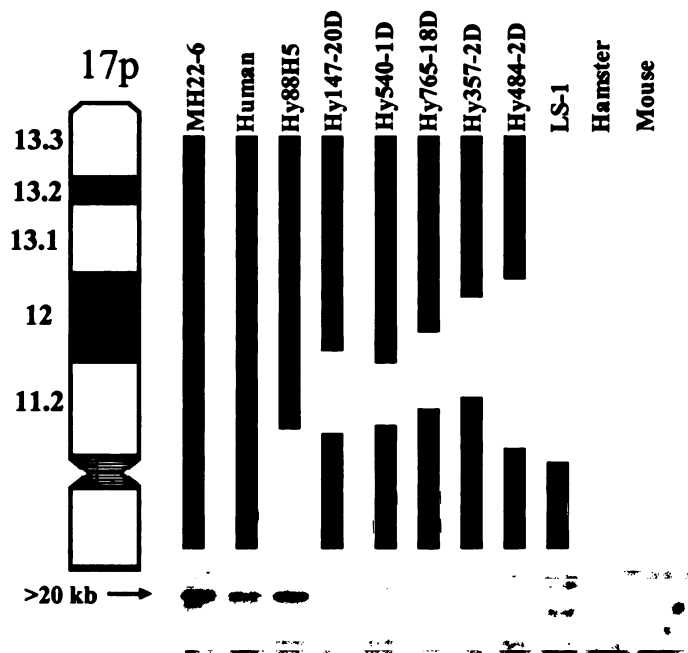
DNA isolated from the rodent:human hybrid cells making up the mapping panel was used in PCR reactions and Southern analysis to fine map newly obtained markers. The panel includes hybrids retaining a deleted human chromosome 17 from SMS patients and non-SMS patients harboring different sized 17p deletions. The following control cell lines were used for mapping along human chromosome 17; MH22-6 retains one entire normal human chromosome 17; LS-1 carries an iso-q 17 chromosome, and 88H5 carries a fused chromosome of 17p:Xq. In order to exclude positive products resulting from the endogenous rodent DNA in the cell lines rodent parental lines a23 and Cl-1D were used.

As the study of the human genome progressed, new markers and ESTs were reported, and their sequences were submitted to public databases at The National

Center for Biotechnology Information (NCBI; <http://www.ncbi.nlm.nih.gov>) including dbEST, dbSTS, and GenBank, as well as to The GDB Human Genome Database (<http://www.gdb.org>). Members of the laboratory regularly searched the databases to identify newly deposited information. The reagents necessary for fine mapping the markers to the SMS critical interval were obtained or created as described below.

To map newly identified genomic markers to the region, specific oligonucleotide primer pairs were obtained for use in PCR using template DNA isolated from the rodent:human somatic cell hybrids. PCR reactions were performed, and products were analyzed for presence or absence of the expected banding pattern based on a positive human control sample. Once markers mapped within the SMS region, they could be used for screening of gridded clone libraries as described below.

Newly identified EST sequences were obtained as clones through The Integrated Molecular Analysis of Genomes and their Expression (I.M.A.G.E.) consortium (<http://image.llnl.gov>) or other international EST consortiums located in Germany and Japan. Plasmid DNA was isolated and sequenced to ensure the correct clone was sent. The clone DNA was used as probe in Southern analysis to filters created through *Eco*RI digestion of the hybrid DNAs and controls. Banding patterns were analyzed to determine map location of the ESTs (Figure 2). Cloned EST DNA was also used in initial characterization of the ESTs as described below.



**Figure 2. Somatic cell mapping of the *RAI1* gene.**

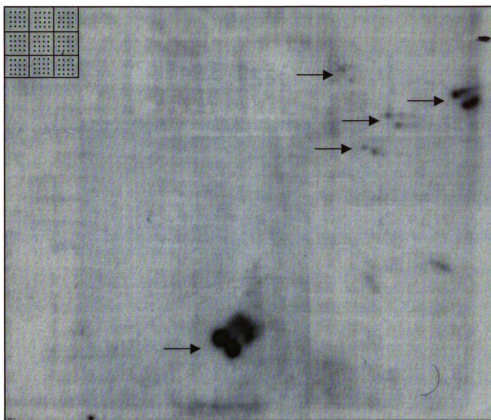
Fine mapping of *RAI1* within the SMS critical interval was performed by hybridizing the ~2.0 kb insert from EST DKFZp434A139Q2 to a panel of *Eco*RI digested somatic cell hybrids. The hybrids carry deleted chromosome 17 from SMS patients (Hy14720-D, Hy540-1D, and Hy484-2D), and non SMS patients with overlapping deletions along 17p (Hy765-1D and Hy357-2D). Proper positive and negative controls are described in the text (MS22-6, Hy88H5, LS-1, a23, and C1-1D). A >20 kb band representing the *RAI1* gene is evident in total human DNA, MH22-6, and Hy88H5 (present but faint because of limited exposure time), demonstrating that *RAI1* is deleted in all SMS patients and maps to the central portion of the SMS critical region.

### **Construction of the physical map and contig of the SMS critical region**

Construction of a continuous path of large insert clones across the critical region was started while the Human Genome Project (HGP) was in its infancy. Sequence data was slow to appear within the SMS critical interval, and sequence data that did appear contained large gaps. Additionally, the markers identified by the HGP that matched those that we had previously fine-mapped to the region usually were in a different order than our data showed. These problems were likely caused by incorrect construction of the sequence fragments produced by shotgun sequencing due to the highly repetitive nature of the DNA in the region.

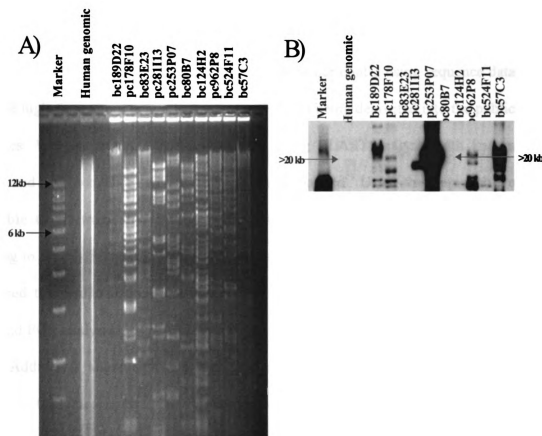
In order to find large insert clones mapping to the region, we repeatedly screened the RP11 BAC library using the markers and ESTs mapping to the SMS critical region. PCR amplified marker, isolated EST DNA, or isolated cosmid DNA was used as probe and hybridized to the gridded arrays using standard Southern protocol. Positive clones are identified by specific spotting patterns displayed on the grid and decoded using a key provided by the library creators (Figure 3).

Positive BACs were obtained and DNA was isolated using a modified Qiagen protocol designed in the lab. An initial contig was created using a combination of *Alu*-PCR and analysis of the *Eco*RI restriction patterns of the BACs. In order to determine the marker order on the individual BACs, Southern analysis was performed using the collection of markers mapped to the region as probes on *Eco*RI digested BAC DNA (Figure 4). Using these methods we were able to construct three contigs



**Figure 3. Screening of the gridded RP11 BAC library with D17S740.**

The gridded RPCI-11 BAC library was screened using ESTs and genomic markers with known map positions inside the SMS critical region. Clones are spotted in duplicate within a 4 x 4 grid on the array. An example of the gridding pattern of the library is indicated in the upper left. Standard Southern technique with radio-labeled DNA probes was used in the screening. In the example above, the library was screened with the genomic marker D17S740. Positive clones are noted with black arrows. Positive clones were ordered and verified for map position by *Eco*RI Southern analysis.



**Figure 4. Mapping of *RAI1* to BACs/PACs.**

**A)** Eight micrograms of DNA from BACs and PACs mapping within the SMS critical region and human genomic DNA (10  $\mu$ g) were digested and then electrophoresed overnight in circulating buffer. The gel was photographed and DNA was transferred to nylon membranes using standard Southern technique. **B)** An autoradiograph of the Southern blot is shown. The insert from plasmid DKFZp434A139Q2 represents the 3' end of the *RAI1* gene and was used as a probe. The blue arrows indicate pc253P07 and a faint corresponding band in the digested human DNA lane.



within the critical region, but were unable to fill the gaps with the BAC resources available at the time. As our project continued, the speed of data coming from the HGP began to increase. We were able to fill gaps in our contig using sequence data from the high-throughput genomic sequences (HTGS) being deposited into the public databases. We searched the HTGS databases using BLAST analysis with known marker and EST sequences which we had already mapped. Using this approach we were able to electronically identify numerous RPCI BACs/PACs and CIT BACs mapping to the SMS critical region. We obtained the positive BACs and PACs and confirmed the results obtained *in silico* using Southern analysis on *EcoRI* digested DNA and PCR analysis.

Additional alignments and contig organization were performed using the UCSC Genome Browser (<http://genome.ucsc.edu>) and Ensembl (<http://www.ensembl.org>). Using these methods, we were able to construct a fluid path of clones with a minimum tiling path of 16 BACs and 2 PACs (Figure 5) (Lucas et al. 2001). In this map several overlapping clones were included to provide greater coverage (Figure 5).

Transcriptional orientation of the genes in the region was determined using the data from the HGP, sequencing of cosmids, and known overlap of genes, such as the 3' ends of the *FLII* and *LLGL1* genes (Campbell et al. 1997). Cosmid genomic clone DNA was isolated from previously mapped cosmids (Juyal et al. 1996; Elsea et al. 1997) as well as from screening of gridded filters of the Los Alamos flow sorted chromosome 17-specific cosmid library (Kallioniemi et al. 1994). At the time of publication, we were able to determine transcriptional orientation of 18 of the

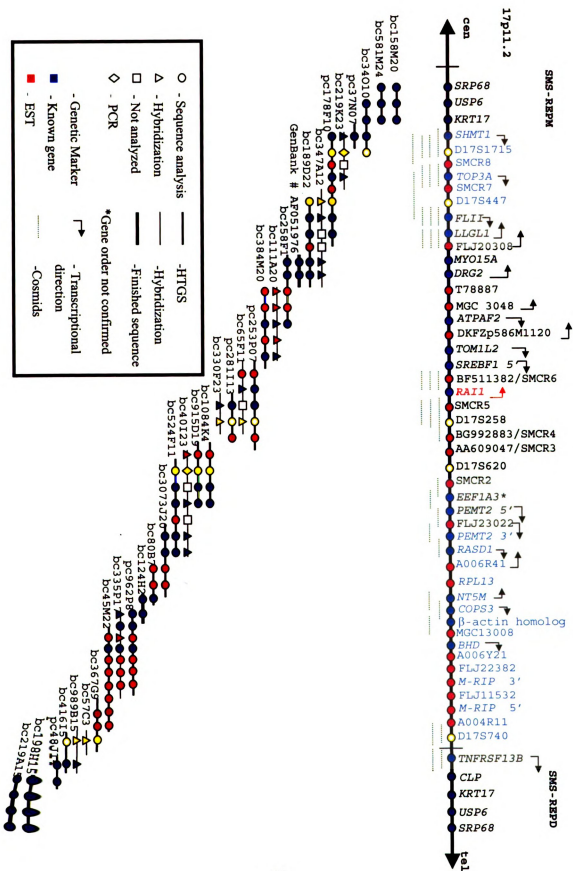
expressed sequences mapping to the critical region (Lucas et al. 2001). The contig as published in 2001 (Lucas et al. 2001) contained 17 known genes, 12 ESTs, and 6 genomic markers (Figure 5).

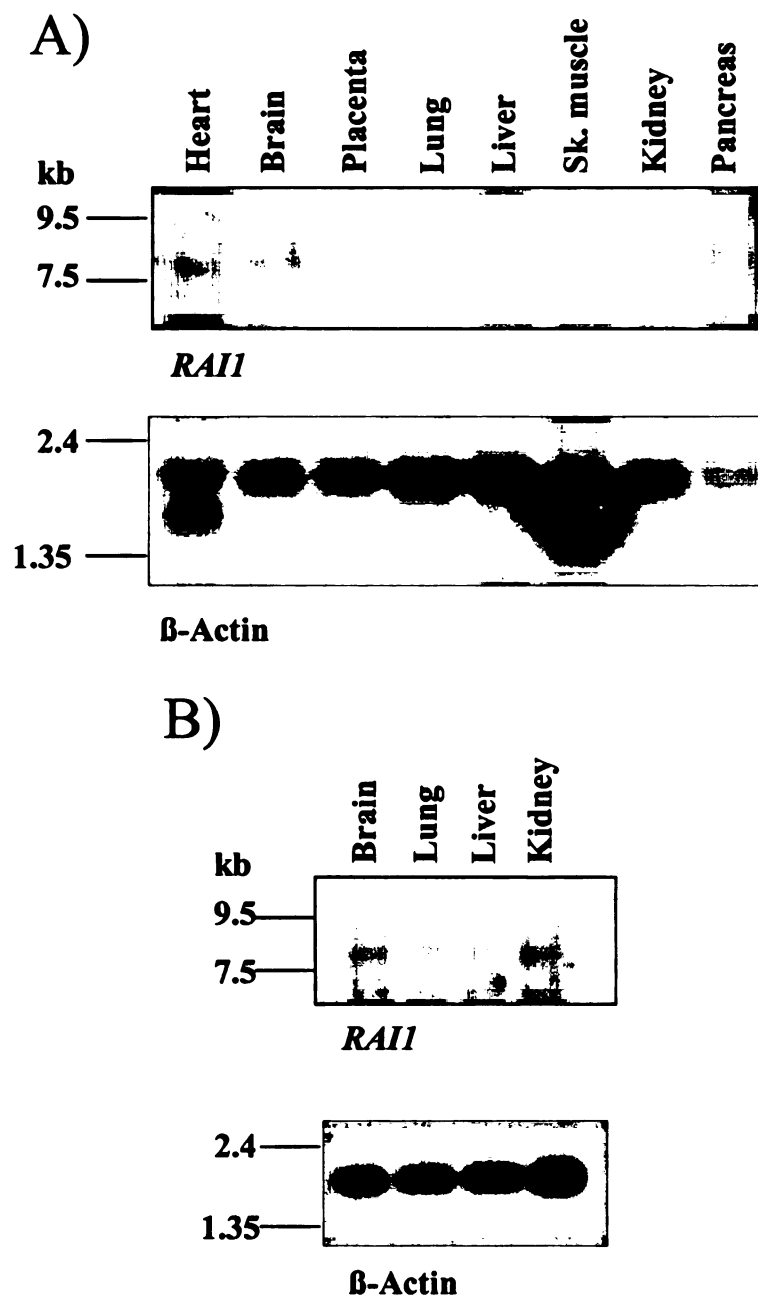
#### **Analysis and characterization ESTs mapped to the SMS critical region**

Through the creation of the physical map of the SMS critical region we were able to identify numerous expressed sequences mapping to the region. These sequences were not only important in creation of the map, but because of their map location likely represented positional candidate genes for SMS. Clones for 13 of the ESTs were commercially obtained and plasmid DNA was isolated. The plasmid DNA was sequenced and database analysis of the DNA sequence was performed against online databases. Further, the DNA sequences of the ESTs were translated in all 6 frames and the amino acid sequences were examined to look for homologous genes and/or sequence motifs. Tissue expression patterns were then explored by probing commercially available multiple tissue northern (MTN) blots of adult and fetal tissues with the ESTs obtained. Northern analysis of EST DKFZp434A139Q2 is shown in Figure 6. The clone represents the 3' end of the human retinoic acid induced 1 (*RAI1*) gene. The analysis shows an ~8.0 kb transcript in all tissues examined. Examination of all ESTs was performed as discussed above. Summary of the genes mapping within the SMS critical region as defined by the deletion carried by HOU142-540 are described in Tables 2-4. Table 1 shows seven ESTs reported to map to the SMCR by our lab in 2001 (Lucas et al. 2001); those indicated in gray no

**Figure 5. Transcription map of the ~1.5 MB SMS critical interval (adapted from Lucas *et al.*, 2001).**

The SMS critical interval as of 2001 (Juyal *et al.*, and Elsea *et al.*, 1997) spanned between the two vertical hatch marks. This distance is roughly between the proximal and middle SMS-REPs. The contig is comprised of overlapping BAC and PAC clones which span the distance from the genomic marker D17S40 to the proximal end of the distal SMS-REP. Markers, ESTs, and known genes were mapped to the contig through a combination of PCR mapping, Southern hybridization to digested BAC and PAC DNA, and database analysis of draft human genome sequence. Overlap of some BAC/PAC sequence confirmed the position of clones in the absence of available markers. Refinement of the SMS critical region has been performed and will be discussed further in Chapter 3. Genes shown in blue no longer map within the SMS critical region.





**Figure 6. Human *RAI1* northern analysis.**

The ~2.0 kb insert from EST DKFZp434A139Q2 represents the 3' end of the *RAI1* gene. The insert was hybridized to A) adult and B) fetal multiple tissue northern blots from Clontech. An ~8.0 kb transcript is evident in all adult and fetal tissues.  $\beta$ -actin is included as an mRNA loading control for both blots.

**Table 1. ESTs mapped by Lucas et al. within the 2001 ~1.5 Mb critical interval which have been identified as known genes.**

EST amplicon(s) (GenBank accession)	Unigene	Tissue expression	Protein motifs	Known gene
WI-13499 (GenBank R41366)	Hs.78582	2.2 kb transcript in all adult and fetal tissues <sup>a</sup>	GTP-binding protein motif	<i>DRG2</i> (GenBank X80754)
D17S2021 (GenBank AA281720)	Hs.13434	1.8 kb transcript in all adult and fetal tissues; highest in heart	Mitochondrial leader sequence	ATPAF2 (GenBank AF052185)
A003A44 (GenBank AA236905) stSG9692 (GenBank AF038192)	Hs.8125 Hs.12537	6.5 kb transcript in all adult tissues, highest in heart, brain, sk. muscle	VHS and GAT domains	TOM1L2 (GenBank NM_144678)
DKFZp434A139Q2 (GenBank AL133649)	Hs.278684	8.0 kb transcript in all adult tissues	CAG repeat, nuclear localization signal, PHD zinc finger domain	<i>RAI1</i> (GenBank AY172136)
NIB1041 (GenBank T16275)	Hs.106359	5 kb transcript in all adult tissues <sup>a</sup>	Member of Ras subfamily	<i>RASD1</i> (GenBank NM_016084)
WI-11472 (GenBank R72633)	Hs.16614	1.6 kb transcript in adult heart, brain, and skeletal muscle <sup>b</sup>	Hydrolase domain	<i>NT5M</i> (GenBank NM_020201)
stSG26124 (GenBank AI570799)	Hs.84883	6.0 kb transcript in all adult and fetal tissues	KOG4807: domain from F-actin binding protein	<i>M-RIP</i> (interim gene symbol) (GenBank NM_015134)

Note: Genes have been ordered proximal to distal and shading indicates gene is located outside of the current SMS critical interval (Vlangos et al., 2003). Protein motifs identified from NCBI conserved domain database and NCBI protein-protein BLAST alignments. <sup>a</sup>Results reported in Vlangos et al., 2000; <sup>b</sup>Northern results similar to those reported in Rampazzo et al., 2000; <sup>c</sup>Northern results similar to those reported in Tu and Wu, 1999.

**Table 2. Novel ESTs mapped by the Elsea lab within the 2001 ~1.5 Mb critical interval.**

EST Amplimer (GenBank Accession)	Unigene	Tissue expression	Protein motifs
T78887 (GenBank AA730163)	Hs.187422	Clontech MTE blot showed expression in heart, GI tract, germ cell, lung, kidney, placenta	378 bp LINE1 repeat
stSG8339 (GenBank H57290)		Clontech MTE blot showed faint expression in heart, GI tract, brain, lung, placenta, germ cell	93 bp MIR repeat
FL20308 (GenBank BC062339)	Hs.356770	3.5 and 4.0 kb transcript in all adult and fetal tissues examined	KOG4176: uncharacterized conserved domain
MGC3048 (GenBank BC000636)	Hs.115437	4.4 kb and 4.0 kb transcript in adult heart, sk. muscle, kidney, pancreas and fetal liver and lung; 4.0 kb transcript in adult and fetal brain, lung and placenta	KOG4635: domain found in vacuolar import and degradation proteins
DKFZp586M1120 (GenBank AL136926)	Hs.159068	Unigene reports 5.8 kb transcript and ESTs from several adult tissues	43% similarity (140 aa alignment) to sds22 protein homolog (GenProt S68209)
FLJ23022 (GenBank NM_025051)	Hs.287717	Unigene reports SAGE tags from several adult tissues	none
A006Y21 (GenBank H50830)	Hs.31652	Unigene reports ESTs from several adult tissues	322 bp Alu repeat; 85.7% alignment to phospholipase D active site motif
FLJ22382 (GenBank AK026035)	Hs.46783	ESTs from several adult tissues	none
FLJ11532 (GenBank AK021594)	Hs.296656	ESTs from brain and embryo	78 bp Alu repeat
A004R11 (GenBank T91728)	Hs.16899	2.4 kb transcript in adult and fetal liver	none
A006R41/FLJ10193 (GenBank AI500641)	Hs.235195	6.0 kb transcript in adult skeletal muscle and fetal brain	Moderately similar to Drosophila protein RH42446p (GenProt AAN71576)
MGC13008 (GenBank NM_032686)	Hs.326732	ESTs from rhabdomyosarcoma	none

Note: Genes have been ordered proximal to distal and shading indicates that this gene is localized outside of the current SMS critical interval (Vlangos et al. 2003). It remains a possibility that some of these unique ESTs can contain genomic contamination or may be cloning artifacts. Protein motifs were identified from NCBI conserved domain database and NCBI protein-protein BLAST alignments.

**Table 3. ESTs recently mapped by other groups within the 2001 ~1.5 Mb critical interval.**

<b>EST Amplimer (GenBank Accession)</b>	<b>Unigene</b>	<b>Tissue Expression</b>	<b>Protein motifs</b>	<b>Reference</b>
SMCR8 (GenBank AF467440)	Hs.513986	2.8, 3.0, and 6.5 kb transcript in all adult and fetal tissues	KOG3715: domain found in LST7 permease Golgi transport protein	Bi et al. (2002)
SMCR7 (GenBank AF467443)	Hs.100448	2.4 and 3.4 kb transcripts in all adult and fetal tissues, predominantly in heart and sk. muscle	KOG3963: motif found in Mab-21- like cell fate proteins	Bi et al. (2002)
SMCR6 (GenBank BF511382)	Hs.443639	RT-PCR demonstrated expression in all tissues	none	Bi et al. (2002)
SMCR5 (GenBank AF467442)	Hs.352643	RT-PCR demonstrated expression in all tissues	none	Bi et al. (2002)
SMCR4 (GenBank BG992883)	none	RT-PCR demonstrated expression in all adult tissues and fetus	none	Bi et al. (2002)
SMCR3 (GenBank AA609047)	Hs.373802	RT-PCR demonstrated expression in all adult tissues and fetus	none	Bi et al. (2002)
SMCR2	none	RT-PCR demonstrated expression in adult brain, and fetus	none	Bi et al. (2002)
SMCR9	none	RT-PCR demonstrated expression in all adult tissues and fetus	Two PDZ domains	Bi et al. (2002)
BHD (GenBank AF517523)	Hs.396533	3.8 kb transcript in all adult tissues	KOG3715: domain found in LST7 aa permease Golgi transport protein	Nickerson et al. (2002)

Following the publication of our physical and transcription map in 2001 (Lucas et al., 2001), several other genes were mapped to this region by other groups. Genes have been ordered proximal to distal and shading indicates that this gene is localized outside of the current SMS critical interval (Vlangos et al., 2003). Domains were identified from NCBI conserved domain database.



longer map within the critical region and will be discussed in Chapter 3. ESTs listed in Table 2 are additional novel ESTs we have since mapped to the SMCR. The current expression and bioinformatic information about these expressed sequences is included. Some of the sequences listed in Table 2 are novel and/or do not show banding patterns upon northern analysis. These sequences may be artifacts from EST cloning in the form of genomic sequence. It is also possible that the sequences are not expressed in tissues represented or are expressed at too low of a level to be detected on northern analysis. Further bioinformatic analysis may reveal more about these possible transcripts as the genomic sequence is polished. Additional groups have mapped EST sequences within the 1.5 Mb region represented by the deletion carried by HOU142-540. These sequences are shown in Table 3. Any of the EST sequences mapping within the SMS critical region may represent genes that play a role in modification or expression of the SMS phenotype. It is also possible that genes mapping outside the critical region may have an effect on the phenotype due to breakpoint location within the expressed sequence or disruption or deletion of yet to be discovered promoter or enhancer sequences. Though, the focus of this work is on genes mapping within the SMS critical region as they are most likely responsible for the characteristics making up the SMS phenotype.

#### **Summary of known candidate genes mapping within the SMS critical region**

Mapping of genes and ESTs to the SMS critical region provided us with numerous potential positional candidate genes for SMS. We were most interested in genes showing expression during fetal development or in brain tissues. Those genes

containing known protein motifs important in cell signaling, DNA binding, or protein-protein interaction domains were given priority. We believed genes with these characteristics were likely to be most susceptible to changes in dosage.

Several genes mapping within the SMS critical region have been well characterized by others and are no longer candidates of high priority for SMS; these include *TOP3A* (Hanai et al. 1996; Fritz et al. 1997; Elsea et al. 1998), *FLII* (Campbell et al. 1993; Chen et al. 1995; Campbell et al. 1997; Campbell et al. 2002), *MYO15A* (Wang et al. 1998; Liang et al. 1999), *SREBF1* (Hua et al. 1995; Smith et al. 2002), *PEMT2* (Sesca et al. 1996; Vance 1996; Vance et al. 1997; Walkey et al. 1997), and *COPS3* (Elsea et al. 1999; Potocki et al. 2000b; Yan et al. 2003). Further, gene targeting of mouse orthologs of the six genes listed above (*Top3a*, *Fliih*, *Myo15*, *Srebf1*, *Pemt*, and *Cops3*) show no phenotype in heterozygous mice (Table 4). This is in contrast to the report indicating an SMS like physical phenotype in mice with heterozygous deletion of the syntenic SMS region on mouse chromosome 11 (Walz et al. 2003). Combined, these data suggest that these six genes are not likely dosage-sensitive in the hemizygous state and are no longer high priority candidates for our study.

The candidate genes of most interest to the lab during my tenure included *LLGL1*, *DRG2*, *ATPAF2*, *TOMIL2*, and *RAI1*. The contribution of these genes to the SMS phenotype was unknown at the beginning of this project. Studies of *RAI1*, *DRG2*, and *TOMIL2* will be discussed at length in chapters 4, 5, and 6, respectively. A summary of the two remaining candidate genes *LLGL1* and *ATPAF2* is given below:

**Table 4. SMS candidate gene disruption in mouse<sup>a</sup>.**

Gene disrupted	Publication	Heterozygous phenotype	Homozygous phenotype
<i>Pemt</i>	Walkey et al. (1997)	No abnormal phenotype	No abnormal phenotype on normal diet
<i>Srebfl</i>	Shimano et al. (1997)	No abnormal phenotype	No abnormal phenotype in live births, ~65% embryonic lethal
<i>Myo15</i>	Probst et al. (1998)	No abnormal phenotype	Profound deafness and vestibular defects
<i>Fliih</i>	Campbell et al (2002)	No abnormal phenotype	Embryonic lethal, can be rescued with human <i>FLII</i> transgenically
<i>Top3a</i>	Li and Wang (1998)	No abnormal phenotype	Embryonic lethal

<sup>a</sup>Table is adapted from Vlangos *et al.*, 2003.

**Lethal giant larvae homolog 1 (*LLGL1*):**

*LLGL1* (NCBI Entrez Gene ID#3996; Unigene Hs.513983; OMIM 600966) is the human homolog of the *Drosophila* tumor suppressor gene, D-lethal (2) giant larvae (D-lgl). When disrupted, D-lgl can produce tumors in the imaginal disks and abnormal transformation of the adult optic centers in larval brains. D-lgl was found to be associated with the cytoskeleton (Strand et al. 1995). In the initial description, human *LLGL1* was reportedly expressed in brain, kidney, and muscle, with little expression reported in heart and placenta (Strand et al. 1995). Current electronic expression profiling of human ESTs in Unigene shows highest expression levels in stomach. Antibodies against *LLGL1* were able to co-immunoprecipitate *LLGL1* and non-muscle myosin heavy chain. This protein interaction is conserved in *Drosophila* (Strand et al. 1995). *LLGL1* and D-lgl are also associated with a serine kinase, which is able to specifically recognize and phosphorylate serine residues within the protein likely allowing for regulation (Strand et al. 1995). *LLGL1* was first mapped to chromosome 17p11.2 with FISH and confirmed with Southern blotting (Strand et al. 1995). It is interesting that the 3' end of *LLGL1* overlaps with the 3' end of the *FLII* gene (Campbell et al. 1997) which helped in orientation of these genes during construction of our transcription map (Lucas et al. 2001).

**ATP synthase mitochondrial F1 complex assembly (*ATPAF2*; formerly *ATP12*):**

Originally, *ATPAF2* (NCBI EntrezGene ID# 91647; Unigene Hs.13434; OMIM 608918) was investigated by our group as the novel EST amplicon D17S2021 (Elsea et al. 1997). The *ATPAF2* gene is the human homolog of the yeast nuclear

gene *Atp12p* as determined via BLAST analysis (Lucas et al. 2001). The role of the *Atp12p* gene is reportedly limited to ATP synthase assembly, and it is required to mediate formation of the  $F1_{\alpha}$  subunit of the yeast mitochondrial ATPase (Wang et al. 2000; Wang et al. 2001). We examined expression patterns of *ATPAF2* via commercially available northern blots where a 1.8 kb transcript was noted in all tissues examined (Table 2). An equivalent function for the yeast and human homologs was proven by showing that transfection of the *ATPAF2* cDNA into an inviable yeast *atpaf2* mutant rescues the phenotype (Wang et al. 2001). Partial deficiency of complex V of the oxidative phosphorylation system has been associated to recessive missense mutation of the *ATPAF2* in one patient (De Meirleir et al. 2004). Heterozygous carriers of the reported mutation appear normal (De Meirleir et al. 2004). It is unclear what percentage of gene function is lost in the reported patients with DNA changes. Complete abolishment of gene function is likely incompatible with life (De Meirleir et al. 2004); thus, the homozygous patient must have some gene function. It is interesting that the affected patient displays features consistent with SMS including infantile hypotonia and poor sucking, a prominent nasal bridge, and hypoplastic kidneys (De Meirleir et al. 2004). At this point it is unclear if haploinsufficiency of *ATPAF2* via hemizygous deletion has any role in the SMS phenotype.

## **SUMMARY**

Using genomic markers and ESTs mapping to chromosome 17p11.2, we were able to create a fluid genomic contig and transcription map across the 1.5 Mb SMS critical region. In constructing of the contig and transcription map we were able to identify many positional SMS candidate genes. Initial characterization of the candidate sequences led us to focus on five potential candidate genes. Additional in depth analysis of these genes will focus on identifying the cellular function of these genes, as well as any potential role they may play in generation of the SMS phenotype.

## **MATERIALS AND METHODS**

### ***DNA isolation***

**BAC and PAC DNA extraction:** BAC and PAC clones were obtained from Research Genetics (now Invitrogen) and BAC/PAC Resources. Plasmid DNA was isolated from *E. coli* cultures using a modified Qiagen very low copy protocol created in our lab. Bacteria containing individual BACs or PACs were streaked on antibiotic selective LB agar plates. Single colonies were picked and grown with continuous shaking for 16 to 18 hours at 37°C in 2 ml of Luria-Bertani broth (LB) containing appropriate selective antibiotic (12.5 µg/ml chloramphenicol for BACs or 100 µg/ml kanamycin for PACs). Starter cultures were diluted 1/500 in 250 ml LB containing selective antibiotic and grown for 16 to 18 hours at 37°C while shaking. Bacteria were pelleted by spinning at 6000 x g for 15 minutes at 4°C. The supernatant was removed, followed by immediate resuspension of the pellet in 60 ml of buffer P1 (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, pH 8.0, 0.1 mg/ml RNase A). Sixty milliliters of lysing buffer P2 (0.2 N NaOH, 1% sodium dodecyl sulfate{SDS}) were added to the cells and solutions were gently mixed via inversion to ensure complete cell lysis. Sixty milliliters of buffer P3 (3 M potassium acetate) were immediately added to the lysed solution and gently mixed. Tubes were placed on ice for thirty minutes to ensure complete precipitation of cellular debris. Following this incubation, the tubes were centrifuged at 4°C at  $\geq 20,000 \times g$ . Supernatant containing plasmid DNA was promptly filtered through filter paper (Whatman #4). The BAC/PAC DNA was precipitated by adding 0.7 volumes of room temperature isopropanol and centrifuged at  $\geq 15,000 \times g$  for 30 minutes at 4°C. The pelleted DNA was allowed to dry slightly

(inverted for 5 minutes) and then resuspended in 500 µl sterile 1x TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA). After the DNA was completely resuspended, 5 ml of buffer QBT (0.75 M NaCl and 50 mM MOPS buffer containing 15% isopropanol and 0.15% Triton X-100, pH 7.0) was added to the plasmid DNA. Resuspended DNA was applied to a buffer QBT equilibrated Qiagen Tip 500. The column was washed three times with 10 ml of buffer QC (1 M NaCl, 50 mM MOPS buffer, pH 7.0 containing 15% isopropanol). The BAC/PAC DNA was precipitated from the column by using 15 ml of buffer QF (1.25 M NaCl, 50 mM Tris, pH 8.5, containing 15% isopropanol) preheated to 65°C. The purified DNA was precipitated by adding 0.7 ml of room temperature isopropanol and centrifuged at  $\geq 15,000 \times g$  at 4°C for 30 minutes. The supernatant was decanted and the plasmid pellet was allowed to air dry for ~10 minutes. The pellet was resuspended in 200 µl sterile 1x TE. Purified BAC/PAC DNA was electrophoresed on a 1% TBE agarose gel and visualized under a UV transilluminator after staining with ethidium bromide. Finally, DNA quantity and quality was determined using a spectrophotometer.

EST plasmid DNA isolation: EST clones representing markers mapping to chromosome 17p11.2 were obtained (Research Genetics, Incyte Genomics, the German Genome Project, or RIKEN) and grown on LB agar plates containing appropriate antibiotic for selection. Individual colonies were then picked and grown in 5 ml LB broth containing antibiotic. DNA was isolated from bacterial cultures using the Qiagen miniprep kit according to manufacturer's instructions. After isolation plasmid DNA was electrophoresed on a 1% TBE agarose gel and visualized under a UV transilluminator after staining with ethidium bromide. DNA quantity and



quality was determined using a spectrophotometer.

***Southern analysis:*** Standard Southern analysis was used to map markers to BACs/PACs within the contig and to the hybrid mapping panel. BAC or PAC DNA (6 µg), human genomic DNA (9 µg), or rodent-human hybrid DNA (15 µg) was digested with 4 U/µg of *EcoRI* and 2.5 mM spermidine for ~16-18 hours at 37°C. Digested DNAs were electrophoresed on 1% TAE agarose gels (TAE; 0.04 M Tris-acetate, 1 mM EDTA) with continuous buffer recirculation at room temperature. Gels were depurinated in two gel volumes of 0.25 N HCl followed by denaturation in 2 gel volumes of 0.4 N NaOH. DNA was transferred to an Amersham Hybond-N+ nylon membrane via wicking with 10x SSC (SSC; 20x stock solution is 3 M NaCl, 0.3 M Na-citrate). After transfer, DNA was UV-crosslinked to the membrane using a Stratalinker (Stratagene). Membranes were prehybridized and hybridized in a solution of 1 M NaCl, 1% SDS, 10% dextran sulfate, and 0.1 mg/ml herring sperm DNA at 65°C with rotation in a hybridization oven. All DNA probes (purified PCR products or plasmid inserts) were labeled with <sup>32</sup>P-dCTP using an Amersham Rediprime II kit. Unincorporated nucleotides were removed with a spermine and herring sperm DNA precipitation followed by radioactivity quantification in a Brinkman scintillation counter. During preassociation, ~10<sup>6</sup> cpm/ml of probe was annealed to 0.25 mg/ml placental DNA to mask repeat sequences and then hybridized to the membrane for ~16-18 hours. Blots were washed for 20-30 minutes in 0.1x SSC, 0.1% SDS at room temperature, followed by a stringent wash in preheated 0.1x SSC, 0.1% SDS at 65°C. The blots were exposed to X-ray film with two intensifying

screens at -80°C for 1-5 days before developing.

**Database searches:** Extensive information about the genes and ESTs in the region was obtained through the National Center for Biotechnology Information (NCBI) website (<http://www.ncbi.nlm.nih.gov>), including Unigene, the EST database (dbEST), and the STS database (dbSTS). Draft assembly of the human genome project was found at the UCSC genome bioinformatics site (<http://www.genome.ucsc.edu>) and through Project ENSEMBL (<http://www.ensembl.org>). Supplementary marker information was derived from the Genome Database (<http://www.gdb.org>). Sequence alignments and unfinished high-throughput genome sequence BLAST searches were conducted through the Baylor College of Medicine Search Launcher (<http://searchlauncher.bcm.tmc.edu>) or NCBI. Conserved sequence domains were identified through databases at NCBI (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>).

**PCR:** The polymerase chain reaction (PCR) was performed in 25 µl volumes. Amplification of 50-100 ng of template DNA was performed using 1 U of *Taq* polymerase in a cocktail consisting of 0.8 µM of each primer, 0.25 mM dNTPs, and 1x PCR buffer (10x buffer contains 100mM Tris-HCl pH 8.3, 15 mM MgCl<sub>2</sub>, 500 mM KCl, 0.01% gelatin). Amplification was performed in an Applied Biosystems (ABI) or MJ Research thermocycler at the following conditions, unless otherwise noted: initial denature at 94°C for 4 minutes, followed by 30 cycles of 94°C for 1 minute, 55°C for 1 minute, and 72°C for 3 minutes, followed by a final extension of

72°C for 10 minutes. PCR products were then electrophoresed on 1% or 2% TBE agarose gels, stained with ethidium bromide, and visualized on a UV transilluminator.

***Northern analysis:*** The expression of various ESTs was visualized using human multiple tissue, fetal, and brain II northern blots purchased from Clontech. Prehybridizations and hybridizations were performed in 5x SSPE (0.75 M NaCl, 0.05 M NaH<sub>2</sub>PO<sub>4</sub>, 5 mM EDTA), 10x Denhardt's solution (0.2% BSA, 0.2% ficoll, 0.2% polyvinylpyrrolidone), 100 µg/ml herring sperm DNA, 2.0% SDS, and 50% deionized formamide at 42°C with rotation in a hybridization oven. Radiolabeled probes were created as described for Southern analysis. Approximately 10<sup>6</sup> cpm/ml of probes was denatured and added to prehybridized filters. Filters were hybridized for ~18 to 20 hours at 42°C, followed by washing in 2x SSC, 0.05% SDS for ~30 minutes at room temperature with 1 to 2 changes of wash solution. Filters were then transferred to a wash of 0.1x SSC, 0.1% SDS for ~30-40 minutes at 50°C. The blots were exposed to X-ray film between two intensifying screens at -80°C for 2-5 days before developing.

## **Chapter III**

### **Analysis of 17p11.2 deletions using fluorescent *in situ* hybridization**

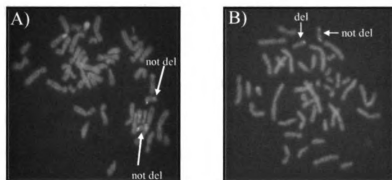
#### **BACKGROUND**

Until the early 1990s detection of chromosomal anomalies, including microdeletion disorders, depended on staining and analysis of the entire chromosome complement. This was usually accomplished by techniques such as Geimsa staining, also known as G-banding. The chromosomes were stained with a mix of the toxic dyes methylene azure and eosin, which bind to the phosphate groups of DNA. The regions of the chromosome where the DNA is more densely packaged (heterochromatin) stains darker than regions where the DNA is more loosely packaged (euchromatin). Thus, staining reveals a recognizable pattern of bands along the chromosomes. Using this standard technique, the entire genome usually shows ~350 distinct bands. Analysis of the number of total chromosomes and the banding patterns uncovers the chromosomal abnormalities.

This standard karyotype analysis works very well for detection of aneusomy (i.e. Down syndrome, trisomy 21) and large segmental anomalies (i.e. translocations, inversions, deletions, or duplications) and is still routinely used today. Though the technique is very reliable, there are limitations. Increasing the number of bands to the 550 or 850 level by preparing chromosomes of increased metaphase length helps intensify the sensitivity of the test. Even so, small anomalies involving less than ~1.5 Mb of DNA cannot be unambiguously detected. Technological advances using fluorescent dyes have aided greatly in the detection of these smaller anomalies.

The use of fluorescent *in situ* hybridization (FISH) for gene mapping dates back to the late 1970s when it was used to map the human globin genes to the long arms of chromosomes 4 and 5 (Cheung et al. 1977). It wasn't until 1986 that FISH was reported as a use for detecting chromosomal anomalies using human genomic DNA as a probe (Pinkel et al. 1986). DNA probes made of genomic DNA representing known loci were labeled with nucleotides conjugated with haptens such as biotin or digoxigenin. The labeled nucleotides are incorporated, at a defined density, into the nucleic acid probes by DNA polymerases. The labeled probes were then hybridized to denatured metaphase chromosomes much in the way radiolabeled probes are hybridized to Southern blots. The labeled probes are detected with fluorescently labeled antibodies to the hapten, and viewed using a properly equipped fluorescent microscope. The number of fluorescent signals detected is a direct measure of DNA copy number (Figure 7). If a microdeletion syndrome is detected via karyotype, the deletion is usually confirmed via FISH. Additionally, FISH may be ordered directly without a prior karyotype if a specific chromosomal anomaly is suspected. The current guidelines for diagnosis of microdeletion syndromes, including SMS, recommend testing and/or confirmation via FISH (ACMG/ASHG positional statement. 2000).

Technology has advanced to make FISH experiments faster and cheaper. Today, the flouorochrome is attached directly to the nucleotide. This negates the need for secondary antibody detection. The probes are more robust and emit a greater amount of light from the same excitation input. Initial FISH experiments in the Elsea Laboratory were performed with indirect labeling methods, but we have since



**Figure 7. Examples of not deleted and deleted FISH experiments.**

Metaphase chromosomes were hybridized with fluorescently labeled BAC probes. Green signals are indicative of clones mapping within the SMS critical region. The red signals represent a BAC mapping to 17q and is used as a control. **A)** Green and red signals of both chromosomes 17 indicate that this patient is not deleted for the test probe. **B)** In this patient sample the green test probe is only present on one chromosome 17, while the red control probe is present on both chromosomes 17. This indicates that this patient is deleted for this test probe.

switched to using directly labeled probes. All FISH experiments described herein have been performed with directly labeled FISH probes.

FISH is not only a wonderful diagnostic tool but also a powerful tool for research of microdeletion syndromes. Using probes mapping in the region of the disorder, it is possible to determine which DNA sequences are absent or present. SMS patients displaying the full phenotype of the disorder may have different sized deletions (Trask et al. 1996). The smallest region of overlap of the deletions carried by these patients defines the syndrome critical region. The genes mapping within this region are the positional candidate genes for the disorder. This method has been used successfully in many microdeletion disorders including Smith-Magenis syndrome (Bettio et al. 1995; Gong et al. 1996; Juyal et al. 1996; Elsea et al. 1997).

## **RATIONALE**

In studying deletions of chromosome 17p11.2 associated with Smith-Magenis syndrome, we were most interested in patients with deletions overlapping the SMS critical region (described in Chapters 1 and 2). Using FISH analysis on SMS patient samples we identified deletions that overlapped the smallest SMS deletion known at the time. Detailed analysis of the unusual deletions carried by these patients allowed us to reduce the size of the SMS critical region and the number of candidate genes. The work presented in this chapter resulted in the following publications; (Vlangos et al. 2003; Vlangos et al. 2004).

## **RESULTS**

### **FISH detection of 17p11.2 deletions and refinement of the SMS critical region**

Using FISH, we characterized the 17p11.2 deletion size in 22 patient samples. Clinical analysis of the patients was obtained from patient medical records, and from a questionnaire developed in house (Appendix A). Because the common SMS deletion was reported to occur in >90% of SMS patients we expected that the majority of our patient cohort would harbor the common SMS deletion (Chen et al. 1997; Lupski 1998; Potocki et al. 2000b; Bi et al. 2002; Shaw et al. 2002). Thorough FISH analysis using probes mapping along chromosome 17p11.2 revealed that 8 of our 22 patient samples carried the common SMS deletion (Figure 8). The number of patients with the SMS common deletion was significantly lower than we expected.

Four of the 22 patient samples analyzed showed deletions that were uncommon, but that did not overlap the 1.5 Mb SMS critical region (Figure 8). These patients all showed the complete spectrum of the SMS phenotype. The unusual deletions these patients carry do give further insight into the mechanisms causing 17p11.2 deletions associated with SMS, as the deletions are not due to NAHR between the SMS-REP sequences. These data indicate that the true incidence of the SMS common deletion may be lower than < 90% previously reported (Chen et al. 1997; Lupski 1998; Potocki et al. 2000b; Bi et al. 2002; Shaw et al. 2002).

Since 1997 the SMS critical region has been defined by a single 17p11.2 deletion spanning ~1.5 Mb of DNA (Elsea et al. 1997). In our analysis of 17p11.2 deletions associated with SMS, we were very interested in finding patients displaying the SMS phenotype whose deletions overlapped the 1.5 Mb carried by HOU142-540.



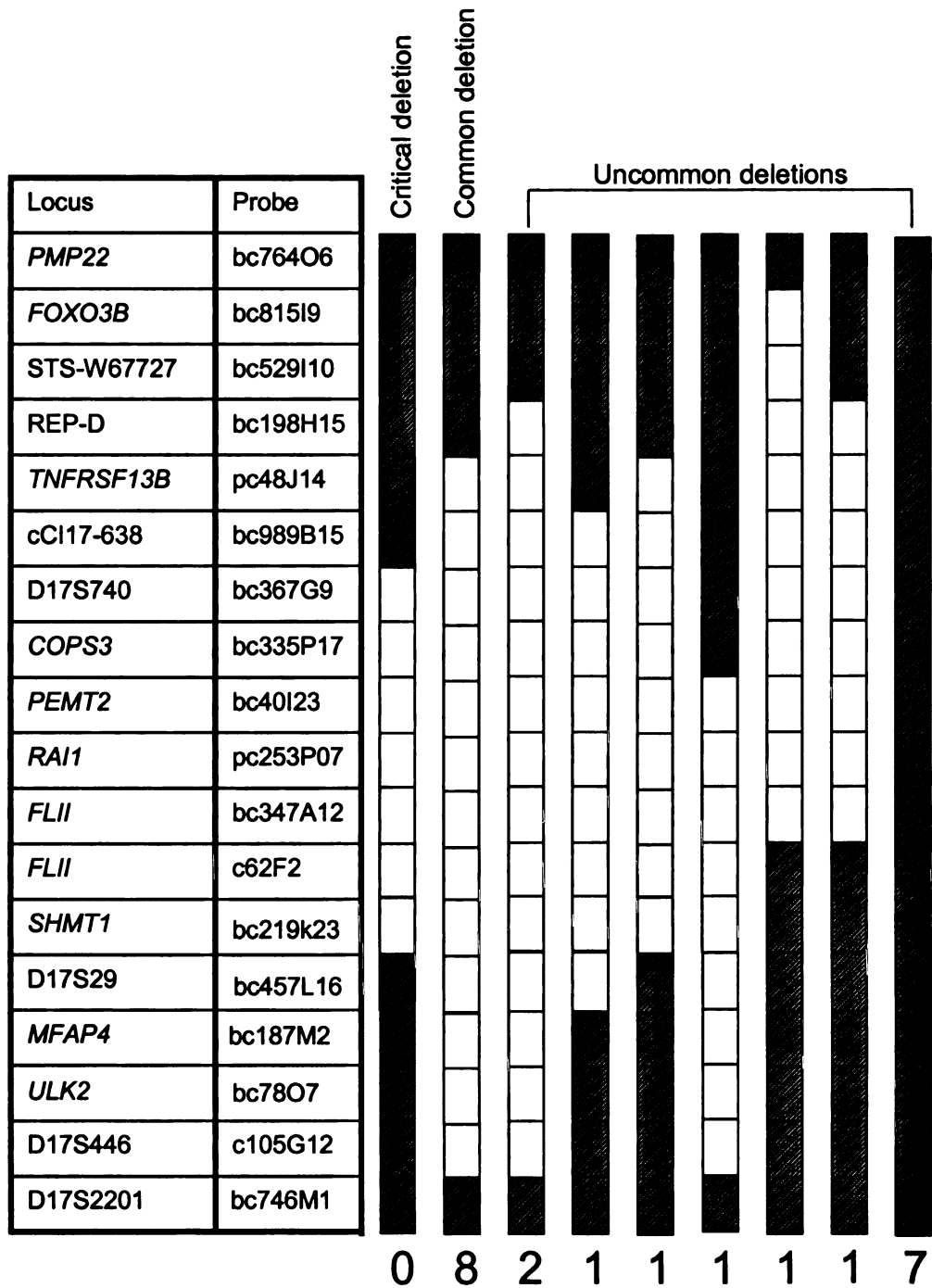
Through thorough FISH analysis on the patient samples received, we found three patients who carry deletions overlapping with the deletion carried by patient HOU142-540 (Figure 8).

SMS135 is the first patient referred to us whose deletion helped narrow the smallest region of overlapping deletion on chromosome 17p11.2 associated with SMS. SMS135 is a 3 year, 8 month-old male of Filipino descent who displays the characteristic SMS phenotype (Figure 9A, Table 5). FISH with clones mapping along chromosome 17p11.2 allowed for identification of his deletion breakpoints. FISH data indicate that the proximal breakpoint of this deletion occurs at or near the proximal SMS-REP. The distal breakpoint, however, occurs within the current SMS critical interval (Figure 8). FISH analysis indicated that this patient is not deleted for genomic clones on 17p11.2 from the SMS-REPD through BAC RP11-335P17 containing the *COPS3* gene (Figure 8). The patient is deleted for BAC CTD-40I23, containing the *PEMT2* gene (Figure 8). Finer mapping with cosmids was performed using clones from the chromosome 17 library (LA17NC01) (Kallioniemi et al. 1994) for *RASD1* (c107A2) and *PEMT2* (c26E10) (Figure 10). The patient is deleted for c26E10 and not deleted for c107A2 which places the distal breakpoint in the ~30 kb between these cosmids.

Two additional patient samples, SMS182 and M2359 (Figure 9A and 9B, Table 5), were sent to us for analysis because high resolution karyotyping showed possible interstitial deletion of chromosome 17p11.2 though diagnostic testing for SMS with commercially available FISH probes was negative in both cases. We strongly suspected SMS in both patients, and performed FISH analysis using probes

**Figure 8. Summary of 17p11.2 deletions identified in this study.**

Probes used in FISH analysis and their respective locus are indicated in the table to the left. Deletions found in the course of this study are indicated by the ideograms. Hatched boxes indicated DNA that is not deleted. Open boxes indicated DNA that is deleted. The critical deletion as reported in Juyal *et al.*, 1996 and Elsea *et al.*, 1997 is indicated for reference. The number of patients determined to carry each deletion is listed under each ideogram.



mapping along 17p11.2.

The proximal breakpoints for both SMS182 and M2359 map within the region between the *FLII* and *LLGL1* genes (Figures 8 and 10). Distally, the deletion carried by SMS182 extends into 17p12 near the *ADORA2B* gene. The distal breakpoint for M2359 is located near the genomic marker STS-W67727 in 17p11.2 (Figures 8 and 10). Using the proximal deletion breakpoint for these two patients and the distal breakpoint for patient SMS135, we determined that the smallest shared region of deletion of studied SMS patients is ~700 kb (Figure 10) and currently contains 11 known genes (Vlangos et al. 2004).

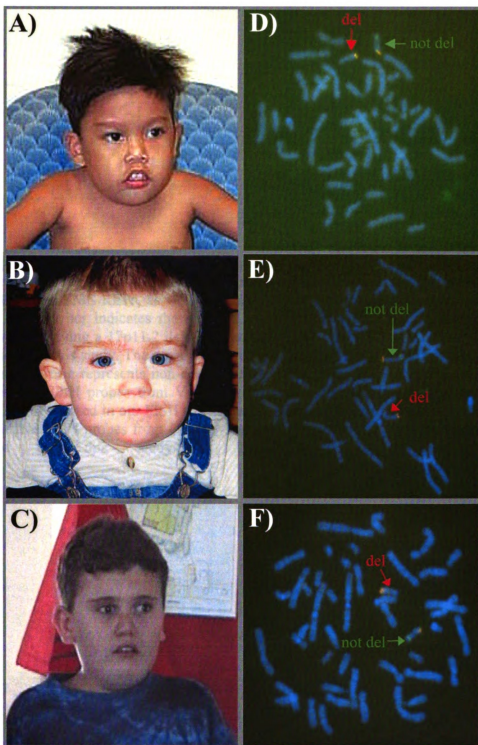
#### **Commercially available diagnostic FISH probes for SMS will not detect all cases**

Both SMS182 and M2359 display characteristics typical of the SMS phenotype for their respective ages. Both patients tested negative for SMS using commercially available SMS FISH probes. Utilizing genomic probes mapping along 17p11.2 for FISH, we were able to determine that both SMS182 and M2359 carry interstitial deletions of 17p11.2 and that both patients are deleted for the *RAI1* gene.

As will be discussed further in Chapter 4, patients carrying mutations in the *RAI1* gene mapping within the SMS critical region display features consistent with the SMS phenotype (Slager et al. 2003; Bi et al. 2004; Girirajan et al. 2005). Characteristics that are not seen patients with mutation in the *RAI1* gene include short stature, infantile hypotonia, and cardiac and/or renal anomalies. These features occur at a greater frequency in SMS patients who harbor a 17p11.2 deletion than in the general population and thus, are included in the SMS phenotype. It is possible that

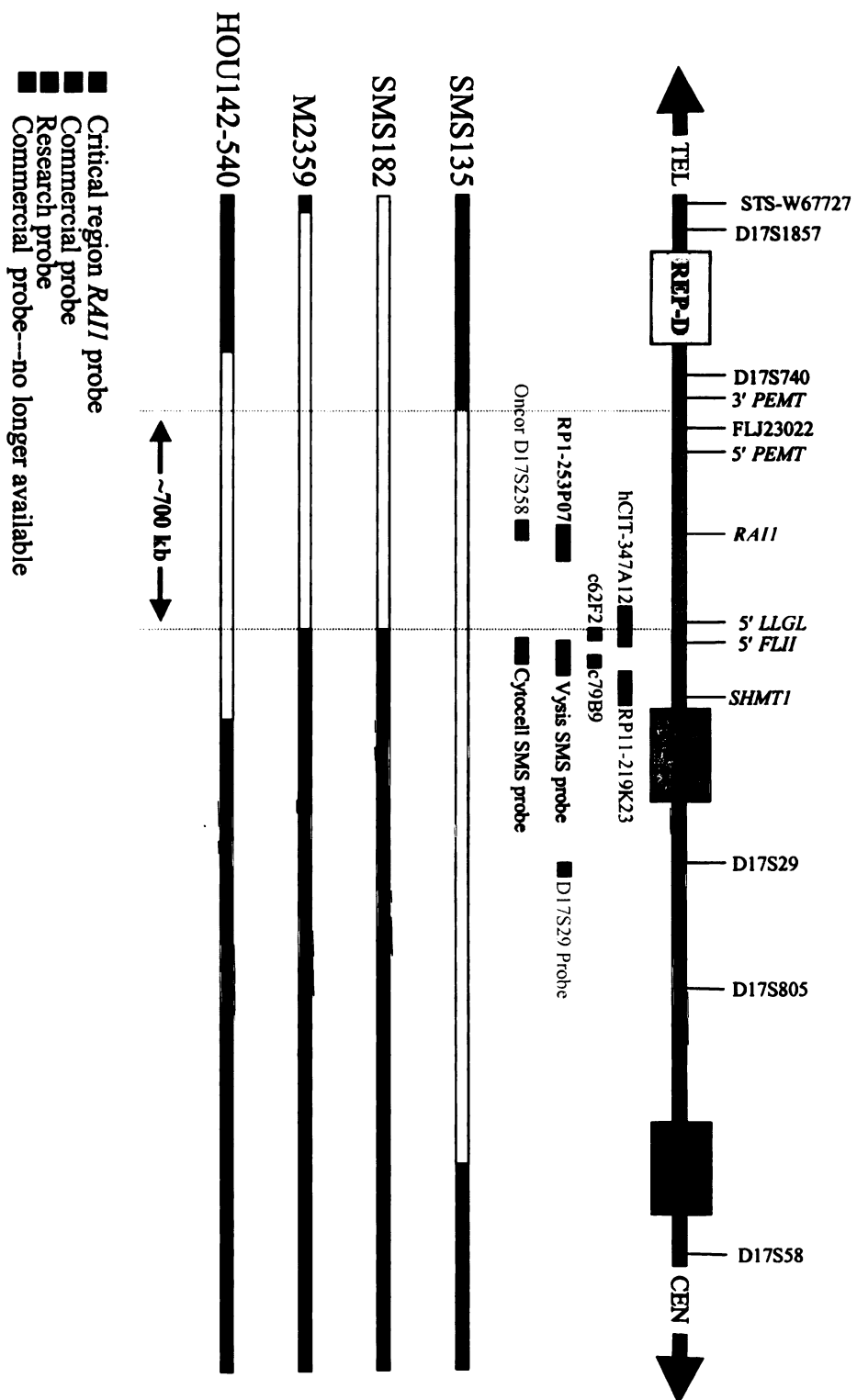
**Figure 9. SMS patients with unusual deletions used to refine the smallest region of overlapping deletion in SMS.**

The three patients who carry deletions allowing for refinement of the SRO in SMS all display typical SMS facial features; **A)** SMS135 at age 3 years, 8 months, **B)** SMS182 at age 19 months, and **C)** M2359 at age 10 years. **D)** FISH analysis of metaphase chromosomes from SMS135 indicate he is not deleted for bc3073J20 as green test signal is only detected on one chromosome 17. This probe is deleted original SMS critical interval described by Elsea *et al.*, 1997. **E)** Metaphase chromosomes from SMS135 indicate he is deleted for the pc253P07 probe containing the *RAI1* gene. **F)** Metaphase chromosomes from M2359 indicate he is also deleted for the pc253P07 probe. Both SMS182 and M2359 both tested negative for SMS using currently available commercially available FISH probes.



**Figure 10. Location of SMS FISH probes in relation to 17p11.2 deletions.**

The genomic region of chromosome 17p11.2 is represented and indicates the location of the three SMS-REPs, as well as genes and markers mapping to this region. The red hatched bar indicates the smallest region of DNA that is deleted in all SMS patients carrying a 17p11.2 deletion, the open/closed bars indicate deletions of SMS patients in this study. The open portion of the bars represents deleted DNA, while the closed portion represents non-deleted DNA. The green bars represent the currently available FISH probes from Vysis and Cytocell. The red bar represents the PAC RP1-253P07 used in our FISH experiments. The purple bars represent the previously available FISH probes from Oncor. The blue bars indicate experimental test probes used in this study. This figure is not to scale.





**Table 5. Phenotypic comparison of SMS patients with 17p deletions<sup>a</sup>**

	SMS126 Common SMS deletion	HOU142 -540 Smallest SMS deletion to date	SMS135 Unusual 17p11.2 deletion	SMS182 Unusual 17p11.2 deletion	M2359 Unusual 17p11.2 deletion
<b>Features common in &gt;75% of SMS Pts.</b>					
<b>Craniofacial/Skeletal</b>					
Brachycephaly	-	-	+	+	+
Midface hypoplasia	+	+	+	+	+
Prognathism (relative to age)	+	+	N	N	+
Tented upper lip	+	+	+	+	+
Broad, square face	+	-	+	+	+
Brachydactyly	+	-	+	+	+
Short stature	+	+	+	-	-
<b>Otolaryngologic abnormalities</b>					
Chronic ear infections/PE tubes	+	+	N	+	+
Hoarse, deep voice	+	+	N	N	+
<b>Neurological</b>					
Mental retardation	+	+	+	+	+
Speech delay	+	+	+	+	+
Motor delay	+	-	+	+	+
Infantile hypotonia	+	-	+	+	+
Sleep disturbance	+	+	+	+	+
<b>Behavioral</b>					
Self-hugging/hand wringing	+	+	+	N	+
Onychotillomania	+	+	N	N	-
Polyembolokoilomania	+	+	N	N	+
Head banging / face slapping	+	+	N	+	+
Hand-biting	+	+	N	N	+
Attention seeking	+	+	N	N	+
<b>Features common in 50 – 75% of SMS patients</b>					
Hearing loss	+	+	-	+	-
Myopia	+	+	-	-	-
Strabismus	-	+	+	+	-
Iris abnormalities	-	+	-	-	-
Synophrys	-	+	-	+	+
Scoliosis	+	-	+	N	+
<b>Features common in &lt; 50% of SMS patients</b>					
Cardiovascular anomalies	+	+	-	-	-
Renal anomalies	-	+	-	-	-
Seizures	-	+	-	+	-
Cleft lip/palate	-	-	-	-	-
Gender/Age at evaluation	Female 41y	Male 10y, 14y, 15y	Male 3y 8m	Male 14m, 2y	Male 10y

<sup>a</sup> Modified from the Gene-Reviews website ([www.genereviews.com](http://www.genereviews.com)) and Slager *et al.*, (2003)

Note: +, feature present; -, feature not present; and N, feature not evaluated/ too young to evaluate

haploinsufficiency of *RAI1* may also cause these characteristics; however, we have not yet identified patients with *RAI1* mutations who display these features. It is also possible that haploinsufficiency of another gene(s) mapping to 17p11.2 is responsible for these characteristics. Thus, studies of patients with unusual deletions will continue in order to narrow SMS region and the number of possible candidate genes to further understand genotype:phenotype correlation in SMS.

Narrowing of critical region in microdeletion syndromes is common and can greatly aid in the search for causative genes as has been seen in Rubinstein-Taybi syndrome [del(16)(p13.3)], due to haploinsufficiency of *CREBBP* (Petrij et al. 1995), and Alagille syndrome [del(20)(p12)] due to haploinsufficiency of *JAG1* (Li et al. 1997b; Oda et al. 1997b). As the critical regions are further reduced and the causative genes for the microdeletion syndromes are located, it is important that FISH probes are moved to represent these causative genes. Clinical FISH for Alagille syndrome associated deletions is now conducted with a probe containing the entire *JAG1* gene, while clinical FISH for Rubinstein-Taybi associated deletions is conducted with a contig of five cosmids covering the *CREBBP* gene (<http://www.genereviews.org>).

It is important to note that so far, all SMS patients who carry a 17p11.2 deletion are deleted for the *RAI1* gene (Figures 8 and 10). Currently available commercial FISH probes do not contain *RAI1*. We have shown here that these currently available commercial FISH probes from Vysis Inc. (A subsidiary of Abbot Laboratories) and Cytocell Technologies will not detect all cases of SMS with a 17p11.2 deletion. We suggest that all future diagnostic FISH tests be performed with

probes containing the *RAII* gene. This will require that companies offering commercial FISH probes re-evaluate the location of their probes based on current data. Certain patients suspected to have SMS who previously tested negative with currently available FISH probes (i.e. those containing *FLII* or D17S29) may need to be re-tested with a probe containing *RAII*, but we strongly recommend careful consideration of the SMS phenotype by thorough clinical re-evaluation before costly re-testing is pursued. It is also important to note that patients tested with the Oncor D17S258 FISH probe, which is no longer commercially available, do not need to be re-tested by FISH. The D17S258 probe is a cosmid, and the genetic marker D17S258 is located within an intron of the *RAII* gene (Figure 10). Some patients may display the typical SMS phenotype yet still test negative with FISH probes containing the *RAII* gene. These patients may be candidates for further testing and sequencing of the *RAII* gene as will be discussed in Chapter 4.

#### **The SMS common deletion occurs less frequently than reported**

The deletion seen with greatest frequency, known as the common deletion, is ~3.5 Mb and was been reported to occur in 90% - 95% of SMS patients (Chen et al. 1997; Lupski 1998; Potocki et al. 2000b; Bi et al. 2002; Shaw et al. 2002). The remaining ~5% are reported to be deletions of differing sizes, though size distribution, larger or smaller, has not been reported. During our FISH analysis we noticed that unusual deletions were occurring more frequently than expected. We reviewed the deletions of patients reported in the literature in addition to those reported here (Table 6). To date, including the 16 new SMS patients with 17p11.2

deletions reported here, there are a total of 85 reported SMS patient deletions fully characterized (Juyal et al. 1996; Chen et al. 1997; Yang et al. 1997; Potocki et al. 2000b; Bi et al. 2002; Park et al. 2002). The common deletion was reported as either the presence of a junction fragment on a PFGE Southern blot resulting from homologous recombination between SMS-REPP and SMS-REPD (Chen et al. 1997) or by FISH analysis. The common deletion has been detected in 60/85 (70.59%) SMS deletions, while unusual deletions have been found in 25/85 (29.41%) SMS deletions (Table 6). Approximately 30% of SMS deletions are likely mediated by other sequences, including the middle SMS-REP and/or other LCR elements, or the location of the SMS region within the pericentromeric region on the chromosome. An analysis of the August 2001 release of the human genome sequence discovered 169 sequences flanked by similar duplications of which only 24 are known to be associated with human disease (Bailey et al. 2001).

It is thought that 5–10% of the human genome is duplicated (Ji et al. 2000) leaving many repetitive sequences uncharacterized. Newly described LCR elements have been characterized which flank the SMS-REP sequences, and these newly characterized LCRs are postulated to mediate recombination in unusual deletions (Park et al. 2002). The data presented here further suggest that LCRs yet to be characterized mediate recombination at a much higher rate than previously thought. Additionally, many microdeletion syndromes occur in the pericentromeric region of the chromosome (Williams syndrome del(7)(q11.2), DiGeorge syndrome del(22)(q11.2), Smith-Magenis syndrome del(17)(p11.2), NF1 del(17)(q11.2), and Prader-Willi/Angelman syndromes del(15)(q11-q13)). The location of these

syndromes in the pericentromeric region of the chromosome may also facilitate the homologous recombination leading to unusually sized deletions.

## **SUMMARY**

Analysis of SMS patient samples has allowed us to narrow the smallest region of overlapping deletion of chromosome 17p11.2 associated with Smith-Magenis syndrome from 1.5 Mb to ~700 kb. This region contains the *RAI1* gene, which likely causes the majority of the SMS phenotype when in the hemizygous/haploinsufficient state. Currently available commercial FISH probes for SMS do not contain the *RAI1* gene, and not all SMS patients with 17p11.2 deletions are detected for the currently available commercial FISH probes. Thus, not all SMS patients with deletions will be diagnosed with these probes. This will require that companies offering commercial FISH probes re-evaluate the location of their probes based on current data.

Further, we have determined that the ~3.5Mb common SMS deletion occurs less frequently than the >90% reported. The common deletion occurs in ~70% of SMS patients. Other newly described, or undiscovered repetitive sequences likely contribute to the recombination of chromosome 17p11.2 leading to SMS.

**Table 6. Frequency and type of SMS patient deletions<sup>a</sup>.**

Publication	Total # complete <sup>b</sup>	Common Deletion	Deletion other than common
Juyal <i>et al.</i> (1996)	59	50	9
Yang <i>et al.</i> (1997)	1	0	1
Chen <i>et al.</i> , (1997)	2	2	0
Park <i>et al.</i> , (1998)	1	0	1
Potocki <i>et al.</i> , (2000)	4	0	4
Bi <i>et al.</i> , (2002)	2	0	2
Vlangos <i>et al.</i> , (2003)	11	7	4
This Report	5	1	4
Total	85	60/85 (70.6%)	25/85 (29.4%)

<sup>a</sup> Modified from Vlangos *et al.*, 2003

<sup>b</sup> Total number complete represents fully characterized deletions. If patients have multiple reports, they are counted only once.

## **MATERIALS AND METHODS**

*Patient and sample ascertainment:* Contact between potential research participants was established between the patient's guardian and/or medical care provider and the project Principal Investigator (Dr. Sarah H. Elsea, Ph.D.). Research samples and completed questionnaire (see appendix A) were collected in accordance with Institutional Review Board approved protocols at Michigan State University, Virginia Commonwealth University, and The National Institutes of Health. Peripheral blood samples were collected and shipped to the Elsea laboratory for preparation of transformed lymphoblastoid cell lines, metaphase chromosomes, and genomic DNA.

### *Clinical descriptions of SMS135, SMS182, and M2359:*

**Patient SMS135:** SMS135 is a male of Filipino descent evaluated at age 3 years, 8 months by collaborator Dwight Yim, M.D. SMS135 displays the common SMS facial features including a distinct tented upper lip, brachycephaly, and midface hypoplasia (Figure 9A, Table 5). SMS135 also has speech and motor delay, short stature, brachydactyly, and was hypotonic as an infant. He is mentally retarded, has stereotypic self-hugging behavior, and experiences sleep disturbance common in SMS patients. SMS135 also has myopia and scoliosis, which are seen in < 50% of SMS patients. SMS135 has not yet developed self-injurious behaviors, though these may develop with further aging. Commercial FISH testing was positive for SMS. The patient was referred because of the family's interest in SMS research.

**Patient SMS182:** SMS182 is healthy, non-verbal male evaluated at 14 months

(Figure 9B) and again at 2 years (Table 5). At age two years his height was 83.82 cm (15%ile), weight 12.93 kg (60%ile), head circumference 47.94 cm (25%ile), and BMI 18.41 (88%ile). He displays the common physical and neurological SMS characteristics expected for his age (Fig. 9B and Table 5). The parents of SMS182 also reported that he developed significant sleep disturbance at ~18 months of age, and head-banging and face slapping upon becoming frustrated at age 3. He displays some features seen less commonly in SMS, including hearing loss, strabismus, synophrys, and seizures (Table 5). High-resolution karyotypes at two different facilities showed an interstitial deletion of one chromosome 17p11.2. Commercial FISH testing, performed using the Vysis SMS probe, was negative (Fig. 13).

Patient M2359: Evaluated by colleagues at NIH and in Australia, M2359 (Figure 9C, Table 5) is a male evaluated at 10 years of age with height 144.15 cm (90%ile), weight 43.4 kg (97%ile), head circumference 55.5 cm (90%ile), and BMI 20.89 (90-95%ile). He has nearly all features most common to SMS, including typical craniofacial anomalies, speech and motor delay, mental retardation, and infantile hypotonia. In addition, he has synophrys, scoliosis, and bifid uvula. He displays a majority of the characteristic behaviors seen in SMS patients including self-hugging, polyembolokoilomania, head banging, hand biting, and attention seeking, and he has evidence of nail damage, although he does not yet display nail yanking as is often seen in older SMS patients (Finucane et al. 2001). Significant sleep disturbance was reported by parents and validated by actigraphy. He does not have short stature or cardiac/renal anomalies. A high-resolution karyotype of M2359 showed a suspected



deletion of one chromosome 17p11.2 region. FISH was performed with SMS probes from Vysis and Cytocell, but both tests were negative.

*Isolation of metaphase chromosomes:* Two hundred and fifty microliters of whole blood collected in a sodium heparin Vacutainer was inoculated in 5 ml of RPMI media (Gibco) supplemented with 10% fetal bovine serum and 1x antibiotic/antimycotic solution (Gibco). Fifty microliters of the mitogen phytohemagglutinin (PHA) were added to a final concentration of 1 x, tubes were mixed well and incubated at 37°C under 5% CO<sub>2</sub> for 72 hours with gentle mixing at least once per 24 hours. After 72 hours incubation, synchronization of the cell cycle was started by the addition of 10 µl of 50 µM methotrexate. After 16–18 hours incubation at 37°C under 5% CO<sub>2</sub>, 50 µl of 1 mM thymidine was added. Cells were incubated another 4 hours at 37°C under 5% CO<sub>2</sub>. Harvest of synchronized cells was started by the addition of 25 µl of 10 µg/ml colchicine. Cells were incubated 30 minutes at 37°C under 5% CO<sub>2</sub> before centrifugation at room temperature for 8 minutes at 1,200 rpm. Supernatant was discarded, followed immediately by the addition of 6 ml of 75 mM KCl. Cells were incubated at room temperature for 15 minutes before the addition of 10–12 drops of ice cold fixative (3:1 methanol:acetic acid) with a Pasteur pipette. Tubes were immediately mixed and centrifuged as above. All but 0.5 ml of the supernatant was removed, and cells were resuspended in 1 ml fixative and gently mixed. Five milliliters of additional fixative was immediately added and tubes were mixed and centrifuged as above. After centrifugation, the supernatant was completely removed and cells were resuspended

in 5 ml of fixative. Centrifugation and washing with fixative was repeated 2 additional times. After final centrifugation, the pellet was resuspended in ~4 ml fixative and incubated overnight at 4°C. After incubation metaphase chromosome spreads were prepared on glass slides using standard cytogenetic methods.

*Fluorescent in situ hybridization:* Probes for FISH were chosen based on mapping information from our recently completed SMS critical interval contig, which was constructed with large insert bacterial artificial chromosomes (BACs) and P1 artificial chromosomes (PACs). Additional probes (BACs, PACs, and cosmids) were obtained using mapping data obtained during construction of our contig, and from online databases, including the Genome Browser at UCSC (<http://www.genome.ucsc.edu>) and GenBank at NCBI (<http://www.ncbi.nlm.nih.gov>). FISH probes were created using BAC, PAC, or cosmid DNA isolated as described in Chapter 2 and labeled by using a commercially available nick translation kit to incorporate Spectrum Green or Spectrum Orange dUTP by following manufacturer instructions (Vysis Inc.). Metaphase chromosomes were prepared for hybridization by incubating at 37°C in 2x SSC for 30 minutes. Slides were dehydrated by placing in an ethanol series (70%, 80%, 95%, and 100%, two minutes each) and allowed to air dry. Chromosomes were then denatured at 72°C for two minutes in a solution of 70% deionized formamide and 2 x SSC at pH 7.0 and immediately dehydrated in ethanol as above. Probe DNA (100 ng BAC/PAC or 180 ng cosmid) was precipitated in the presence of 1 µg COT-1 DNA and 2 µg human placental DNA by adding 0.1 volume of 3 M sodium acetate and 2.5 volumes 100% ethanol. The probes were placed at -80°C for 15 minutes then

immediately centrifuged at 4°C at > 12,000 x g for 30 minutes. The supernatant was discarded and the precipitated probe was dried to completion in a SpeedVac. The probe was then resuspended in 3 µl H<sub>2</sub>O and 7 µL hybrisol VII (Oncor Inc.). The probes were denatured for five minutes at 72°C and hybridized to chromosomes at 37°C for 14–16 hours. Slides were washed per manufacturer recommendations (Vysis Inc.) and counterstained using Vectashield antifade with DAPI (Vector Labs). Analysis of the FISH experiments was carried out on a Zeiss Axioplan2 microscope and photographed with a Hamamatsu black and white camera using Zeiss AxioVision software version 2.0 (Carl Zeiss Inc).

## **Chapter IV**

### **Analysis of Smith-Magenis syndrome patients without a FISH detectable**

#### **17p11.2 deletion**

#### **BACKGROUND AND RATIONALE**

Analysis of patient deletion sizes by FISH reduced the SMS critical region (SMCR) from 1.5 Mb to ~700 kb. This narrowing of the SMCR greatly reduced the number of candidate genes from >30 to 11. However, of 22 patient samples analyzed using FISH, we were unable to detect a 17p11.2 deletion in six samples even though thorough clinical analysis of these patients indicated a strong suspicion of SMS. We believed that these patients indeed had SMS and hypothesized that there may be dominant mutations in a gene within the SMS critical region that could cause SMS. We therefore undertook direct sequencing of three SMS candidate genes on samples from patients without a detectable SMS deletion. Sequencing was performed on exonic DNA sequence from the RAS-dexamethasone induced 1 (*RASD1*) gene, the developmentally regulated GTP-binding protein 2 (*DRG2*) gene, and the retinoic acid induced 1 (*RAI1*) gene. In our investigation of these 3 genes, we found dominant frameshift mutations in the *RAI1* gene in 4 of the 6 patients samples tested resulting in two publications from the lab (Slager et al. 2003; Girirajan et al. 2005).

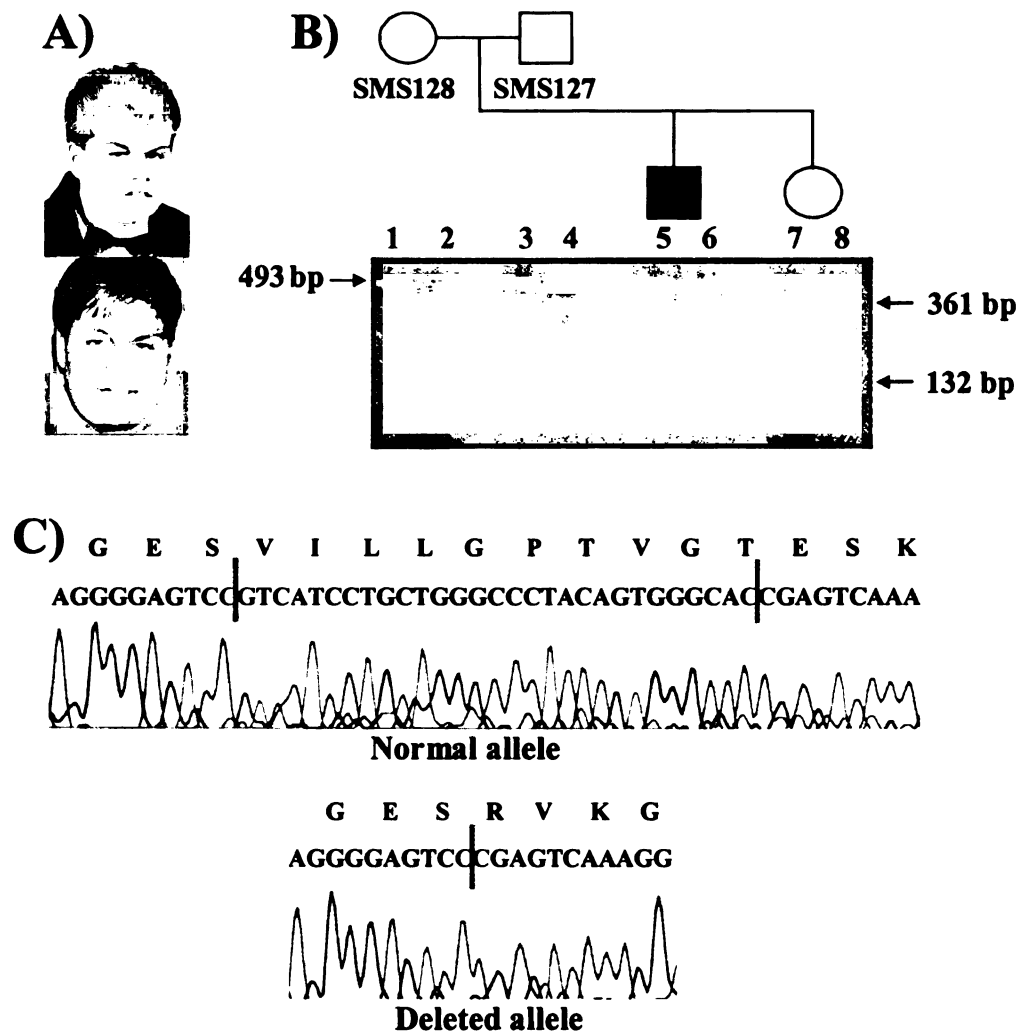
## **RESULTS**

### **Mutation analysis in SMS patients without a FISH detectable deletion**

As described in Chapter 3, thorough FISH analysis was performed on metaphase chromosomes from suspected SMS patients. In six patient samples we were unable to detect a deletion of chromosome 17p11.2. Though FISH can detect smaller deletions than G-banding, it does have limitations. Standard FISH analysis will not detect deletions smaller than ~40 kb. Nor will FISH detect small chromosomal anomalies (i.e. inversions or translocations) occurring intrachromosomally within the same sub-band, though these would be extremely rare.

The clinical description of many of our patient samples without FISH detectable deletions strongly suggested Smith-Magenis syndrome. The patients displayed craniofacial abnormalities, sleep disturbances and characteristic behaviors, including onychotillomania, polyembolokoilomania, spasmodic self-hugging, and explosive and aggressive episodes (Table 7). Because of these features, we decided that further analysis of individual genes within the SMS critical region using these patient samples was warranted.

We theorized that mutations in a gene, or genes, mapping within the SMS critical region might cause the phenotype seen in these patients. We started this study by systematically sequencing three candidate genes, *RASD1*, *DRG2*, and *RAI1* on DNA isolated from patient SMS129 (Figure 11A). These genes were three of our best candidates based on the initial characterization methods described in Chapter 2. Using PCR amplification and sequencing of all exons from the *RASD1*, gene we detected several previously reported SNPs, but no deleterious DNA changes were



**Figure 11. SMS129 *RAI1* mutation analysis.**

**A)** SMS129 pictured at age 19 and 30. **B)** Pedigree analysis of the SMS129, his parents (SMS127 and SMS128) and one sibling (SMS164). The RA13/14 PCR product is present in lanes 1, 3, 5, and 7; note doublet in affected individual SMS129. The *Psp*OMI digestion of the RA13/14 PCR product is shown in lanes 2, 4, 6, and 8; note undigested mutant allele in affected individual SMS129 which is not present in the parents or sibling. The 29 bp mutation was not detected in 102 Caucasian control samples which were analyzed by PCR amplification and *Psp*OMI restriction digestion. **C)** The sequence tracing represents the the 29 bp deletion detected following direct sequencing of the RA13/14 amplicon of *RAI1* from SMS129, as well as the wild type sequence detected on the other allele; this deletion eliminates a *Psp*OMI restriction site, misincorporates 8 amino acids (4 of which are shown in the diagram) and produces a downstream stop codon.

found. Amplification and sequencing of the exons of the *DRG2* gene showed no changes, polymorphic or mutagenic, from the consensus sequence available from the human genome project. This was not a surprise as the *DRG2* gene is highly conserved through evolution as will be discussed in Chapter 5.

Interestingly, a deletion of 29 base pairs was found in exon three of the *RAI1* gene on one allele in individual SMS129 (Slager et al. 2003). This deletion was clearly demonstrated as two distinct bands representing the normal and deleted allele when the PCR product was run on a 2% agarose gel (Figure 11B). Both bands were extracted and purified from the agarose gel, cloned, and then sequenced. The chromatograms from the cloned PCR products revealed one normal allele, and one allele missing 29 base pairs (Figure 11C). This 29 base pair deletion produces a frameshift that introduces 8 incorrect amino acids, followed by a stop codon, truncating the protein produced by this allele (Slager et al. 2003). Bioinformatic sequence analysis using web based programming of the deleted allele revealed the abolishment of a *PspOMI* restriction site (Figure 11B). Using direct sequencing and restriction digest analysis of DNA isolated from the parents and sibling of SMS129 (SMS127, SMS128, and SMS164) we determined that the 29 base pair deletion occurred *de novo* and is not familial (Figure 11B). Though, low level and/or gametic mosaicism cannot be ruled out using this methodology<sup>1</sup>. The *PspOMI* restriction analysis allowed for screening of 100 normal individuals (200 chromosomes) for this deletion (Slager et al. 2003). The deletion was not detected in any of the normal

---

<sup>1</sup> DHPLC analysis of DNA isolated from blood revealed that the mother of SMS129 is ~20% mosaic for the same 29 base pair mutation. The mother is phenotypically normal.

chromosomes analyzed and thus, does not represent a normal polymorphism. Haploinsufficiency of the *RAI1* protein due to the 29 base pair deletion likely causes the SMS phenotype seen in SMS129.

With the discovery of a dominant frameshift mutation in SMS129, we undertook sequencing of additional patients in whom a 17p11.2 deletion could not be detected by FISH analysis. Dominant frameshift mutations were found in SMS153, SMS156, and SMS159. As reported by the lab in 2005, amplification of exon 3 of the *RAI1* gene in SMS153 (Figure 12B) shows a 19 base pair deletion beginning at nucleotide 253 (Figure 12) (Girirajan et al. 2005). Similar to the mutation carried by SMS129, the deleted allele of SMS153 incorporates 60 incorrect amino acids into the *RAI1* protein followed by a premature stop codon. The deletion was not seen in parental DNA samples, nor in >110 normal chromosomes analyzed (Girirajan et al. 2005).

Analysis of *RAI1* exon 3 amplified from template DNA from SMS156 and SMS159 revealed single base pair deletions causing introduction of premature stop codons (Slager et al. 2003). In the case of SMS156 (Figure 14A, Table 7) a single cytosine within a run of 6 C's ending at nucleotide position 5265 of the *RAI1* mRNA was deleted on one allele (Figure 14B). The 5265delC causes a frameshift on the coding DNA strand introducing 74 incorrect amino acids before the addition of a premature stop codon. Bioinformatic analysis of the DNA sequence indicated that this single nucleotide deletion abolishes a *BglII* restriction enzyme recognition site. The 5265delC mutation was not found in either parent of SMS156 (SMS154 and SMS155) or in 100 normal control samples (equal to 200 chromosomes).

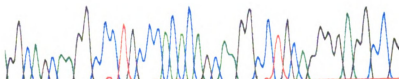


A)

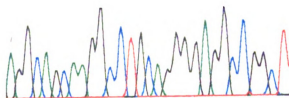


B)

G S K A L P T Q Q G L Q G R P  
 241 GG CAGCA AGGC CCT GCC CAC ACAGCA AGGC CTGC AGGGGA GGCC G Normal allele



G S K A C R G G R L  
 240 A GGC A GCA AGGC C TGC A GGG GAGGC C GGCT Deleted allele



253del 19

**Figure 12. SMS153 *RAI1* mutation analysis**

A) SMS153 pictured at age 19. B) The sequence tracing represents the 19 bp deletion detected following direct sequencing *RAI1* exon 3 from SMS153, as well as the wild type sequence detected on the other allele.

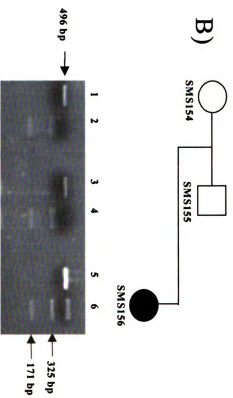
**Figure 13. SMS156 *RAI1* mutation analysis.**

**A)** SMS156 pictured at age 31. **B)** Pedigree analysis of the SMS156, her parents (SMS154 and SMS155). The RA25/26 PCR product is present in lanes 1, 3, and 5 and the *Bgl*I digestion of the RA25/26 PCR product is shown in lanes 2, 4, 6, and 8 (note undigested mutant band present in SMS156 which is not present in parental samples). The 5265delC mutation was not detected in 110 Caucasian control DNA samples which were analyzed by PCR amplification and *Bgl*I restriction digestion. **C)** The sequence tracing from the RA25/26 *RAI1* amplicon reveals the 5265 delC in exon 3 on one allele; this mutation eliminates a *Bgl*I restriction site, misincorporates 74 amino acids, eliminates a *Bgl*I restriction site, and produces a downstream stop codon

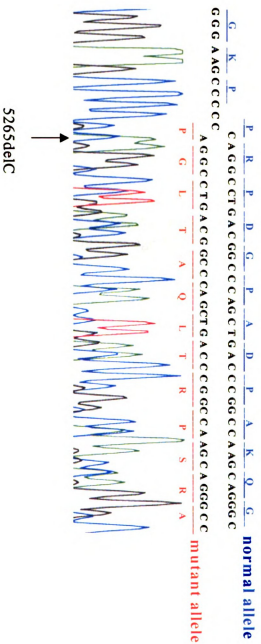
A)



B)



C)



**Figure 14. SMS159 *RAI1* mutation analysis.** A) SMS159 pictured at age 11. B) SMS159 pictured at age 19. C) The 1449delC mutation in *RAI1* exon 3 on one allele is shown, which misincorporates 34 (12 of which are shown in the sequence tracing) amino acids and produces a premature stop codon. As no known restriction site was altered by the 1449delC mutation, amplified and sequenced the amplicon containing the 1449delC mutation from >100 Caucasian samples (including the parents of SMS159), and this mutation was not detected in this population.

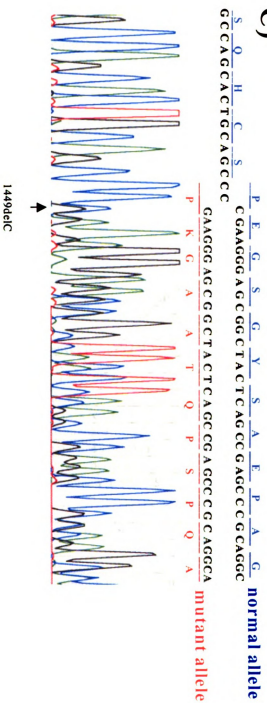
A)



B)



C)



**Table 7. Phenotypic comparison of SMS patients with 17p deletions and *RAI1* mutations**

	SMS 126 typical deletion	HOU14 2-540 small deletion	SMS 129	SMS 153	SMS 156	SMS1 59
<b>Features common in &gt;75% of SMS Patients</b>						
<b>Craniofacial/Skeletal</b>						
Brachycephaly	-	-	-	+	+	-
Midface hypoplasia	+	+	-	+	+	-
Prognathism (relative to age)	+	+	+	+	+	+
Tented upper lip	+	+	+	+	+	+
Broad, square face	+	-	+	+	+	+
Brachydactyly	+	-	-	+	+	-
Short stature	+	+/-	-	-	-	-
<b>Otolaryngologic abnormalities</b>						
Middle ear/laryngeal anomalies	+	+	-	+	+	+
Hoarse, deep voice	+	+	+	+	+	+
<b>Neurological/Behavioral</b>						
Mental retardation	+	+	+	+	+	+
Speech delay	+	+	+	+	+	-
Motor delay	+	-	-	+	-	+
Infantile hypotonia	+	-	-	+	-	-
Sleep disturbance	+	+	+	+	+/-	+
Stereotypic SMS behavior	+	+	+	+	+	+
Self-injurious behavior	+	+	+	+	+	+
<b>Features common in 50 – 75% of SMS patients</b>						
Hearing loss	+	+	-	-	+	-
Ocular abnormalities	+	+	+	+	+	+
Synophrys	-	+	+	-	-	-
Scoliosis	+	-	+	N	-	-
<b>Features common in &lt; 50% of SMS patients</b>				-		
Cardiovascular anomalies	+	+	-	-	-	-
Renal anomalies	-	+	-	-	-	-
Seizures	-	+	-	-	-	+
Cleft lip/palate	-	-	-	-	-	-
Gender/Age at evaluation	Female 41	Male 10,14, and 16	Male 14	Female 14	Female 17, 21	Male 13

Similar to the mutation in SMS156, DNA isolated and amplified from SMS159 (Figure 14 and Table 7) showed a deletion of a cytosine in a run of 4 C's on one allele starting at base pair 1449 of the *RAI1* mRNA (Figure 14) (Slager et al. 2003). The deletion of the cytosine at base pair 1449 causes incorporation of 34 incorrect amino acids before addition of a premature stop codon. Bioinformatic analysis of this mutation did not reveal abolishment or addition of a restriction site. Thus, template DNA from the family of SMS159 and normal control DNAs was amplified from exon 3 of the *RAI1* gene and sequenced. The 1449delC deletion was not detected in the parental or control samples analyzed.

Additional dominant mutations in the *RAI1* gene have now been reported from DNA isolated from patients without a FISH detectable 17p11.2 (Bi et al. 2004; Girirajan et al. 2005). The total number of SMS patients carrying dominant mutations in the *RAI1* gene totals 9 (Bi et al. 2004; Girirajan et al. 2005).

Though the frame shift and nonsense mutations reported have been theorized to prematurely truncate the RAI1 protein (Slager et al. 2003; Bi et al. 2005; Girirajan et al. 2005). Because SMS is also associated with microdeletion of 17p11.2 including the *RAI1* gene, a dominant gain of function mutation in the *RAI1* gene is unlikely. Research reported in the last 10 years indicates that premature stop codons caused by a frameshift or nonsense mutation may trigger decay of the mRNA before it is translated into a protein product (Cheng and Maquat 1993; Belgrader et al. 1994). Currently, research in the Elsea Laboratory is underway to determine whether premature stop codons in the *RAI1* mRNA trigger the nonsense mediated decay pathway. It is likely that nonsense mediated decay of the mutant *RAI1* mRNAs

causes haploinsufficiency of the RAI1 protein.

### **The retinoic acid induced 1 (*RAI1*) gene**

The mouse ortholog of *RAI1* was cloned prior to the discovery of the human gene. In 1995 a Japanese group identified *Rail*, then called *Gt1*, as a transcript that was up-regulated in mouse embryonal carcinoma cells following treatment with retinoic acid (Imai et al. 1995). Mouse embryonal carcinoma cells are known to differentiate into neuronal and glial cells when treated with retinoic acid. The original study aimed to understand the genetic factors underlying this differentiation (Imai et al. 1995). RNA *in situ* hybridization and immunohistochemistry revealed that *Rail* gene expression is localized throughout the adult mouse brain, specifically in neurons (Imai et al. 1995). The role, if any, the *Rail* gene plays in neuronal development has yet to be reported.

As discussed in Chapter 2, our group mapped the human *RAI1* gene to the middle of the SMS critical region as the EST clone DKFZp434A139 (Figures 3, 5, and 6) (Lucas et al. 2001). Further, using commercially available adult and fetal tissue northern blots we determined that the *RAI1* gene is expressed as an ~8.0 kb transcript in all tissues examined (Figure 6). Upon publication of these data in 2001 (Lucas et al. 2001), we were unable to determine any possible cellular role for the *RAI1* gene due to lack of any homology of our EST to known proteins.

Further complicating analysis of the gene was the number of different ESTs deposited into the databases that were thought to represent the *RAI1/Rail* genes. In 2003, the major human *RAI1* transcript was cloned from SKNSH neuroblastoma cells



treated with retinoic acid, the sequence was deposited as GenBank accession AY172136 (Toulouse et al. 2003). This group identified both the 5' and 3' UTR sequences, as well as the presence of an upstream CpG island and a putative retinoic acid response element immediately upstream of exon 1 (Toulouse et al. 2003). The mouse *Rail* gene is represented by Genbank accession AY5481752. Bioinformatic analysis indicates that both mouse and human sequences contain a bibartite nuclear localization signal, a polyglutamine tract, a polyserine tract, and a conserved extended PHD domain at the carboxy end of the amino acid sequence (Girirajan et al. 2005).

In depth bioinformatic analyses were performed in the lab by fellow graduate student Santhosh Girirajan on the human and mouse *RAI1/Rai1* protein sequences. The human and mouse sequences are well conserved with 84% identity and 89% similarity. It was determined that the polyserine tracts seen in these proteins are similar to the human *DRPLA* and *Drosophila* hairless gene. Consistent with the original *RAI1/Rai1* cloning experiments, both *DRPLA* and hairless are reported to play a role in neuronal development (Maier et al. 1992; Onodera et al. 1995; Seranski et al. 2001; Girirajan et al. 2005). Additionally, polyserine and polyglutamine stretches are known to modulate transcriptional activation (Gerber et al. 1994). The extended PHD domain in the carboxy terminal of the protein is similar to that seen in the trithorax nuclear proteins. In *Drosophila* these proteins maintain stable expression of homebox-like genes (Paro 1993; Aasland et al. 1995). The PHD-finger domains likely bind zinc ions and act similarly to Zn-RING finger containing proteins (Aasland et al. 1995). Currently, no evidence has been reported on the binding of *RAI1/Rai1* to DNA, RNA, or other proteins. The only other protein sharing

homology to *RAI1/Rai1* is the human *AR1* gene also known as transcription factor 20, or *TCF20* (Seranski et al. 2001; Girirajan et al. 2005).

The protein homology of *RAI1/Rai1* indicates that the gene is involved in transcription at some level, either as a bonefide transcription factor or a transcriptional co-activator. No reports of the *in vivo* function of the *RAI1/Rai1* gene currently exist. Members of the lab have shown that a GFP-Rai1 fusion protein is localized within the nucleus. Further study of the *RAI1/Rai1* genes will help understand their role in mammalian development and in expression of the SMS phenotype.

## **SUMMARY**

The discovery of 9 dominant *RAI1* mutations together with data analyzed from heterozygous knock-out of mouse *Rai1* (Bi et al. 2005) indicate that haploinsufficiency of the *RAI1* gene is the cause of the craniofacial, neurological, otolaryngological, and behavioral characteristics seen in SMS. Additional in depth studies are underway to better understand the cellular and biochemical role of the *RAI1* gene.

Currently, we postulate that other genes in the region may play a role in the more variable characteristics seen in SMS patients with 17p11.2 deletions, including cleft lip/palate, and visceral anomalies including the kidney and heart. Due to the frequency of hearing loss in SMS patients we also believe that additional genes affecting hearing may map to the SMS critical region. We are continuing to study the role that additional genes mapping within the SMS critical region play toward expression of the SMS phenotype.

## **MATERIALS AND METHODS**

*Patient ascertainment:* SMS129, SMS156, and SMS159 were evaluated by Brenda Finucane, CGC at the Elwyn Inc. (formerly Elwyn Institute). Clinical data for SMS153 were collected by Dr. Sarah H. Elsea, Ph.D. via contact with the patient's health care provider, and through our in house questionnaire (Appendix A). All clinical data and samples were collected in accordance with Institutional Review Boards at Elwyn Inc., Michigan State University (MSU), and Virginia Commonwealth University (VCU).

### *Clinical description of SMS129, SMS153, SMS156, and SMS159:*

**SMS129:** SMS129 is a 30-year-old male who was admitted to residential placement at age 14 because of aggressive and disruptive behaviors which could no longer be managed at home. He was the product of an uncomplicated full term pregnancy, weighing 7 lbs., 1 oz. Motor milestones were normal, with no history of hypotonia, but speech development was significantly delayed. From an early age he exhibited aggressive and self-injurious behaviors, including onychotillomania and polyembolokoilomania, as well as sleep disturbance, and frequent self-hugging. He is obese and has spina bifida occulta. His behavior is currently stable, although he continues to require residential placement and psychotropic medications because of aggression. Results of fragile X and chromosome analyses at age 14 were normal. At ages 21 and 30, he was reevaluated in the genetics clinic and felt to have many behavioral and physical characteristics of SMS. A repeat cytogenetic study (650 band resolution) was normal, as were results of chromosome analysis on skin

fibroblasts and FISH for del(17)(p11.2).

SMS153: SMS153 is a 19-year-old obese, white female. Pregnancy was normal, though delivery was by emergency cesarean section because of failure to progress and a decreased fetal heart rate upon inducing contractions. SMS153 was diagnosed in the operating room with Down syndrome because of floppy muscle tone, upslanting palpebral fissures, midface hypoplasia and presence of a simian crease. Laboratory studies performed were negative for Down syndrome.

Motor and speech were significantly delayed. When she did begin speaking her voice was noted as hoarse. Her facial features are consistent with the SMS phenotype and include brachycephaly, tented upper lip, and broad square face. SMS153 suffers from the self-injurious behaviors commonly seen in SMS patients. These have been particularly difficult on the family. SMS153 also suffers from the characteristic SMS sleep disturbance.

SMS156: SMS156 is a 31-year-old obese, white female. Pregnancy and delivery were normal; birth weight was 8 lbs. Motor milestones were achieved on time, and there was no history of hypotonia. Speech development was mildly delayed compared to that of her siblings. The patient has a history of mild mental retardation and emotional disturbance, including aggressive and defiant behaviors, which prompted residential placement during adolescence. She had self-injurious behaviors, including onychotillomania and polyembolokoilomania, as well as significant sleep disturbance. Her behavior improved in the residential setting, and she was able to return home to live with her parents at age 21. On reevaluation at age 31, her behaviors had stabilized, and she required no psychotropic medications. She

continues to exhibit skin and nail picking, and when excited, the self-hugging stereotypy typical of people with SMS. She also has a persistent habit of inserting string into her nose. Her facial features are subtly similar to those seen in patients with a 17p11.2 cytogenetic deletion. Chromosome and fragile X analyses at age 17 were normal, as were results of subsequent FISH studies for deletion 17p11.2.

SMS159: SMS159 is a 19-year-old male patient who has a history of mild mental retardation, self-injury, and aggressive behaviors which resulted in residential placement at age 13. Pregnancy and delivery were uncomplicated; birth weight was 8 lbs.1 oz. A functional heart murmur was detected in childhood and later resolved. Motor milestones were significantly delayed (walked at 21 months) due to infantile hypotonia. Speech developed within normal limits. Since early childhood, the patient has exhibited self-hugging behavior when excited. Macrocephaly was noted in infancy, and MRI studies of his head and spine at age 11 revealed mild hydrocephalus, Arnold-Chiari malformation, and spina bifida occulta. On physical exam at age 13, he was obese and had facial features suggestive of SMS. The patient also had gynecomastia and hypogonadism. Results of fragile X and cytogenetic studies were normal, including FISH analysis for deletion 17p11.2. At age 19, the patient lived at home and received homebound school services. He suffers from recurrent, debilitating syncopal episodes likely related to his Arnold-Chiari malformation. These episodes were temporarily alleviated by brainstem decompression surgery at age 15 but have gradually increased since then. The patient continues to present behavioral problems, including onychotillomania, which are treated with psychotropic medications. Polyembolokoilomania has not occurred for

several years.

*Fluorescent in situ hybridization (FISH):* FISH on patient chromosomes without detectable 17p11.2 deletions was performed as described in Chapter 3.

*Isolation of patient genomic DNA:* In order to isolate template suitable for PCR amplification, DNA was isolated from peripheral blood or buccal cells from all putative SMS patients, parental and sibling controls, and Caucasian control samples.

*DNA isolation from whole blood:* Approximately 15-50 ml of whole blood was collected and centrifuged at 2500 RPM plasma was removed, and the remaining cells were mixed with “red blood cell lysis solution A” (0.32 M sucrose, 10 mM Tris, pH 7.5, 5 mM MgCl<sub>2</sub> and 1% Triton X-100) and placed on ice for 30 minutes. The mixture was then centrifuged at 2500 RPM, the supernatant was removed, and 50 ml of solution A were added to the pellet. This solution was placed on ice for 20 minutes, centrifuged as above, and the supernatant was removed. The pellet was resuspended in 5 ml of solution B (10 mM Tris, pH 7.5, 400 mM NaCl, and 2 mM EDTA, pH 8) and digested overnight at 37°C with the addition of 100 µl of 20% SDS and 50 µl of 20 mg/ml proteinase K solution. The following day, 3 ml of saturated phenol pH 8.0 were added to the solution while rocking, and then the samples were centrifuged at 2500 RPM for 15 minutes. The upper phase was removed with a Pasteur pipet and transferred to a new 15 ml polypropylene tube and 3 ml of chloroform:isoamyl alcohol (24:1) were added to the aqueous phase with rocking for

15 minutes. Following another centrifugation at 2500 rpm, the DNA upper phase was removed and the DNA was precipitated with 2 volumes of 95% ethanol. The DNA was removed with a Pasteur pipet and allowed to sit 70% ethanol for 5 minutes, then placed in 200  $\mu$ l of 1x TE, pH 7.5.

*DNA isolation from buccal cells:* Genomic DNA was isolated from buccal cells by boiling the cheek brushes in 400  $\mu$ l of 50 mM NaOH at 95°C for 10 minutes. The brush was then discarded and the sample was placed on ice for 10 minutes. This solution was neutralized with 40  $\mu$ l of 1 M Tris, pH 8.0.

*RAI1 PCR amplification and sequencing reactions:* Analysis of *RAI1* in patient and control DNA was performed by PCR and subsequent sequencing and analysis of PCR products was performed by Rebecca Slager in the Elsea Lab. PCR primers covering the entire *RAI1* coding sequence, 5' and 3' untranslated regions (UTR), and alternative splice variants were generously provided by Dr. Laura Schmidt of NCI-Frederick or were designed by this laboratory and synthesized at the Michigan State University Macromolecular Structure Sequencing and Synthesis Facility. PCR was performed in a 25  $\mu$ L volume with 50-200 ng DNA template as described in Chapter 2. Informative primer sequences and conditions are indicated in Appendix B.

In order to check each PCR amplification 5  $\mu$ L of the reaction was electrophoresed in 2% agarose gels containing ethidium bromide. Successful reactions were then purified using the Qiagen Gel Extraction Kit according to manufacturer's instructions, or treated enzymatically in the following manner: 2  $\mu$ L of USB shrimp alkaline phosphatase (1 units/ $\mu$ L) and 1  $\mu$ L USB exonuclease I (10

units/ $\mu$ L) were added to 5  $\mu$ L of PCR amplification mixture, the solution was mixed and incubated at 37°C for 15 minutes in a thermocycler, and then inactivated at 80°C for 15 minutes. A sequencing reaction containing at least 10-40 ng of purified PCR product template in distilled water and 30 pmol of sequencing primer (the forward or reverse PCR primer or an internal primer) was then prepared and sequencing was conducted at the Michigan State University Genomics Technology Support Facility or the Virginia Commonwealth University Nucleic Acid Research Facility using an ABI PRISM® 3100 Genetic Analyzer or ABI PRISM® 3730xl DNA Analyzer.

*Sequence analysis:* PCR amplified *RAI1* fragments were analyzed for mutation by pairwise BLAST analysis at NCBI (<http://www.ncbi.nlm.nih.gov/BLAST>) against the published GenBank sequence for the *RAI1* mRNA (GenBank Accession: AJ271790 and AY172136) and the genomic sequence data from the *RAI1* locus (GenBank Accession: AJ271791 and NT\_010718). SNP databases at NCBI (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=snp>) were also queried in order to identify polymorphic DNA changes. Sequencing chromatograms were also thoroughly visually inspected in order to identify sequence anomalies missed in the automated sequence analysis.

*Restriction enzyme analysis of RAI1 mutations:* The 29 base pair deletion detected in SMS129 abolishes a *Psp*OMI restriction site. Parental and 100 normal control DNA samples were tested for this deletion via restriction analysis. *Psp*OMI digestion of the amplicon from primer pair RA13/14 (Appendix B) normally cleaves the 493 base



pair product into 361 base pair and 132 base pair fragments. *PspOMI* does not cleave the allele with the 29 base pair deletion. PCR products were digested for 2 hours at 37°C with 1-2 U of *PspOMI*, then resolved on a 2% agarose gel in 1x TBE buffer containing ethidium bromide. Agarose gels were photographed on an AlphaImager using software version 5.5.

The 5265delC mutation detected in SMS156 abolishes a *BglI* restriction site. Parental and 100 normal control DNA samples were tested for this deletion via restriction analysis. *BglI* digestion of the amplicon from primer pair RA25/26 (Appendix B) normally cleaves the 496 base pair product into 171 base pair and 325 base pair fragments. *BglI* does not cleave the 5265delC allele. PCR products were digested for 2 hours at 37°C with 4-5 U of *BglI*, then resolved on a 2% agarose gel in 1x TBE buffer containing ethidium bromide. Agarose gels were photographed on an AlphaImager using software version 5.5. Restriction analysis was performed by Rebecca Slager in the Elsea Lab.

*Analysis of control samples:* Analysis normal control samples for mutations that could not be resolved with restriction analysis were directly sequenced and analyzed as described in the methods above.

## **Chapter V**

### **Developmentally regulated GTP-binding protein 2 (*DRG2*)**

#### **BACKGROUND**

Guanine triphosphate hydrolases (also known as GTPases or G-proteins) make up a superfamily of proteins that act as regulatory molecules. The GTPases are known as molecular switches because they are activated while they are bound to GTP and switched off upon hydrolysis of the GTP to GDP. This group of proteins is expansive and is made up of three well-known sub-families; the monomeric GTP-binding protein Ras and its homologs, bacterial translation elongation factors, and the heterotrimeric G proteins (Bourne et al. 1991).

All G-proteins work in the same mechanistic manner even though the primary structures are extremely divergent. The proteins maintain their function in the absence of long stretches of homology through conservation of four distinct domains that allow for conserved folding of the GTP binding core. The differences in primary structure allow for the extreme diversity seen in overall protein function, while conservation of the short GTP binding domains create the very well conserved secondary structure necessary for GTP binding and hydrolysis.

As additional expressed sequences became available additional sub-families of GTPases were described. One such family includes the developmentally regulated GTP-binding proteins (DRG) proteins. The DRG proteins retain the G1-G4 binding motifs necessary for GTP-binding, though they do not share similarity to any of the

well-known GTPase sub-families. One member of the DRG family maps within the SMS critical region on chromosome 17 (Schenker and Trueb 1997; Vlangos et al. 2000). The map location of the developmentally regulated GTP-binding protein 2 (*DRG2*) gene and limited information about its possible role in development, made it one of our best SMS candidate genes. In addition to fine mapping of the gene, we undertook characterization of *DRG2* through expression patterning at the RNA and protein levels.

### **Cloning and identification of mammalian DRG genes**

In 1992 a subtractive cloning approach was used to identify genes that are expressed at high levels in developing and differentiating mouse brain when compared to adult brain. This approach identified many potential transcripts including the *Nedd3* gene (neural precursor cell expressed, developmentally downregulated-3) (Kumar et al. 1992). Identification and characterization of a longer cloned sequence for *Nedd3* showed that it contained short regions of homology to *Drosophila* and *Halobacterium* sequences deposited in public databases. Early bioinformatic analysis of the three related sequences revealed that they all shared domains of GTP-binding. *In vitro* analysis of the Nedd3 protein showed its ability to bind GTP. Based on the results from the *in vitro* GTP-binding assay and the subtractive cloning method enriching for developmental genes used in its original identification the *Nedd3* gene was assigned the name developmentally regulated GTP-binding protein (Sazuka et al. 1992b).

The putative human homolog of *Drg* was also identified by subtractive cDNA

cloning where it was shown to be expressed in human fibroblasts but repressed in virally transformed cells (Schenker et al. 1994). Using FISH techniques, the human gene was mapped to chromosome 17p12 → 13 (Schenker and Trueb 1997). Searches of DNA databases revealed that there were at least 2 *DRG* genes in humans, *DRG1* and *DRG2*. In 2000 we reported that the human gene identified by Schenker *et al* is actually *DRG2*, while original murine *Drg* gene is actually homologous to human *DRG1* (Vlangos et al. 2000).

## **RATIONALE**

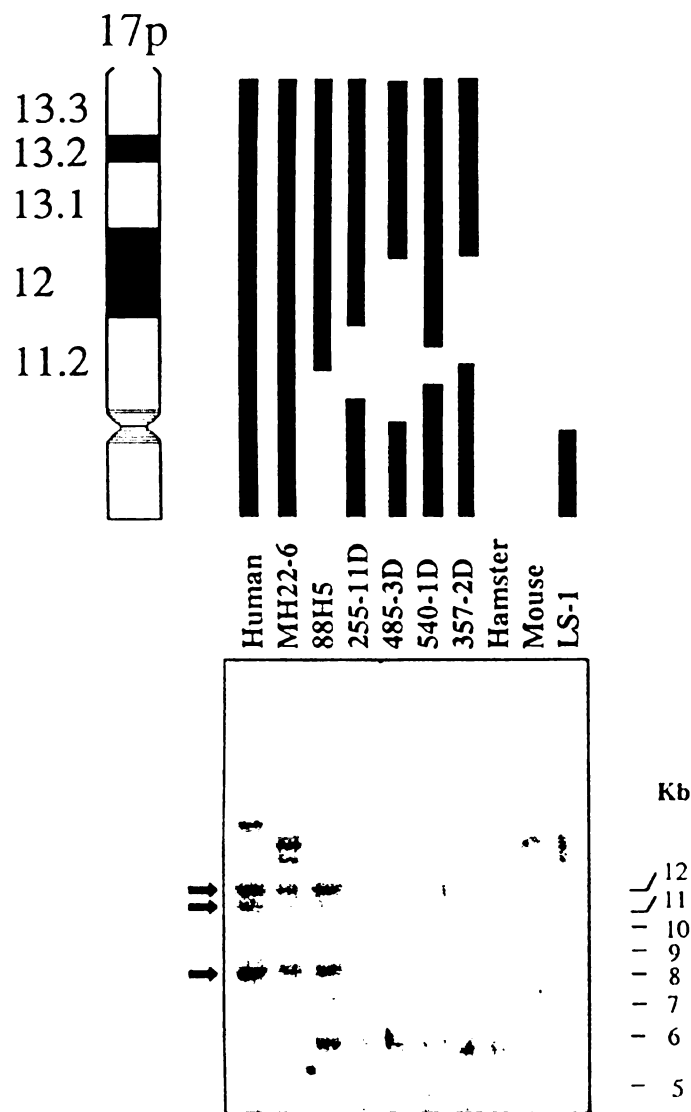
Our lab mapped the *DRG2* gene to chromosome 17p11.2 within the SMS critical region. In order to begin to understand the function of *DRG2* and any role it might play in SMS, we looked at RNA expression patterns of the gene in adult and developing human and developing mouse. We employed immunohistochemistry techniques to determine spatial and temporal expression through mouse development. At the cellular level, localization of the protein was performed using green fluorescent protein technique. The data presented in this chapter are currently being drafted into a manuscript for publication.

## **RESULTS**

### **Identification of the STS WI-13499 as *DRG2* and mapping to the SMS critical region**

In order to clarify the map location and identification of *DRG2* to chromosome 17p11.2, we used the sequence tagged site (STS) WI-13499. We and others previously mapped STS WI-13499 to chromosome 17p11.2 (Elsea et al. 1997; Liang et al. 1999). Subsequent BLAST searched and Unigene analysis of the published WI-13499 sequence indicated 94% identity to the *DRG2* gene. The STS WI-13499 was created from the I.M.A.G.E. cDNA clone 29328. Analysis of the DNA sequence of clone 29328 revealed that it bore 99% identity to *DRG2*. Lower identity of the *DRG2* sequence to WI-13499 was likely due to one-pass sequencing errors. To ensure that we were truly looking at a single-copy gene and that clone 29328 (WI-13499) and the reported *DRG2* sequences were representative of the same gene, we fully sequenced clone 29328 and the PCR product WI-13499 and aligned them to the reported *DRG2* sequences from GenBank. The results indicate that the sequences are truly identical and representative of the same locus.

*DRG2* was originally mapped to chromosome 17p13 → p12 by FISH (Schenker and Trueb 1997). Experiments using Southern analysis (Figure 15) and PCR mapping with somatic cell hybrids retaining portions of chromosome 17 indicate that *DRG2* maps more proximally on chromosome 17, to band 11.2 and the SMS critical region. In order to confirm our mapping data, we obtained a 1,300 base pair *DRG2* clone from Dr. B. Trueb (originally reported in Schenker and Trueb, 1997) and



**Figure 15. Somatic cell mapping of *DRG2*.**

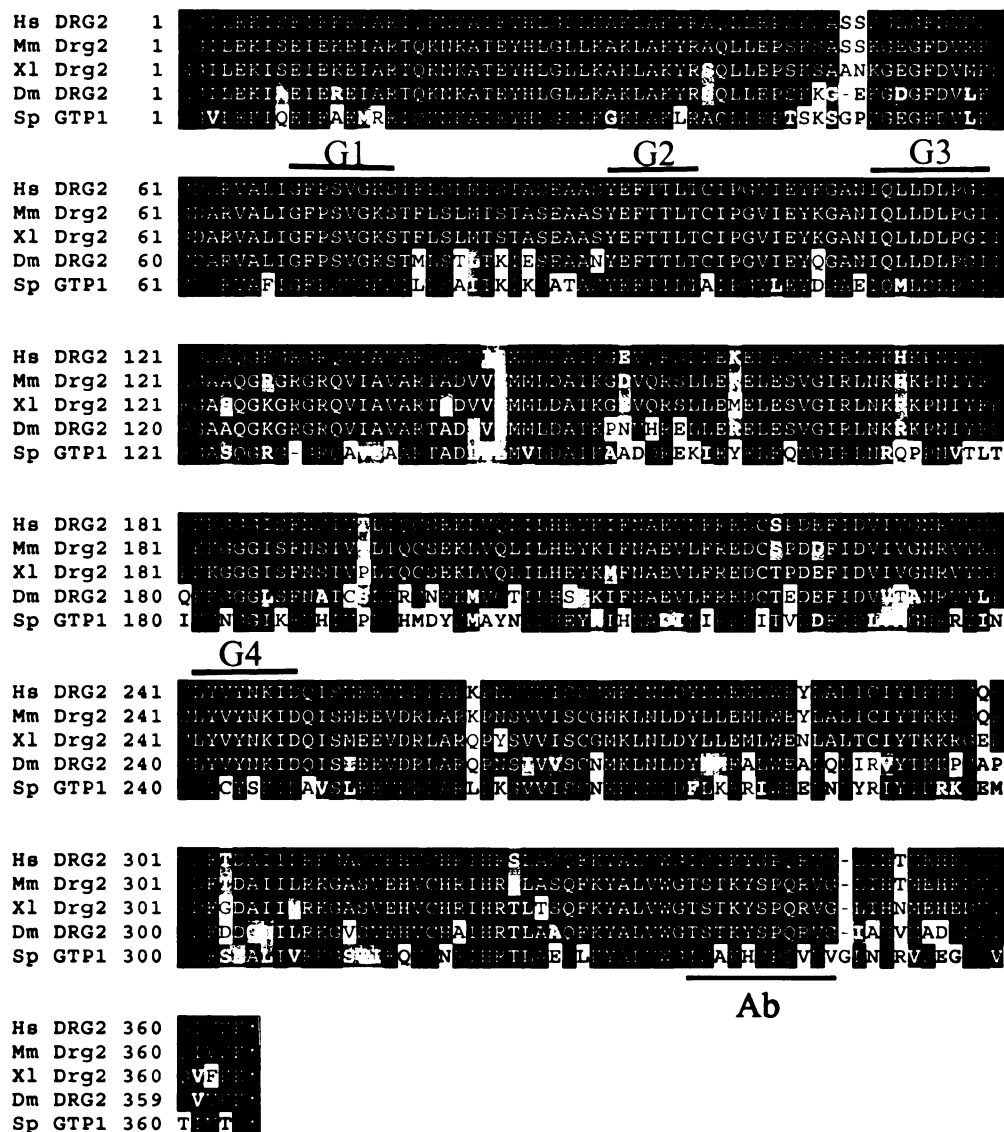
*Eco*RI digests of human, rodent, and hybrid genomic DNAs are shown. The entire ~1.9 kb insert from IMAGE clone 29328 was PCR amplified and used as a probe. Bars above the blot indicate relative regions of chromosome 17 deleted in human and hybrid DNAs. Arrows to the left indicate chromosome 17 specific genomic bands.

obtained the same mapping results using the somatic cell hybrid mapping panel. These data confirmed that *DRG2* maps within the SMS critical region (Vlangos et al. 2000).

### **Bioinformatic analysis of *DRG2***

By direct sequencing we determined that clone 29328 contained an EST insert of 1,917 base pairs. Our southern analysis showed that the *DRG2* gene likely spans a minimum genomic region of 32 kb (Vlangos et al. 2000). Alignment of the cloned EST sequence to the latest version of the human genome sequence shows that *DRG2* is composed of 13 coding exons spanning 20 kb of DNA. A potential ATG start site begins at base number 76 and fits the Kozak consensus sequence of ACCATGG (Kozak 1986). Six-frame translation of the DNA sequence showed an open reading frame resulting in a product of 364 amino acids. This protein has an estimated molecular weight of 40.5 kDa. Four conserved GTP binding domains remain intact in the DRG2 protein (Figure 16). The G1 domain is unique to this subfamily of G-proteins and has been termed OBG/GTP1 based on the orthologous *B.subtilis* and *S.pombe* genes.

Alignment of DRG2 amino acid sequences reveals striking homology through evolution (Figure 16). The human sequence is 98.6% identical to the mouse sequence, 93.9% identical to the *Xenopus*, 80.5% identical to *Drosophila* sequence, and 65.9% identical to the fission yeast sequence. Though the DRG2 sequences are well conserved through evolution, DRG1 and DRG2 only share 57% identity at the amino acid level. It is interesting that human DRG2 is more similar to the fission



**Figure 16. Alignment of DRG2 amino acid sequences.** Amino acid sequences from human DRG2 (Hs DRG2), mouse DRG2 (Mm DRG2), Xenopus (Xl Drg2), *Drosophila* DRG2 (Dm DRG2), and fission yeast gtp1 (Sp gtp1) were obtained and aligned using the CLUSTALW 1.8 program at the Baylor College of Medicine Search Launcher web site (<http://searchlauncher.bcm.tmc.edu>). The black shading indicates identical amino acids, the grey shading indicates similar amino acids, the white shading indicates differing amino acids. The star indicates stop codon, and the – indicates a gap in alignment. After alignment, the figure was created using the BOXSHADE program ([http://http://www.ch.embnet.org/software/BOX\\_form.html](http://http://www.ch.embnet.org/software/BOX_form.html)).



yeast DRG2 than to the human DRG1 sequence.

Previously published phylogenetic analysis of the DRG sequences available in 2000 revealed that 2 genes seen in higher organisms likely evolved from a gene duplication event (Li and Trueb 2000). Only a single DRG gene is present in bacteria. These data reveal that the DRG2 proteins are more similar to the DRG gene seen in bacteria and that the DRG1 proteins are likely the result of gene duplication.

### **Expression levels of *DRG2/Drg2* using northern analysis**

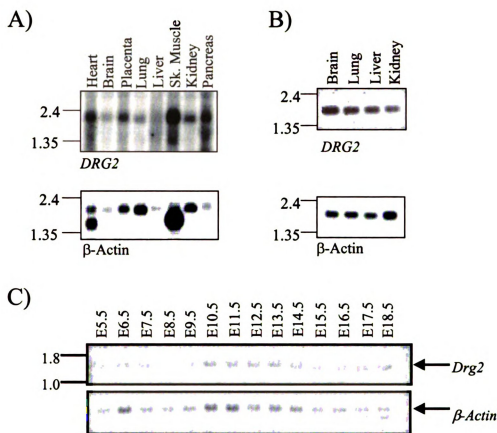
Since *DRG2* is thought to be developmentally regulated, we characterized both human and mouse *DRG2/Drg2* using northern analysis on adult and fetal tissues. Northern analysis of human tissues on a commercially available blot showed an ~2.2 kb transcript in all adult and fetal tissues represented (Figure 17A and 17B). Expression was not significantly elevated in any human tissue analyzed. Rather, the gene seems to be expressed at fairly equal levels in all tissues represented on the northern blots. Electronic expression profiling of human samples using the Unigene database at NCBI currently shows highest levels of *Drg2* expression in cervix and small intestine.

In order to characterize the expression pattern of *Drg2* in the developing mouse, we utilized a northern blot representing embryonic day 5.5 to day 18.5. The blot does not contain RNA from maternal or placental tissues. Using a full-length EST clone for *Drg2* as a probe a ~2.0 kb transcript representing *Drg2* expression was seen in all developmental days at a consistent level (Figure 17C). The blot did not reveal any drastic peaks or depression in gene expression during development. This

is in contrast to reports of mouse *Drg1* where there was a significant decline in gene expression starting after embryonic day 10 (Sazuka et al. 1992a). Electronic expression profiling of mouse *Drg2* indicate expression in all tissues examined and at all developmental stages. Our studies, thus far, do not indicate that *DRG2/Drg2* is developmentally regulated at the RNA level.

### **Spatial and temporal expression of Drg2 using immunohistochemistry**

In order to understand the expression pattern of the Drg2 protein in the developing mouse embryo we undertook immunohistochemistry (IHC) experiments on sectioned C57Bl/6J embryos. The IHC experiments were performed using a polyclonal antibody designed to an 11 amino acid sequence near the carboxy end of the DRG2 protein (Figure 16). The sequence of DRG2 was extensively analyzed in the design of the peptide used to create the antibody. The sequence was run through computer programs generating a hydropathy plot showing regions of hydrophilic amino acids most likely to reside on the outside of the folded protein. BLAST analysis of all deposited amino acid sequences revealed no known homology to any other proteins (known or predicted) except DRG2. The 11 amino acid sequence does not occur in the DRG1 protein. The antibody produced by this peptide is recognized in mouse *Drg2*, *Xenopus drg2*, and *Drosophila* DRG2 allowing for experimental use in multiple organisms. The peptide was injected into two rabbits for polyclonal antibody production.

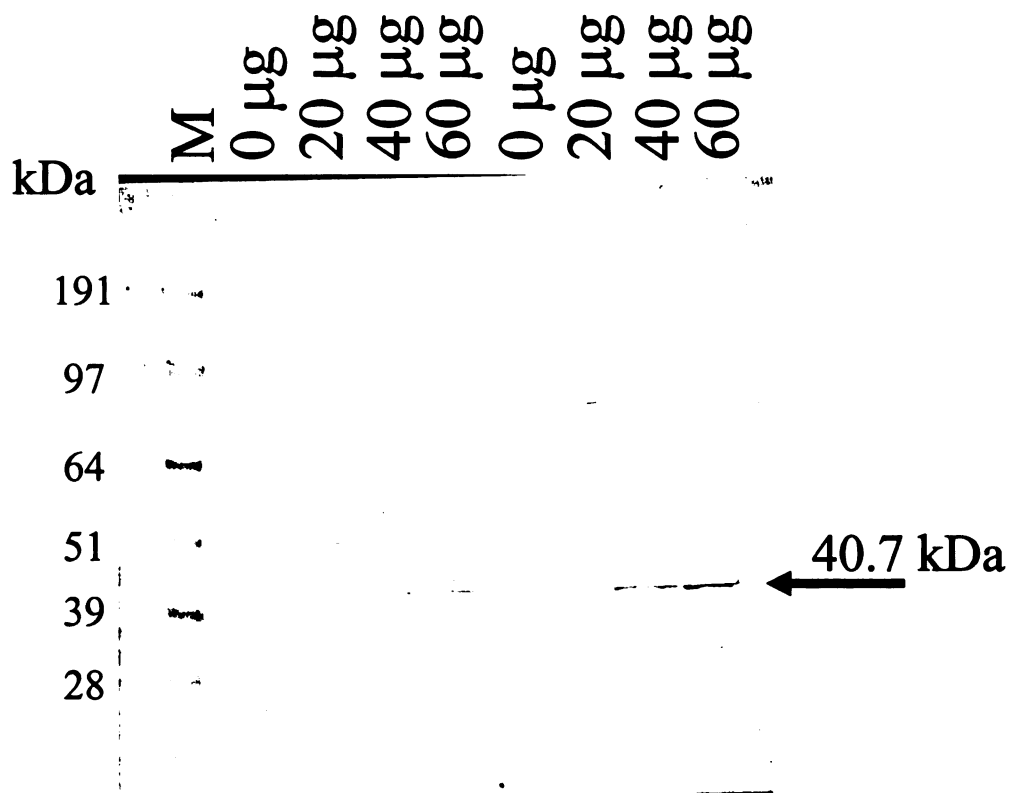


**Figure 17. Northern analysis of *DRG2/Drp2*.**

**A)** Human *DRG2* expression analysis is seen throughout all tissues in this Clontech adult multiple tissue northern blot. **B)** Human *DRG2* is seen in all tissues present in this fetal Clontech multiple tissue blot. **C)** A consistent expression of an ~1.7 kb transcript throughout mouse development is visualized. Northern analysis was performed on a mouse embryonic blot purchased from SeeGene USA (Rockville, MD). IMAGE clone 4910172 representing mouse *Drp2* was used as a probe.  $\beta$ -actin was used as a loading control.

ELISA analysis of the polyclonal antisera showed that the antibody could detect the DRG2 peptide at 1:100,000 dilution. The polyclonal antiserum from rabbit 8139 was further tested in order to determine whether or not it had the ability to recognize the mouse *Drg2* protein. The mouse *Drg2* gene was cloned into a vector containing a cytomegalovirus (CMV) promoter. The plasmid was transfected into COS-7 cells allowing for *Drg2* expression from the CMV promoter. Lysate from cells transfected with varying amounts of DNA was used to create a western blot. The DRG2 polyclonal antiserum from rabbit 8139 was used to detect the *Drg2* protein expressed in the transfected cells. The western analysis revealed an ~40 kDa band of varying intensity which correlated directly with the amount of EST DNA transfected into the COS-7 cells (Figure 18). The *Drg2* banding pattern was not detected in untransfected cells. These data show that the antiserum has the ability to detect the proper epitope.

Sectioned mouse embryos representing embryonic day 8 through birth were examined via IHC to determine the spatial expression pattern of the *Drg2* gene through development. Gestation in the mouse is a fairly rapid process lasting ~21 days. Dissection of embryos prior to day 8 was not attempted due to the small size of the embryo. At embryonic day 8 *Drg2* is highly expressed within the neural tube, migrating neural crest, and somites (Figure 19A), while low level staining is present throughout the rest of the embryonic section.

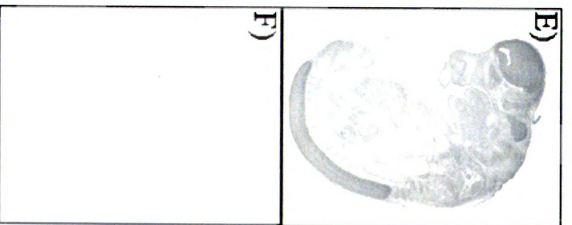
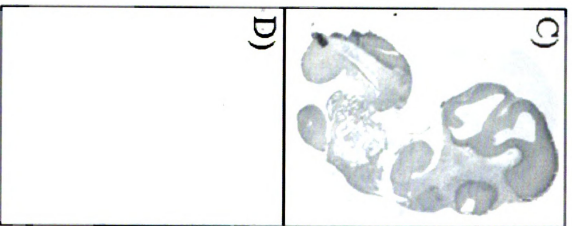
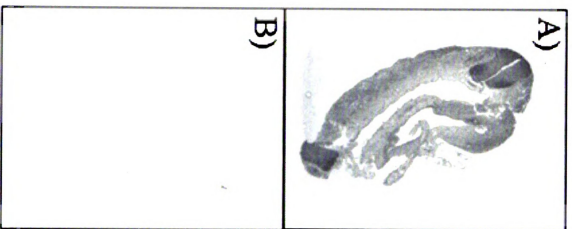


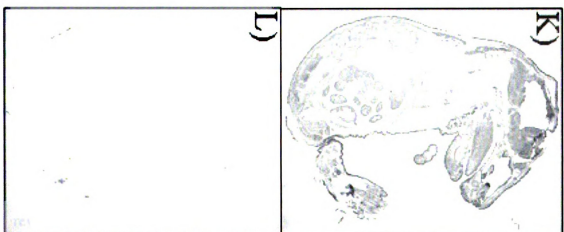
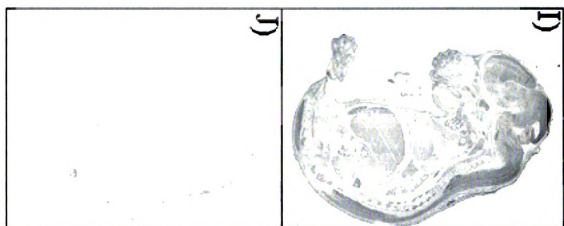
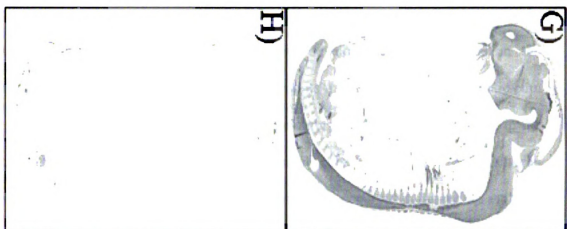
**Figure 18. Transfection of Drg2 into COS-7 cells.**

COS-7 cells were transfected with the full length *Drg2* sequence under control of the CMV promoter. Cell lysate from transfected cells was run on a western blot and analyzed with a polyclonal antibody created to the Drg2 protein. A 40.7 kDa band is seen in transfected cells. The band intensity correlates directly with the amount of DNA added to the cells. The amount of DNA input into the cells is indicated at the top of the figure.

**Figure 19. Protein expression of Drg2 during mouse development.**

**A)** An 8.5 day mouse embryo transverse sectioned is stained with anti-Drg2. Expression is seen throughout the embryo, and is highest in the neural tube and neural crest. **C)** A 10 day sagittal section shows Drg2 expression throughout the embryo with highest expression in the developing brain and neural tube. **E)** and **G)** A day 12 and day 14 embryo respectively stained with Drg2 antibody showing expression in the developing nervous system. **I)** and **K)** show a 16 day and 19 day mouse embryo respectively and show expression in the neural tissue and spinal ganglia. **B D, F, H, J, and L)** Show corresponding pre-immune serum controls for respective embryos.







By embryonic day 10, the most impressive development has occurred in the embryonic brain. The optic vesicle is beginning its transformation into the optic cup. Expression at this time is seen throughout the embryo with higher expression in the optic stalk, forebrain, hindbrain, and neural tube (Figure 19C). Between developmental day 10 and 12 significant maturing has occurred in the mouse embryo. The heart, vascular system, digestive system, and external features are more clearly defined. *Drg2* expression at embryonic day 12 is clearly seen in the developing spinal cord, optic recess, and the developing brain (Figure 19E). At developmental days 14 and 15 the most drastic changes are in the skeletal system where the pre-cartilage has developed into mature cartilage. Additionally, the heart has continued to develop, and the components have reached their post-natal configuration. *Drg2* expression at day 14 is very clear in the developing neural tissues of the entire brain and spinal cord. Additionally, expression is seen in the cervico-thoracic ganglia (Figure 19G). The expression in these ganglia mirrors the reported expression of *Drg1* in day 14 embryos as measured by RNA *in situ* hybridization (Sazuka et al. 1992a). Expression of *Drg2* at day 15 is seen diffusely throughout the developing mouse, with higher expression still being seen in the developing brain and neural tissues (Figure 19I). At embryonic day 20, *Drg2* expression is seen diffusely throughout the entire developing animal. In Figure 19K slightly higher expression can still be noted in the brain and spinal ganglia.

Analysis of the RNA expression patterns of *Drg2* in the developing mouse did not reveal developmental regulation as reported with *Drg1*. As with the RNA expression, analysis of the spatial and temporal expression of the gene using IHC also

reveals expression at all developmental stages evaluated. Greater expression was noted in the developing neural tissue throughout development. These results are similar to those reported for *Drg1* expression as reported using RNA *in situ* experiments. Both human *DRG* homologs may play an interesting role in neural development.

### **Cellular localization of Drg2**

The spatial and temporal expression patterns of *Drg2* revealed in the developing mouse embryo show that the gene is expressed throughout development, especially in developing neural tissue. In order to further understand the biochemical role of the *Drg2* gene, we used green fluorescent protein (GFP) techniques to evaluate the spatial expression of the gene at the cellular level.

The *Drg2* gene was cloned into a vector containing the sequence of the GFP gene. Cloning was performed such that the *Drg2* gene was cloned directionally and in frame with the GFP gene. Expression of the GFP-*Drg2* fusion protein was performed by transfection of COS-7 cells. Examination of cells transfected with the GFP-*Drg2* fusion protein revealed localization to the endoplasmic reticulum (ER) and/or Golgi apparatus (Figure 20).

Transfection of COS-7 cells was repeated in order to verify the cellular localization with ER and Golgi control stains. After transfection but before visualization, cells were treated with a dye that localizes to the ER and fluoresces a blue-white color under a long-pass DAPI filter. Ceramide stain localizing to the Golgi apparatus fluoresces red under a rhodamine or Texas red filter. The two

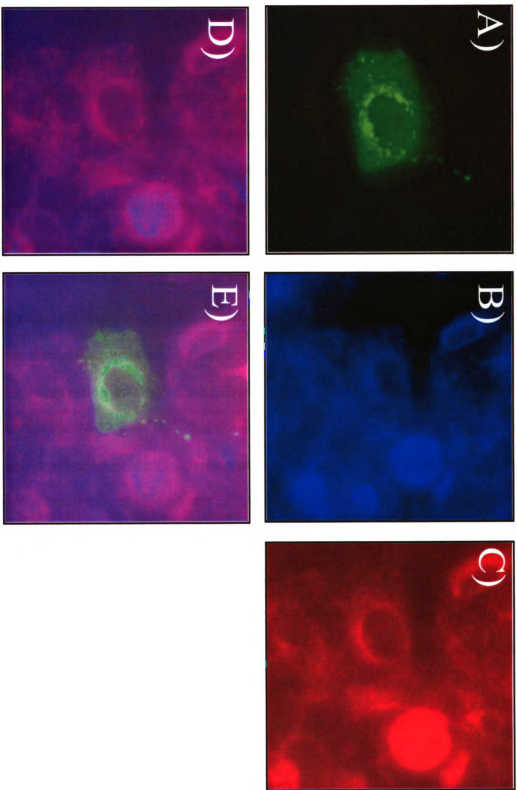
control stains co-localize with the GFP-Drg2 fusion protein indicating that Drg2 likely functions in the ER and Golgi apparatus.

The best studied GTPases associated with the Golgi apparatus and the ER are the ADP-ribosylation factor (ARF) genes and the (Rab) genes. ARF activity has been shown to regulate intracellular vesicular membrane trafficking and stimulate a phospholipase D isoform. Phospholipase D may be involved in cell morphology alterations through intracellular protein trafficking (Hammond et al. 1997). Homozygous mutations in a homologous ARF like gene have been shown to be one possible cause of Bardet-Biedl syndrome type 3 (Fan et al. 2004).

The Rab proteins regulate vesicular transport in endocytosis and exocytosis. They are small GTPases that are distantly related to the Ras proteins. The role of the Rab proteins seems to be as a regulator of the sequential events that must occur for proper vesicular transport (Rodman and Wadinger-Ness 2000). Rab genes have been associated with Griscelli syndrome which manifests with immunological anomalies and partial albinism. Recently, it has been suggested that Rab genes may be one of the many causes of Batten disease (Luiro et al. 2004). These genes are also currently being investigated as the cause of many cancers (Cheng et al. 2005). Many additional GTPases including Rabs, ARFs, and DRGs (and their associated proteins) are being investigated for roles in human genetic diseases.

**Figure 20. Cellular localization of Drg2.**

COS-7 monkey fibroblast cells were transfected with a Drg2-pcDNADEST47 construct, allowing expression of a Drg2-GFP fusion protein. Cells were treated with Bodipy TR Ceramide and ER-Tracker Blue-White to stain the golgi apparatus and endoplasmic reticulum, respectively. **A)** Drg2-GFP fluorescence is visualized under a FITC filter, **B)** ER fluorescence under a DAPI filter and **C)** Golgi fluorescence under a rhodamine filter, **D)** A merged image of the DAPI and rhodamine filter showing colocalization of the ER and Golgi, and **E)** A merged image of the DAPI, FITC, and rhodamine filters showing co-localization of Drg2-GFP to the ER and Golgi.



## **SUMMARY**

Developmentally regulated GTP-binding protein-2 was named because of its homology to *Drg1*, a possible developmentally regulated gene. The *Drg1* gene was originally reported as being developmentally regulated because of the subtractive cloning method used in its original isolation (Sazuka et al. 1992b). Further, at the RNA level *Drg1* is reported to be highly expressed at embryonic day 10 with decreasing expression seen through birth and into adulthood (Sazuka et al. 1992a). Unfortunately, the authors of this paper failed to include positive controls in their analysis of gene expression.

Expression of *Drg2* during development is highest in the developing neural tissues. The *Drg2* gene may function in the developing neural tissue by regulating protein trafficking in the Golgi apparatus and the ER. It may directly interact with proteins being moved through the cell, or similar to its relatives *Rab* and *Arf* it may regulate the formation of the vesicles used in the trafficking system. Further biochemical analysis of the protein will help reveal its function in the developing neural tissue and its role at the cellular level. Any role that the gene may play in the SMS phenotype will need to be further evaluated upon resolution of the protein biochemistry, and/or creation of a *Drg2* targeted mouse.

## **MATERIALS AND METHODS**

*Alignment of orthologous DRG2 proteins:* DRG2 protein sequence from *Homo sapien* (I.M.A.G.E. clone 29328 and Genbank accession NP\_001379), *Mus musculus* (I.M.A.G.E. clone 4910172 and Genbank accession NP\_067329), *Drosophila melanogaster* (Genbank accession NP\_536733), and *Schizosaccharomyces pombe* (Genbank accession JT0741) were aligned using ClustalW 1.8 at the Baylor College of Medicine Search Launcher (<http://searchlauncher.bcm.tmc.edu>). Aligned sequence was graphically transformed using Boxshade 3.21 ([http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html))

*Northern Analysis:* A mouse embryonic northern blot was obtained commercially (SeeGene USA, Rockville, MD). Northern analysis was performed as described in Chapter 2. Human  $\beta$ -actin DNA (Clontech Inc.) cross reacts with all mammalian  $\beta$ -actins and was used as a control probe on the mouse blot. Northern analysis of  $\beta$ -Actin was also carried out as described in Chapter 2, except that  $5 \times 10^5$  CPM/ml was used, and the hybridization temperature was reduced to 42°C.

*Cloning of Drg2 into a GFP vector:* The Gateway cloning system from Invitrogen was employed to clone the *Drg2* EST sequence into a vector containing the GFP sequence. The Gateway system uses site-specific recombination properties of bacteriophage lambda to shuttle cloned sequences between various vectors. The entire

coding sequence for mouse *Drg2* was amplified from I.M.A.G.E clone 4910172 using primer pair SHE242/243. Primer SHE242 anneals to the 5' end of the *Drg2* coding sequence and contains a **Kozak consensus sequence** upstream of the ATG start site for the *Drg2* gene allowing for proper gene expression upon transfection (SHE242: 5'-**CACCATG**GGGATCTTGGACA-3'). SHE243 anneals to the 5' end of the *Drg2* coding sequence and was designed without the *Drg2* stop codon to allow for proper in-frame read-through transcription to the GFP gene (SHE243: 5'-CTTCACAATCTGCATGA). PCR product from SHE242/SHE343 amplification was cloned into the pENTR vector from Invitrogen per manufacturer instructions.

Using the Gateway recombination technology, the *Drg2* sequence was moved from the entry vector (pENTR) to a vector containing C-terminus GFP expression (pDEST47) per manufacturer instructions (<http://www.invitrogen.com>). All cloning reactions were confirmed via direct sequencing of the product. Plasmid DNA was isolated using the protocols described in Chapter 2.

*Calcium-phosphate transfection of COS-7 cells:* COS-7 cells were grown to confluence in T-25 flasks for *Drg2* expression or on sterile cover glass in 6 well plates for GFP-*Drg2* fusion protein expression. Cells were grown to ~60% confluence in commercially available RPMI media (Invitrogen) supplemented with 10% fetal bovine serum, and 1x antibiotic/antimycotic solution (Invitrogen). Media on cells was changed 2 hours prior to transfection. DNA to be transfected was isolated and diluted to a total volume of 500  $\mu$ l in ddH<sub>2</sub>O. To the DNA 50  $\mu$ l of 2.5 M CaCl<sub>2</sub> was added. DNA solution was briefly vortexed. Five hundred microliters of



2x HeBS (2x HeBS is 50 mM HEPES [pH 7.1], 280 mM NaCl, 1.5 mM Na<sub>2</sub>HPO<sub>4</sub>) was placed in a 15 ml conical tube. Air was bubbled through the HeBS solution with a sterile 5 ml pipette connected to a battery operated pipette aid. While bubbling HeBS, the DNA solution was added dropwise. The 15 ml tubes were vortexed briefly, and incubated at room temperature for 15 minutes. After 20 minutes the DNA solution was added to the cells. Cells were incubated for 7 hours at 37°C under 5% CO<sub>2</sub>. After incubation, media was replaced and cells were allowed to grow for 2 days.

After 2 days, the coverglass with cells transfected with GFP-Drg2 plasmid DNA was removed and rinsed with sterile 1x Hank's balanced salt solution (HBSS; Invitrogen). Cells not stained with Golgi or ER control stains were mounted onto glass slides with Vectashield with DAPI (Vector Labs) and viewed on a fluorescent AxioPlan2 microscope (Zeiss). Fluorescing cells were photographed using an AxioCam and AxioVision software version 4 (Zeiss). Cells grown to be stained prior to viewing were rinsed with HBSS and a mixture of 1 nM ER tracker blue white-solution (Invitrogen), 5 µM BODIPY TR ceramide (Invitrogen) in RPMI media or HBSS was added. Staining mixture was allowed to incubate for 30 minutes before cells were rinsed with HBSS and mounted to glass slides with Vectashield without DAPI (Vector Labs). Cells were viewed and photographed as above.

Cells transfected for Drg2 expression were removed from the T-25 flasks using 2 ml 1x Trypsin-EDTA (Invitrogen). Cells were collected in 15 ml tubes via centrifugation, rinsed in 1x HBSS, and pelleted via centrifugation. Cell pellets were lysed in 200 µl 1x Laemmli buffer (2x Laemmli buffer is 4% SDS, 120 mM Tris pH

6.8), 1x protease inhibitor (Boehringer) for lysis. Lysate was vortexed briefly and run through a 20-gauge needle in order to shear the DNA. Lysate was stored at -20°C until use.

*Western blotting:* Protein samples were prepared for western blotting in 1x loading dye (Invitrogen) and 1x reducing agent (Invitrogen) to a total volume of 25 µl. Samples were heated to 65°C for 10 minutes prior to separation on polyacrylamide gels. Invitrogen's western blotting system includes precast 5-12% Bis-Tris acrylamide gel (Invitrogen) run and transferred in an XCell western blotting apparatus. Protein gels were run and transferred to PVDF membrane per manufacturer instructions.

Western analysis was performed using antibodies in a 1:1000 dilution. A chromogenic WesternBreeze kit from Invitrogen was used for western analysis following manufacturer instructions (<http://www.invitrogen.com>).

*Immunohistochemistry:* IHC was performed using a polyclonal antibody designed and created by Alpha Diagnostics. Timed matings were performed using C57Bl/6J (The Jackson Laboratory). Embryos were dissected and fixed in 4% paraformaldehyde overnight. Sectioning in paraffin blocks was performed at the Michigan State University Histology Lab. Sections were de-waxed in xylene, and re-hydrated through an ethanol series before use. The sections were exposed to antigen unmasking by heating to boiling in a microwave, boiling continued for 5 minutes in antigen unmasking solution (Vector Labs). Blocking was performed overnight at

room temperature in a humid chamber. Blocking reagent was 1x BM blocking reagent (Roche), 5% goat serum, 0.2% Tween 20, 1x PBS. Sera were diluted 1:1000 in blocking solution and incubated to sections overnight at room temperature. Sections were washed and stained using the Vectastain Elite ABC Kit and DAB/Ni detection (Vector Labs) per manufacturer instructions.

## **Chapter VI**

### **Target of myb-1 (chicken) like 2 (*TOM1L2*)**

#### **BACKGROUND**

In addition to studies involving the *DRG2* gene, I also embarked on characterization of the novel EST stSG9692. The stSG9692 EST was previously mapped to the SMS critical interval (Elsea et al. 1997), though no additional information was available upon beginning the study. Through sequence analysis we determined that stSG9692 represents the 3' UTR of a gene known as target of myb-1 (chicken) like 2 (*TOM1L2*). Initial characterization using northern analysis revealed low-level expression in all adult human tissues studied, including the brain. At this point in the study, we had identified dominant frame shift mutations in the *RAI1* gene in persons with SMS. Study of the *TOM1L2* gene was given less priority in favor of studies of the *RAI1* gene.

#### **RATIONALE**

Bioinformatic and northern analysis of *TOM1L2* was performed in our initial characterization of all the known genes and ESTs mapping to the SMS critical region. Upon finding mutations in SMS patients in the *RAI1* gene research of *TOM1L2* was given less priority. Study of *TOM1L2/Tom1l2* resumed upon discovery of a mouse embryonic stem cell line freely available to academic researchers carrying a gene-trapping vector inserted within the *Tom1l2* gene. The gene trapping method disrupts genes by random insertion in the genome. The ES cells were obtained and mice

carrying the targeted *Tom1l2* were generated. In addition to the initial characterization of the *TOM1L2* gene, we also analyzed the *Tom1l2* mice in order to determine if any of the SMS characteristics could be recapitulated in the gene-trapped mice.

## **RESULTS**

### **Bioinformatic analysis of *TOM1L2***

Initial database searching of the small amount of sequence available for EST stSG9692 revealed only that it was represented by IMAGE clone 23808. This clone was commercially obtained and DNA was isolated and sequenced. When compared to nucleotide databases at NCBI, the sequence only aligned with the small amount of sequence available for itself and the now completely sequenced PAC RP1-253P07. Six frame translation of the clone sequence did not reveal an open reading frame.

As newer gene prediction software became available we were able to match the cloned EST sequences to predicted genes reported online (<http://www.genome.ucsc.edu>). Searching the Gene Scan program at the Golden Path website (<http://genome.ucsc.edu>), we found that clone 23808 was predicted to lie within the 3' untranslated region of a predicted gene with accession ctg17005.291. This predicted gene spans 53.1kb and is made up of 15 exons. BLAST analysis at NCBI showed sequence homology to the gene target of myb1 (chicken) homolog (*TOM1*). *TOM1* was originally identified in chicken cells as a gene target of the proto-oncogene v-myb when chicken myelomonocytic cells were transformed with avian myeloblastosis virus or avian leukemia virus E26 (Burk et al. 1997). The

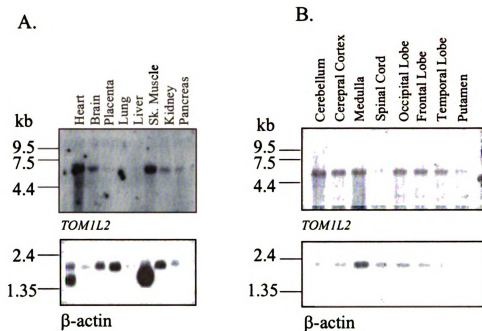
human *TOM1* gene was mapped to chromosome 22q13.1, expression analysis showed a 2.3 kb transcript in all adult tissues. *In situ* hybridization was performed on mouse embryo and demonstrated ubiquitous expression (Seroussi et al. 1999).

We wanted to confirm the predicted *Tom1l2* sequence *in vivo* in preparation for further experiments on mouse *Tom1l2*. RT-PCR using RNA isolated from mouse tail was used to amplify the *Tom1l2* cDNA. The resulting cDNA was isolated and sequenced. The *Tom1l2* gene in mouse contains 15 exons and spans 122 kb (Figure 22). Six frame translation of the cDNA sequence revealed an open reading frame resulting in a 507 amino acid product.

Sequence analysis of the *TOM1L2/Tom1l2* sequences using BLAST programs at NCBI show two conserved domains within the sequence. There is a VHS domain present, which is seen in proteins involved in vesicular trafficking and endocytosis (Lohi and Lehto 1998). Also present within the predicted sequence is a GAT domain thought to be responsible for stabilizing membrane-bound ARF1 in the GTP state (Takatsu et al. 2001).

### **Northern Analysis of *TOM1L2***

Expression analysis was carried out using commercially available northern blots from Clontech (Figure 21). Hybridization was carried out on the multiple tissue and brain 2 northern blots. The expression pattern on both northern blots revealed that the transcript of this EST is ~6.5 kb. Expression pattern in the adult tissues (Figure 19A) shows approximately equal expression levels in heart, brain, skeletal muscle and kidney. Lower levels of expression are seen in placenta, lung, liver, and pancreas.



**Figure 21. *TOMIL2* northern analysis.**

The ~ 2 kb insert from clone 23808, representing the 3' end of human *TOMIL2*, was hybridized to Clontech northern blots. Panel A) shows hybridization signal of ~6.5 kb to the Clontech multiple tissue northern blot. Panel B) shows hybridization of ~6.5 kb to the Clontech Brain Two northern blot.  $\beta$ -actin blots are shown as controls for both blots.

Expression in the brain (Figure 21B) is fairly equal in the cerebellum, cerebral cortex, medulla, occipital pole, frontal lobe, and temporal lobe. There is slightly lower expression in the putamen and a very low level signal is seen in the spinal cord.

### ***TOM1L2* gene trapped ES cell lines**

Gene trapping is a method of disrupting gene function that is rapid and fairly inexpensive when compared to traditional gene knockout methods. The gene-trap relies on a vector that randomly inserts itself into the genome. The vector is equipped with a reporter gene flanked by a splice acceptor and polyadenylation sites. If the vector lands within a gene it disrupts the mRNA by creating a fusion RNA of the endogenous gene and sequence from the vector which codes for reporter and selector genes. An antibiotic resistance gene within the gene vector allows for positive selection of cells in which the vector has properly inserted. The insertion site of the gene-trap vector is determined using 5' RACE experiments. Amplification from vector sequence in the fusion mRNA into the endogenous RNA allows for determination of the insertion site. The gene-trap method has been verified and used successfully to knock-out gene expression in a multitude of genes (Skarnes et al. 1992; Mitchell et al. 2001).

Access to databases of gene-trapped ES cell lines are available for searching on the Internet. The cells are freely available to academic researchers from gene-trapping centers around the world. Using BLAST analysis for SMS candidate genes we determined that a cell line was available with a gene-trap in *Tom1l2* from the BayGenomics consortium (<http://www.baygenomics.ucsf.edu>). Bioinformatic



analysis of the 5' RACE performed from the XG909 cell line indicated that the gene-trapping vector inserted between exons 10 and 11 in the mouse *Tom1l2* gene (Figure 22A and 22B). The gene product resulting for this gene-trapped allele in the *Tom1l2* gene would be missing 147 amino acids from the carboxy end.

### **XG909 genotyping and creation of XG909 mice**

Genotyping of the XG909 cell line (and mice subsequently developed from the cell line) was performed by Southern analysis. Digestion of mouse genomic DNA was performed with *Bgl*I. When Southern blots are probed with a PCR product from the intronic sequence between *Tom1l2* exons the wild type allele is indicated by a band of 2812 base pair (Figures 22B and 22C). The insertion of the gene-trap vector between exons 10 and 11 shifts the native 2812 base pair band up to an ~3800 base pair band (Figures 22B and 22C). Wildtype animals only have the ~2800 base pair band. Heterozygous animals display both the ~2800 base pair and ~3800 base pair bands. Homozygous targeted animals show only the 3800 base pair band.

The XG909 cell line was obtained and sent to the Transgenic Animal Model Core (TAMC) at the University of Michigan for creation of gene-trapped founder mice. The XG909 cell line was expanded and then injected into C57Bl/6J for creation of a chimeric (founder) line of mice. The XG909 cell line was created from 129/OlaHsd embryonic stem cells. These mice have white coat color, while the C57Bl/6J mice have black coats. Chimeric animals are a mix of both cell lines and display a mottled coat with patches of black or white fur (Figure 22D). Male chimera animals were mated with C57Bl/6J females to produce an F1 line. Chimeras whose

germ cells were of the 129/OlaHsd lineage produced agouti colored offspring when mated with C57Bl/6J animals. Agouti color results from crossing of white and black coated animals. One-half of the agouti animals were expected to carry the gene-trapped *Tom112* allele. The F1 agouti animals were genotyped, and brother and sister carriers were mated in order to obtain an F2 generation for phenotypic analysis.

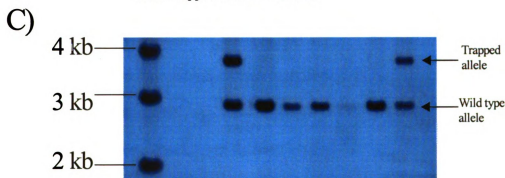
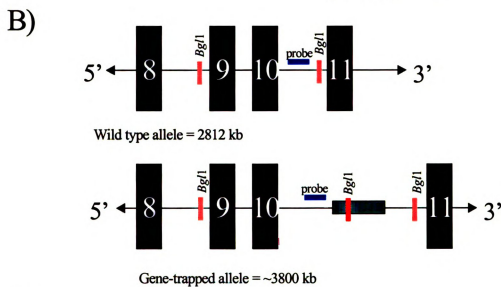
#### **Analysis of the F2 generation of *Tom112* gene-trapped mice**

Brother sister mating of carrier F1 gene-targeted *Tom112* mice resulted in 15 F2 offspring. The gene-target occurs at an autosomal locus and should follow simple Mendelian inheritance;  $\frac{1}{4}$  should be XG909<sup>+/+</sup>,  $\frac{1}{2}$  should be XG909<sup>+/-</sup>, and  $\frac{1}{4}$  should be XG909<sup>-/-</sup>. Of the 15 F2 mice 3 XG909<sup>+/+</sup>, 3 XG909<sup>+/-</sup>, and 9 XG909<sup>-/-</sup> were obtained, which is not in a Mendelian fashion. Increasing the number of offspring may bring the inheritance back to agreement with Mendelian inheritance.

The F2 offspring were physically assessed using the guide in Appendix B. Mice were thoroughly evaluated for physical, neurological, and behavioral anomalies at age 5 and 13 weeks. All physical measurements were within normal range. Total body length, body weight, and cage hang time measurements were statistically analyzed as described below.

**Figure 22. Genetic analysis of the XG909 cell line.**

**A)** A schematic representation of the *Tom1l2* gene with location of the pGT1lxf gene-targeting vector. The *Tom1l2* gene is composed of 15 exons. The gene-trapping vector inserted into the intron between exons 10 and 11. **B)** A schematic of the wild type and targeted alleles in the XG909 cell line. The insert causes a *Bgl*I fragment to increase from 2812 bp to ~3800 bp. The *Bgl*I sites are indicated in red, the probe used in Southern analysis is indicated in blue, the pGT1lxf vector is indicated in green. **C)** Southern analysis of F1 mice generated from the XG909 cell line. Mice heterozygous for the trapped allele are seen in the first and last lanes as indicated by the presence of both the 2812 bp and ~3800 bp bands. **D)** Photograph of the chimeric animals generated from the XG909 cell line.



The data obtained from the mice studies were grouped based on the age at evaluation and their genotype. The means and standard deviations of the groups were compared using the student t-test. Also, comparisons within the groups were performed by the analysis of variance (ANOVA). Comparisons of all the three groups at 5 weeks and 13 weeks indicated statistically significant difference between homozygous trapped and wild-type littermates only in length at 13 weeks.

Total body length was measured from the tip of the nose to the tip of the tail while the mice were under light anesthetic. At 5 weeks, there is no statistically significant difference in body length between groups (variance of 0.78 and mean comparisons showed a student t test value of 2.7 ( $p=0.08$ ))(Figure 21A). A significant difference was seen while comparing the lengths of 13 week-old mice.

A significant p-value of 0.02, at  $Df=3$  was obtained when comparing the means and standard deviation (student t-test value of 2.3). A significant difference was seen between the means from the homozygous trapped and wild type ( $p=0.03$ ) and between normal and the heterozygous mice ( $p=0.01$ ), though no difference can be deduced between the transgenic and the heterozygous mice (Figure 23B). This cannot be a conclusive evidence of the difference as power calculations were not performed for adequacy of the sample size for evaluation.

Weight assessments performed at 5 weeks and 13 weeks showed no significant differences between the genotypes ( $p=0.48$  and  $0.315$  and variance of  $0.28$  and  $0.24$  respectively)(Figures 23C and 23D). Comparison of mean values of each group by student t-test, gave a value of 2.36.

Analysis of the hang time of the mice to the cage tops did not show statistical

significance ( $p=0.55$  at 5 weeks and  $0.0505$  at 13 weeks). The ANOVA showed an  $R^2$  value of  $0.24$  and  $0.56$ . (Figures 23E and 23F)

## **SUMMARY**

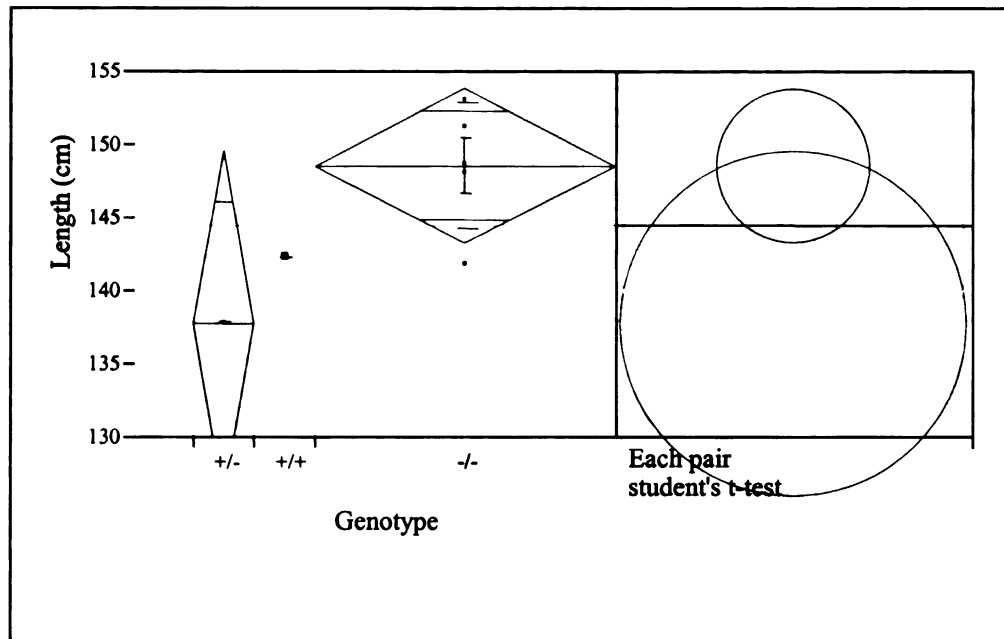
The *TOM1L2* gene was first mapped as an EST during definition of the SMS critical region (Elsea et al. 1997). Using bioinformatic analysis and RT-PCR we were able to identify the full length sequence of the human and mouse *TOM1L2/Tom1L2* genes. Northern analysis revealed expression in all human adult tissues studies. Work on the *TOM1L2* gene was placed on hold upon discovery of mutations in the *RAI1* gene in some SMS patients.

Mouse analysis of the *Tom1l2* gene began when we discovered that embryonic stem cells with a gene-trapping vector in the *Tom1l2* gene were available at no cost. We created mice from these embryonic stem cells in order to try to identify any phenotype due to dosage changes in *Tom1l2* expression. Initial evaluation of the F2 generation of gene-trapped only showed statistical differences in total body length at 13 weeks of age between wild-type and homozygous trapped mice. Though, due to the small number of animals analyzed the statistical analysis lacks the power needed to be reliable. In order to determine whether the gene trap is having an effect on phenotype additional F2 mice are currently being generated.

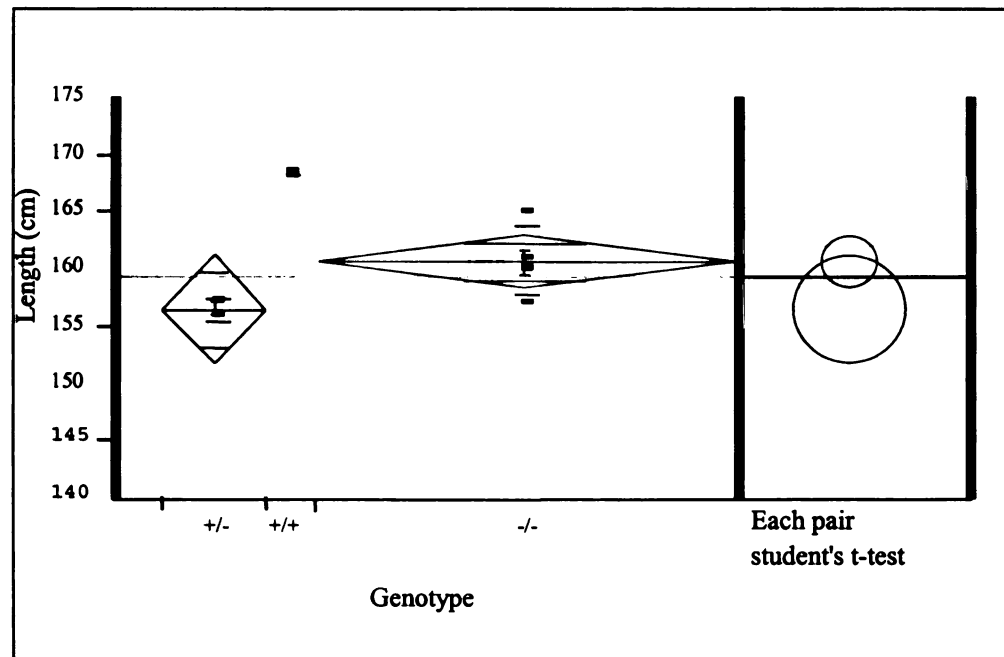
**Figure 23. Statistical analysis of XG909 mouse measurements.**

Mice were measured at 5 and 13 weeks for length and weight. Hang time from the cage top was also assessed. Data was analyzed using the student t-test and ANOVA with JMP software. Student t-test results are indicated by overlapping circles. ANOVA analysis is represented by the diamonds. Measurements for heterozygous mice are colored green, wild type measurements are colored yellow, and homozygous targeted are colored red. **A)** Comparison of length at 5 weeks. **B)** Comparison of length at 13 weeks. **C)** Comparison of weight at 5 weeks. **D)** Comparison of length at 13 weeks. **E)** Comparison of cage top hang time at 5 weeks. **F)** Comparison of cage top hang time at 13 weeks.

A)



B)



Group 1	Group 2	Difference	Lower CL	Upper CL	p-Value	Difference
+/	+/-	11.80000	3.6426	19.95736	0.0102961	
+/	-/-	7.66250	0.5980	14.72698	0.0368724	
-/-	+/-	4.13750	-1.1281	9.40305	0.1075639	

**Figure 23**



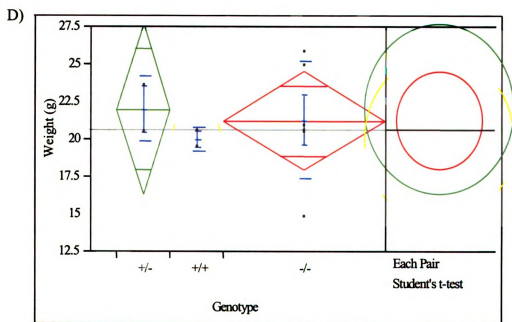
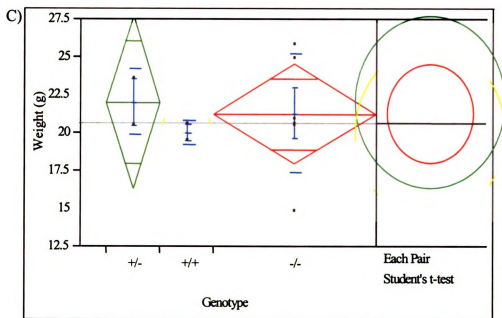
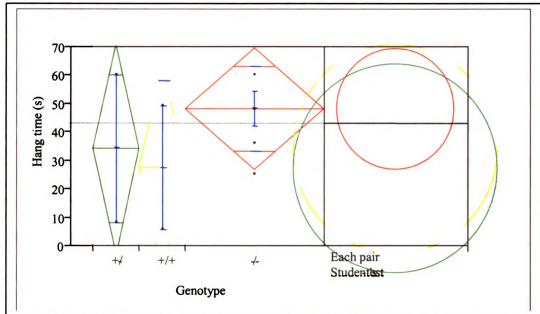


Figure 23 continued

E)



F)

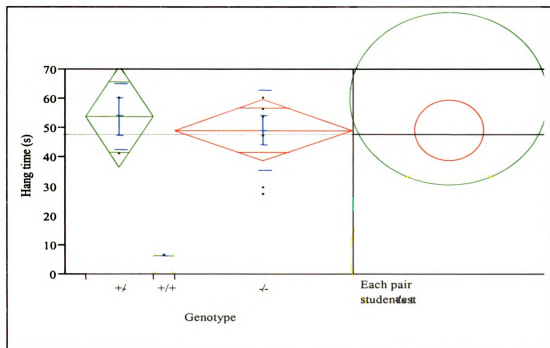


Figure 23 continued

## **MATERIALS AND METHODS**

*Northern analysis:* Northern analysis was performed as described in Chapter 2.

*Mouse DNA isolation from tail biopsies:* Tail biopsies were obtained from 21 day old mice after weaning. Biopsies were digested/lysed over night in 600 µl 1x TNES (10 mM Tris, pH 7.5, 400 mM NaCl, 10 mM EDTA, and 0.6% SDS), and 35 µl of 10 mg/ml overnight at 55°C. After digestion/lysis DNA was isolated using P:C:I as described in Chapter 2.

*Southern analysis:* Eight micrograms of mouse DNA was digested with 3-4 U *Bgl*I overnight at 37°C. Southern blots were created from digested DNA as described in Chapter 2. Probe for southern analysis was generated by PCR amplification of the intron between *Tom1l2* exons 10 and 11 was performed as described in chapter 2. Southern hybridization, washes, and detection were performed exactly as described in Chapter 2.

*Mouse assessment:* Assessments of the F2 mouse generation was performed using the guide attached as Appendix C. All protocols for work with mice were approved by Animal Use Committees at MSU and VCU.

*Statistical analysis:* Analyses were performed using the JMP software, version 5.1 with the significance value set at 0.05.

## Chapter VII

### Discussion

Through the research presented here significant progress has been made in the understanding of the genes that play a role in Smith-Magenis syndrome. Our efforts have resulted in construction of a contig of large insert clones across the 1.5 Mb SMS critical interval that included a transcription map of the region (Lucas et al. 2001). Analysis of SMS patients with 17p11.2 deletions allowed for two refinements of the smallest deletion region of overlap in SMS (Vlangos et al. 2003; Vlangos et al. 2004). The SMS critical region has been reduced from 1.5 Mb to ~700 kb. Correspondingly, the number of genes mapping within this interval has also been reduced from ~40 to 11 (Vlangos et al. 2004).

Research on samples from patients who display the SMS phenotype but in whom a deletion could not be detected with extensive FISH analysis yielded the most insight into the genetic basis of SMS. Systematic sequencing of genes mapping within the SMS critical region led to discovery of dominant frameshift mutations in the *RAI1* gene. Nine total SMS patients with *RAI1* mutations have been reported since our initial findings in three patients (Slager et al. 2003; Bi et al. 2004; Girirajan et al. 2005).

Research on the expression pattern of the *DRG2* gene revealed dynamic expression in the developing mouse nervous system. Though *DRG2* is no longer thought to be developmentally regulated, it likely does play an important role in the proper development of the nervous system. Though its role in SMS is still

inconclusive, the results of the IHC experiments presented here make it tempting to hypothesize that haploinsufficiency of *DRG2* may play some role in infantile hypotonia seen in SMS patients with a deletion but not seen in patients with mutations in *RAI1*.

### **Future studies in the Elsea Laboratory**

My work in the Elsea Laboratory has helped develop an environment full of opportunity with regard to research on SMS and the genes mapping to chromosome 17p11.2. Additional research on *TOM1L2*, *DRG2*, and *RAI1* will yield additional insight into human genetic disease. In order to further understand the function of these genes, and any possible role they may play in human disease, the following research is proposed.

### ***TOM1L2/Tom1l2***

Mice with gene trapped *Tom1l2* alleles did not show an overt phenotype in either the heterozygous or homozygous state in the C57Bl/6J background. Many questions remain with respect to the gene-targeted *Tom1l2* allele. Although we were able to determine the genotype of the *Tom1l2* mice, we did not yet study whether expression levels of the *Tom1l2* gene were affected by the gene-trapping. Because the ES cells cannot survive the positive antibiotic selection in propagation of the cell line it is unlikely, but not impossible, that the gene-trapping vector inserted incorrectly. PCR analysis and direct sequencing of the vector insertion location could be used to measure the accuracy and location of the insert site. PCR primers have

been designed throughout the vector and flanking genomic sequence to amplify across the insertion breakpoint. The PCR reactions to accomplish this task need to be optimized.

Unlike traditional gene targeting via “knocking-out,” gene-trapping does not guarantee a null allele. Genes that have been gene-targeted can be leaky and result in a hypomorphic allele instead of a true null allele. Though no publications are reported regarding the efficiency of null alleles in gene-trapping, the groups promoting gene-trapping report that ~10% of gene-trap lines are leaky (<http://genetrap.gsf.de>). One report characterizing two gene trap lines that were leaky report that the actual percentage of leaky lines may be upwards of 50% (Voss et al. 1998). To determine whether a null or hypomorphic allele has been generated measurement of gene expression in *Tom112* gene-trapped mice should be investigated.

*Tom112* expression could be analyzed from the normal and trapped alleles at both the RNA and protein levels. Northern analysis of RNA isolated from the *Tom112* mouse line probed with 5' sequence from the *Tom112* gene should reveal the normal ~6.5 kb message but should also show a shifted band of RNA containing the *Tom112*-vector fusion RNA. Some studies using this technique have found that the normal RNA is preferentially expressed over the trapping exons resulting in normal gene expression in the homozygous trapped mice (McClive et al. 1998; Voss et al. 1998). Further complicating the analysis is that the varying expression can be tissue specific (McClive et al. 1998).

Additional analysis could be performed at the protein level. We have created

a polyclonal antibody near the amino terminal of TOM1L2/Tom112. Protein expression of the trapped allele could be analyzed via western analysis in order to determine whether protein levels are reduced in the gene-trapped mice. The recent reports of leaky gene trapping in mice and previous reports of gene trapping in *Drosophila* suggest that the method may not reduce gene expression from trapped alleles as much as predicted.

Even if the *Tom112* gene-trapped mice are hypomorphic they may be very useful in studying *Tom112* gene expression in mice. In addition to positive selection by antibiotic resistance, the gene-trapping vector allows for expression analysis by expression of the *lacZ* reporter gene. Gene expression can be studied by staining mouse sections or whole mount mice for *LacZ* expression. Expression analysis could be performed both in developing mouse embryos and adult tissues. *LacZ* expression could be verified by RNA ISH and/or IHC to the *Tom112* gene.

Conserved domains within the *Tom112* gene predict that it plays a role in trafficking of proteins from the trans-Golgi network to the lysosome. I have cloned the full length *Tom112* sequence into the Gateway system entry vector in proper orientation for use in GFP studies. GFP studies would provide insight into the cellular localization of the Tom112 protein. Human genetic disease related to trafficking is not rare, and includes mannosidosis, Neiman-Pick disease, and Zellweger syndrome. Though the role *TOM1L2* plays in human genetic disease remains unknown, it is a good candidate gene that should be explored further.

## ***DRG2/Drg2***

From the data presented we hypothesize that the *DRG2/Drg2* genes likely play an important role in mammalian development. The highly conserved evolutionary sequence and the expression patterns seen in the developing mouse make *DRG2/Drg2* an interesting gene for further research as the role that *DRG2* may play in SMS or other genetic disorders is still unknown.

With the discovery of novel genes comes the job of characterizing function, regulation, and possible interactions with other proteins. In order to determine the biochemical function of DRG2 the lab plans to use techniques to pursue any protein-protein interactions involving DRG2. Proposed experiments involve the use of yeast two hybrid assays to study possible protein-protein interactions. The yeast two hybrid assay was first reported in 1989 as a screen for binary protein interactions (Fields and Song 1989). The assay involves cloning your protein of interest into a plasmid with a DNA binding domain such as GAL4. The GAL4 fusion protein is the “bait” that allows for screening for interactions with “prey” proteins. Prey proteins are created from cDNA libraries where the clones are created fused to a transcriptional activating domain. Fusion of the bait and prey allow for expression of a reporter gene such as  $\beta$ -galactosidase, or for genes allowing for growth in selective media. Positive yeast colonies are screened to determine which cDNAs interact with the bait protein. The yeast two hybrid assay is a powerful tool for determination of protein-protein interactions. The technique has been very successful in identification of proteins in cell signaling pathways (Fashena et al. 2000).

The yeast two hybrid is a good assay for finding protein-protein interactions in



novel genes, but it is not without faults. The false positive rate can be very high, though newer techniques are being developed to reduce the number of false hits. Further, many interactions that naturally occur *in vivo* may not occur in this system. The yeast two hybrid system operates in the nucleus in order to activate the reporter gene. Transmembrane proteins, an extremely important class of signaling molecules, cannot be studied effectively with this assay. Interactions with proteins that require post-translational modification such as glycosylation before becoming active might also be missed in the traditional yeast two hybrid assay (Stagljar 2003). With these limitations in mind, the lab also plans to use coimmunoprecipitation techniques to identify interactions that may be missed with the yeast two hybrid assay and to verify any positive interactions identified. We hypothesize that if we are able to determine the protein partners of *DRG2*, we may further understand its role in mammalian development.

The expression patterning in the developing mouse shows that *Drg2* likely plays a role in the developing nervous system. Generation of a *Drg2* knock-out mouse may reveal any phenotype associated with dosage changes in the *Drg2* gene.

### ***RAI1/Rai1***

Our finding of SMS patients with dominant frameshift mutations in the *RAI1* gene unraveled some of the mystery behind the etiology of SMS; however, it also created many more questions that need to be addressed. Though research will continue on other candidate genes in the SMS region, the thrust of the work in the lab will focus on research of the *RAI1* gene.

Mouse modeling to study the effect of dosage changes in *Rail* has begun. With the hypothesis that *Rail* is a dosage sensitive gene, the lab has generated a line of transgenic *Rail* mice with multiple copies of a BAC containing the *Rail* gene. Current work in the lab involves assessing whether increased dosage of the *Rail* gene results in a phenotypic effect. Mice in the line generated carry between 2 and 5 copies of a BAC containing the *Rail* gene. Phenotypic measurements taken thus far have not revealed a difference between normal littermates and those with increased *Rail* copy number (Rebecca Slager, MSU Ph.D. Dissertation, 2004). Though increased copy number at the DNA level has been confirmed, it is unknown if the level of *Rail* expression is also increased. Experiments to determine expression levels at the RNA level will likely begin soon.

The transgenic mice were created in conjunction with an *Rail* targeted knockout project. At this time, *Rail* has been successfully knocked-out in embryonic stem cells. The targeted cells are in queue for injection into blastocysts for creation of chimeric animals. Thorough physical and behavioral analysis of the *Rail* knockout animals is planned.

As discussed in Chapter 1, mice deficient for the portion of mouse chromosome 11 syntenic to human chromosome 17p11.2 show a distinct phenotype overlapping with some of the features seen in the SMS phenotype (Walz et al. 2003; Walz et al. 2004). Recently, it has been shown that inactivation of the *Rail* gene recapitulates the physical features seen in the deletion mice (Bi et al. 2005). Behavioral studies of the *Rail* targeted mice were not reported. Both mouse studies reported were performed in a mixed C57BL/6J x 129SvEv background (Walz et al.

2003; Walz et al. 2004; Bi et al. 2005). The *Rail* knockout mouse that the Elsea lab is creating will be in a full C57BL/6J background. There are many documented reports showing that phenotypic differences in knockout animals can vary greatly in different background strains (Dobkin et al. 2000; Fleming et al. 2001; Humphries et al. 2001). It is preferential to analyze behavior and circadian rhythm of mutant mice in a full C57BL/6J background. Differences in phenotype may be found between the knockout mice produced in different labs.

Not much is known about the biochemistry of *RAI1/Rail*. By sequence analysis it is postulated to act in transcriptional regulation (Seranski et al. 2001; Girirajan et al. 2005). Studies utilizing GFP show that *Rail* is localized to the nucleus (Elsea lab unpublished data and Bi et al. 2005). Initial studies of *Rail* in yeast show that there are two functional transcriptional transactivation domains present near the N-terminal end of the protein (Bi et al. 2005). It is unknown if the protein interacts directly with DNA or acts through association with other transcription factors.

Vitamin A (retinol) and its derivatives are extremely important in embryonic development (Ross et al. 2000). Retinoic acid (RA) is the product of oxidation of vitamin A. *Rail* was originally cloned because of increased expression in response to retinoic acid (Imai et al. 1995). However, the correlation between RA, *RAI1*, and development is unknown. It has been shown that RA plays a direct role in the proper regulation of the *Hox* genes (Langston and Gudas 1994). Correct expression of the ~38 known *Hox* genes is crucial for proper development (Langston and Gudas 1994; Ross et al. 2000). The lab may choose to investigate whether *RAI1* acts in the

signaling pathway in response to RA resulting in *Hox* gene regulation. This could be performed by testing to see if RAI1 binds to the promoters of any of the large number of *Hox* genes. The spatial and temporal expression of most of the *Hox* genes during development is known (Langston and Gudas 1994; Ross et al. 2000).

In addition to identification of the potential DNA binding function of RAI1, determination of any protein-protein interactions would help determine potential partners in transcriptional activation and help determine the transcriptional targets of RAI1/Rai1. The yeast two hybrid assay or coimmunoprecipitation are two of the best techniques to study protein interactions. These are techniques that would also be of use in studying other novel genes in the lab. Determining which proteins act with RAI1/Rai1 might lead to investigation of a potential second locus for SMS.

The lab has a cohort of samples from patients without detectable 17p11.2 deletions in whom mutations in *RAI1* have not been located. It is possible that dosage changes in genes that interact with RAI1 could phenocopy the characteristics seen with haploinsufficiency of RAI1. Heterogeneity in Rubinstein-Taybi syndrome (RSTS) has recently been reported (Roelfsema et al. 2005). RSTS is characterized by mental retardation, craniofacial and skeletal anomalies, EEG abnormalities and seizures, and an increased incidence of tumor formation (Petrij et al. 1995). Rubinstein-Taybi syndrome was originally reported associated with a hemizygous deletion of chromosome 16p13.3. Research then revealed that heterozygous mutations in the *CREBBP* gene mapping to 16p13.3 also caused Rubinstein-Taybi (Petrij et al. 1995). It is interesting to note that the *CREBBP* and *RAI1* genes are likely transcriptional co-activators that both contain a PHD domain (Kalkhoven et al.

2003). The PHD domain in *CREBBP* has been shown to be extremely important in acetyltransferase activity towards histones and CBP. This reduction in acetyltransferase activity led directly to reduction in coactivator function for the transcription factor CREB (Kalkhoven et al. 2003). The authors of this study postulate that disruption of the PHD domain in *CREBBP* alone may be enough to cause RSTS (Kalkhoven et al. 2003).

Research of the *CREBBP* gene found that it sometimes interacts with the *EP300* gene as a transcriptional coactivator in the regulation of gene expression through various signal transduction pathways (Weaver et al. 1998; Roelfsema et al. 2005). Some patients who were clinically diagnosed with RSTS do not have detectable deletions of chromosome 16p13.3 or mutations in *CREBBP*. Mutations in the *EP300* gene were found in samples from this cohort of patients by direct DNA sequencing (Roelfsema et al. 2005). It is possible that SMS may display heterogeneity similar to that seen in RSTS. Research on our cohort of SMS patients without 17p11.2 deletion or *RAI1* mutation may reveal a second locus for SMS.

## **Conclusion**

The research presented here sets the groundwork for continuing study of the genes mapping to chromosome 17p11.2. We now know that the *RAI1* gene is dosage sensitive and that haploinsufficiency of *RAI1* leads to most of the characteristics in the SMS phenotype. Currently, we can only speculate on the many developmental pathways that *RAI1* could potentially play a role. Further research using *in vivo* and *in vitro* techniques, as well as physical and behavioral analysis of *Rail* mouse models, will help determine the true biochemical role of *RAI1*.

Though *RAI1* is the cause of the majority of the characteristics seen in the SMS phenotype we cannot yet say that it is the only gene involved. Patients with mutations in *RAI1* do not display short stature, infantile hypotonia, or visceral anomalies seen in some SMS patients with a 17p11.2 deletion. Further research into other genes including *DRG2* and *TOM1L2* may show that these genes are responsible for these characteristics.

Clinically, we have improved the diagnostic testing for SMS by identification of a more efficient FISH probe and by helping implement commercially available *RAI1* mutation screening. With improved testing and knowledge of SMS, patients can begin to receive intervention through proper therapies at an earlier age. Early intervention greatly increases the quality of life for SMS patients and their families. Further work using the research presented here will no doubt answer even more questions with regard to SMS and proper mammalian development.

## Appendix A

### Features of Smith-Magenis Syndrome A Questionnaire for Parents and Guardians

This questionnaire is designed for parents or guardians of children diagnosed with Smith-Magenis syndrome (SMS). The information provided in this survey will be used to better understand the correlation of the syndrome phenotype (the “symptoms”) and the genotype (the DNA content). We will use the information collected to try to determine the gene or genes that play a role in SMS. Participation in this survey is completely voluntary. All answers will be kept strictly confidential. Survey participants will not gain any direct benefit for participating in the survey; however, this information will be very important for understanding the genes involved in SMS.

If at anytime you have additional questions during the course of this study about the research or your rights as a research subject, you may address them to The Michigan State University Committee on Research Involving Human Subjects at (517) 355-2180. In the event that any problem or question arises you may contact Dr. Sarah Elsea at (517) 355-5597, or email at [elsea@msu.edu](mailto:elsea@msu.edu).

I agree to voluntarily participate in this survey titled “Features of Smith-Magenis syndrome: a questionnaire for parents and guardians.” Participants may refuse to participate or discontinue involvement AT ANY TIME without penalty.

Signature \_\_\_\_\_ Date \_\_\_\_\_

Please Print Name \_\_\_\_\_

This Consent form is not valid without The Institutional Review Board stamp of certification.

\_\_\_\_\_

For Office Use Only:

Survey #	
Family ID	
SMS ID	
R/I Req?	
R/I Sent?	
Date Survey Sent	
Date Survey Rec'd	

**Personal Information:**

Your name \_\_\_\_\_

Your relationship to person with  
SMS \_\_\_\_\_

Patient's  
name \_\_\_\_\_

Patient's sex \_\_\_\_\_ Patient's date of  
birth \_\_\_\_\_

Mother's  
Name \_\_\_\_\_

Father's  
Name \_\_\_\_\_

Mother's date of birth \_\_\_\_\_ Mother's Ethnic Background \_\_\_\_\_

Father's date of birth \_\_\_\_\_ Father's Ethnic Background \_\_\_\_\_

Address information is optional (see below):

Street  
address \_\_\_\_\_

City \_\_\_\_\_  
State \_\_\_\_\_

Zip/Postal Code \_\_\_\_\_  
Country \_\_\_\_\_

Telephone Number (with area  
code) \_\_\_\_\_

e-mail  
address \_\_\_\_\_

**Pregnancy/Birth/Diagnosis:**



**Was the pregnancy considered normal?**

\_\_\_\_\_  
**(If not, please explain any complications in the space below)**

**Was your child carried to term?** \_\_\_\_\_

**If not, how premature was your child?** \_\_\_\_\_

**What was your child's birth weight?** \_\_\_\_\_

**What was your child's length at birth?** \_\_\_\_\_

**At what age was your child diagnosed with SMS?** \_\_\_\_\_

**How was the diagnosis confirmed (e.g. FISH, karyotype)?** \_\_\_\_\_

**In the space below please explain any other features that pertain to the pregnancy/birth of your child:**

### Physical Features:

If your child displays or has been diagnosed with any of the following features, please circle yes or no. If you can recall when the feature first became apparent or was diagnosed, please indicate in the space to the right. Additional comments may be made in space to the right or at the end of this section.

<i>Feature/Question</i>	<i>Circle</i>	<i>Age of onset/Additional comments</i>
<b>Short stature</b>	<b>Yes</b> <b>No</b>	
<b>Smaller than average midface (hypoplasia)</b>	<b>Yes</b> <b>No</b>	
<b>General shortness of the head (brachycephaly)</b>	<b>Yes</b> <b>No</b>	
<b>Large forehead (frontal bossing)</b>	<b>Yes</b> <b>No</b>	
<b>Broad nasal bridge</b>	<b>Yes</b> <b>No</b>	
<b>Abnormal ear shape</b>	<b>Yes</b> <b>No</b>	
<b>Down-turned upper lip</b>	<b>Yes</b> <b>No</b>	
<b>Projecting of the jaw (prognathism)</b>	<b>Yes</b> <b>No</b>	
<b>Cleft lip</b>	<b>Yes</b> <b>No</b>	<b>If yes, has it been corrected? and when?</b>
<b>Cleft palate</b>	<b>Yes</b> <b>No</b>	<b>If yes, has it been corrected? and when?</b>
<b>Dental abnormalities</b>	<b>Yes</b> <b>No</b>	<b>If yes please explain in space below</b>
<b>Unison of the eyebrow into one large brow (synophrys)</b>	<b>Yes</b> <b>No</b>	
<b>Upslanting eyes</b>	<b>Yes</b> <b>No</b>	
<b>Iris anomalies</b>	<b>Yes</b> <b>No</b>	
<b>Presence of a wandering eye (strabismus)</b> <b>Does the eye wander toward the nose (esotropia)</b> <b>or is wandering away from the nose (exotropia)</b>	<b>Yes</b> <b>No</b> <b>Yes</b> <b>No</b> <b>Yes</b> <b>No</b>	
<b>Micro cornea</b>	<b>Yes</b>	

	No	
<b>Hoarse Voice</b>	Yes No	
<b>Short, broad hands</b>	Yes No	
<b>Decreased use of hands</b>	Yes No	
<b>Fusion of the fingers or toes (syndactyly)</b>	Yes No	
<b>Cold hands/feet</b>	Yes No	
<b>Difficulty finding shoes that fit</b>	Yes No	
<b>Hammer toes</b>	Yes No	
<b>High arched feet (pes cavus)</b>	Yes No	
<b>Flat feet (pes planus)</b>	Yes No	
<b>Curvature of the spine (scoliosis)</b> <b>Has surgery been performed to correct/prevent?</b>	Yes No  Yes No	<b>If yes, at what age?</b>
<b>Skeletal/vertebral changes</b>	Yes No	<b>If yes, please explain below</b>
<b>Elbow limitations (decreased ability to extend arm fully)</b>	Yes No	
<b>Abnormal gait</b>	Yes No	
<b>Dry skin or psoriasis</b>	Yes No	
<b>Muscle weakness</b>	Yes No	
<b>Decreased tolerance to exercise</b>	Yes No	

In the space below please explain any other features that may pertain to the physical appearance of your child.

**Development:**

Please indicate the age at which your child performed the following features.

<i>When did your child first...</i>	<i>Age</i>
<b>Laugh</b>	
<b>Smile</b>	
<b>Roll over</b>	
<b>Sit up without aid</b>	
<b>Crawl</b>	
<b>Pull themselves up to standing</b>	
<b>Stand without aid</b>	
<b>Take first steps</b>	
<b>Begin to walk</b>	
<b>Speak first word</b>	
<b>Speak in complete sentences</b>	

If your child displays or has been diagnosed with any of the following features with regard to development, please circle yes or no. If you can recall when the feature first became apparent or was diagnosed, please indicate in the space to the right. Additional comments may be made in space to the right or at the end of this section.

<i>Feature/Question</i>	<i>Circle</i>	<i>Age of onset or diagnosis/additional comments</i>
Was your child hypotonic ("floppy")	<b>Yes</b> <b>No</b>	
Do you know the developmental age of your child	<b>Yes</b> <b>No</b>	Age?
Has your child had their IQ measured	<b>Yes</b> <b>No</b>	Measurement:  Age of measurement:
Does your child attend school  Is it a special school/program  Does your child have an aide at	<b>Yes</b> <b>No</b> <b>Yes</b> <b>No</b> <b>Yes</b> <b>No</b>	

school		
Does (did) your child use sign language	<b>Yes</b> <b>No</b>	Age used?
Did you notice a change in frustration level when your child began to communicate more effectively	<b>Yes</b> <b>No</b>	

#### Medical History:

If your child displays or has been diagnosed with any of the following features with regard to their medical history, please circle yes or no. If you can recall when the feature first became apparent or was diagnosed, please indicate in the space to the right. Additional comments may be made in space to the right or at the end of this section.

<i>Feature/Question</i>	<i>Circle</i>	<i>Age of onset or diagnosis/additional comments</i>
Vision problems due to myopia (nearsightedness)	<b>Yes</b> <b>No</b>	
Wear glasses	<b>Yes</b> <b>No</b>	
Retinal detachment	<b>Yes</b> <b>No</b>	
Macular degeneration	<b>Yes</b> <b>No</b>	
Hearing impairment	<b>Yes</b> <b>No</b>	Specify:

If yes, has the impairment been corrects	<b>Yes</b> <b>No</b>	How?
Frequent ear infections (Otitis media)	<b>Yes</b> <b>No</b>	# per year
Have ear tubes been inserted	<b>Yes</b> <b>No</b>	
Sinus infection	<b>Yes</b> <b>No</b>	
Peripheral neuropathy	<b>Yes</b> <b>No</b>	
Decreased sensitivity to pain	<b>Yes</b> <b>No</b>	If yes, please explain which parts of body below
Muscle cramps		
Seizures	<b>Yes</b> <b>No</b>	Types:
Is medication used to control?	<b>Yes</b> <b>No</b>	Explain:
Tremors	<b>Yes</b> <b>No</b>	
Congenital heart defects	<b>Yes</b> <b>No</b>	
Breathing problems	<b>Yes</b> <b>No</b>	
Feeding problems during infancy	<b>Yes</b> <b>No</b>	
Swallowing problems	<b>Yes</b> <b>No</b>	
Excessive choking	<b>Yes</b> <b>No</b>	
Excessive drooling	<b>Yes</b> <b>No</b>	
Difficulty digesting foods	<b>Yes</b> <b>No</b>	
Regular constipation problems	<b>Yes</b> <b>No</b>	
Bowel obstruction	<b>Yes</b> <b>No</b>	
Regular loose bowel movements	<b>Yes</b> <b>No</b>	

Frequent urinary tract infections	Yes No	
Kidney abnormality	Yes No	
Diabetes	Yes No	
Obesity	Yes No	
High cholesterol	Yes No	

**Please list any medications your child receives in the space below:**

**Please list other items you believe would be helpful below:**

## Appendix B

### *RAI1* gene amplification primers

<i>RAI1</i> Exon	Forward Primer	Reverse Primer	Annealing Temp
1	RA1:CCTTCCCTCCCTCCCTCCC TTCC	RA2:CACCCCTGCAGGTAGTGG CTG	65°C
1+2	RA3:CGCTATGCTGGTGAGGAG AGCC	RA4:CCGACTGGTAGGCATGAA GATTC	64 °C
2	RA5:CCATGACAGGCCGCTGAC TGC	RA6:CAGGGAGCTTGTCTTCTG AAG	62 °C
*2+3	RA7:CTGACCACAGCCACTTCA TGCC	RA8:CACGGACTCGGGCTTGGC CTTCG	63 °C
3	RA9:CAGCTTCCTCTACTGCAAC CAG	RA10:GCGAAGGCCACGGAAGG GTCTTC	60 °C
3	RA11:GCCCCGACTCCTTGCAGCT GGAC	RA12:CCGGTCAGCCTTGGCCAC CTCGG	65 °C
3	RA13:GGACTTCAAGCAGGAGG AGGTGG	RA14:CAGAGAGGCGTCCGAGG TGGTG	64 °C
*3	RA15:CACATGAAGCAGGTGAA GAGG	RA16:CTGGAGGCAGCCTTGGG TGAG	65 °C
*3	RA17:CGTTCTCTCACGGCCCTG AGTG	RA18:GCCACTGGCGTTGCTGCT GCTGC	68 °C
*3	RA19:GCGCTCAAAAGGAAGTC GGCC	RA20:CCACATTTACCAGGCCTT CTTCC	64 °C
*3	RA21:CCCTTTCCGACAAAGACC GTGG	RA22:GTGTGGCCTGGCTGTTTC TGTG	64 °C
*3	RA23:GGAGGAGCTGGGCCTGG CCTC	RA24:CAAGTGCATCGTGGAGG AGAGG	64 °C
4	RA25:CCTGGCCACACTCCCTGG AGG	RA26:CTGCCGGAGCCTCCTTGC TGCAC	68 °C
5	SHE1:TGTGCAGCTGCCGCCACT	SHE2:ACTCTGCAGATTGTCCCG AGA	57 °C
6	SHE5:GCACACACCACCAACCC TCACT	SHE6:AATGCCTCATTTCCATGT CC	62 °C
7	SHE7:GGTGACAGGTTGCCCTT AAT	SHE8:GCTGTCCTTGCTGTGGGT TCT	59 °C
8	RA45:GGACTGTGAAGGAGGTG CGAGG	RA46:GGAGTGGAGTGGAGTGT GGAGG	66 °C
9	RA35:GAGGCTCCTGTGCTACTT TGCC	RA36:GTTGACACAGCCCAACC ATGTGC	64 °C
9	RA37:GCACATGGTTGGGCTGTG TCAAC	RA38b:GTCAATAAAGATACAAC GATTG	62 °C
9	RA39:CAGCTCGATACACACAA TCTTC	RA40:CCGTTGTGCACCACCAGG GACC	64 °C
9	RA41:GGTCCCTGGTGGTGACA ACGG	RA42:GTGGGAGACGGCTTTGTC CTGG	64 °C
9	RA43:CCAGGACAAAGCCGTCT CCCAC	RA44:GACTGTGAAGTCCGAGG TCGTC	57 °C

Primer pairs noted with \* require addition of Q solution of 5 M betaine



## APPENDIX C

This mouse assessment was developed by the project PI, Dr. Sarah H. Elsea. The form usually has entries for mice up to one year old.

### Elsea lab physical and behavioral assessment for transgenic/gene-targeted mice

**Name of investigator:**

**Mouse number:**

**Sex:**

**DOB:**

**Coat color:**

**Strain:**

**Genotype:**

*\*Please note any physical or behavioral abnormalities, even if not specifically indicated on this form\**

Directions for performing most of the physical and behavioral tests are given at the end of this form

Please perform the assays requiring anesthesia last!

<b>Weight (grams)</b>			
<b>Whisker appearance and length (in mm)</b>			
<b>Skin or fur abnormalities?</b> If so, where?	yes/no	yes/no	yes/no
<b>Bald patches?</b> If so, where?	yes/no	yes/no	yes/no
<b>Condition of genital/rectal areas</b>			
<b>Condition of nails</b>			
<b>Condition of teeth</b>			
<b>Cage movement?</b> Briefly describe	yes/no	yes/no	yes/no
<b>Righting?</b>	Did mouse right itself?  time in secs:	Did mouse right itself?  time in secs:	Did mouse right itself? yes/no  time in secs:
<b>Sound orienting?</b>	Right: yes/no  Left: yes/no	Right: yes/no  Left: yes/no	Right: yes/no  Left: yes/no
<b>Pupil constriction/dilation?</b>	Right eye: Constriction? yes/no Dilation? yes/no Left eye: Constriction? yes/no Dilation? yes/no	Right eye: Constriction? yes/no Dilation? yes/no Left eye: Constriction? yes/no Dilation? yes/no	Right eye: Constriction? yes/no Dilation? yes/no Left eye: Constriction? yes/no Dilation? yes/no

Whisker response? Normal/abnormal If abnormal, please comment	Normal/ abnormal	Normal/ abnormal	Normal/ abnormal
Eye blink? If no, please comment	Right eye: yes/no Left eye: yes/no	Right eye: yes/no Left eye: yes/no	Right eye: yes/no Left eye: yes/no
Ear twitch? If no, please comment	Right ear: yes/no Left ear: yes/no	Right ear: yes/no Left ear: yes/no	Right ear: yes/no Left ear: yes/no
Distance from outer ear to outer ear (mm)			
Distance from outer eye to outer eye (mm)			
Distance from top of head to tip of nose (mm)			
Distance from tip of nose to tip of tail (mm)			
Length of trunk from mandible (mm)			
Length of limb -front left (mm)	From toe to knee:  From knee to hip:	From toe to knee:  From knee to hip:	From toe to knee:  From knee to hip:
Length of limb-front right (mm)	From toe to knee:  From knee to hip:	From toe to knee:  From knee to hip:	From toe to knee:  From knee to hip:
Length of limb-back left (mm)	From toe to knee:  From knee to hip:	From toe to knee:  From knee to hip:	From toe to knee:  From knee to hip:
Length of limb-back right (mm)	From toe to knee:  From knee to thigh:	From toe to knee:  From knee to thigh:	From toe to knee:  From knee to thigh:
Wild running? If so, please comment	yes/no	yes/no	yes/no

Freezing? If so, please comment	yes/no	yes/no	yes/no
Sniffing?	yes/no	yes/no	yes/no
Licking?	yes/no	yes/no	yes/no
Rearing?	yes/no	yes/no	yes/no
Defecation?	yes/no	yes/no	yes/no
Urination?	yes/no	yes/no	yes/no
Movement around entire cage?	yes/no	yes/no	yes/no
<b>Sensorimotor reflexes and strength</b>			
Postural reflex Was animal able to stay upright for 10s test? If no, please comment	Cage shake from side to side: yes/no  Cage shake up and down: yes/no		Cage shake from side to side: yes/no  Cage shake up and down: Yes/no
Response to being picked up by tail Normal/abnormal If abnormal, please comment	Normal/ abnormal		Normal/ abnormal
Cage top hang test (seconds) Test three times (max 60 seconds)	Test 1:  Test 2:  Test 3:		Test 1:  Test 2:  Test 3:
Gaiting test Normal/abnormal If unusual gait, please comment	Normal		Normal/ abnormal



### **Description of physical tests**

- Righting:** turn the mouse onto its back and measure time in seconds for mouse to right itself; comment if mouse is unable to right itself or unusual response is noted
- Sound orienting:** make brief, sharp sound to the right and left of mouse; note if mouse turns head in appropriate direction
- Pupil constriction/dilation:** with a pen-light, shine a beam of light in the direction of the mouse's eye; constriction should occur when light is shone and dilation should occur when light is removed
- Whisker response:** lightly brush the whiskers of freely moving animal with a small paint brush and note response (normal mice will stop moving their whiskers and may turn head)
- Eye blink:** approach the eye with the tip of a clean cotton swab and note if eye blinks (test both eyes)
- Ear twitch:** touch with ear with the tip of a clean cotton swab and note if ear twitches (test both ears)

### **Bodily measurements**

Anesthetize the mouse with isoflurane prior to conducting measurements. A cotton ball at the bottom of the jar is soaked with 300  $\mu$ L of isoflurane and the mice are placed in the chamber. A cover is placed on the jar but oxygen flow is retained by not sealing the jar completely. At this isoflurane concentration, the mice take approximately 3 minutes to lose consciousness. Once the mice are asleep, we carry out the measurements, which takes <5 minutes. If the mice wake up during the procedure, they are placed back in the chamber. After every third mouse, an additional 300  $\mu$ L of isoflurane is added. The mice tolerate this dosage of anesthetic well and wake up and move about within 5 minutes of performing the measurements.

### **Sensorimotor and reflexes**

- Postural reflex:** place mouse in empty cage shake from side to side and up and down for 10 seconds each; note mouse's ability to maintain upright position
- Response to being picked up by tail:** note response when mouse is picked up by tail for 10 s; normal response is to raise head, extend extremities and reach for ground when lowered
- Cage top hang test:** hang mouse in empty cage with a cage top; measure time in seconds for mouse to remain hanging and note crawling movement along cage top; repeat test three times (max 60 seconds)
- Gaiting test:** gaiting is measured by coloring each mouse's foot with one color (using non-toxic materials) and walking mouse through a small tunnel

on white paper

*Hot-plate test:* place mouse on analgesic hot plate set at 60°C and measure time in seconds for mouse to display a common response (jump, raise paw, paw lick or paw shake) or an unusual response (note and comment); max trial time is 30 secs.

## REFERENCES

- (2000) Technical and clinical assessment of fluorescence in situ hybridization: an ACMG/ASHG position statement. I. Technical considerations. Test and Technology Transfer Committee. *Genet Med* 2:356-361
- Aasland R, Gibson TJ, Stewart AF (1995) The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends Biochem Sci* 20:56-59
- Alagille D, Odievre M, Gautier M, Dommergues JP (1975) Hepatic ductular hypoplasia associated with characteristic facies, vertebral malformations, retarded physical, mental, and sexual development, and cardiac murmur. *J Pediatr* 86:63-71
- Allanson JE, Greenberg F, Smith AC (1999) The face of Smith-Magenis syndrome: a subjective and objective study. *J Med Genet* 36:394-397
- Bailey JA, Yavor AM, Massa HF, Trask BJ, Eichler EE (2001) Segmental duplications: organization and impact within the current human genome project assembly. *Genome Res* 11:1005-1017
- Belgrader P, Cheng J, Zhou X, Stephenson LS, Maquat LE (1994) Mammalian nonsense codons can be cis effectors of nuclear mRNA half-life. *Mol Cell Biol* 14:8219-8228
- Bettio D, Rizzi N, Giardino D, Grugni G, Briscioli V, Selicorni A, Carnevale F, Larizza L (1995) FISH analysis in Prader-Willi and Angelman syndrome patients. *Am J Med Genet* 56:224-228
- Bi W, Ohyama T, Nakamura H, Yan J, Visvanathan J, Justice MJ, Lupski JR (2005) Inactivation of *Rai1* in mice recapitulates phenotypes observed in chromosome engineered mouse models for Smith-Magenis syndrome. *Hum Mol Genet*
- Bi W, Saifi GM, Shaw CJ, Walz K, Fonseca P, Wilson M, Potocki L, Lupski JR (2004) Mutations of *RAI1*, a PHD-containing protein, in nondeletion patients with Smith-Magenis syndrome. *Hum Genet* 115:515-524
- Bi W, Yan J, Stankiewicz P, Park SS, Walz K, Boerkoel CF, Potocki L, Shaffer LG, Devriendt K, Nowaczyk MJ, Inoue K, Lupski JR (2002) Genes in a refined Smith-Magenis syndrome critical deletion interval on chromosome 17p11.2 and the syntenic region of the mouse. *Genome Res* 12:713-728
- Botta A, Novelli G, Mari A, Novelli A, Sabani M, Korenberg J, Osborne LR, Digilio MC, Giannotti A, Dallapiccola B (1999) Detection of an atypical 7q11.23

- deletion in Williams syndrome patients which does not include the STX1A and FZD3 genes. *J Med Genet* 36:478-480
- Bourne HR, Sanders DA, McCormick F (1991) The GTPase superfamily: conserved structure and molecular mechanism. *Nature* 349:117-127
- Brzezinski A (1997) Melatonin in humans. *N Engl J Med* 336:186-195
- Burk O, Worpenberg S, Haenig B, Klempnauer KH (1997) tom-1, a novel v-Myb target gene expressed in AMV- and E26-transformed myelomonocytic cells. *Embo J* 16:1371-1380
- Campbell HD, Fountain S, McLennan IS, Berven LA, Crouch MF, Davy DA, Hooper JA, Waterford K, Chen KS, Lupski JR, Ledermann B, Young IG, Matthaei KI (2002) Fliih, a gelsolin-related cytoskeletal regulator essential for early mammalian embryonic development. *Mol Cell Biol* 22:3518-3526
- Campbell HD, Fountain S, Young IG, Claudianos C, Hoheisel JD, Chen KS, Lupski JR (1997) Genomic structure, evolution, and expression of human FLII, a gelsolin and leucine-rich-repeat family member: overlap with LLGL. *Genomics* 42:46-54
- Campbell HD, Schimansky T, Claudianos C, Ozsarac N, Kasprzak AB, Cotsell JN, Young IG, de Couet HG, Miklos GL (1993) The *Drosophila melanogaster* flightless-I gene involved in gastrulation and muscle degeneration encodes gelsolin-like and leucine-rich repeat domains and is conserved in *Caenorhabditis elegans* and humans. *Proc Natl Acad Sci U S A* 90:11386-11390
- Cassidy SB, Allanson JE (2001) Management of genetic syndromes / edited by Suzanne B. Cassidy, Judith E. Allanson. Wiley-Liss, New York
- Chen KS, Gunaratne PH, Hoheisel JD, Young IG, Miklos GL, Greenberg F, Shaffer LG, Campbell HD, Lupski JR (1995) The human homologue of the *Drosophila melanogaster* flightless-I gene (flil) maps within the Smith-Magenis microdeletion critical region in 17p11.2. *Am J Hum Genet* 56:175-182
- Chen KS, Manian P, Koeuth T, Potocki L, Zhao Q, Chinault AC, Lee CC, Lupski JR (1997) Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nat Genet* 17:154-163
- Chen KS, Potocki L, Lupski JR (1996) The Smith-Magenis syndrome [del(17)p11.2]: clinical review and molecular advances. *Ment Retard Dev Disabil Res Rev* 2:122-129
- Cheng J, Maquat LE (1993) Nonsense codons can reduce the abundance of nuclear mRNA without affecting the abundance of pre-mRNA or the half-life of



- cytoplasmic mRNA. *Mol Cell Biol* 13:1892-1902
- Cheng KW, Lahad JP, Gray JW, Mills GB (2005) Emerging role of RAB GTPases in cancer and human disease. *Cancer Res* 65:2516-2519
- Cheung SW, Tishler PV, Atkins L, Sengupta SK, Modest EJ, Forget BG (1977) Gene mapping by fluorescent in situ hybridization. *Cell Biol Int Rep* 1:255-262
- Curran ME, Atkinson DL, Ewart AK, Morris CA, Leppert MF, Keating MT (1993) The elastin gene is disrupted by a translocation associated with supravalvular aortic stenosis. *Cell* 73:159-168
- de Almeida JC, Reis DF, Martins RR (1989) Interstitial deletion of (17)(p11.2). A microdeletion syndrome. Another example. *Ann Genet* 32:184-186
- De Leersnyder H, Bresson JL, de Blois MC, Souberbielle JC, Mogenet A, Delhotal-Landes B, Salefranque F, Munnich A (2003) Beta 1-adrenergic antagonists and melatonin reset the clock and restore sleep in a circadian disorder, Smith-Magenis syndrome. *J Med Genet* 40:74-78
- De Leersnyder H, De Blois MC, Claustat B, Romana S, Albrecht U, Von Kleist-Retzow JC, Delobel B, Viot G, Lyonnet S, Vekemans M, Munnich A (2001a) Inversion of the circadian rhythm of melatonin in the Smith-Magenis syndrome. *J Pediatr* 139:111-116
- De Leersnyder H, de Blois MC, Vekemans M, Sidi D, Villain E, Kindermans C, Munnich A (2001b) beta(1)-adrenergic antagonists improve sleep and behavioural disturbances in a circadian disorder, Smith-Magenis syndrome. *J Med Genet* 38:586-590
- De Meirleir L, Seneca S, Lissens W, De Clercq I, Eyskens F, Gerlo E, Smet J, Van Coster R (2004) Respiratory chain complex V deficiency due to a mutation in the assembly gene ATP12. *J Med Genet* 41:120-124
- Dobkin C, Rabe A, Dumas R, El Idrissi A, Haubenstock H, Brown WT (2000) Fmr1 knockout mouse has a distinctive strain-specific learning impairment. *Neuroscience* 100:423-429
- Donnai D, Karmiloff-Smith A (2000) Williams syndrome: from genotype through to the cognitive phenotype. *Am J Med Genet* 97:164-171
- Dykens EM, Finucane BM, Gayley C (1997) Brief report: cognitive and behavioral profiles in persons with Smith-Magenis syndrome. *J Autism Dev Disord* 27:203-211
- Eichler EE (1998) Masquerading repeats: paralogous pitfalls of the human genome. *Genome Res* 8:758-762
- Elsea SH, Fritz E, Schoener-Scott R, Meyn MS, Patel PI (1998) Gene for topoisomerase III maps within the Smith-Magenis syndrome critical region:

- analysis of cell-cycle distribution and radiation sensitivity. *Am J Med Genet* 75:104-108
- Elsea SH, Mykytyn K, Ferrell K, Coulter KL, Das P, Dubiel W, Patel PI, Metherall JE (1999) Hemizygosity for the COP9 signalosome subunit gene, SGN3, in the Smith-Magenis syndrome. *Am J Med Genet* 87:342-348
- Elsea SH, Purandare SM, Adell RA, Juyal RC, Davis JG, Finucane B, Magenis RE, Patel PI (1997) Definition of the critical interval for Smith-Magenis syndrome. *Cytogenet Cell Genet* 79:276-281
- Ewart AK, Morris CA, Ensing GJ, Loker J, Moore C, Leppert M, Keating M (1993) A human vascular disorder, supravalvular aortic stenosis, maps to chromosome 7. *Proc Natl Acad Sci U S A* 90:3226-3230
- Fashena SJ, Serebriiskii I, Golemis EA (2000) The continued evolution of two-hybrid screening approaches in yeast: how to outwit different preys with different baits. *Gene* 250:1-14
- Fields S, Song O (1989) A novel genetic system to detect protein-protein interactions. *Nature* 340:245-246
- Finucane B, Dirrigl KH, Simon EW (2001) Characterization of self-injurious behaviors in children and adults with Smith-Magenis syndrome. *Am J Ment Retard* 106:52-58
- Finucane BM, Jaeger ER, Kurtz MB, Weinstein M, Scott CI, Jr. (1993) Eye abnormalities in the Smith-Magenis contiguous gene deletion syndrome. *Am J Med Genet* 45:443-446
- Fleming RE, Holden CC, Tomatsu S, Waheed A, Brunt EM, Britton RS, Bacon BR, Roopenian DC, Sly WS (2001) Mouse strain differences determine severity of iron accumulation in Hfe knockout model of hereditary hemochromatosis. *Proc Natl Acad Sci U S A* 98:2707-2711
- Fritz E, Elsea SH, Patel PI, Meyn MS (1997) Overexpression of a truncated human topoisomerase III partially corrects multiple aspects of the ataxia-telangiectasia phenotype. *Proc Natl Acad Sci U S A* 94:4538-4542
- Gerber HP, Seipel K, Georgiev O, Hofferer M, Hug M, Rusconi S, Schaffner W (1994) Transcriptional activation modulated by homopolymeric glutamine and proline stretches. *Science* 263:808-811
- Girirajan S, Elsas LJ, Devriendt KH, Elsea SH (2005) RAI1 variations in Smith-Magenis syndrome patients without 17p11.2 deletions. *J Med Genet*
- Gong W, Emanuel BS, Collins J, Kim DH, Wang Z, Chen F, Zhang G, Roe B, Budarf ML (1996) A transcription map of the DiGeorge and velo-cardio-facial syndrome minimal critical region on 22q11. *Hum Mol Genet* 5:789-800

- Greenberg F, Guzzetta V, Montes de Oca-Luna R, Magenis RE, Smith AC, Richter SF, Kondo I, Dobyns WB, Patel PI, Lupski JR (1991) Molecular analysis of the Smith-Magenis syndrome: a possible contiguous-gene syndrome associated with del(17)(p11.2). *Am J Hum Genet* 49:1207-1218
- Greenberg F, Lewis RA, Potocki L, Glaze D, Parke J, Killian J, Murphy MA, Williamson D, Brown F, Dutton R, McCluggage C, Friedman E, Sulek M, Lupski JR (1996) Multi-disciplinary clinical study of Smith-Magenis syndrome (deletion 17p11.2). *Am J Med Genet* 62:247-254
- Guzzetta V, Franco B, Trask BJ, Zhang H, Saucedo-Cardenas O, Montes de Oca-Luna R, Greenberg F, Chinault AC, Lupski JR, Patel PI (1992) Somatic cell hybrids, sequence-tagged sites, simple repeat polymorphisms, and yeast artificial chromosomes for physical and genetic mapping of proximal 17p. *Genomics* 13:551-559
- Hanai R, Caron PR, Wang JC (1996) Human TOP3: a single-copy gene encoding DNA topoisomerase III. *Proc Natl Acad Sci U S A* 93:3653-3657
- Hol FA, Hamel BC, Geurds MP, Hansmann I, Nabben FA, Daniels O, Mariman EC (1995) Localization of Alagille syndrome to 20p11.2-p12 by linkage analysis of a three-generation family. *Hum Genet* 95:687-690
- Hua X, Wu J, Goldstein JL, Brown MS, Hobbs HH (1995) Structure of the human gene encoding sterol regulatory element binding protein-1 (SREBF1) and localization of SREBF1 and SREBF2 to chromosomes 17p11.2 and 22q13. *Genomics* 25:667-673
- Humphries MM, Kiang S, McNally N, Donovan MA, Sieving PA, Bush RA, Machida S, Cotter T, Hobson A, Farrar J, Humphries P, Kenna P (2001) Comparative structural and functional analysis of photoreceptor neurons of Rho<sup>-/-</sup> mice reveal increased survival on C57BL/6J in comparison to 129Sv genetic background. *Vis Neurosci* 18:437-443
- Imai Y, Suzuki Y, Matsui T, Tohyama M, Wanaka A, Takagi T (1995) Cloning of a retinoic acid-induced gene, GT1, in the embryonal carcinoma cell line P19: neuron-specific expression in the mouse brain. *Brain Res Mol Brain Res* 31:1-9
- Ji Y, Eichler EE, Schwartz S, Nicholls RD (2000) Structure of chromosomal duplicons and their role in mediating human genomic disorders. *Genome Res* 10:597-610
- Juyal RC, Figuera LE, Hauge X, Elsea SH, Lupski JR, Greenberg F, Baldini A, Patel PI (1996) Molecular analyses of 17p11.2 deletions in 62 Smith-Magenis syndrome patients. *Am J Hum Genet* 58:998-1007
- Kalkhoven E, Roelfsema JH, Teunissen H, den Boer A, Ariyurek Y, Zantema A, Breuning MH, Hennekam RC, Peters DJ (2003) Loss of CBP

- acetyltransferase activity by PHD finger mutations in Rubinstein-Taybi syndrome. *Hum Mol Genet* 12:441-450
- Kallioniemi OP, Kallioniemi A, Mascio L, Sudar D, Pinkel D, Deaven L, Gray J (1994) Physical mapping of chromosome 17 cosmids by fluorescence in situ hybridization and digital image analysis. *Genomics* 20:125-128
- Kozak M (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44:283-292
- Krantz ID, Piccoli DA, Spinner NB (1999) Clinical and molecular genetics of Alagille syndrome. *Curr Opin Pediatr* 11:558-564
- Kumar S, Tomooka Y, Noda M (1992) Identification of a set of genes with developmentally down-regulated expression in the mouse brain. *Biochem Biophys Res Commun* 185:1155-1161
- Langston AW, Gudas LJ (1994) Retinoic acid and homeobox gene regulation. *Curr Opin Genet Dev* 4:550-555
- Li B, Trueb B (2000) DRG represents a family of two closely related GTP-binding proteins. *Biochim Biophys Acta* 1491:196-204
- Li DY, Toland AE, Boak BB, Atkinson DL, Ensing GJ, Morris CA, Keating MT (1997a) Elastin point mutations cause an obstructive vascular disease, supravalvular aortic stenosis. *Hum Mol Genet* 6:1021-1028
- Li L, Krantz ID, Deng Y, Genin A, Banta AB, Collins CC, Qi M, Trask BJ, Kuo WL, Cochran J, Costa T, Pierpont ME, Rand EB, Piccoli DA, Hood L, Spinner NB (1997b) Alagille syndrome is caused by mutations in human Jagged1, which encodes a ligand for Notch1. *Nat Genet* 16:243-251
- Liang Y, Wang A, Belyantseva IA, Anderson DW, Probst FJ, Barber TD, Miller W, Touchman JW, Jin L, Sullivan SL, Sellers JR, Camper SA, Lloyd RV, Kachar B, Friedman TB, Fridell RA (1999) Characterization of the human and mouse unconventional myosin XV genes responsible for hereditary deafness DFNB3 and shaker 2. *Genomics* 61:243-258
- Lockwood D, Hecht F, Dowman C, Hecht BK, Rizkallah TH, Goodwin TM, Allanson J (1988) Chromosome subband 17p11.2 deletion: a minute deletion syndrome. *J Med Genet* 25:732-737
- Lohi O, Lehto VP (1998) VHS domain marks a group of proteins involved in endocytosis and vesicular trafficking. *FEBS Lett* 440:255-257
- Lowery MC, Morris CA, Ewart A, Brothman LJ, Zhu XL, Leonard CO, Carey JC, Keating M, Brothman AR (1995) Strong correlation of elastin deletions, detected by FISH, with Williams syndrome: evaluation of 235 patients. *Am J Hum Genet* 57:49-53

- Lucas RE, Vlangos CN, Das P, Patel PI, Elsea SH (2001) Genomic organisation of the approximately 1.5 Mb Smith-Magenis syndrome critical interval: transcription map, genomic contig, and candidate gene analysis. *Eur J Hum Genet* 9:892-902
- Luiro K, Yliannala K, Ahtiainen L, Maunu H, Jarvela I, Kyttala A, Jalanko A (2004) Interconnections of CLN3, Hook1 and Rab proteins link Batten disease to defects in the endocytic pathway. *Hum Mol Genet* 13:3017-3027
- Luo B, Aster JC, Hasserjian RP, Kuo F, Sklar J (1997) Isolation and functional analysis of a cDNA for human Jagged2, a gene encoding a ligand for the Notch1 receptor. *Mol Cell Biol* 17:6057-6067
- Lupski JR (1998) Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet* 14:417-422
- Lynch HJ, Wurtman RJ, Moskowitz MA, Archer MC, Ho MH (1975) Daily rhythm in human urinary melatonin. *Science* 187:169-171
- Maier D, Stumm G, Kuhn K, Preiss A (1992) Hairless, a *Drosophila* gene involved in neural development, encodes a novel, serine rich protein. *Mech Dev* 38:143-156
- McClive P, Pall G, Newton K, Lee M, Mullins J, Forrester L (1998) Gene trap integrations expressed in the developing heart: insertion site affects splicing of the PT1-ATG vector. *Dev Dyn* 212:267-276
- Meng X, Lu X, Li Z, Green ED, Massa H, Trask BJ, Morris CA, Keating MT (1998) Complete physical map of the common deletion region in Williams syndrome and identification and characterization of three novel genes. *Hum Genet* 103:590-599
- Mitchell KJ, Pinson KI, Kelly OG, Brennan J, Zupicich J, Scherz P, Leighton PA, Goodrich LV, Lu X, Avery BJ, Tate P, Dill K, Pangilinan E, Wakenight P, Tessier-Lavigne M, Skarnes WC (2001) Functional analysis of secreted and transmembrane proteins critical to mouse development. *Nat Genet* 28:241-249
- Morris CA, Demsey SA, Leonard CO, Dilts C, Blackburn BL (1988) Natural history of Williams syndrome: physical characteristics. *J Pediatr* 113:318-326
- Oda T, Elkahoul AG, Meltzer PS, Chandrasekharappa SC (1997a) Identification and cloning of the human homolog (JAG1) of the rat Jagged1 gene from the Alagille syndrome critical region at 20p12. *Genomics* 43:376-379
- Oda T, Elkahoul AG, Pike BL, Okajima K, Krantz ID, Genin A, Piccoli DA, Meltzer PS, Spinner NB, Collins FS, Chandrasekharappa SC (1997b) Mutations in the human Jagged1 gene are responsible for Alagille syndrome. *Nat Genet* 16:235-242
- Onodera O, Oyake M, Takano H, Ikeuchi T, Igarashi S, Tsuji S (1995) Molecular

- cloning of a full-length cDNA for dentatorubral-pallidoluysian atrophy and regional expressions of the expanded alleles in the CNS. *Am J Hum Genet* 57:1050-1060
- Osborne LR (1999) Williams-Beuren syndrome: unraveling the mysteries of a microdeletion disorder. *Mol Genet Metab* 67:1-10
- Park SS, Stankiewicz P, Bi W, Shaw C, Lehoczký J, Dewar K, Birren B, Lupski JR (2002) Structure and evolution of the Smith-Magenis syndrome repeat gene clusters, SMS-REPs. *Genome Res* 12:729-738
- Paro R (1993) Mechanisms of heritable gene repression during development of *Drosophila*. *Curr Opin Cell Biol* 5:999-1005
- Patel PI, Roa BB, Welcher AA, Schoener-Scott R, Trask BJ, Pentao L, Snipes GJ, Garcia CA, Francke U, Shooter EM, et al. (1992) The gene for the peripheral myelin protein PMP-22 is a candidate for Charcot-Marie-Tooth disease type 1A. *Nat Genet* 1:159-165
- Petrij F, Giles RH, Dauwerse HG, Saris JJ, Hennekam RC, Masuno M, Tommerup N, van Ommen GJ, Goodman RH, Peters DJ, et al. (1995) Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature* 376:348-351
- Pinkel D, Straume T, Gray JW (1986) Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proc Natl Acad Sci U S A* 83:2934-2938
- Pollet N, Dhome-Pollet S, Deleuze JF, Boccaccio C, Driancourt C, Raynaud N, Le Paslier D, Hadchouel M, Meunier-Rotival M (1995) Construction of a 3.7-Mb physical map within human chromosome 20p12 ordering 18 markers in the Alagille syndrome locus. *Genomics* 27:467-474
- Potocki L, Chen KS, Park SS, Osterholm DE, Withers MA, Kimonis V, Summers AM, Meschino WS, Anyane-Yeboah K, Kashork CD, Shaffer LG, Lupski JR (2000a) Molecular mechanism for duplication 17p11.2- the homologous recombination reciprocal of the Smith-Magenis microdeletion. *Nat Genet* 24:84-87
- Potocki L, Glaze D, Tan DX, Park SS, Kashork CD, Shaffer LG, Reiter RJ, Lupski JR (2000b) Circadian rhythm abnormalities of melatonin in Smith-Magenis syndrome. *J Med Genet* 37:428-433
- Rodman S, Wadinger-Ness A (2000) Rab GTPases coordinate endocytosis. *J Cell Sci* 113:183-192
- Roelfsema JH, White SJ, Ariyurek Y, Bartholdi D, Niedrist D, Papadia F, Bacino CA, den Dunnen JT, van Ommen GJ, Breuning MH, Hennekam RC, Peters DJ (2005) Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in

- both the CBP and EP300 genes cause disease. *Am J Hum Genet* 76:572-580
- Ross SA, McCaffery PJ, Drager UC, De Luca LM (2000) Retinoids in embryonal development. *Physiol Rev* 80:1021-1054
- Sazuka T, Kinoshita M, Tomooka Y, Ikawa Y, Noda M, Kumar S (1992a) Expression of DRG during murine embryonic development. *Biochem Biophys Res Commun* 189:371-377
- Sazuka T, Tomooka Y, Ikawa Y, Noda M, Kumar S (1992b) DRG: a novel developmentally regulated GTP-binding protein. *Biochem Biophys Res Commun* 189:363-370
- Schenker T, Lach C, Kessler B, Calderara S, Trueb B (1994) A novel GTP-binding protein which is selectively repressed in SV40 transformed fibroblasts. *J Biol Chem* 269:25447-25453
- Schenker T, Trueb B (1997) Assignment of the gene for a developmentally regulated GTP-binding protein (DRG2) to human chromosome bands 17p13-->p12 by in situ hybridization. *Cytogenet Cell Genet* 79:274-275
- Schmickel RD (1986) Contiguous gene syndromes: a component of recognizable syndromes. *J Pediatr* 109:231-241
- Seranski P, Hoff C, Radelof U, Hennig S, Reinhardt R, Schwartz CE, Heiss NS, Poustka A (2001) RAI1 is a novel polyglutamine encoding gene that is deleted in Smith-Magenis syndrome patients. *Gene* 270:69-76
- Seroussi E, Kedra D, Kost-Alimova M, Sandberg-Nordqvist AC, Fransson I, Jacobs JF, Fu Y, Pan HQ, Roe BA, Imreh S, Dumanski JP (1999) TOM1 genes map to human chromosome 22q13.1 and mouse chromosome 8C1 and encode proteins similar to the endosomal proteins HGS and STAM. *Genomics* 57:380-388
- Sesca E, Perletti GP, Binasco V, Chiara M, Tessitore L (1996) Phosphatidylethanolamine N-methyltransferase 2 and CTP-phosphocholine cytidyltransferase expressions are related with protein kinase C isozymes in developmental liver growth. *Biochem Biophys Res Commun* 229:158-162
- Shaw CJ, Bi W, Lupski JR (2002) Genetic proof of unequal meiotic crossovers in reciprocal deletion and duplication of 17p11.2. *Am J Hum Genet* 71:1072-1081
- Skarnes WC, Auerbach BA, Joyner AL (1992) A gene trap approach in mouse embryonic stem cells: the lacZ reported is activated by splicing, reflects endogenous gene expression, and is mutagenic in mice. *Genes Dev* 6:903-918
- Slager RE, Newton TL, Vlangos CN, Finucane B, Elsea SH (2003) Mutations in RAI1 associated with Smith-Magenis syndrome. *Nat Genet* 33:466-468
- Smith AC, Dykens E, Greenberg F (1998a) Behavioral phenotype of Smith-Magenis

- syndrome (del 17p11.2). *Am J Med Genet* 81:179-185
- Smith AC, Dykens E, Greenberg F (1998b) Sleep disturbance in Smith-Magenis syndrome (del 17 p11.2). *Am J Med Genet* 81:186-191
- Smith AC, Gropman AL, Bailey-Wilson JE, Goker-Alpan O, Elsea SH, Blancato J, Lupski JR, Potocki L (2002) Hypercholesterolemia in children with Smith-Magenis syndrome: del (17) (p11.2p11.2). *Genet Med* 4:118-125
- Smith AC, McGavran L, Robinson J, Waldstein G, Macfarlane J, Zonona J, Reiss J, Lahr M, Allen L, Magenis E (1986) Interstitial deletion of (17)(p11.2p11.2) in nine patients. *Am J Med Genet* 24:393-414
- Stagljär I (2003) Finding partners: emerging protein interaction technologies applied to signaling networks. *Sci STKE* 2003:pe56
- Stankiewicz P, Lupski JR (2002) Genome architecture, rearrangements and genomic disorders. *Trends Genet* 18:74-82
- Stankiewicz P, Shaw CJ, Dapper JD, Wakui K, Shaffer LG, Withers M, Elizondo L, Park SS, Lupski JR (2003) Genome architecture catalyzes nonrecurrent chromosomal rearrangements. *Am J Hum Genet* 72:1101-1116
- Strand D, Unger S, Corvi R, Hartenstein K, Schenkel H, Kalmes A, Merdes G, Neumann B, Krieg-Schneider F, Coy JF, et al. (1995) A human homologue of the *Drosophila* tumour suppressor gene *l(2)gl* maps to 17p11.2-12 and codes for a cytoskeletal protein that associates with nonmuscle myosin II heavy chain. *Oncogene* 11:291-301
- Stratton RF, Dobyns WB, Greenberg F, DeSana JB, Moore C, Fidone G, Runge GH, Feldman P, Sekhon GS, Pauli RM, et al. (1986) Interstitial deletion of (17)(p11.2p11.2): report of six additional patients with a new chromosome deletion syndrome. *Am J Med Genet* 24:421-432
- Takatsu H, Katoh Y, Shiba Y, Nakayama K (2001) Golgi-localizing, gamma-adaptin ear homology domain, ADP-ribosylation factor-binding (GGA) proteins interact with acidic dileucine sequences within the cytoplasmic domains of sorting receptors through their Vps27p/Hrs/STAM (VHS) domains. *J Biol Chem* 276:28541-28545
- Tassabehji M, Metcalfe K, Donnai D, Hurst J, Reardon W, Burch M, Read AP (1997) Elastin: genomic structure and point mutations in patients with supravalvular aortic stenosis. *Hum Mol Genet* 6:1029-1036
- Toulouse A, Rochefort D, Roussel J, Joobert R, Rouleau GA (2003) Molecular cloning and characterization of human *RAI1*, a gene associated with schizophrenia. *Genomics* 82:162-171
- Trask BJ, Mefford H, van den Engh G, Massa HF, Juyal RC, Potocki L, Finucane B, Abuelo DN, Witt DR, Magenis E, Baldini A, Greenberg F, Lupski JR, Patel



- PI (1996) Quantification by flow cytometry of chromosome-17 deletions in Smith-Magenis syndrome patients. *Hum Genet* 98:710-718
- Udwin O, Yule W (1991) A cognitive and behavioural phenotype in Williams syndrome. *J Clin Exp Neuropsychol* 13:232-244
- Uitto J, Christiano AM, Kahari VM, Bashir MM, Rosenbloom J (1991) Molecular biology and pathology of human elastin. *Biochem Soc Trans* 19:824-829
- Urban Z, Helms C, Fekete G, Csiszar K, Bonnet D, Munnich A, Donis-Keller H, Boyd CD (1996) 7q11.23 deletions in Williams syndrome arise as a consequence of unequal meiotic crossover. *Am J Hum Genet* 59:958-962
- Vance DE (1996) Phosphatidylethanolamine N-methyltransferase: unexpected findings from curiosity-driven research. *Eur J Med Res* 1:182-188
- Vance DE, Walkey CJ, Cui Z (1997) Phosphatidylethanolamine N-methyltransferase from liver. *Biochim Biophys Acta* 1348:142-150
- Vlangos CN, Das P, Patel PI, Elsea SH (2000) Assignment of developmentally regulated GTP-binding protein (DRG2) to human chromosome band 17p11.2 with somatic cell hybrids and localization to the Smith-Magenis syndrome critical interval. *Cytogenet Cell Genet* 88:283-285
- Vlangos CN, Wilson M, Blancato J, Smith AC, Elsea SH (2004) Diagnostic FISH probes for del(17)(p11.2p11.2) associated with Smith-Magenis syndrome should contain the RAI1 gene. *Am J Med Genet A* 132A:278-282
- Vlangos CN, Yim DK, Elsea SH (2003) Refinement of the Smith-Magenis syndrome critical region to approximately 950kb and assessment of 17p11.2 deletions. Are all deletions created equally? *Mol Genet Metab* 79:134-141
- Voss AK, Thomas T, Gruss P (1998) Compensation for a gene trap mutation in the murine microtubule-associated protein 4 locus by alternative polyadenylation and alternative splicing. *Dev Dyn* 212:258-266
- Walkey CJ, Donohue LR, Bronson R, Agellon LB, Vance DE (1997) Disruption of the murine gene encoding phosphatidylethanolamine N-methyltransferase. *Proc Natl Acad Sci U S A* 94:12880-12885
- Walz K, Caratini-Rivera S, Bi W, Fonseca P, Mansouri DL, Lynch J, Vogel H, Noebels JL, Bradley A, Lupski JR (2003) Modeling del(17)(p11.2p11.2) and dup(17)(p11.2p11.2) contiguous gene syndromes by chromosome engineering in mice: phenotypic consequences of gene dosage imbalance. *Mol Cell Biol* 23:3646-3655
- Walz K, Spencer C, Kaasik K, Lee CC, Lupski JR, Paylor R (2004) Behavioral characterization of mouse models for Smith-Magenis syndrome and dup(17)(p11.2p11.2). *Hum Mol Genet* 13:367-378

- Wang A, Liang Y, Fridell RA, Probst FJ, Wilcox ER, Touchman JW, Morton CC, Morell RJ, Noben-Trauth K, Camper SA, Friedman TB (1998) Association of unconventional myosin MYO15 mutations with human nonsyndromic deafness DFNB3. *Science* 280:1447-1451
- Wang ZG, Sheluho D, Gatti DL, Ackerman SH (2000) The alpha-subunit of the mitochondrial F(1) ATPase interacts directly with the assembly factor Atp12p. *Embo J* 19:1486-1493
- Wang ZG, White PS, Ackerman SH (2001) Atp11p and Atp12p are assembly factors for the F(1)-ATPase in human mitochondria. *J Biol Chem* 276:30773-30778
- Weaver BK, Kumar KP, Reich NC (1998) Interferon regulatory factor 3 and CREB-binding protein/p300 are subunits of double-stranded RNA-activated transcription factor DRAF1. *Mol Cell Biol* 18:1359-1368
- Yan J, Walz K, Nakamura H, Carattini-Rivera S, Zhao Q, Vogel H, Wei N, Justice MJ, Bradley A, Lupski JR (2003) COP9 signalosome subunit 3 is essential for maintenance of cell proliferation in the mouse embryonic epiblast. *Mol Cell Biol* 23:6798-6808
- Yang SP, Bidichandani SI, Figuera LE, Juyal RC, Saxon PJ, Baldini A, Patel PI (1997) Molecular analysis of deletion (17)(p11.2p11.2) in a family segregating a 17p paracentric inversion: implications for carriers of paracentric inversions. *Am J Hum Genet* 60:1184-1193

