Virginia Commonwealth University
VCU Scholars Compass

Theses and Dissertations Graduate School

2010

IDENTIFICATION OF LOCI CONTRIBUTING TO THE SMITH-MAGENIS SYNDROME-LIKE PHENOTYPE AND MOLECULAR EVALUATION OF THE RETINOIC ACID INDUCED 1 GENE

Stephen Williams
Virginia Commonwealth University

Follow this and additional works at: https://scholarscompass.vcu.edu/etd

Part of the Medical Genetics Commons

© The Author

Downloaded from
https://scholarscompass.vcu.edu/etd/65

This Dissertation is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

المنارة للاستشارات

© Stephen Richardson Williams May 2010
All Rights Reserved

IDENTIFICATION OF LOCI CONTRIBUTING TO THE SMITH-MAGENIS SYNDROME-LIKE PHENOTYPE AND MOLECULAR EVALUATION OF THE RETINOIC ACID INDUCED 1 GENE

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

By

STEPHEN RICHARDSON WILLIAMS
Bachelor of Science (B.S.)
James Madison University, Harrisonburg, Virginia, 2001

Director: Sarah H. Elsea, Ph.D., F.A.C.M.G.
Associate Professor, Departments of Pediatrics and Human and Molecular Genetics

Virginia Commonwealth University
Richmond, Virginia
May, 2010

Acknowledgements

There are many people who contributed to my success at VCU both directly and indirectly and I would like to thank those that had the largest impact on my pursuit of scientific knowledge. First and foremost, I thank my mentor Dr. Sarah H. Elsea. Her knowledge and guidance are unmatched and I feel lucky to have work with and under her. Dr. Elsea provided the motivation to perform well in the classroom and at the bench. She is a forward-thinking scientist and human geneticist who rarely receives the recognition she truly deserves. I think my greatest thanks to her is for driving me to publish solid scientific work and participate in any and all conferences and symposiums. Through these efforts I have been given the opportunity to succeed in the future.

Next, I would like to thank all the members of Elsea lab, both past and present, who aided in my success. I would especially like to thank Dr. Santhosh Girirajan and Dr. Lily H. Truong who are friends and colleagues that encouraged me to think outside the realm of what is probable into that which is possible. I know we will be aiding each other for years to come. I thank Dr. Paula Hauck, Barbara Szomju, Nisha Patel, and David Shin for their guidance when I first entered the Elsea lab. I thank Sara Kohal, Heather Wright, Conschetta Wright, Ria Vyas, Eri Kamura, Brooke Burns, Kristie Schmidt, Sun Kim, and Sureni Mullegama for always supporting me in the lab.

I also thank committee members Dr. Jill Bettinger, Dr. Richard Moran, Dr. Rita Shiang, and Dr. Timothy York. Their guidance and motivation have contributed greatly to the direction and quality of these studies.

I am thankful to my teachers in the Department of Human and Molecular Genetics, Dr. Arti Pandya, Dr. Joyce Lloyd, Dr. Colleen Jackson-Cook, Dr. Mike Grotewiel, Dr. Shawn Holt, Dr. Jim Lister, Dr. Peter O'Connell for their knowledge and willingness to teach.

I also thank my friends Erik Ballagh, Jon Mott, Casey and Jennifer Bohn, Jun Kim, Sami Amr, Nathan Rea and my girlfriend Kelly Shircliff who always supported me in my time at VCU.

Lastly and most importantly, I have to thank my family, Fred R. Williams, Kate S. Williams, Ellen W. McCue, and all the Williams and Sparks for their support. They have always encouraged me to make the best out of every situation and I can only hope to make them proud in the future.

Technical Assistance

Chapter 2 assistance was as follows: Whole genome array comparative genomic hybridization was performed either by Dr. Eli Hatchwell or Empire Genomics. Assistance with collection of clinical information was provided by Dr. Sarah Elsea and Emily Edelman.

Chapter 3 assistance was as follows: Clinical information was gathered in part by Dr. Sarah Elsea, Sureni Mullegama, Dr. William Allen and Dr. Charles Williams. Real time qPCR assistance was provided by Sureni Mullegama. Oligonucleotide array CGH and FISH analysis were performed by Signature Genomics. Jill Rosenfeld assisted with aCGH data analysis.

Chapter 4 assistance was as follows: Collection of clinical information was aided by Dr. Sarah Elsea, Dr. Micheala Aldred, and Dr. Ellen Magenis. MLPA probes were designed by Dr. Micheala Aldred. Arrays were performed by Empire genomics or at the Cleveland Clinic.

Chapter 5 assistance was as follows: Microarray was performed by Nimblegen. Mouse tissues were collected by Brooke Burns or Kristie Schmidt.

Table of Contents

Acknowledgments iii

Technical Assistance v

List of Tables xi

List of Figures xii

Abstract xiv

Chapter **Page**

1 Background and literature review 1

- Introduction 1
- Copy Number Variation 4
- Smith-Magenis syndrome 4
- Intellectual disability 6
- Sleep and circadian rhythm abnormalities 6
- Behavioral abnormalities 8
- Craniofacial and skeletal abnormalities 8
- Other abnormalities 10
- Variation in SMS deletion 10
- Duplication 17p11.2 syndrome (Potoki-Lupski syndrome, PTLS) 12
- Mouse models for Smith-Magenis and Duplication 17p11.2 13
- Unexplained intellectual disability 20

2 Identification of loci contributing to the Smith-Magenis-like phenotype

Introduction 21
Materials and methods 22
    Subject ascertainment and samples 23
    Whole genome array-based comparative genomic hybridization (aCGH) 23
    Fluorescence *in situ* hybridization 25
    Sequencing of *RAI1* 25
    Bioinformatic and statistical analyses 25
Results 26
    SMS vs. SMS-like phenotypes 26
    Whole genome array comparative genomic hybridization 29
    Pathogenic copy number variants 28
    Non-pathogenic copy number variants 37
Discussion 41

3 Haploinsufficiency of *MBD5* associated with a syndrome involving microcephaly, intellectual disabilities, severe speech impairment, and seizures 45
  Introduction 45
  Materials and methods 46
    Patient ascertainment 46
    Cell lines 46
    RNA extraction 47

Real-time qPCR 47

RAI1 sequencing 48

Whole-genome array comparative genomic hybridization............48

Oligonucleotide array comparative genomic hybridization...........48

Fluorescence *in situ* hybridization (FISH) analysis 49

Results...49

Case reports...49

SMS185..49

SMS361..55

Molecular analysis 56

aCGH 56

Quantitative real-time PCR...58

Discussion...61

4 Haploinsufficiency of *HDAC4* results in brachydactyly mental retardation Syndrome 63

Introduction...63

Materials and methods 64

Subject ascertainment 64

Genomic DNA sequence analysis...64

Multiplex ligation-dependent probe amplification and SNP array analysis...65

Results...69

BDMR clinical phenotype 69
Refinement of the BDMR critical region....75
SMS117 clinical phenotype 80
Mutation analysis 81
Discussion 83
HDAC4....83
Hdac4-/- and related mouse models 84
Clinical implications 85

5 RAI1 transcriptionally regulates *CLOCK* and is a major contributor to circadian homeostasis 88
Introduction....88
Circadian Rhythm 91
Materials and methods 91
Creation of Plasmids 91
Transfections....93
Luciferase assay 93
Calculation of relative luciferase activity 94
Statistical analysis....94
Mouse tissue collection....95
Quantitative PCR 95
Chromatin immunoprecipitation with microarray (ChIP-Chip) 95
Candidate gene filtration....96

Results 97

Chromatin immunoprecipitation with microarray 97

Genes identified for further evaluation 101

CLOCK 101

SNRPN 101

UBE3A 102

BDNF 102

RAI1 transcriptionally activates *CLOCK* via an intron 1 enhancer element 103

Dysregulation of circadian genes in *Rai1+/-* mice 105

Dysregulation of circadian genes in SMS fibroblasts 108

Discussion 111

Summary 114

Development 115

Behavior and sleep 116

Future directions 118

References 121

Appendix A ....

Appendix B ....

Appendix C ....

Appendix D ....

List of Tables

Page

Table 1. Clinical features of chromosome 17p11.2 duplication cases 14

Table 2. Phenotypes of *Rai1* engineered mice 19

Table 3. Phenotypic features observed in Smith-Magenis syndrome-like cohort 27

Table 4. Pathogenic copy number variants identified in the SMS-like cohort 32

Table 5. Disorders previously associated with the deletions and duplications identified in this study 36

Table 6. Non-pathogenic copy number variations identified in the SMS-like cohort 38

Table 7. Clinical features of cases with 2q23.1 microdeletion syndrome 52

Table 8. *HDAC4* amplification and sequencing oligonucleotides 66

Table 9. *HDAC4* cDNA oligonucleotides for amplification and sequencing 67

Table 10. MLPA probes for 2q37.3 and *HDAC4* 68

Table 11. Clinical features of cases with deletions or mutations involving 2q37.3 71

Table 12. List of transcripts which overlap between ChIP-Chip arrays 106

Table 13. Ranked order of top 10 gene regulatory regions identified in ChIP-Chip experiments 100

List of Figures

Page

Figure 1. Mechanism for non-allelic homologous recombination....................................3

Figure 2. Circadian sleep pattern of SMS subject..7

Figure 3. Craniofacial features observed in Smith-Magenis syndrome.............................9

Figure 4. Deletion variation identified in SMS...

Figure 5. Mouse synteny to chromosome 17p11.2 14

Figure 6. *Rai1+/- mice* growth rates compared to WT litter mates 17

Figure 7. CGH array plots for pathogenic CNVs 30

Figure 8. Del2q37 clinical phenotype...35

Figure 9. Non-pathogenic copy number variants..43

Figure 10. Clinical features of the 2q23.1 deletion syndrome 56

Figure 11. Array analysis of individuals with microdeletions of 2q23.1.........................64

Figure 12. FISH confirmation of 2q23 deletions in patients SMS361 and SMS185........66

Figure 13. Quantitative real-time PCR expression analysis of *MBD5* and *EPC2* in patient cells 67

Figure 14. Delineation of the BDMR critical region 80

Figure 15. MLPA analysis of *HDAC4* deletion breakpoints in key 2q37.3 deletions 81

Figure 16. Skeletal anomalies observed in BDMR...86

Figure 17. SMS117 *HDAC4* mutation and impact on HDAC4 protein...........................91

Figure 18. RAI1 localization and genomic and protein structure of RAI1.......................99

Figure 19. ChIP-Chip arrays...107

Figure 20. RAI1 impact on *CLOCK* transcription 112

Figure 21. *Rai1* and *Clock* expression in mouse tissues 114

Figure 22. qPCR of circadian genes expressed in mouse hypothalamus........................115

Figure 23. qPCR of core circadian gene expression in the mouse hypothalamus during day light and night time periods...117

Figure 24. Core circadian gene expression in unaffected and SMS fibroblasts 118

Figure 25. Proposed schematic of *RAI1* contribution to proper skeletal development...125

Figure 26. RAI1 regulates *CLOCK* transcription and is a major element in the circadian cycle 128

Abstract

IDENTIFICATION OF LOCI CONTRIBUTING TO THE SMITH-MAGENIS SYNDROME-LIKE PHENOTYPE AND MOLECULAR EVALUATION OF THE RETINOIC ACID INDUCED 1 GENE

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

By

STEPHEN RICHARDSON WILLIAMS
Bachelor of Science (B.S.)
James Madison University, Harrisonburg, Virginia, 2001

Smith-Magenis syndrome (SMS) is a multiple congenital abnormalities intellectual disability syndrome that results from a deletion of chromosome 17p11.2 or mutation of the *retinoic acid inducted one* gene (*RAI1*). SMS is characterized by a multitude of phenotypic features including craniofacial defects, short stature, obesity, intellectual disability, self-abusive behavior, sleep disturbance and behavioral abnormalities. Interestingly, although SMS is a clearly defined syndrome with a known molecular change at its foundation, ~40% of all candidate cases sent to the Elsea lab for evaluation do not have a mutation or deletion of *RAI1*. We hypothesize that at least one other locus must be responsible for this Smith-Magenis-like (SMS-like) phenotype.

To address this hypothesis, we first compiled a cohort of 52 subjects who had been referred to the Elsea lab for a clinical diagnosis of SMS. Once these individuals were confirmed to not have an *RAI1* mutation or deletion, their phenotypes were compiled and statically analyzed to distinguish whether SMS and SMS-like cohorts are different in the prevalence of the core phenotypes of SMS such as, but not limited to, sleep disturbance, self-abusive behavior and

developmental delay. SMS-like and SMS cohorts are not different in prevalence for these core features. Next, all SMS-like subjects were sent for whole genome array comparative genomic hybridization (aCGH) to identify duplications or deletions of each individual's genome which contribute to the phenotype observed. We identified 6 pathogenic copy number variants (CNVs) in six individuals which contribute directly to the clinical phenotype, including two del(2)(q37). This study enabled us to draw relationships between SMS and other syndromes that had never been appreciated before and helped to identify pathways in which *RAI1* may function.

Using the data from our SMS-like study we were able to further characterize two known syndromes; Deletion 2q37 syndrome (brachydactyly mental retardation syndrome) and deletion 2q23 syndrome. With regard to deletion 2q37, syndrome we used genomic data from known and new deletion 2q37 subjects to refine the critical region to one gene: the *histone deacetylase 4* gene (*HDAC4*). Using both clinical and molecular clues, we were able to identify one subject from our SMS-like cohort who has an insertion in *HDAC4* which results in a premature stop codon. We conclude from this study that mutation of *HDAC4* results in brachydactyly mental retardation syndrome.

With regard to deletion 2q23 syndrome there were only five known cases in the published literature to which we were able to add two more. Using as similar approach to our del2q37 study we refined the critical region for this syndrome to one gene, the *methyl binding domain 5* gene (*MBD5*). Using a molecular and clinical approach we were able to conclude that haploinsufficiency of *MBD5* results in the core phenotypes seen in del2q23 syndrome including microcephaly, intellectual disabilities, severe speech impairment, and seizures**.**

Using all the data generated from the three previous studies we set out to characterize the molecular function of *RAI1*. We hypothesize that RAI1 is a transcription factor that regulates gene expression of core genes involved in development, neurological function, and circadian rhythm. Using a ChIP-chip based approach we identified 257 transcripts we believe *RAI1* regulates. Following up on these transcripts, using *in vitro* and *in vivo* methods, we have been able to conclude that RAI1 is a positive regulator of *CLOCK*, the master regulator of the central circadian cycle.

Taken together, these studies have given us insight into the specific molecular changes that contribute to SMS and SMS-like syndromes. We have unveiled pathways and genes which are important to normal human development and behavior and identified novel functions of *RAI1*. These studies will provide the foundation for the future discovery of the pathways affected.

Chapter 1: Background and literature review

Introduction

Smith-Magenis syndrome (SMS, OMIM#182290) is a congenital abnormalities, intellectual disability syndrome that can present with sleep disturbance, self-abusive and aggressive behavior, craniofacial abnormalities, neurological abnormalities, and obesity. First described by Ann Smith and Dr. Ellen Magenis, SMS was initially thought to be a contiguous gene syndrome resulting from a deletion of the 17p11.2 region [3]. However, later molecular evidence revealed one gene contributes to the majority of the phenotypes observed. Data revealed that mutation or deletion of the *retinoic acid induced 1* gene (*RAI1*) resulted in SMS [4]. Thus, SMS is a sporadic, dominant syndrome defined by haploinsufficiency of *RAI1*.

Duplications and deletions of the genome have long been thought to contribute to evolution of all species. While there are rare cases for these copy number changes to confer an advantage to said species, these types of changes are typically detrimental. Deletion and duplication syndromes are a major contributor to the overall number of recognized intellectual disability syndromes which are being identified and characterized at a rapid pace with the advent of new technologies such as whole genome array comparative genomic hybridization (aCGH) [5]. SMS, duplication 17p11.2 syndrome (also known as Potocki-Lupski syndrome), fragile X-syndrome, and Williams syndrome all result from a duplication or deletion of specific regions of the genome.

Typically, deletion syndromes are more common and phenotypically more severe than duplication syndromes, due to haploinsufficiency, one copy not being sufficient, of one or more core genes (a gene imperative to the proper function of the disrupted pathway). When a chromosomal region is duplicated, the resulting phenotype can be more subtle and is thus more likely to be missed by a clinician, not warranting a molecular evaluation [6]. So, it is more appropriate to state that the incidence of deletion and duplication events in the population is likely very similar, however, the identification of these events is more likely to yield the unveiling of a deletion as the cause of the phenotype observed [7, 8]. This can be seen in SMS versus duplication 17p11.2 syndrome where the intellectual disability, craniofacial, sleep, and metabolic abnormalities are much more severe in SMS and the number of subjects identified with a 17p11.2 duplication is ~75 whereas more than 300 SMS subjects have been confirmed [7]. There is, however, some evidence showing that in the SMS region, on 17p11.2, deletions are twice as likely to occur than duplications [8]. This being said, the molecular events that take place which results in these two outcomes, duplication vs. deletion, are similar and due to non-allelic homologous recombination (NAHR) which is mediated by segmental duplications in the 17p11.2 region [9] (Figure 1).

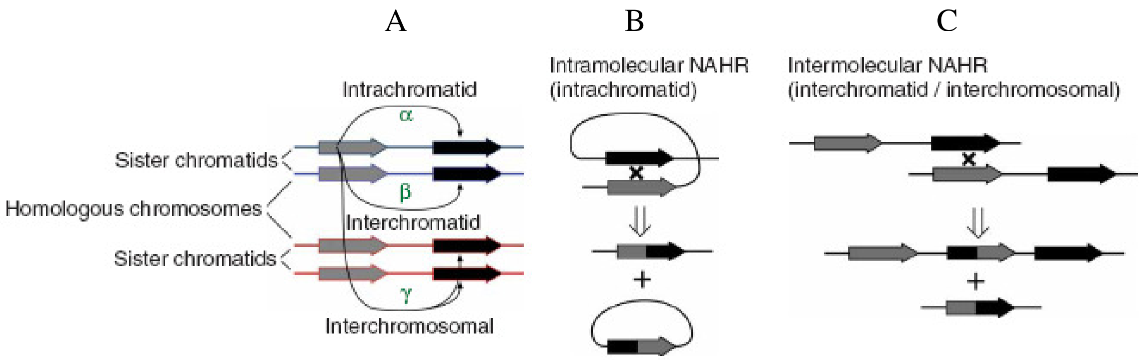

A
B
C
Intrachromatid
α
Sister chromatids
β
Homologous chromosomes
Interchromatid
Sister chromatids
γ
Interchromosomal
Intramolecular NAHR
(intrachromatid)
Intermolecular NAHR
(interchromatid / interchromosomal)
+
+

Figure 1. Mechanism for non-allelic homologous recombination. A) Grey and black arrows indicate regions of homology between sister chromatids. NAHR can occur in two forms: B) Intramolecular, wherein sister chromatids misalign leading to deletion of intervening sequences. C) Intermolecular, wherein chromatids from different chromosomes misalign leading to a duplication or deletion allele. Modified from Turner et al. 2008 [8].

Copy Number Variation (CNV)

Copy number variations are defined as segments of DNA ranging from 1 kilobase (kb) to several megabases (Mb), for which copy-number differences have been revealed by comparison of two or more genomes [10]. Investigation into the importance of copy number variations has become an integral part in understanding human evolution and disease. Various molecular events can result from CNV, including gene dosage, gene disruption, gene fusion and position effects, CNVs can cause Mendelian or sporadic traits [11], or be associated with complex diseases [12]. CNVs can also represent singularly benign polymorphic variants that contribute to a combinatorial effect in variation. In fact, it is thought that CNVs contribute more to human variation than do single nucleotide polymorphisms [13]. Additionally, it is thought that CNVs have a much higher *de novo*, locus specific mutation rate than SNPs [11]. Given these facts, it is important to study disorders and syndromes which result from CNV to better understand and recognize regions of the genome that may contribute in large part to disorder or variation.

Smith-Magenis syndrome

Smith-Magenis syndrome is a recognizable syndrome characterized by physical, developmental, neurological, and behavioral features. SMS is typically a *de novo* disorder with an estimated prevalence of 1:15,000–25,000 live births [14] however; there is one report of an SMS subject having a mother who is mosaic for a del(17)(p11.2) [15]. SMS has many clinical phenotypes that overlap with other intellectual disability syndromes, such as Prader-Willi and Williams syndromes, and interestingly only ~46% of those individuals referred the Elsea lab for 17p11.2/*RAI1* molecular evaluation are positive for molecular changes leading to SMS (Elsea, unpublished data).

Physical features consist of distinctive craniofacial and skeletal anomalies including brachycephaly, frontal bossing, hypertelorism, synophrys, up slanting palpebral fissures, midface hypoplasia, a broad square-shaped face flat nasal bridge and a tented upper lip [16]. Dental anomalies include tooth agenesis, especially of premolars, and taurodontism [17]. Cleft lip or palate is also reported at a higher rate than seen in the normal population [18] and hearing loss is present at 60-68%.

The neurological and behavioral phenotypes seen in SMS consists of hypotonia, stereotypies such as self-hugging and hand twirling, self-abusive behavior, hyperactivity, and oral and/or motor dysfunction [18, 19]. An inverted circadian rhythm which results in waking periods during the night and napping periods during the day, and times of activity, complicate the neurological and behavioral phenotypes observed [20].

Skeletal and developmental abnormalities include short stature (<5th percentile), brachycephaly, brachydactyly, scoliosis, fifth-finger clinodactyly, fore arm and elbow limitations, and polydactyly. Additionally obesity is observed and cardiac defects are present at ~30% [18, 20].

Otolaryngological abnormalities such as hearing loss, a hoarse deep voice, and vocal cord nodules and polyps are also common [21, 22, 23]. Additionally, ophthalmologic features are present in 64% of SMS patients and include myopia, iris anomalies such as heterochromic irides or Wolfflin-truckmann spots (iris hamartomas), strabismus, microcornea, and rarely, retinal detachment (which can result from violent behaviors) [18, 24, 25].

Intellectual disability

Intellectual disability (ID) is defined as a person having an IQ <70 and is present in 1–3% of the general population. The cause of ID is unknown in more than half of cases identified [26].

Most SMS individuals have mild to moderate intellectual disability with IQ ranging between 20-78 and most children falling between 40-54 [22]. School-age children with IQs in the low normal range have been identified; however, IQ decreases as the child ages ultimately placing the individual in the mild intellectual disability range by adulthood (Elsea, unpublished data). Delayed speech occurs in ~84% of SMS patients [27]. These individuals have better receptive language skills than expressive language, while their social-emotional functions remain within the normal range [28, 29]. In addition, delayed fine/gross motor skills, problems with sensory integration, and poor adaptive function are seen. Other neurological features include peripheral neuropathy, pes cavus or pes planus, abnormal gait, and decreased sensitivity to pain, which may lead to an increased risk for self-injury.

Sleep and circadian rhythm abnormalities

Sleep disturbance is a hallmark of Smith-Magenis syndrome. It is reported that 75%-100% of all confirmed SMS cases present with sleep disturbance and this is one of the early indicators that an individual may have SMS [19, 22]. Infants typically present with hypersomnolence early in life, and sleep disturbance in older children include difficulty falling asleep, inability to enter or maintain REM sleep, reduced night sleep, shortened and broken sleep cycles with frequent night time and early-morning awakenings and excessive daytime sleepiness [20, 30] (Figure 2).

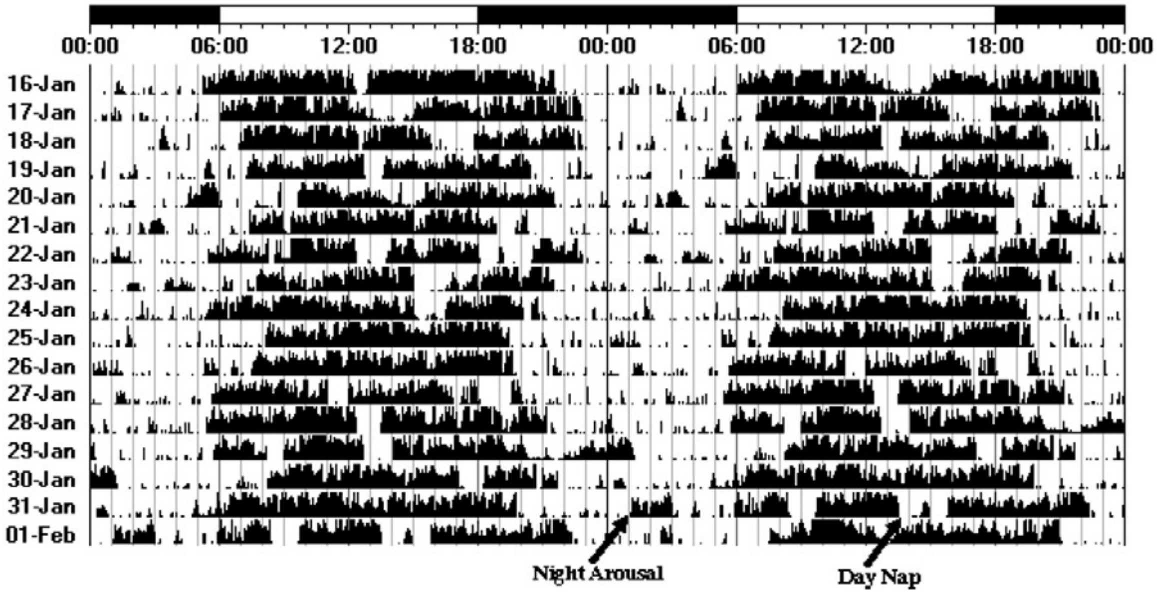


Figure 2. Circadian sleep pattern of SMS subject. 5 year old confirmed SMS subject. Data collected over a 16 day period using a wrist activity meter. Top bar indicates wake periods (white) and sleep periods (black). Vertical black lines indicate periods of activity. Note consistent night time arousals and day time naps. Modified from Gropman et al. 2006 [28].

Thought to be caused by an inverted rhythm of melatonin secretion, these circadian rhythm difficulties can complicate behavioral and learning difficulties (Elsea, unpublished data). However, multiple cases have been reported that have normal melatonin secretion but still have sleep disturbance, indicating a possible molecular role for *RAI1* in the sleep cycle and the core molecular pathway of circadian rhythm. [30]

Behavioral abnormalities

The SMS phenotype consists of core behavioral abnormalities. Core phenotypes include attention-seeking, aggression, disobedience, a lack of respect for personal space, distraction, stereotypical and self-injurious behaviors. Interestingly, self-injurious behaviors are seen in ~90% of cases and include head-banging and skin picking, two features unique to SMS, onychotillomania, and polyembolokoilamania are more often seen in older children [31, 32]. Stereotypical behaviors include self-hugging, hand twirling, and body rocking [31, 33]. It is noteworthy that although some SMS subjects are diagnosed with autism early in life (because of delayed speech), this diagnosis typically changes with the acquisition of language skills.

Craniofacial and skeletal abnormalities

Craniofacial and skeletal abnormalities include short stature, broad face, flat nasal bridge, brachycephaly, brachydactyly, scoliosis, crainostenosis, prognathasism, midface hypoplasia, tented upper lip, and cleft lip and/or palate. The range of craniofacial phenotypes is variable early in life but as subjects age there is a “typical” appearance that becomes identifiable (Figure 3)

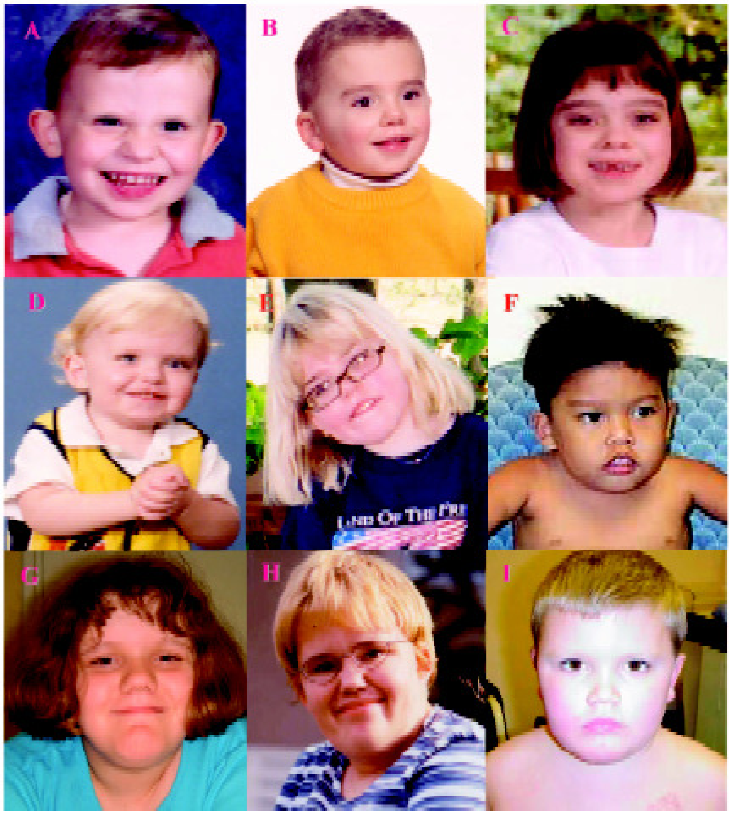


Figure 3. Craniofacial features observed in Smith-Magenis syndrome. A-C Common deletion. D-F Atypical deletion. G-I *RAI1* mutation. Modified from Girirajan et al. 2006 [34]

Other abnormalities

Along with the typical behavioral, neurological, and craniofacial features there are some phenotypes in SMS that are more variable. Dental abnormalities are observed in ~50% of subjects and include malocclusion, tooth agenesis, taurodontism, gingavitus, and poor overall dental hygiene [17].

Cardiac abnormalities are also seen in a subset of SMS subjects and include ventral septal defect, atrial septal defect, tricuspid stenosis, mitral stenosis, tricuspid and mitral regurgitation, aortic stenosis, pulmonary stenosis, mitral valve prolapse, tetralogy of Fallot, and total anomalous pulmonary venous return [18].

Seizures occur in up to 30% of cases and affect those with a common deletion at the highest rate. However, ~50% of diagnosed SMS cases present with abnormal EEGs [18]. This suggests that up to 20% of individuals with SMS do not reach the threshold for seizure.

Variation in SMS deletion

17p11.2 deletions leading to SMS can be broken down into four main categories; common, large, small and atypical deletions (Figure 4). All deletions encompass the *RAI1* gene but size and position of the deletion can account for phenotypic variance present in the individual. For example, short stature is seen at a higher prevalence in individuals with common deletions than those found to have large deletions [18]. Those with small deletions are more likely to have an abnormal EEG than those with common deletions [18]. Additionally, those with small deletions have a more similar phenotype to individuals with *RAI1* mutations [14].

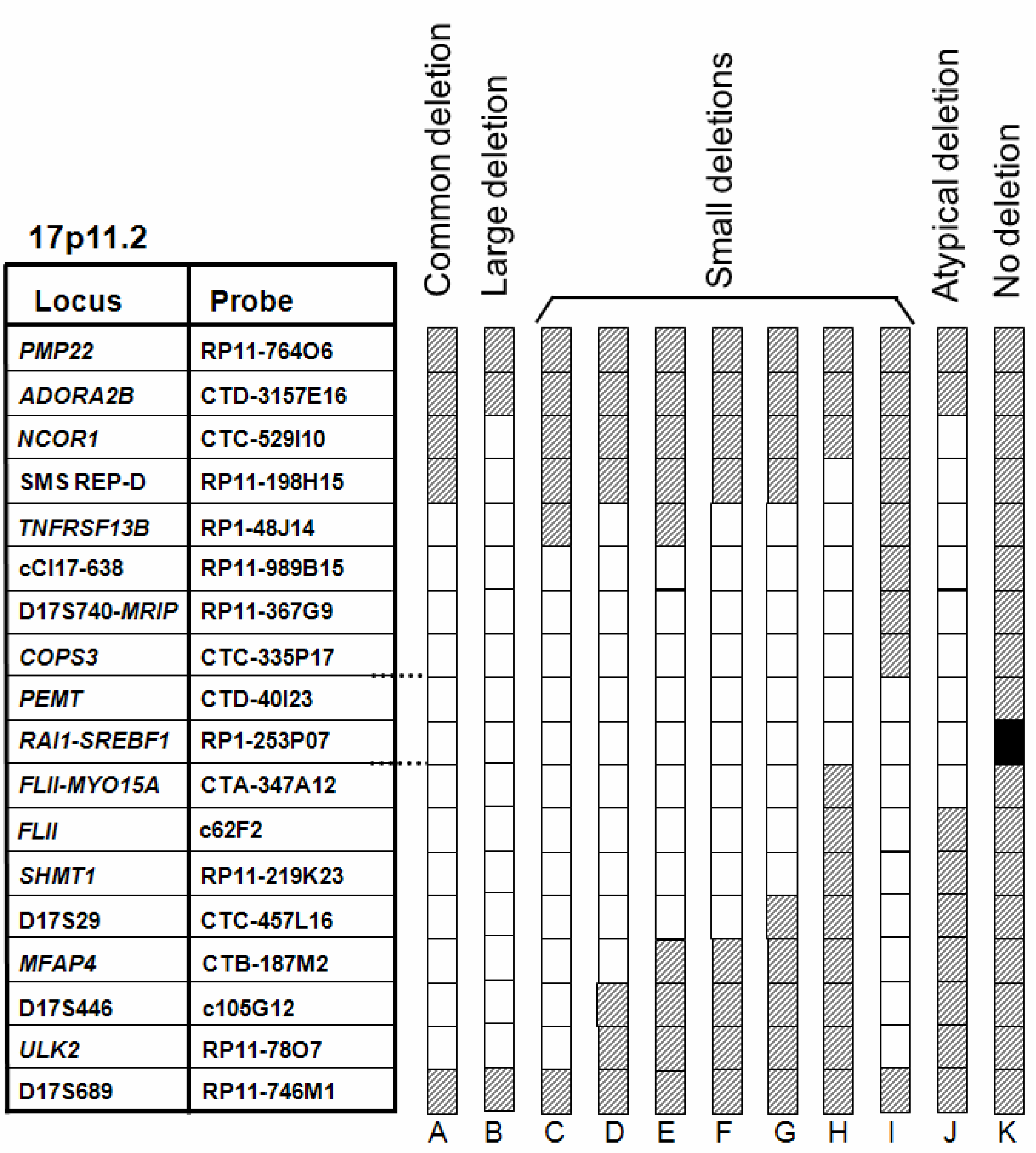

17p11.2
Locus
Probe
PMP22
RP11-764O6
ADORA2B
CTD-3157E16
NCOR1
CTC-529I10
SMS REP-D
RP11-198H15
TNFRSF13B
RP1-48J14
cCI17-638
RP11-989B15
D17S740-MRIP
RP11-367G9
COPS3
CTC-335P17
PEMT
CTD-40I23
RAI1-SREBF1
RP1-253P07
FLII-MYO15A
CTA-347A12
FLII
c62F2
SHMT1
RP11-219K23
D17S29
CTC-457L16
MFAP4
CTB-187M2
D17S446
c105G12
ULK2
RP11-78O7
D17S689
RP11-746M1
Common deletion
Large deletion
Small deletions
Atypical deletion
No deletion
A
B
C
D
E
F
G
H
I
J
K

Figure 4. **Deletion variation identified in SMS.** Loci and FISH probes used to test for presence or absence of these loci are represented in the first two columns. Shaded boxes represent regions not deleted and empty boxes indicate regions that are deleted. Shown are common, large, small and atypical deletions seen in SMS. Note that *RAI1* is found in all deletions. Modified from Girirajan et al. 2006 [34].

Duplication 17p11.2 syndrome (Potoki-Lupski syndrome, PTLS)

Interstitial duplication of 17p11.2 results in duplication 17p11.2 syndrome, also known as Potoki-Lupski syndrome (PTLS). Dup(17)(p11.2) results in mild to severe intellectual disability, infantile hypotonia, failure to thrive, congenital cardiovascular anomalies, sleep-disordered breathing, developmental delay, intellectual disability, hyperactivity, and autism [35, 36]. Craniofacial features include a triangular face, microcephaly, micrognathia, broad nasal bridge, high arched palate, and hypertelorism, while dental anomalies in the form of malocclusion of the teeth have also been reported [37]. Limb abnormalities including flexion deformity of the fingers and club foot have also been described. Other features include cognitive and language impairment, pharyngeal dysphasia, obstructive and central sleep apnea, structural cardiovascular abnormalities, and electroencephalogram (EEG) abnormalities [35, 36]

Less commonly reported features are hearing impairment, structural otolaryngological defects, and ophthalmic abnormalities, such as myopia and iris hamartoma. Other rarely reported features are genitourinary and/or renal anomalies, scoliosis, and hypercholesterolemia [36].

The predicted incidence of dup(17)(p11.2) syndrome are the same as that of SMS (~1:25,000) but is likely under diagnosed because the phenotype can be much more subtle than that of SMS. Only ~75 cases of dup(17)(p11.2) syndrome have been reported [7]. However, a recent report indicates that this region may be subject to germline rates of deletion to duplication of 2.14:1, indicating a molecular preference for deletion over duplication in 17p11.2 [8].

Duplication of this 17p11.2 region can range between a 0.4 Mb to 13.3 Mb with 3.7 Mb being the most common duplication (between distal and proximal repeats flanking RAI1) and 5 Mb being the least common duplication (located between LCR17pA and LCR17pD) having been recently identified in two individuals [7]. A list of features with regard to the different classifications are seen in Table 1.

Mouse models for Smith-Magenis and duplication 17p11.2 syndromes

Mouse models that mimic human conditions are invaluable tools to help understand the molecular, developmental, and behavioral role of a gene or genes. The sequencing of the mouse genome has allowed for regions of synteny to human genomic sequence to be mapped and utilized for research [38].

Three types of genetically engineered mice are typically used for research and are knockout, gene-trap, and transgenic. Knockout and gene-trap serve the purpose to eliminate a given gene or region from expression and function whereas transgenic mouse models contain genomic material used to over express, or simply express foreign DNA.

Murine chromosome 11qB1.3-B2 is syntenic to human chromosome 17p11.2. An approximate 34 cM region surrounding *RAI1/Rai1* is conserved between mouse and human genomic structure making this region fit for creation of knockout and transgenic mice (Figure 5). Mice lacking *Df(11)17*/+ or duplicated for *Dp(11)17*/+ a 2 Mb region, containing seven genes, syntenic to human chromosome 17p11.2 were originally created by Waltz et al. 2003 [39] using chromosome engineering technology. These deletion/duplication mice were created to mimic the common deletion seen in SMS and duplication seen in dup(17)(p11.2) syndrome. These mice recapitulate many of the core features observed in SMS and duplication 17p11.2 syndrome.

Table 1. Clinical features of chromosome 17p11.2 duplication cases. Data collected from Girirajan et al. 2007 [40] and Zang et al. 2010 [7].

Clinical features	**dup(17)(p10p12) N=4**	**dup(17)(p11.2p12) N=7**	**dup(17)(p11.2p11.2) N=8**	**Uncommon dup(17)(p11.2p12) N=2**
Mental retardation	4/4	7/7	8/8	2/2
Pre/postnatal growth retardation	3/4	7/7	3/3	0/2
Motor delay	4/4	7/7	2/3	N
Speech delay	4/4	5/5	2/4	N
History of spontaneous abortions	1/3	N	N	N
Hypotonia	4/4	2/3	2/4	1/1
Feeding difficulties	4/4	3/4	1/3	2/2
Hyperactivity/behavioral abnormalities	1/3	2/5	4/6	2/2
Seizures	2/4	1/6	3/7	0/2
Decreased nerve conduction	¼	5/6	0/8	0/2
Delayed deep tendon reflexes	1/4	6/7	0/8	N
Craniofacial features				
Triangular face	3/3	2/4	1/2	2/2
Microcephaly	4/4	2/3	4/6	N
Micrognathia	3/4	1/4	2/3	N
Hypertelorism	2/4	3/3	1/3	N
Broad nasal bridge	4/4	4/5	2/4	0/2
Low-set/malformed ears	2/4	4/4	3/3	0/1
High arched palate	1/4	3/5	3/4	N
Cleft lip/palate	0/4	0/7	1/4	N
Ophthalmologic disorders	2/4			2/2
Short broad neck	N	1/7	N	0/2
Hearing loss	1/4	1/7	1/3	0/1
Cardiac anomalies	0/4	2/7	0/8	0/2
Renal/urinary tract anomalies	1/4	1/7	2/8	1/1
Short stature	1/4	6/6	5/8	N
Decreased weight	1/4	4/4	3/5	0/2
Bone and limb defects				
Clinodactyly of fifth finger	N	4/7	2/8	N
Flexion contractures	1/4	3/7	0/8	N
Foot anomalies	2/4	6/7	0/8	N
Short sternum	N	1/6	N	N
Dental anomalies	1/4	4/7	5/8	N

Df(11)17/+ mice have features consistent with SMS including craniofacial abnormalities, obesity, seizures and neurobehavioral abnormalities including an abnormal circadian rhythm [39]. Neurological abnormalities were observed including overt clinical seizures and abnormal EEGs which are also seen in SMS [39]. Histological examinations were performed with brains, hearts, spleens, and kidneys, and no obvious defects were identified in any of these tissues [39] which is also consistent with the SMS phenotype.

Dp(11)17/+ (duplication) mice displayed growth retardation, hyperactivity, increased anxiety-related responses in a light-dark test, increased conditioned fear, and decreased startle response with normal sensorimotor gating on a prepulse inhibition assay [39, 41]. No craniofacial anomalies or altered circadian rhythm were observed in the *Dp(11)17/+* mice [39, 41].

Later, using a targeted approach, mice haploinsufficient for *Rai1* were created [42]. These mice were developed to test the phenotypic consequences of *Rai1* inactivation and to provide insight into the phenotypes Rai1 contributes directly to in SMS. Young *Rai1+/-* mice were under weight as compared to wild-type littermates but by age 23 weeks became significantly larger with progressive weight gain [42] (Figure 6). Eighteen percent of *Rai1+/-* mice had craniofacial abnormalities, including short curved snouts caused by the malformation of the craniofacial skeletal elements [42].

Rai1-/- mice largely were not viable and Mendelian ratios of litters (+/+ 36.2%, +/- 59.3%, -/-4.5%) were skewed indicating significant embryonic lethality. *Rai1-/-* mice that did survive beyond the embryonic stage had growth retardation and craniofacial abnormalities as compared to WT litter mates [42]. Additionally, these showed tone and context dependent impaired fear conditioning, as well as overt seizures [41, 42].

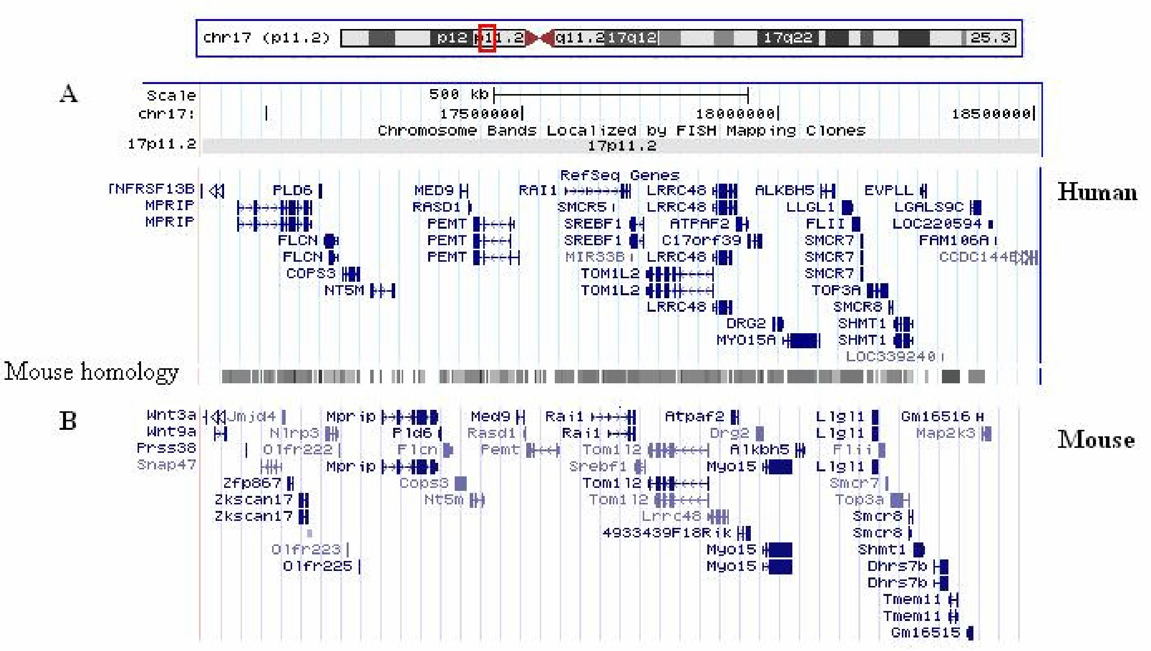

chr17 (p11.2)
p12
p11.2
q11.2
17q12
17q22
25.3
A
Scale
chr17:
500 kb
17500000
18000000
18500000
Chromosome Bands Localized by FISH Mapping Clones
17p11.2
17p11.2
RefSeq Genes
TNFRSF13B
MPRIP
MPRIP
PLD6
FLCN
FLCN
COPS3
NT5M
MED9
RASD1
PEMT
PEMT
PEMT
RAI1
SMCR5
SREBF1
SREBF1
MIR33B
TOM1L2
TOM1L2
LRRC48
LRRC48
ATPAF2
C17orf39
LRRC48
LRRC48
DRG2
MYO15A
ALKBH5
LLGL1
FLII
SMCR7
SMCR7
SMCR7
TOP3A
SMCR8
SHMT1
SHMT1
LOC339240
EVPLL
LGALS9C
LOC220594
FAM106A
CCDC144B
Human
Mouse homology
B
Wnt3a
Wnt9a
Prss38
Snap47
Jmjd4
Nlrp3
Olfr222
Zfp867
Zkscan17
Zkscan17
Olfr223
Olfr225
Mprip
Mprip
Pld6
Flcn
Cops3
Nt5m
Med9
Rasd1
Pemt
Rai1
Rai1
Tom1l2
Srebf1
Tom1l2
Tom1l2
Lrrc48
4933439F18Rik
Atpaf2
Drg2
Alkbh5
Myo15
Myo15
Myo15
Llgl1
Llgl1
Flii
Llgl1
Smcr7
Top3a
Smcr8
Smcr8
Shmt1
Dhrs7b
Dhrs7b
Tmem11
Tmem11
Gm16515
Gm16516
Map2k3
Mouse

Figure 5. **Mouse synteny to chromosome 17p11.2**. A) Human 17p11.2. RefSeq Genes (Build: February 2009, GRCh37/hg19). B) Mouse 11qB1.3-B2. RefSeq Genes (Build: July 2007, NCBI37/mm9). Grey bars between sections A and B represent regions of synteny. Modified from UCSC Genome Browser (http://www.genome.ucsc.edu/)

Recently, Girirajan et al. [43, 44] generated and performed an in-depth behavioral analysis of bacterial artificial chromosome(BAC) transgenic, *Rai1* overexpression mice (*Rai1-Tg*) harboring three copies of the *Rai1* gene. The authors found that these mice exhibited a variety of defects. Growth retardation was seen with these mice being undersized as compared to WT litter mates which continued as development progressed. However, transgenic mice were able to achieve normal size by 20 weeks [43]. Abnormal maternal behavior was observed wherein *Rai1-Tg* mothers could not properly take care of their pups (1/5 survival average, WT or transgenic pups) [44]. When these pups were placed with a WT foster mother (all, 5/5 average) pups survived. Additionally, *Rai1* overexpressing mice exhibit a reduction in the levels of 5-HIAA suggesting a defect in the enzyme monoamine oxidase (MAO), which converts serotonin to 5-HIAA. Further, these mice have altered social behavior with impaired nesting behavior, social dominance, aggression, reduced social memory, hyperactivity, anxiety-related behavior, and altered sociability, all 5-HIAA mediated behaviors.

Taken together, the *Df(11)17/+*, *Rai1 +/-, Dp(11)17/+*, and *Rai1-Tg* mice have proven to be good models for their respective syndromes. The *Rai1+/-* mice have craniofacial, behavioral and developmental abnormalities consistent with those of SMS. Additionally, *Dp(11)17)/+* duplication mice have behavioral, neurological and developmental abnormalities consistent with that of dup17p11.2 syndrome. These data reinforce the dosage sensitivity of the *retinoic acid induced 1* gene. A summary of the phenotypes seen in these mice can be seen in Table 2

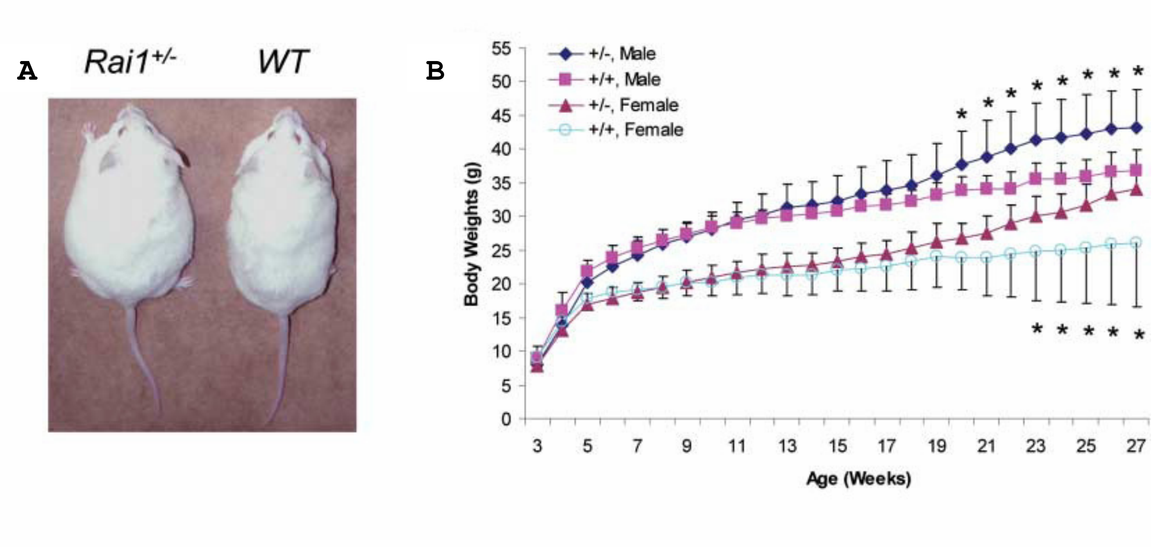


Figure 6. ***Rai1*** $^{+/-}$ **mice growth rates compared to WT litter mate**s. Note obesity and shortened snout. A) *Rai1* $^{+/-}$ and WT size and craniofacial phenotypes. B) Growth rates of *Rai1* $^{+/-}$ mice. Modified from Bi et al. 2005 [42].

Table 2. Phenotypes of *Rai1* engineered mice. Bi et al. 2005, 2007, Waltz et al. 2003, 2004, 2006, Girirajan et al. 2009.

Mouse	*Df(11)17/+* [39, 41]	*Rai1+/-* [42, 45]	*Rai1-/-* [45]	*Dp(11)1)/+)* [39, 41, 46]	*Rai1-Tg* [44]
Feature					
Craniofacial Defect	+	+	+	-	-
Developmental Defect	+	+	+	+	+
Overweight	+	+	-	-	-
Underweight	-	-	+	+	+
Circadian Rhythm Abnormality	+	N/A	N/A	-	N/A
Impaired Conditioned Fear	+	-	+	+	N/A
Hyperactivity	-	-	-	+	+
Seizure	+	+	+	-	-
Abnormal EEG	+	+	+	-	N/A
Abnormal Maternal Behavior	N/A	N/A	N/A	N/A	+
Altered Mendelian Transmission	+	+	-	+	+

"+" indicates presence of feature. "-" indicates absence of feature. "N/A" indicates feature not assessed.

Unexplained intellectual disability

Understanding the molecular basis for intellectual disability continues to be a problematic hurdle to jump in the field of human genetics. ID has a 1-3% prevalence rate in the general population [47, 48] and although there are many well-described ID syndromes associated with known molecular changes such as SMS (del(17)(p11.2)/*RAI1* mutation), Williams syndrome (del(7)(q11.23)), Rett syndrome (*MECP2* mutation), Prader-Willi (paternal del(15)(q11-q13), the molecular basis for ID in many individuals remains unclear.

The major genomic and molecular contributors to ID are aneuploidy, chromosomal microdeletions or microduplications, translocations, inversions, uniparental disomy (UPD), exonic mutation, splice-site or intronic mutation, and epigenetic or imprinting changes. Chromosomal deletions and duplications account for the vast majority of ID, however in up to 80% of all ID cases the molecular change cannot be identified [49, 50, 51]. The most common clinical diagnoses of individuals with ID are Down syndrome (9.2%), microdeletion 22q11.2 (2.4%), Williams–Beuren syndrome(1.3%), fragile-X syndrome (1.2%), Cohen syndrome (0.7%), and monosomy 1p36.3 (0.6%) [52].

New technologies have evolved to help characterize these ID syndromes, including whole genome array comparative genomic hybridization (aCGH), targeted array, high resolution fluorescence *in situ* hybridization (FISH), and high throughput sequencing (both whole genome and whole exome) have allowed for faster, and more accurate molecular diagnosis the probability of a multiple hit theory exists making the majority of ID fall under the "complex" arena of genetic diagnosis [5].

Chapter 2

Identification of loci contributing to the SMS-like phenotype.

Introduction

Smith-Magenis syndrome (SMS, OMIM#182290) is a complex congenital abnormalities/mental retardation disorder caused by a deletion of chromosome 17p11.2 that includes the *RAI1* gene or a mutation in *RAI1* [4]. Phenotypically, SMS overlaps with other syndromes such as Prader-Willi, Williams, and Down syndromes. While certain features overlap between these syndromes, the phenotype associated with SMS is characterized by a specific combination of traits. In short, SMS can be characterized by developmental delays, craniofacial abnormalities, speech and motor delay, neurological abnormalities, sleep disturbance, and self-injurious behaviors [3]. Due to the dense overlap between mental retardation syndromes, it is imperative that the genetic basis for each clinical phenotype is unraveled to provide the most complete information for families, to target the most appropriate services for the child, and to provide the most accurate recurrence risk assessment.

We have acquired a cohort of individuals with clinical features of Smith-Magenis syndrome in whom a deletion or mutation of *RAI1* cannot be identified. We refer to these cases as “Smith-Magenis syndrome-like” (SMS-like), since they are not phenotypically distinguishable from individuals with a molecularly-confirmed *RAI1* mutation or 17p11.2 deletion. Each case is evaluated on an individual basis, and if the subject meets the phenotypic criteria for a clinical diagnosis of SMS, the coding region of the *RAI1* gene is sequenced. Each subject referred for molecular evaluation had a normal karyotype and 17p11.2 FISH analysis, followed by sequencing of the *RAI1* coding region [53, 54]. Of patients referred to our laboratory for *RAI1*mutation or deletion analysis, 46% (n = 52/112) were confirmed to carry a heterozygous deletion or mutation of *RAI1*. Although the remaining cases (n = 60/112) have significant phenotypic overlap with SMS, they do not have a molecular diagnosis. Thus, we hypothesized that at least one additional locus, possibly functioning in a common pathway with RAI1, is contributing to the high frequency of this phenotype. Eight of the remaining cases were not included in this study because of the lack of genetic material for analysis or insufficient clinical data were available for the subject.

Array comparative genomic hybridization (aCGH) has become an invaluable tool in the genomic evaluation of subjects with mental retardation, developmental delay, and congenital anomalies, as well as more complex disorders such as autism and schizophrenia [55, 56]. Using aCGH as our primary tool, we set out to identify genomic regions potentially containing dosage sensitive genes that when deleted or duplicated, lead to an SMS-like phenotype. Analysis of these 52 cases revealed copy number abnormalities of chromosomal regions which contain genes that contribute to neurological integrity, cognition, and development, all of which when disturbed could lead to the phenotypes observed in the SMS and SMS-like groups. Given the phenotypic and genetic information presented herein and the loci identified, these data will improve diagnosis, provide insight into the etiology of mental retardation/congenital abnormalities syndromes, and one day could lead to better treatments and therapy for individuals with these genomic disorders.

Materials and Methods

Subject ascertainment and samples

Subjects were referred to the Elsea laboratory for molecular evaluation of SMS at Michigan State University or Virginia Commonwealth University. Samples were collected in accordance with Institutional Review Board approved protocols from the appropriate institution. Peripheral blood was collected, and DNA and metaphase chromosomes prepared following standard methods. All samples then evaluated for mutations in the *RAI1* gene. Phenotypic information was collected using medical records, geneticist reports, patient photos, and in some cases a clinical checklist sent to the referring geneticist. All information was collected in accordance with IRB approved protocols.

Whole genome array-based comparative genomic hybridization (aCGH)

aCGH was performed as previously described [40]. The procedures for the construction, hybridization, and data analysis is described below.

Construction of the human BAC CGH array**:** DNA printing solutions were prepared from sequence connected RPCI-11 BAC by ligation-mediated PCR as described previously [57,[58,[59]. The minimal tiling RPCI BAC array contains ~19,000 BAC clones that were chosen by virtue of their STS content, paired BAC end-sequence and association with heritable disorders and cancer. The backbone of the array consists of ~4600 BAC clones that were directly mapped to specific, single chromosomal positions by fluorescent in situ hybridization (FISH) [57]. Each clone is printed in duplicate on amino-silanated glass slides (Schott Nexterion typeA+) using aMicroGrid ll TAS arrayer (Genomic Solutions, Inc.). The BAC DNA products have ~80 µm diameter spots with 150µm center to center spacing creating an array of ~39,000 elements. The printed slides dry overnight and are UV-crosslinked (350mJ) in a Stratalinker 2400 (Stratagene) immediately before hybridization. A complete list of the RPCI-11 BAC clones spotted on the 19K array can be found at: http://microarrays.roswellpark.org.

Labeling and hybridization of DNA: Reference and test sample genomic DNAs (1 µg each)are individually fluorescently labeled using the BioArray CGH Labeling System (Enzo Life Sciences). Initially, DNA is denatured in the presence of a random primer at 99° C for 10 min in a thermocycler, and then quickly cooled to 4°C. The tubes are transferred to ice and labeling occurs with the addition of dNTP-cyanine 3 mix (or dNTP-cyanine 5) and Klenow fragment. Incubation takes place for 4 h at 37°C in a thermocycler. The unincorporated nucleotides were removed using a QIAquick PCR purification column (Qiagen), and the labeled probe is eluted with 2 x 25 ul washes. Prior to hybridization, the test and reference probes were combined with 100 ug human Cot-1 DNA (Invitrogen) and precipitated for one hour with sodium acetate and ethanol. The probes are pelleted, resuspended in 110 µl SlideHyb Buffer #3 (Ambion) containing 5 µl of 100 µg/µl yeast tRNA (Invitrogen), heated to 95°C for 5 min, then incubated at 37°C for 30 min. Hybridizations to the 19K BAC arrays were performed for 16 h at 55°C using a GeneMachine hybridization station (Genomic Solutions, Inc.) as described [60]. After hybridization, the slides are automatically washed in the GeneTAC station with reducing concentrations of SSC and SDS. aCGH was also performed under similar conditions as a fee for service at Empire Genomics using the same BAC CGH array as previously described [40].

Image and data analyses: The hybridized aCGH slides are scanned using a GenePix 4200AL Scanner (Molecular Devices) to generate high-resolution (5 µm) images for both Cy3 (test) and Cy5 (control) channels. Image analysis was performed using the ImaGene (version 8.0.0) software from BioDiscovery, Inc. The log2 test/control ratios were normalized using a sub-grid loess correction. Mapping information was added to the resulting log2 test/control values. The mapping data for each BAC are found by querying the human genome sequence at http://genome.ucsc.edu, and BACs in regions of segmental duplication or large-scale variation are flagged.

Fluorescence *in situ* hybridization

In order to determine the presence of a 17p11.2 deletion (which would indicate Smith-Magenis syndrome), metaphase chromosomes were prepared for each patient in whom clinical FISH had not been previously reported. *RAI1* FISH was performed as previously described [53]. A minimum of 10 metaphase spreads and 10 interphase nuclei were observed for each case evaluated.

Sequencing of *RAI1*

Overlapping *RAI1* primers covering the entire coding region (exons 3-6, NM_030665) and the intron-exon junctions were designed to PCR-amplify the patient DNA samples. Sequencing of *RAI1* was performed as previously described [34].

Bioinformatic and statistical analyses

Deleted or duplicated BACs were located on the UCSC Genome Browser (http://www.genome.ucsc.edu/), Human Mar. 2006 assembly, and genomic position determined for BACs that flanked the deletion or duplication (and were included in the aberration). Prevalence of features in SMS and SMS-like cohorts were analyzed by Fisher's exact test utilizing a web-based program located at http://www.langsrud.com/fisher.htm. P-values ≤ 0.05 were considered statistically significant.

Results

SMS vs. SMS-like phenotypes

We have assessed the phenotypes of 52 SMS-like cases and compared the overall features of this cohort to those compiled in a meta-analysis of all the known and reported SMS cases (Table 3) [18]. As evident in the SMS-like cohort, the most common phenotypic features are not different in prevalence between this group and the SMS cohort (Table 3). We propose that even though *RAI1* is the major contributor to SMS, other players in the "SMS pathway", when duplicated or deleted, result in a similar phenotype. The SMS and SMS-like cohorts are very similar (Table 3). Sleep disturbance, developmental delay, self-injurious behavior, hyperactivity, oral/motor dysfunction, hypotonia, dental anomalies, chronic ear infections, digestive problems, obesity, and neurological abnormalities were not different in prevalence between the two groups and were present in the majority of subjects (Table 3). These commonalities led us to hypothesize that at least one other locus exists and contributes to the global SMS phenotype and/or disruption of a common pathway leads to this phenotype.

Whole genome array comparative genomic hybridization

Once this SMS-like cohort was collected and all testing was negative for classical SMS by molecular techniques (see Methods), DNAs were then evaluated by whole genome array comparative genomic hybridization using a 19,000 RPCI-11 BAC array. We identified 15 copy number variants (CNVs) not previously described in these subjects (Table 4, 5). These variants occur in 21% (11 out of 52) of our population. Within these 15 CNVs, 8 chromosomes are represented, and similar variants are found in multiple cases (Table 4, 5). These CNVs are of

Table 3. Phenotypic features observed in Smith-Magenis syndrome-like cohort.

Common feature	**Cohort**		
	SMS[1] (del17p11.2 or *RAI1* mut)	**SMS-like**[2] (negative for *RAI1* del or mut)	
Developmental delay	100%	100%	
Stereotypies	59%	84%	*
Self-injury	90%	90%	
Sleep disturbance	88%	97%	
Self-hugging/tics	63%	55%	
Hyperactivity	74%	77%	
Seizures	28%	56%	*
Hypotonia	71%	72%	
Signs of peripheral neuropathy	86%	72%	
Oral/motor dysfunction	85%	70%	
Craniofacial abnormalities	89%	70%	*
Eye abnormalities	87%	59%	*
Ear abnormalities	63%	64%	
Hearing loss	60%	39%	
Ear infections	90%	69%	*
Dental anomalies	54%	73%	
Cleft lip/palate	15%	2%	*
Hoarse voice	86%	50%	*
Overweight	33%	53%	
Hypercholesterolemia	41%	23%	
Heart defect	31%	11%	*
Digestive problems	95%	67%	*
Short stature	67%	43%	*
Brachydactyly	79%	47%	*
Scoliosis	32%	34%	

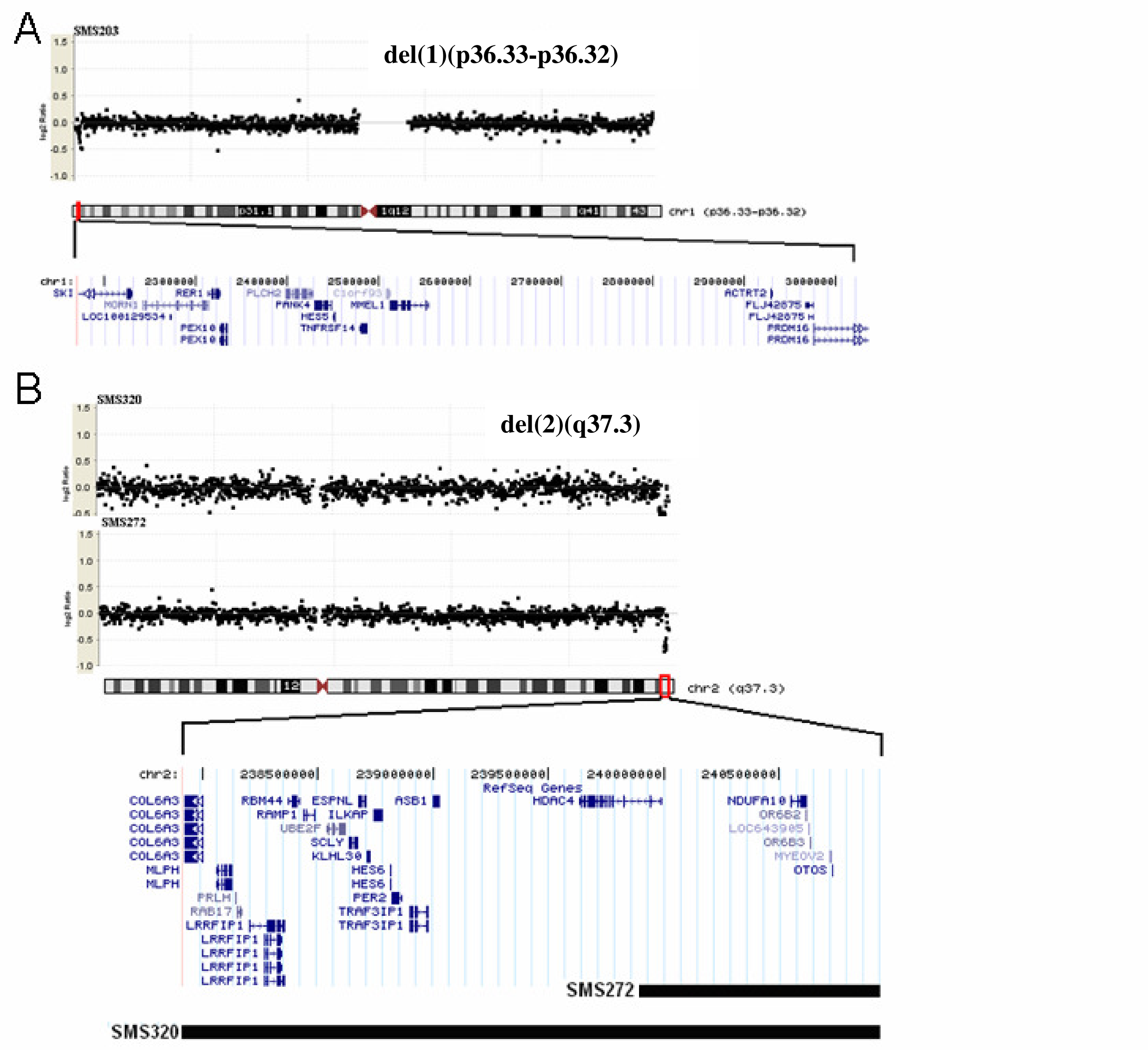
[1] Values for SMS cohort taken from Edelman et al 2007 (n=105). [2] Cases reported in this study (n=52). Statistics calculated with Fisher's Exact Test where *=p-value <0.05. Range of age at evaluation: 2-44 years for SMS-like cases and 4 months -72 years for SMS cases (Edelman et al 2007 [18]).

particular interest because of the potential to further define regions of variation that contribute to the SMS-like phenotype, given common genes that are either deleted or duplicated. Common and possibly polymorphic BACs were excluded by referencing a control group containing 372 individuals from the National Institute of Mental Health (NIMH) Genetics Initiative as control samples [61].

Pathogenic copy number variants

Of the 15 CNVs identified in this cohort, some have been previously described but none have been reported in cases with phenotypic overlap to SMS. Described below are the CNVs believed to be pathogenic and thus, contribute directly to the observed phenotype (Table 4).

- Monosomy 1p36.32-p36.33, identified in SMS203 (Figure. 7, Table 4), is associated with mental retardation, developmental delay, hearing impairment, seizures, growth delay, hypotonia, hyperphagia, brachydactyly, self-abusive behavior, obesity, and a Prader-Willi-associated phenotype [62, 63]. SMS203 is a 44 year old adult female referred for lifelong sleep difficulties and learning and behavioral problems who presented with short stature, obesity, prognathism, dental abnormalities, brachydactyly, scoliosis, eye abnormalities, chronic ear and respiratory infections, and self-injurious behavior. A strong overlap between phenotypes can be seen between both 1p36.33 syndrome and SMS; however, the clinical overlap between these well-described syndromes has not previously been appreciated.
- A single case of del(2)(q23.1) (Figure 7, Table 4) was identified in this cohort. This deletion was previously reported in association with severe psychomotor retardation, speech impairment, epilepsy, microcephaly, ataxia and behavioral disabilities, and described as a pseudo-Angelman


A
SMS203
del(1)(p36.33-p36.32)
log2 Ratio
chr1 (p36.33-p36.32)
chr1:
SKI
MORN1
LOC100129534
RER1
PEX10
PEX10
PLCH2
PANK4
HES5
TNFRSF14
C1orf93
MMEL1
ACTRT2
FLJ42875
FLJ42875
PRDM16
PRDM16
B
SMS320
del(2)(q37.3)
SMS272
log2 Ratio
chr2 (q37.3)
chr2:
RefSeq Genes
COL6A3
MLPH
PRLH
RAB17
LRRFIP1
RBM44
RAMP1
UBE2F
ESPNL
ILKAP
SCLY
KLHL30
HES6
PER2
TRAF3IP1
ASB1
HDAC4
NDUFA10
OR6B2
LOC643905
OR6B3
MYEOV2
OTOS
SMS272
SMS320

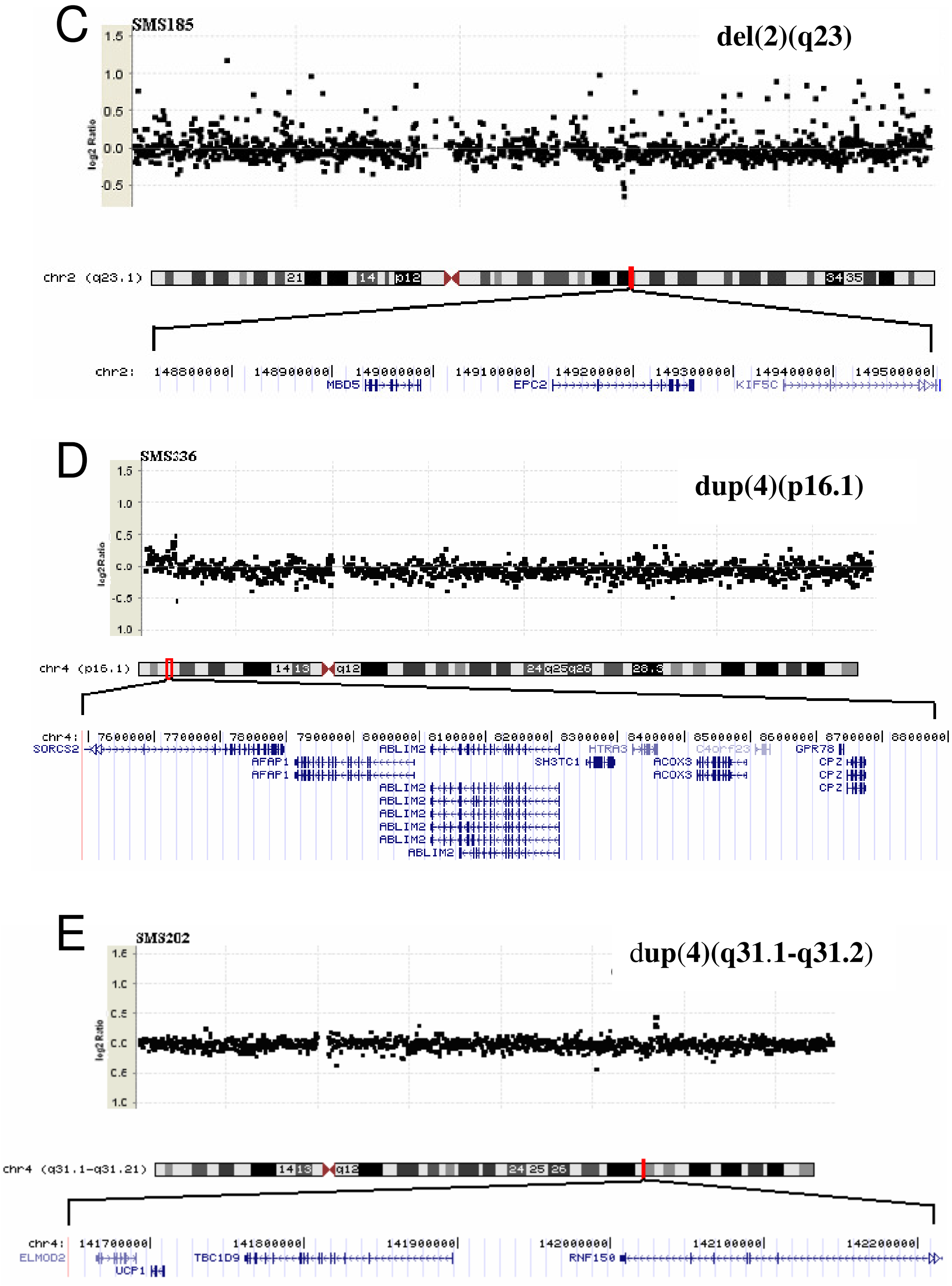

C
SMS185
del(2)(q23)
log2 Ratio
1.5
1.0
0.5
0.0
-0.5
chr2 (q23.1)
21
14
p12
34
35
chr2:
148800000|
148900000|
149000000|
149100000|
149200000|
149300000|
149400000|
149500000|
MBD5
EPC2
KIF5C
D
SMS336
dup(4)(p16.1)
log2Ratio
1.5
1.0
0.5
0.0
-0.5
1.0
chr4 (p16.1)
14
13
q12
24
q25
q26
28.3
chr4:
7600000|
7700000|
7800000|
7900000|
8000000|
8100000|
8200000|
8300000|
8400000|
8500000|
8600000|
8700000|
8800000|
SORCS2
AFAP1
AFAP1
ABLIM2
SH3TC1
HTRA3
ACOX3
ACOX3
C4orf23
GPR78
CPZ
CPZ
CPZ
ABLIM2
ABLIM2
ABLIM2
ABLIM2
ABLIM2
ABLIM2
E
SMS202
dup(4)(q31.1-q31.2)
log2 Ratio
1.5
1.0
0.5
0.0
0.5
1.0
chr4 (q31.1-q31.21)
14
13
q12
24
25
26
chr4:
141700000|
141800000|
141900000|
142000000|
142100000|
142200000|
ELMOD2
UCP1
TBC1D9
RNF150

Figure 7. CGH array plots for pathogenic CNVs A) SMS203, del(1)(p36.32-36.33). B) SMS272 and SMS320, del(2)(q37.3) C) SMS185, del(2)(q23.1) D) SMS236, dup(4)(p16.1) and E) SMS202, dup(4)(q31.1-q31.2). Log2 ratio of subject to control shown on y-axis. CNVs were considered significant for duplication or deletion, when flanking BACs held values >0.3 or <-0.3, respectively, with 3 consecutive BACs involved in a given abnormality. Plots were created using aCGH viewer (http://falcon.roswellpark.org/aCGHview/). Genomic data illustrating the extent of each deletion or duplication were modified from UCSC Genome Browser (http://www.genome.ucsc.edu/) and RefSeq genes are indicated. Figure not to scale.

Table 4. Pathogenic copy number variants identified in the SMS-like cohort.

Case #	Copy Number Variant (CNV)	Genomic Location (bp)	Size (Mb)
SMS203	del(1)(p36.32p36.33)*	Chr1:2,171,401-3,037,274	0.86
SMS185	del(2)(q23.1)*	Chr2:148,543,058-149,504,469	0.96
SMS272	del(2)(q37.3)$^+$	Chr2:239,764,593-240,938,545	1.17
SMS320	del(2)(q37.3)*	Chr2:237,920,956-240,938,547	3.02
SMS336	dup(4)(p16.1)*	Chr4:7,493,733-8,829,348	1.34
SMS202	dup(4)(q31.1q31.2)*	Chr4:141,648,098-142,218,495	0.57

Deletions contained flanking BACs below -3.0 $\log_2$ ratio. Duplications contained flanking BACs above 3.0 $\log_2$ ratio. CNV location taken from UCSC genome browser, Mar. 2006 build, with flanking BACs as reference. CNV confirmation by: *clinical BAC array, $^+$G-Banding.

phenotype [64]. SMS185 is a female who presented at age 3.5 y with mental retardation, short stature, midface hypoplasia, microcephaly, brachydactyly, hoarse voice, heart abnormalities, and hyperactivity but did not exhibit sleep disturbance or self-injurious behavior. Previous methylation testing for Angelman syndrome was negative.

- SMS202 was shown to have a duplication of chromosome 4q31.1-q31.21 (Figure 7, Table 4), which has been associated with mild mental retardation and poor language acquisition [65], features commonly seen in SMS. SMS202 is a 6 year
old female with mental retardation, craniofacial abnormalities, 2-3 toe syndactyly, pes planus, elbow limitations, sleep disturbance, self-injurious behavior, hyperactivity, and stereotypical behaviors.
- SMS336 is a 3 year old male who carries a novel CNV, dup(4)(p16.1) (Figure 7, Table 4), not previously reported in the literature. This region contains 8 known genes and 1 predicted gene, including the VPS10 domain containing receptor 2 gene (*SORCS2),* which is highly expressed in the mouse central nervous system. SMS336 presented with a typical SMS phenotype, including developmental delays, cognitive impairment, short stature, craniofacial, dental, and eye abnormalities, scoliosis, hypotonia, difficulty communicating, hearing loss, chronic ear and respiratory infections, decreased sensitivity to pain, heart abnormalities, obesity, sleep disturbance, hyperactivity, and self-injurious behaviors. This individual also carries a del(14)(q11.2) [66] (Figure 9, Table 6), which is a reported polymorphic CNV.

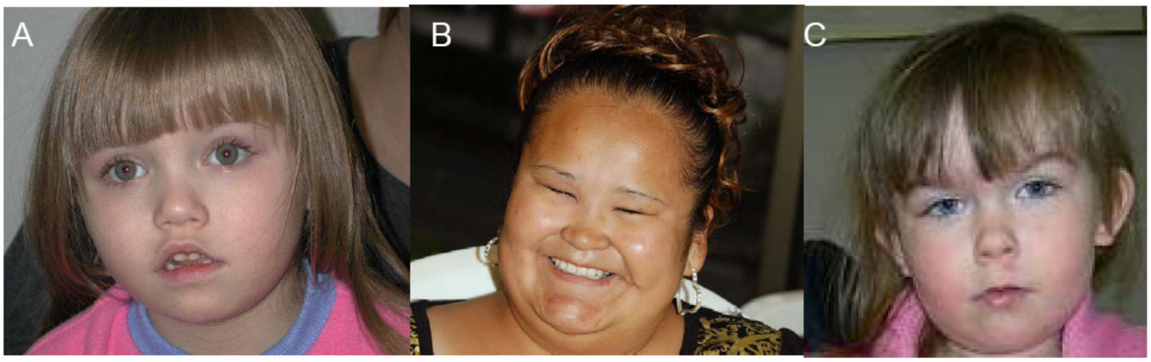


Figure 8. del2q37 clinical phenotype. A) SMS185 with del(2)(q23.1), age 2 years; B) SMS272 with del(2)(q37.3), age 15 years; and C) SMS320 with del(2)(q37.3), age 3 years.

- A deletion involving 2q37.3 was identified in 2 cases (Figure 7, 8, Table 4):
    - o SMS272 is a 15 year old female with midface hypoplasia, brachycephaly, and brachydactyly (Figure 8b). Her neurobehavioral phenotype consists of stereotypies, and aggressive behavior. She is also obese but does not exhibit sleep disturbance or self-abusive behavior, both of which are typically seen in SMS. A known polymorphic CNV, del(16)(p11.2), was also identified in this case (Figure 7c, 9, Table 6).
    - o SMS320 is a 2 year old female with mental retardation, speech and motor delay, sleep disturbance, stereotypies, attention-seeking, and self-injurious behaviors (Figure 8c). Her craniofacial and skeletal phenotype consists of brachydactyly, brachycephaly, midface hypoplasia, tented upper lip, broad, square face, and synophrys.
- A maternally inherited duplication of 17p12 was identified in SMS348 (Figure 9, Table 6), a 10 year old male who presented with typical features of SMS. Charcot-Marie Tooth disease type 1A (CMT1A) is caused by dup(17)(p12) that includes the *PMP22* gene. The associated demyelinization of the peripheral nervous system is well-described and does not involve developmental delays or sleep disturbance [67, 68]. This finding in SMS348 (Figure 9, Table 6), is not causative of the child’s developmental phenotype, with the possible exception that in severe cases of CMT1A, peripheral neuropathy may be evident, even in a young child. However, del(17)(p12) was recently shown to confer a 10-fold increased risk for schizophrenia [69] indicating that gene dosage may predispose an individual to a mild increase in risk (Table 5). Dup(17)(p12) was also identified in an aCGH study involving 100 mentally retarded individuals [70].

Table 5. Disorders previously associated with the deletions and duplications identified in this study.

CNV	Associated phenotype
del(1)(p36.33)	Obesity, hyperphagia, dysmorphism, 1p36 deletion syndrome [62,[63]
del(2)(q23.1)	Pseudo-Angelman syndrome [64]
del(2)(q37)	Brachydactyly-mental retardation syndrome/AHO-like [71]
del(8)(p23.1)	Cornelia de Lange alternate locus [72]
dup(15)(q11.2)	Autism [73,[74]
dup(17)(p12)	Schizophrenia [70]/autism [75]/mental retardation [70]/Charcot-Marie-Tooth disease Type 1A [67,[68]

Non-pathogenic copy number variants

Several previously reported non-pathogenic CNVs were detected in this cohort as described in Table 6. Two of these CNVs were detected in more than one case but are within regions of known polymorphisms and thus, not likely to be pathogenic (Figure 9, Table 6). The 8p23.1 region of deletion contains a cluster of defensin genes and *SPAG11B*. While these host defense genes are not excellent candidates for the syndromes described, predicted genes within the deletion region and genes just outside the deleted region, possibly influenced by position effects, cannot be ignored. Additionally, SMS279 also carries a dup(16)(p11.2) which flanks the reported region associated with autism, neurodevelopmental delay, and dysmorphism [76].

The 2p11.2 region, deleted in 2 cases in this cohort (Figure 9, Table 6), includes several genes that hold potential importance for antibody variation, including *IgA*, *IGVK-A2*, and *IGV3-15*. Position effects could also be altering gene expression in the regions just outside the common deletion, as possible candidates include *FABP1* which is associated with age-dependant obesity in female mice [77] and *FOXI3* which is important in craniofacial development in zebrafish [78].

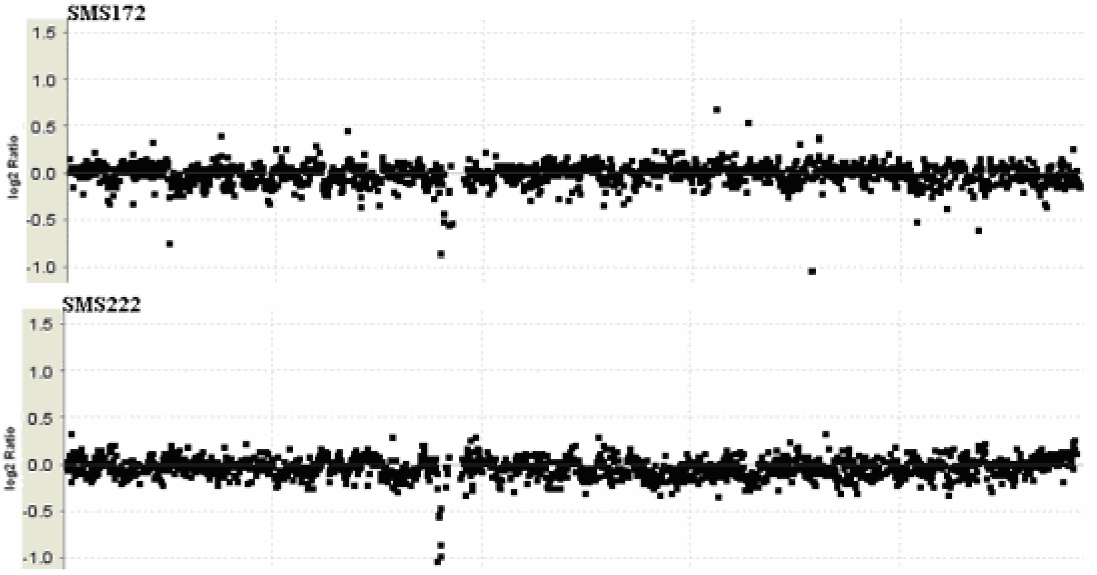
SMS272, who carries a del(2)(q37), also carries a del(16)(p11.2) distal to the recently reported autism-susceptibility region (Figure 9, Table 6) [79, 80]. A dup(15)(q11.2) was found in SMS202, who also carries a dup(4)(q31.1-q31.2). This deletion region has been associated with autism and may contribute to combinatorial effects along with the dup(4)(q31.1-31.2) that is thought to be pathogenic.

Table 6. Non-pathogenic copy number variations identified in the SMS-like cohort.

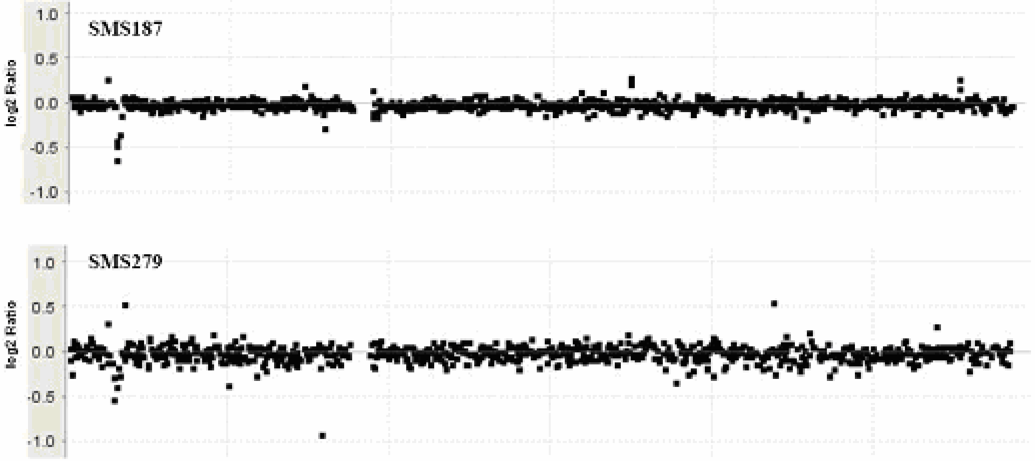
Case #	Copy Number Variation	Genomic Location (bp)	Size (Mb)
SMS172	del(2)(p11.2)	Chr2:89,068,208-91,746,640	1.89
SMS222	del(2)(p11.2)	Chr2:89,068,208-90,016,961	0.95
SMS272	del(16)(p11.2)	Chr16:32,456,421-33,593,137	1.14
SMS336	del(14)(q11.2)	Chr14:21,636,643-22,120,718	0.50
SMS202	dup(15)(q11.2)	Chr15:18,617,196-18,882,916	0.27
SMS187	del(8)(p23.1)	Chr8:7,156,827-7,936,168	0.78
SMS279	del(8)(p23.1) dup(16)(p11.2)	Chr8:7,156,827-7,936,168 Chr16:31,855,107-33,295,070	0.78 1.44
SMS348	dup(17)(p12)#	Chr17:14,080,569-15,223,450	1.14

Deletions contained flanking BACs below -3.0 log2 ratio. Duplications contained flanking BACs above 3.0 log2 ratio. CNV locations taken from UCSC genome browser, Mar. 2006 build, with flanking BACs as reference. #Mother also affected with CMT1a.

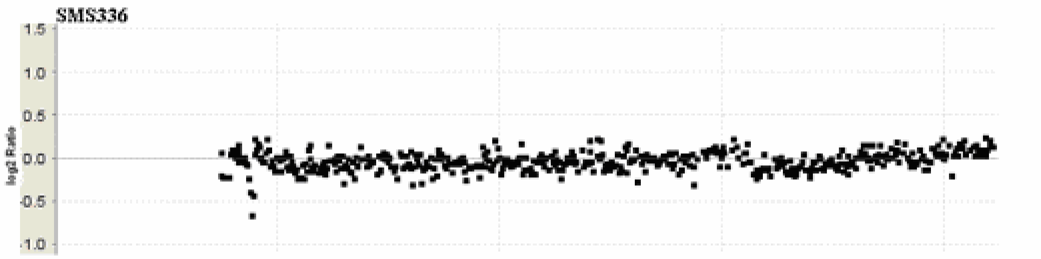
Del(2)(p11.2)



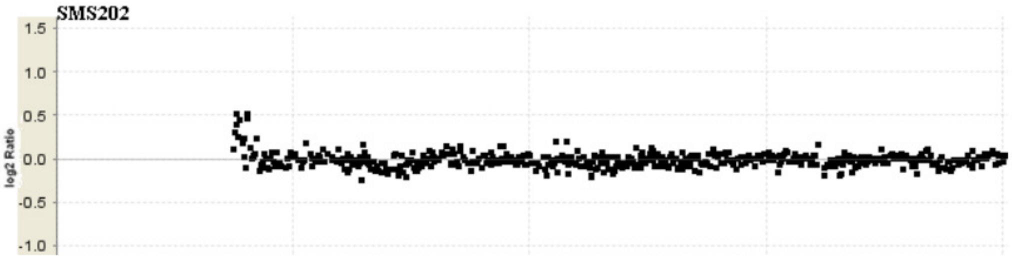
Del(8)(p23.1)



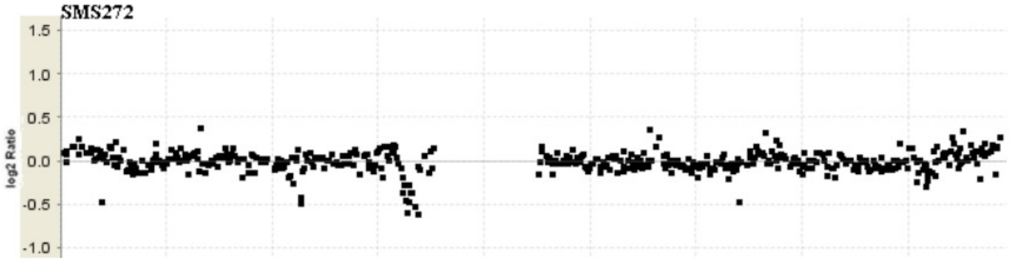
Del(14)(q11.2)



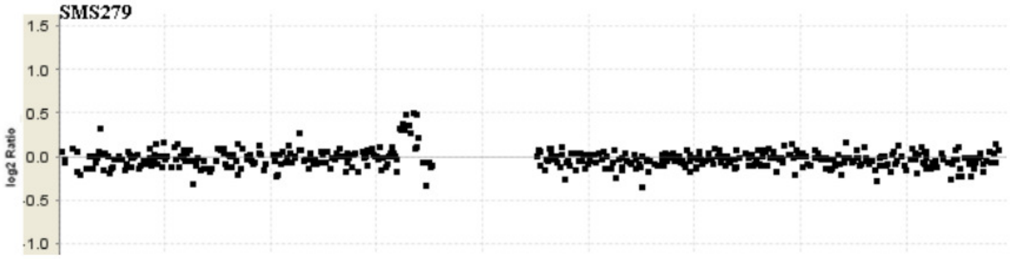
Dup(15)(q11.2)



Del(16)(p11.2)



Dup(16)(p11.2)



Dup(17)(p12)

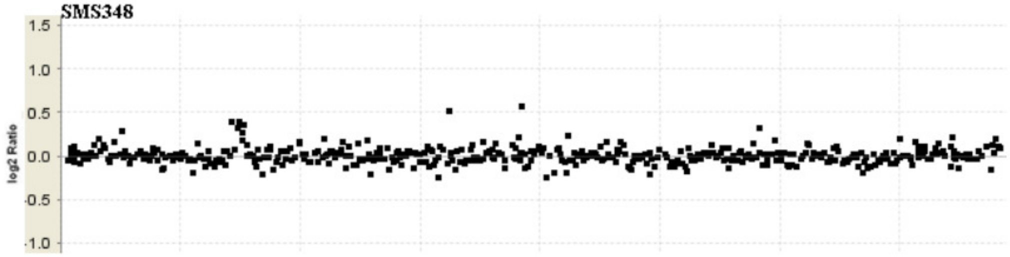


Figure 9. Non-Pathogenic copy number variants. Array plots for each subject are shown. Log2 Ratio of subject to control shown on y-axis. Plots were created using aCGH viewer http://falcon.roswellpark.org/aCGHview/). Genomic data illustrating the extent of each case deletion or duplication were modified from UCSC Genome Browser (http://www.genome.ucsc.edu/). Figure not to scale.

Discussion

In this study, we utilized whole genome aCGH to assess a cohort of 52 clinically well characterized patients referred to our laboratory for molecular evaluation of Smith-Magenis syndrome. We identified duplications or deletions (15 CNVs in 11 subjects) not previously described in these cases, with CNVs in 6 cases that are likely pathogenic and associated with the SMS-like phenotype present in the patient (Tables 4, 6). Some of the same genomic abnormalities can be found in multiple cases, supporting a common etiology for the similar phenotypes. These chromosomal regions include genes that are responsible for fatty acid metabolism, neurological development, cell signaling, circadian rhythm, transport, development, mental health, neuromuscular function, morphology, and cell cycle control. Because of the phenotypic overlap between the SMS and SMS-like cases, a possible indirect role of *RAI1* in the etiology of these phenotypically similar disorders should be explored [81]. Data from this study may give us some insight into the pathways in which RAI1 functions. RAI1 is a putative transcription factor that spans ~130 Kb on 17p11.2 and contains 6 exons, 4 of which are coding. Bioinformatic analysis of RAI1 has identified a bipartite nuclear localization signal and a PHD domain, both of which are consistent with the proposition that RAI1 functions as a transcription factor. Although the cases were ascertained based on stringent SMS-phenotypic criteria, 52 cases from our cohort neither carried a 17p11.2 lesion nor any pathogenic CNVs. It is possible that these cases carry alterations in the yet-uncharacterized regulatory region of *RAI1* or a point mutation in another critical gene.

We have also identified multiple variants that are associated with reported syndromes (Table 5). This fact speaks to two points: one, there are likely common pathways that lead to the phenotypic overlap between SMS and these known syndromes and two, better delineation of the phenotypes and training for diagnosis of these syndromes is needed. In addition, there were CNVs that were missed by clinical arrays that did not cover these regions, which supports the fact that diagnostic arrays need better coverage over regions that are associated with known syndromes.

One region of variation, del(2)(q37.3), was found in more than one subject in this cohort (Figure 9, Table 4). Deletion 2q37 syndrome, also known as brachydactyly-mental retardation syndrome or Albright's hereditary osteodystrophy-like syndrome, is well-described and characterized by craniofacial abnormalities, mental retardation, stereotypies, aggressive/self-injurious behavior, distinct brachydactyly, short stature, obesity, eczema, and sparse hair [82, 83]. Although the typical craniofacial and skeletal abnormalities previously reported in association with deletion 2q37 syndrome do not overlap with SMS [16], the neurological, behavioral, and growth abnormalities are similar [18, 24, 34]. Further, SMS272 and SMS320 exhibit the characteristic brachydactyly type E seen in del(2)(q37) syndrome, which is different than the brachydactyly commonly seen in SMS [71, 84]. Both of these subjects also exhibit similar craniofacial abnormalities as those in SMS, as well as many of the other features. These data emphasize not only the phenotypic overlap between these two syndromes but also expand the phenotype of the well-described 2q37 deletion syndrome. These phenotypic and molecular overlaps may lead us to candidate genes and pathways that contribute to these complex features.

Several excellent candidate genes have been identified and warrant further study. Potential candidate genes localized to the overlapping deleted region in 2q37.3 include *HES6,* which has been implicated in mRNA transcription regulation and neurogenesis [85], as well as *SCLY* which is implicated in fatty acid metabolism and *HDAC4* which is involved in chondrogenesis. Given the significant overlapping behavioral and physical phenotype, we propose that one or more genes present in del(2)(q37) syndrome fall in the same pathway(s) as those leading to SMS.

The 1p36.32-p36.33 region deleted in SMS203 contains more than 10 genes. Deleted in this region is *HES5,* a gene important in neuronal development that when reduced in expression leads to premature depletion of neuronal stem cells [86], which could also lead to some of the neurological phenotypes seen in this patient.

On chromosome 2q23.1, the *MBD5* gene was found to be deleted in a single case using SNP array-based technology to assess subjects with mental retardation [1]. MBD5 is a methyl-CpG binding domain protein. Interestingly this was the only gene deleted in this patient who presented with mental retardation, seizures, retarded motor development, and limited social interactions, similar to the phenotype of SMS185 in this study.

Duplicated in the chromosome 4p16.1 region (in case SMS336), *SORCS2* is expressed throughout the murine central nervous system, including the olfactory bulb, piriform cortex, amygdala, hippocampus, median geniculate nucleus, interpeduncular nucleus, and some cerebellar Purkinje cells [87]. This gene is thought to be involved with intracellular sorting, and altered dosage may contribute to some of the neurological findings such as hypotonia, decreased sensitivity to pain, sleep disturbance, self-abusive behavior, hyperactivity, and attention deficit disorder.

UPC1 is duplicated in SMS202, who carries a dup(4)(q31.1-q31.3). Deletion of this gene has been shown to cause obesity in mice, whereas duplication of *UPC1* is thought to have a protective effect [88]. SMS202 is not overweight; however, because of the *UPC1* pattern of expression in the central nervous system [89], a potential dosage effect should be considered regarding neurological abnormalities.

Candidate gene sequencing in the subjects wherein no CNV was found is a logical next step toward identifying causative genes for the syndromes mapped to the genomic regions discussed above. Further, the documented interaction of RAI1 with candidate genes will aid in the delineation of molecular pathways causative of the phenotypes described. In conclusion, this study is unique in that we have taken a group of patients clinically diagnosed with a very specific phenotype, Smith-Magenis syndrome, and identified potentially alternate causative loci. These findings will help to improve the diagnosis of individuals with phenotypic overlap to SMS, including those with known syndromes that previously were not known to overlap phenotypically, such as del(2)(q37) and del(1)(p36) syndromes. The results presented will help to further improve diagnosis (both molecular and phenotypic) and potentially treatment of microdeletion/microduplication syndromes and will provide insight into the molecular pathways involved in these phenotypically complex disorders.

We propose that, together, Smith-Magenis syndrome and the cases described above, who do not have SMS, encompass a group of disorders that are molecularly related and that patients with developmental delays, self-injurious behaviors, sleep disturbance, stereotypies, and sudden violent outbursts who would otherwise be clinically appropriate for a diagnosis of Smith-Magenis syndrome be considered for whole genome (not targeted) aCGH in order to identify the molecular etiology of the patient's phenotype.

Chapter 3

Haploinsufficiency of *MBD5* associated with a syndrome involving microcephaly, intellectual disabilities, severe speech impairment, and seizures

Introduction

Owing to the phenotypic complexity of many genomic disorders and the fact that most of these disorders result in developmental delays and behavioral problems, overlapping physical and behavioral features between syndromes are commonly observed [90]. Although there are clear phenotypic differences between most syndromes, Smith–Magenis, Down, and Prader–Willi syndromes have many major features in common, including craniofacial features, obesity, and hypotonia, whereas Angelman and Rett syndromes feature seizures and severe developmental delays. Correct identification and diagnosis of specific syndromes are more difficult because of these phenotypic similarities. Ultimately, when comparative genomic hybridization (CGH) data are considered and new syndromic regions of the genome are identified, sorting these cases by phenotype alone may be even more difficult. Microdeletion of chromosome 2q23.1 results in a novel syndrome previously reported in five individuals [64, 70, 91, 92]. In this study, we report two new cases with overlapping 2q23.1 deletions and draw attention to possible candidate genes that could be causative for the characteristic phenotype observed in our cases and in the other five patients reported in literature. A comparison of phenotypes to the other five known cases revealed similar features and confirmed a consistent phenotypic pattern associated with 2q23.1 microdeletion. Although this new microdeletion syndrome has certain phenotypic similarities to various known disorders such as Angelman, Smith–Magenis, and Rett syndromes, it is not a classic phenocopy for any of these conditions. On the basis of our evaluation of the spectrum of clinical features in our cases and in previously reported ones, we have further delineated the clinical characterization of this new microdeletion syndrome. On the basis of the similar features observed, we established a clinical characterization of the novel microdeletion syndrome. Furthermore, an analysis of all reported cases involving a deletion of 2q23.1 revealed a minimal critical region consisting of a single gene, *MBD5*, as the only deleted gene in one case, and *MBD5* and *EPC2* deleted in all other cases. In this study, we report two new cases of the 2q23.1 deletion syndrome, describe the syndrome phenotype, define the minimal critical region, and analyze the expression of critical region genes toward identification of the causative gene(s) for the disorder.

Materials and Methods

Patient ascertainment

Subjects were referred for molecular evaluation of the Smith–Magenis syndrome (SMS185) and the Prader–Willi syndrome (SMS361). The Institutional Review Boards at Virginia Commonwealth University and Michigan State University approved this study.

Cell lines

Lymphoblastoid cell lines (Epstein–Barr virus-transformed human lymphocytes) from SMS185, her mother (SMS184), her father (SMS183), and from an unaffected control were cultured in RPMI-1640 with 2 mM L-glutamine and 25 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethane-sulfonic acid) (Invitrogen, Carlsbad, CA USA), 10% fetal bovine serum, and 1% penicillin/streptomycin antibiotic-antimycotic solution (Invitrogen) at 37°C in a humidified 5% CO_2 chamber. Cells were grown at the same time period for a span of 3 weeks and supplemented with fresh media every 2–3 days as needed. Cells were monitored for growth using bright field microscopy and were counted using Trypan blue exclusion with a hemocytometer. After a 3-week growth period, the cells were counted, pelleted, and used immediately for RNA extraction or stored at -80°C for RNA extraction at a later date.

RNA extraction

RNA was extracted from either 5–10x10^6 lymphoblast cells using the standard TRIzol protocol (Invitrogen) or from 4 ml whole blood, as described here. The whole blood was mixed with 46 ml Puregene RBC lysis solution, mixed, and let to stand at room temperature for 15 min. White blood cells were pelleted at 600g for 10min at 41°C. The pellet was resuspended in 1 ml RBC lysis solution and incubated at room temperature for 5min. Cells were pelleted at 3000 RPM at 41°C for 2min resuspended in 1 ml cold PBS, and further pelleted at 3000 RPM at 41°C for 2 min. RNA was then isolated using the standard TRIzol protocol (Invitrogen). The concentration and purity of RNA were measured at an absorbance of 260 and 280 nm. RNA was stored at -80°C until ready for use.

Real-time qRT-PCR

First-strand cDNA synthesis using SuperScript II RT (Invitrogen) was prepared with 3 mg of total RNA using Olgio(dT)12-18, and optional RNase OUT treatment, followed by RNA degradation with 2.5 units RNase H (Applied Biosystems Inc., Foster City, CA, USA). For quantitative real-time PCR, predesigned assays on Demand Gene Expression Products (Fermentus, Glen Burnie,MD USA) Taqman MGB probes for MBD5, EPC2, KIF5C, and GAPDH were used (Applied Biosystems Inc.). GAPDH was used as the endogenous control. All samples of cDNA were run in triplicate in 10 ml reaction volumes. Taqman Universal PCR Master Mix (Applied Biosystems Inc.), the probe, and deionized water were mixed together in a fixed ratio, and 8 ml was added to each well. Diluted cDNA (1:5) was then added to each well. PCR conditions were the Standard 7500 Run mode of the ABI Prism 7900 HT Sequence Detection System (Applied Biosystems Inc.). The cycle threshold (CT) was determined during the geometric phase of the PCR amplification plots, as recommended by the manufacturer. Relative differences in transcript levels were quantified using the DDCt method. Acquired data were analyzed using 7500 Fast System SDS Software (Applied Biosystems Inc.).

RAI1 sequencing

RAI1 sequencing was performed as previously described [34] for SMS185. A novel familial polymorphism was found, but no pathogenic mutations were identified in this case. Peripheral blood was collected, from which DNA was prepared immediately or from cell lines established previously. DNA was extracted following standard phenol chloroform methods.

Whole-genome array comparative genomic hybridization (aCGH)

BAC aCGH was performed as previously described [40, 93].

Oligonucleotide array CGH

Oligonucleotide-based microarray analysis was performed for SMS361 using a 105K-feature whole-genome microarray (SignatureChip Oligo Solution made for Signature Genomic Laboratories by Agilent Technologies, Santa Clara, CA,USA) as previously described [94]. Oligonucleotide-based microarray analyses were performed using a 244K feature whole-genome microarray made for Signature Genomic Laboratories by Agilent Technologies to characterize the extent of the abnormalities in SMS185 and SMS361. Genomic DNA labeling was performed as described for BAC arrays, whereas array hybridization and washing were performed as specified by the manufacturer (Agilent Technologies). Arrays were scanned using an Axon 4000B scanner (Molecular Devices, Sunnyvale, CA, USA) and analyzed using Agilent Feature Extraction software v9.5.1 and Agilent CGH Analytics software v3.5.14. Results were then displayed using custom oligonucleotide array CGH analysis software (Oligoglyphix, Signature Genomic Laboratories, Spokane, WA, USA).

Fluorescence *in situ* hybridization (FISH) analysis

Metaphase FISH was performed using bacterial artificial chromosome (BAC) clone RP11–951G8 from the 2q deleted region as previously described [95]. FISH data are shown in Figure 12.

RESULTS

Case reports

Case 1 (SMS185)

SMS185 is a 13-year-old Caucasian girl first seen for a clinical genetic evaluation when she was 4 years old because of concerns regarding developmental delays, short stature, and microcephaly (Figure 10, Table 7). She was the first child of phenotypically normal and nonconsanguineous parents (mother 25 years and father 19 years). No remarkable detail of pregnancy was mentioned. Birth weight was 3.0 kg and length was 48.26 cm (both between the 25th and 50th percentile) with no perinatal problems. As an infant, she had significant feeding difficulties due to hypotonia. An atrial septal defect was identified at 1 year of age, with spontaneous closure by 3 years of age. By 16 months of age, microcephaly, global developmental delays, and short stature were evident.

A pediatric developmental evaluation was pursued by 3 years for sleep difficulties and hyperactivity. The initial physical examination at age 4, 5 and 12 years revealed a weight of 16.1 kg (50th percentile), a height of 92.25 cm (just below the third percentile), and a head circumference of 46 cm (~three SD below the mean). She was quite active, with a hoarse voice and no understandable words. Craniofacial examination revealed midface hypoplasia, an upturned nose, apparently widely spaced eyes, and a tented upper lip (Figure 10, Table 7). Skeletal examination revealed short, thick, and tapered fingers, which measured 9.5, 5.5, and 4 cm for total hand, palm, and middle finger lengths, respectively (all below the third percentiles). X-rays of the hands revealed shortened and widened metacarpals and phalanges in a symmetric manner, with no evidence of focally shortened bones.

Follow-up exam at 9, 3, and 12 years revealed a history of the onset of complex-partial seizures at 8 years associated with an abnormal EEG with “mild, diffuse encephalopathy” changes noted, but no areas of focal changes or seizure activity. Her brain MRI scan was normal, and after beginning trileptal, there was no recurrence of seizure activity. A further evaluation at 10, 9, and 12 years noted behavioral difficulties related to significant hyperactivity, minimal expressive speech, a weight of 33.1 kg (30th percentile), a height of 123.2 cm (approximately three SD below the mean), and a head circumference of 50 cm (just below the third percentile). Her physical features were unchanged from the initial examination.

Results of karyotype and FISH analyses to detect a Smith–Magenis syndrome deletion were normal at 4 years of age. Sequencing of the *RAI1* gene did not reveal any pathogenic

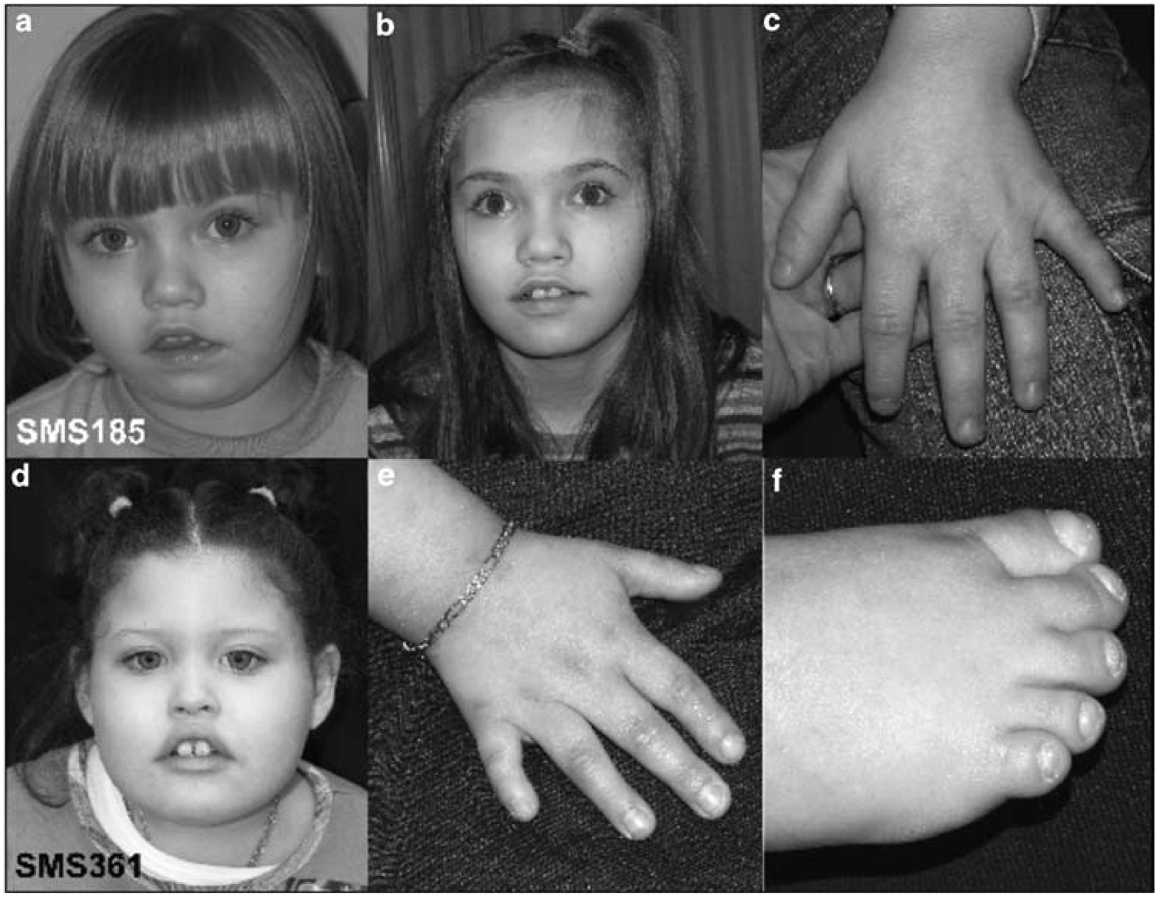


Figure 10. Clinical features of the 2q23.1 deletion syndrome. A) Case 1, SMS185 aged 2 years. Note tented upper lip, open mouth, and microcephaly. B) SMS185, 13 years of age, note tented upper lip and prominent incisors. C) Hands of SMS185, 13 years of age; note brachydactyly. D) Case 2, SMS361 aged 8 years. Note tented upper lip with prominent upper incisors and open mouth. E) Hand of SMS361. Note generalized brachydactyly. F) Foot of SMS361, note small foot and short toes with bulbous tips.

Table 7. Clinical features of cases with 2q23.1 microdeletion syndrome. In part created by Sureni Mullegama.

PHENOTYPE	SMS185 Case 1, This Study	SMS361 Case 2, This Study	Patient 29195 Wagenstaller et al 2007	Patient 1 Jaillard et al 2008	Patient 2 Jaillard et al 2008	Patient 3 Vissers et al 2003 Koolen et al 2004	Patient 2 De Vries et al 2005
Sex	F	F	M	M	M	F	F
Ages of Evaluation	4 & 13 years	3 & 9 years	1.5 years	2 & 10 years	3 & 10 years	2, 12 &14 years	N/A
Mental Retardation	+	+	+	+	+	+	+
Motor Delay	+	+	+	+	+	+	+
Language Impairment	Severe	50 words	+	Less than 6 words	Severe	Severe	N/A
Behavioral Problems	+ Hyperactivity	+ Social contact seeking behavior Hyperactivity Impulsiveness	+ Limited social interactions	+ Inappropriate laughter	+ Inappropriate laughter Autistic component Self-biting of hands/forearm Decreased sensitivity to pain Stereotypies	N/A	+ Picking of the eyes Hypernea Full hands into mouth
Short Attention Span	+	+	N/A	+	N/A	N/A	N/A

SLEEPING DIFFICULTIES	+	-	N/A	-	+	N/A	N/A
POST-NATAL GROWTH RETARDATION	+	-	N/A	+	+	+	+
SHORT STATURE	+	+	N/A	-	-	+	+
SEIZURES	+	+	+	+	+	+	+
ATAXIA	-	+	N/A	+	+	N/A	N/A
FEEDING DIFFICULTIES	+	+	N/A	+	+	N/A	N/A
INFANTILE HYPOTONIA	+	+	N/A	+	+	N/A	N/A
CONSTIPATION	+	+	N/A	+	+	N/A	N/A
	+	+		+	+	+	+
CRANIOFACIAL MANIFESTATIONS							
▪ MICROCEPHALY	+	15%	N/A	+	+	+	+
▪ BRACHYCEPHALY	-	-	N/A	-	+	-	-
▪ MIDFACE HYPOPLASIA	+	+	N/A	-	+	-	-
▪ HYPERTELORISM	+/- Apparently widely-spaced eyes	-	N/A	+	-	N/A	N/A
▪ FLAT NOSE	-	-	+	+	-	Small nose	N/A

▪ Wide Mouth	+	+	+	+	+	N/A	+
▪ Tented Upper Lip	+	+	N/A	+	+	N/A	N/A
▪ Open Mouth	+	+	+	+	+	N/A	N/A
▪ Small Chin	-	+	N/A	+	+	-	-
▪ Dental Anomalies	+	+	N/A	-	+	+	N/A
▪ Macroglossia	-	-	N/A	-	+	-	-
EYE FINDINGS		Optic nerve hypoplasia Mild myopia 30% right intermittent esotropia		Hypermetropia Astigmatism			
HAND/FOOT ABNORMALTIES	+	+	+	+	+	+	+
▪ 5th Finger Clinodactyly	-	+	N/A	-	+	+	+
• Brachydactyly	- Small hands and feet	- Small hands and feet	N/A	-	+	N/A	N/A
OTHER		Lumbar lordosis Toes bulbous at tips Bitemporal narrowing Relative obesity	Hypoactive Sandal gap	Micropenis Scoliosis Valgus feet Recurrent ear infections	Calm baby Hypogenitalism Pseudoarthrosis of the clavicle	Sandal gap Large ears	

mutations (data not shown). DNA methylation for the Angelman syndrome was normal at 9 years.

Case 2 (SMS361)

SMS361 is a 10-year-old Hispanic girl who was first seen at 26 months of age for developmental delay (Table 7). She was born to an 18-year-old gravida 1 mother. She was thought to have Down syndrome at birth, but chromosome testing was negative by report. Her family history was unremarkable.

She had a generalized tonic-clonic seizure at 8 months of age, but her EEG and CT scans were normal. She was treated with valproic acid and has had only 1–2 seizures/year. At 4 years of age, a repeat EEG showed spike and wave epileptiform discharges from the frontal lobes. This pattern has been observed on and off during successive EEGs, with the latest EEG at 7 years of age demonstrating a left-sided spike and wave pattern from the left frontal lobe. Several brain MRIs were normal. She walked after 2 years of age. Her speech was delayed, but she did learn new words. At the age of 2 years and 11 months, she knew about a total of 50 words in English and Spanish. She did not start putting two words together until after 9 years of age. Her speech was unclear, although she continues to add words to her vocabulary. She can follow simple commands, but her attention span is limited.

A physical examination at 9 years of age showed truncal obesity, with a height of 122 cm (5th percentile), weight of 41.0 kg (90th percentile), and an occipitofrontal circumference of 50.5 cm (15th percentile). She was hyperactive, impulsive, and had a short attention span, moving from object to object in the examination room

(Table 7). She had social contact-seeking behavior, maintained good eye contact, and had a happy demeanor. She had an unusual facial appearance with a tented upper lip and bitemporal narrowing. She had light-colored irides. Her hands were small (total hand length of 13 cm, less than third percentile) and showed clinodactyly of the fifth digits (although her hand X-ray showed no shortened fourth metacarpal). Her total foot length was 17.5 cm (less than third percentile) and the toes were bulbous at the tips (Figure 10). Ophthalmologic examination revealed mild optic nerve hypoplasia, mild myopia, and a 30% right intermittent esotropia. Her caregiver reports a healthy appetite but no food seeking was described.

Molecular analyses

aCGH

Array comparative genomic hybridization: SMS185 was evaluated by whole-genome array comparative hybridization (aCGH BAC array) [96], revealing a ~700kb deletion of 2q23.1 and confirmed by a clinical BAC array (data not shown). To further refine the break points of this deletion, a high-resolution oligo array (Agilent 244K) was performed, defining the deletion region to be ~930 kb at chr2:148,447,496–149,377,297 (NCBI36/hg18 coordinates, Figure 11), encompassing four known genes, *ORC4L, MBD5, EPC2*, and *KIF5C*. SMS361 was evaluated by both 105K and 244K oligo arrays (Figure 11), as described in the methods section.

Array data revealed a 3.51-Mb deletion of chromosome 2q22.3–q23.3 (chr2:146,798,229–150,310,317, NCBI36/hg18 coordinates) encompassing nine known or predicted genes, *PABPCP2, ACVR2A, ORC4L, MBD5, EPC2, KIF5C, LYPD6B, LYPD6,* and *MMADHC*.

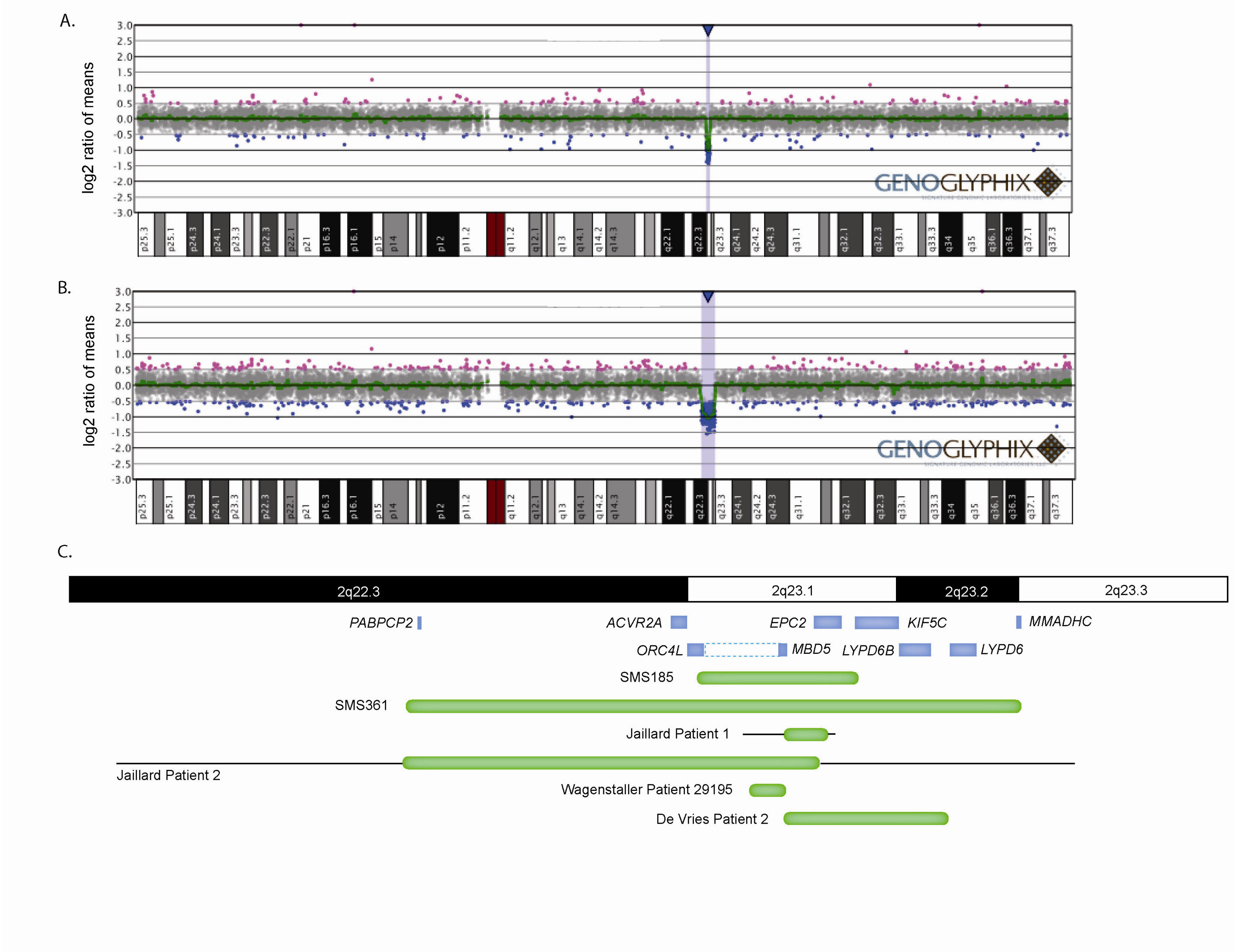


Figure 11. **Array analysis of individuals with microdeletions of 2q23.1**. A, B) Oligonucleotide microarray profiles for (a) a single-copy loss of 930 kb at chr(2)(q23.1); (chr2: 148,447,496–149,377,297) in SMS185 and (b) a single-copy loss of 3.51 Mb at chr(2)(q22.3-q23.3) (chr2: 146,798,229– 150,310,317) in SMS361. For microarray plots, probes are ordered on the x axis according to physical mapping positions, with distal 2p to the left and distal 2q to the right. C) Summary of the deletion sizes in individuals with microdeletions encompassing 2q23.1. Green bars indicate the minimum deletion sizes in individuals in this study and those in the literature, and black lines extend to show maximum deletion sizes. Genes in the region are indicated by blue boxes. Coding exons for *MBD5* are included in the solid blue box, whereas 5’ noncoding exons described by Wagensteller et al [1] are within the dashed blue box. Modified from Williams et al. 2009 [2]. Arrays performed by Signature Genomics.

Metaphase FISH with RP11–951G8 confirmed a deletion in five out of five cells examined for both SMS185 and SMS361 (Figure 12). Array data indicate a common, overlapping region between these two cases that includes *ORC4L, MBD5, EPC2, and KIF5C* (Figure 11).

Quantitative real-time PCR

Loss of genes or gene function due to chromosomal deletion does not necessarily correlate to a reduction in gene expression; however, haploinsufficiency of one or more genes is typically the pathological defect in microdeletion syndromes. To determine the genes that have a reduced expression in the overlapping 2q23.1 region, both subjects were assessed for *MBD5, EPC2, and KIF5C* mRNA levels in white blood cells. Correlating directly to the confirmed deletions, both SMS185 and SMS361 have an ~50% reduced expression of *MBD5* and *EPC2* mRNA compared with normal controls (Figure 13).

Normal levels of expression were observed for *KIF5C*; however, *KIF5C* has a very low expression in peripheral blood, and thus, levels may not be reflective of the expression in the brain (data not shown). *KIF5C* expression may be affected in both the larger and smaller deletion cases, as the smaller deletion includes the 5' region of the *KIF5C* gene, which may alter gene expression. The reduction in expression of *MBD5* and *EPC2* supports the potential role of each of these genes in the pathophysiology of this disorder.

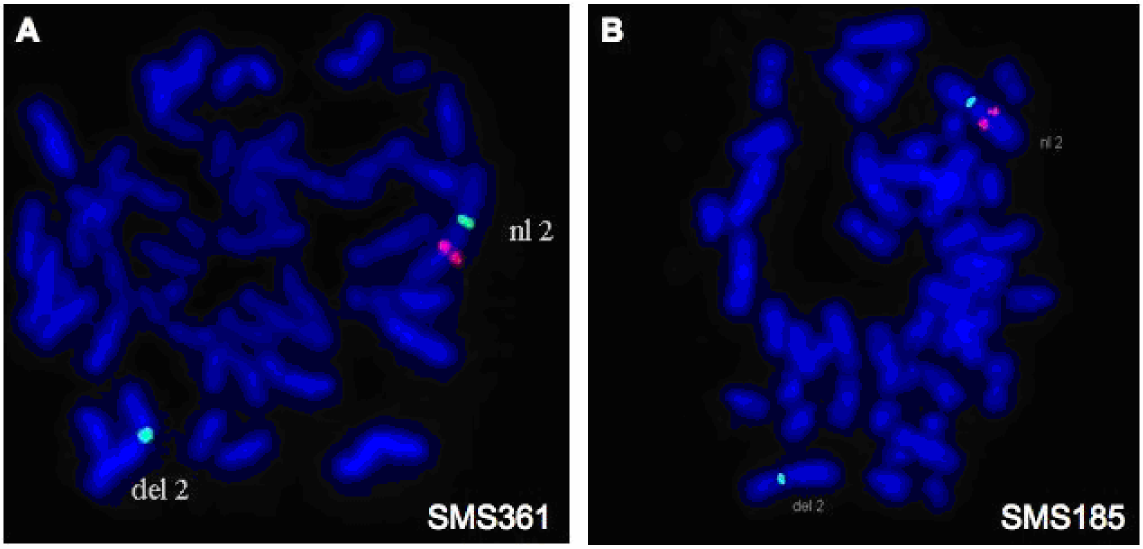


Figure 12. FISH confirmation of 2q23 deletions in patients SMS361 (A) and SMS185 (B). for both images, BAC clone RP11-951G8 from 2q23.1 is labeled in red and a chromosome 2 centromere probe is labeled in green as a control. The red signal is missing, while the green signal is retained on one chromosome 2 homologue, indicating a deletion. Performed by Signature Genomics.

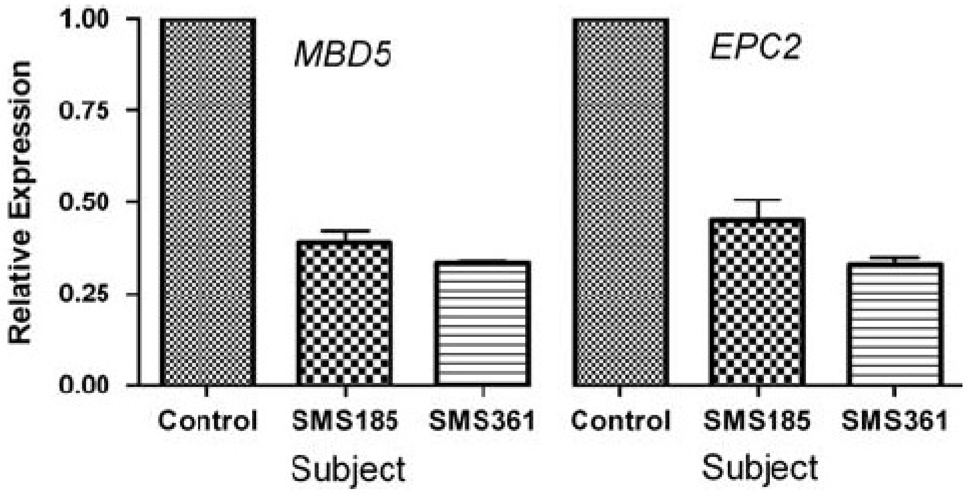


Figure 13. Quantitative real-time PCR expression analysis of *MBD5* and *EPC2* in patient cells. *MBD5* and *EPC2* relative gene expressions in SMS185 (EBV-derived lymphoblastoid cell line) and SMS361 (fresh white blood cells) are shown. Gene expression is shown relative to *GAPDH* expression and control cells, with relative expression values based on the ΔΔCt value with normal control set to 1. Control cells were cultured or isolated at the same time and with the same methods as test cells. All experiments were performed in triplicate and average values for all combined experiments are shown (n=2–5). Data generated, in part, by Sureni Mullegama.

DISCUSSION

In this study, we present two subjects, SMS185 and SMS361, with a del(2)(q23.1). These individuals have a distinct clinical phenotype that includes severe cognitive impairment, seizures, relative obesity, small hands and feet, and mild craniofacial dysmorphism, including one child with microcephaly.

A recent report described two patients with overlapping deletions of this region on chromosome 2 exhibiting a 'pseudo-Angelman' phenotype [64]. Both the subjects in this study tested negative for an abnormal *SNRPN* DNA methylation patterning of the 15q11.2 region, which rules out a diagnosis of Angelman syndrome. However, they do have features consistent with the 'pseudo-Angelman' phenotype described by Jaillard et al (2008) [64]. In addition to the report by Jaillard et al (2008) [64], three other cases of 2q23.1 deletion have been previously described [70, 91, 92] (Figure 11). The common phenotypes observed in these subjects are severe mental retardation, motor delay, severe language impairment (sometimes accompanied by hoarse voice), behavioral abnormalities (including hyperactivity and inappropriate laughter), postnatal growth retardation, relative obesity, and seizures (Table 7). In addition, a distinctive craniofacial phenotype is also apparent that includes small head size, wide and open mouth, a tented upper lip, and prominent incisors (Figure 10). Skeletal abnormalities include generalized brachydactyly (Figure 10, Table 7).

The more variable phenotypes between the cases reported here and those in previous studies include craniofacial abnormalities such as brachycephaly and relative hypertelorism (likely associated with microcephaly), as well as eye findings and genital abnormalities. Overall, these individuals with varying overlapping deletions of 2q23.1 present with a very similar global phenotype, which leads us to believe that a common gene(s) might be involved.

Additionally, after completion of this study a study by van Bon et al. 2010 [97] confirmed the findings reported in the present study. All subjects fell within the global phenotype described and it was concluded that haploinsufficiency of *MBD5* was responsible for the clinical manifestations observed. It was also mentioned that *EPC2* and *KIF5C* are likely modifiers of overall variability [97].

It is worthwhile to mention that reported duplications of the 2q23 region are rare and tend to be much larger in size than the deletions described in this and previous reports [98, 99, 100]. The reported duplication 2q23 cases presented with craniofacial abnormalities, low set ears, and heart, kidney, genital, skeletal, and neurological abnormalities, with death resulting in early infancy [101]. The phenotype observed in dup(2)(q23) further supports a gene dosage as a mechanism and the importance of the genes in this region. When considering all of the reported deletions involving 2q23.1 (Figure 11), a critical region emerges that includes a single gene, *MBD5*. *MBD5* (methyl-CpG binding domain protein 5) is expressed neuronally and likely functions in the regulation of gene expression [102, 103]. MBD5 contains a methyl-binding domain, sharing sequence homology to MECP2, which when mutated or deleted, results in Rett syndrome. We have shown in this study that *MBD5* is haploinsufficient in patient white blood cells and lymphoblastoid cells (Figure 13), providing further evidence of its role in this syndrome and the need for proper gene dosage for normal development and behavior. We do not, however, rule out the possibility that KIF5C and/or EPC2 may contribute to variability or severity of the syndrome. These data, taken together with the previous report of a small deletion within *MBD5* [1, 64] and the recent study by van Bon et al. [97], indicate that haploinsufficiency of *MBD5* is the likely common pathological defect for most features of this syndrome.

Chapter 4

Haploinsufficiency of *HDAC4* results in brachydactyly mental retardation syndrome

Introduction

Chromosome deletions involving the 2q37 region result in brachydactyly mental retardation syndrome (BDMR, OMIM 600430), also known as Albright hereditary osteodystrophy-like syndrome (AHO-like). BDMR is a complex disorder that presents with a spectrum of clinical features, including developmental delay, autism spectrum disorder, craniofacial and skeletal abnormalities, cardiac defects, behavioral problems, and neurological anomalies. Approximately 100 individuals have been reported with a deletion involving the chromosome 2q37 region [83]. To date, the critical region for this syndrome has not been clearly defined, and a variety of phenotypes have been mapped to the 2q37.1→2qter region [83]. Variable expressivity and reduced penetrance for most major features complicate genotype:phenotype correlations in this syndrome and further impair diagnosis.

In a study of 52 individuals referred for a phenotype consistent with Smith-Magenis syndrome (SMS) but for whom no 17p11.2 deletion or *RAI1* mutation could be found, we identified by whole genome array comparative genomic hybridization (aCGH) several copy number variants responsible for the SMS-like phenotypes observed in these individuals [96]. Interestingly, we identified 2 cases that carried overlapping deletions of the 2q37 region [96]. After a literature review, we discovered that the phenotypic overlap between these two distinct syndromes was striking. Utilizing this unique aCGH data set and additional cases, we refined the critical region for BDMR, reducing the likely candidates to a single gene, *HDAC4*. Given the apparent phenotypic overlap between SMS and BDMR, we then investigated more closely the specific phenotypes in our cohorts with Smith-Magenis syndrome-like and/or AHO-like features and identified one subject who lacked a 2q37 deletion but had striking phenotypic overlap with known 2q37 deletion cases. Sequencing of the coding region of *HDAC4* in this case revealed a *de novo* single base insertion that truncates the protein. *HDAC4* (histone deacetylase 4) is known to be critical for proper skeletogenesis and chondrogenesis [104], as well as neuronal survival [105]. Taking into consideration the overlapping deletion data and *de novo* mutation presented here, we assert that haploinsufficiency of *HDAC4* results in the brachydactyly mental retardation syndrome.

Materials and methods

Subject ascertainment

Subjects were referred to the Elsea laboratory for molecular evaluation of SMS at Michigan State University, or Virginia Commonwealth University, or to the Aldred Laboratory at the Cleveland Clinic for AHO-like phenotype. Samples and medical records were collected in accordance with Institutional Review Board approved protocols from the appropriate institution. Peripheral blood was collected, and DNA and metaphase chromosomes were prepared following standard methods. Phenotypic information was collected from medical records, geneticist reports, and patient photos.

Genomic DNA sequence analysis

Total DNA was isolated from cultured peripheral white blood cells using the QIAamp DNA Minikit (QIAGEN, GmbH, Hilden, Germany). PCR was performed using intronic primers from the *HDAC4* gene generated using Primer3 (v. 0.4.0) (http://frodo.wi.mit.edu/primer3/), which flank both the 5’ and 3’ regions of each *HDAC4* exon (Table X). PCR products were treated with ExoSAP-IT (USB, Cleveland, OH) to digest PCR primers. Sequencing and analysis of electropherograms were done as previously described [54]. PCR products were sequenced and analyzed on both the forward and reverse DNA strands. Sequencing primers are provided in the Tables 8, 9.

Multiplex ligation-dependent probe amplification and SNP array analysis

MLPA was performed as previously described [106] using five custom-designed probe sets that were approximately equally spaced across the genomic extent of *HDAC4* (Table 10). Breakpoints outside of the *HDAC4* gene were defined using Illumina CNV370 single nucleotide polymorphism arrays and Beadstudio software to identify copy number changes.

Table 8. *HDAC4* amplification and sequencing oligonucleotides.

Exon	Primer	Primer Sequence	Tm (C)	Amplicon
1	1F	GGCTCGGCGCTTGAACGTCT	64	220
1	1R	TGGGCAAAGAAAGCCCCGCT	64	
2	2F	TGCGCGCAGTTTCTGAAGCC	64	214
2	2R	CCCCTCGCCCTCTCTGCACT	64	
3	3F	CCAGGGACAGCAAAGGGCGG	66	376
3	3R	CGGAGGCAGGGCTGGAGTCA	66	
4	4F	AGCCCGGAATGGCCCTGACT	63	654
4	4R	TACTCCCCGTGTGCTGCCCC	63	
5	5F	TAGCCGTCCCCAGCCCTTCC	64	412
5	5R	CTGCCGCACGAGTGTACCCC	64	
6	6F	GGGGCAGTGCGCTGGGTAAA	64	285
6	6R	CTGCAGGTGATTCCTTCTCTAAGTGG	64	
7	7F	TGAGCTCCCTGCGCTCTCCC	64	400
7	7R	GGGGGTTGACAGCGTGAGGC	64	
8	8F	GGGCATTCGGGCCACAGGTC	64	320
8	8R	AGGCCACTTTCCCTCACCCCA	64	
9	9F	ATGTTTGGCCGTGACAGACT	60	219
9	9R	AAGGACCCATCACCACCAC	60	
10	10F	AGCATCCTGGCTGTGCTTT	58	242
10	10R	CCAGGCCCATTTGTGCTC	58	
11	11F	TTCCCCTCTGCTGTTTCTTC	58	398
11	11R	GTTCCCTCTTTCTGCCTCCT	58	
12	12F	GACCCAGCTCTCTGTGCTTC	60	377
12	12R	ACCACAGAAGATGCCACCTG	60	
13	13F	CAACACGGCCGTTTCTTC	60	243
13	13R	ACCCTCAGGCTGCACAAA	60	
14	14F	ATGACACGCTGATGCTGAAG	58	237
14	14R	TAAGCCCAAAGAACCACCTG	58	
15	15F	CTGTCTGTGGAGCTGAAGCA	60	222
15	15R	ACCCAATATGGGAGGAAAGG	60	
16	16F	CCTCGTTGTCCCACAAATG	58	220
16	16R	ACCACTGGGACTCGAGAAGA	58	
17	17F	TCACTGTGGGGTGTTGTTTC	58	232
17	17R	CAGCCTGATGAGAGGGAGAC	58	
18	18F	AGGGTGCAGCAAGAAACTGT	58	245
18	18R	CCTAAGGGAGGGAAGGAAGA	58	
19	19F	CTCCAGCGTCAGTTCTCTCC	60	223
19	19R	CCTAAGCTTCCCACATCCAA	60	
20	20F	TGCCTCAGCCCTGAAGTAGT	60	179
20	20R	GGCCCTTATATACCCCACCT	60	
21	21F	CGTGTGTTTCTCTCCTTCTGG	60	215
21	21R	GACACGCTCATCTCCAACAA	60	
22	22F	ACCCAGTAACGCCTTCTCCT	60	234
22	22R	TAAAAAGGGGACCTGACACG	60	
23	23F	TCTTACGATGCCATGAGACG	60	242
23	23R	GGGTCTCTGGGGTCTTCCTA	60	
24	24F	GTCTCGGAACACCCGTCTAA	60	240
24	24R	GTATAGGGGGACAGGGATGG	60	
25	25F	ACTTTCCTCACCCCACCAC	60	246
25	25R	GGTTCTGACCCTGAATAGTGTG	60	

26	26F	CACAGCCTTTAACCCACGTT	60	183
26	26R	TGGCTGAGCTTCAAGACAGA	60	

Table 9. ***HDAC4*** **cDNA oligonucleotides for amplification and sequencing.**

Coding exon (cDNA)	Primer	Sequence	Tm	Amplicon (bp)
1-8	cDNA 1F	GAGTTTGGAGCTCGTTGGAG	56	1135
1-8	cDNA 1R	AAGGATGGCGATGTGTAGAGG	56	
1-8	Sequencing Primer 1F	GCAGCTCAAGAACAAGGAG	NA	NA
1-8	Sequencing Primer 1R	CTCTTTGCCCTTCTCCTTGTTC	NA	NA

Table 10. MLPA probes for 2q37.3 and *HDAC4*. Designed by Dr. Michaela Aldred

Name	**Sequence**	**Ensembl Build 41**
rs870790A	**GGGTTCCCTAAGGGTTGGA**TTCTTAGGACCTGAAAGTCTGAACCAGCATTCCAA	239980739
rs870790B	GTGGGAGTATTGTTCAAGCGGTGATGGAATTG**TCTAGATTGGATCTTGCTGGCAC**	239980805
HDAC4-2A	**GGGTTCCCTAAGGGTTGGA**GAGGCTCGGCGCTTGAACGTCTG	239939431
HDAC4-2B	TGACCCAGCCCTCACCGTCCCGGTAC**TCTAGATTGGATCTTGCTGGCAC**	239939479
HDAC4-3A	**GGGTTCCCTAAGGGTTGGA**TGGCCGAGACCAGCCAGTGGAGCTG	239823237
HDAC4-3B	CTGAATCCTGCCCGCGTGAACCACAT**TCTAGATTGGATCTTGCTGGCAC**	239823287
HDAC4-i9A	**GGGTTCCCTAAGGGTTGGA**GCACTTGCCCTTCACTCTTCACCTTCCA	239726184
HDAC4-i9B	ATTTGGGGTGAGGGAAAGTGGCCTGTG**TCTAGATTGGATCTTGCTGGCAC**	239726238
HDAC4-i26A	**GGGTTCCCTAAGGGTTGGA**CTCCTTCCAGTGCCAAAGCCCCTTAGAGAC	239639900
HDAC4-i26B	GCATGAGGAGCATTAGATCCTGAACAGATGGA**TCTAGATTGGATCTTGCTGGCAC**	239639961

Results

Utilizing recent studies of SMS-like [96] and known deletion 2q37 subjects [107], we were able to make genotype:phenotype correlations allowing refinement of the critical region for BDMR to ~chr2: 239,639,900-240,938,547 within the 2q37.3 band, including the *HDAC4* gene.

BDMR clinical phenotype

The deletion 2q37 syndrome phenotype consists of a variety of findings that overlap with other syndromes such as Smith-Magenis, Prader-Willi, Angelman, and fragile-X syndromes (Table 11). Common features include mild facial dysmorphism, congenital heart defects, distinct brachydactyly type E, intellectual disabilities, developmental delay, seizures, autism spectrum disorder, and obesity [83, 107]. Genotype:phenotype correlations have been attempted in the past but have failed to identify a single gene that may contribute to the core findings observed in BDMR.

Deletion cases we previously reported showed the classical BDMR phenotype (Table 11). SMS272 is obese Native American with cognitive and developmental delays and was originally referred for features of Smith-Magenis syndrome (SMS). Skeletal features include brachydactyly type E (shortened 3rd, 4th, and 5th, metacarpals with shortened and proximately paced 4th left toe), midface hypoplasia, and brachycephaly (Figure 16, Table 11). She also exhibits stereotypies and aggressive behavior but lacks two of the key features, sleep disturbance and self-abusive behavior, which are common in SMS. She carries a small, 1.17 Mb, 2q37.3 deletion which was later confirmed on karyotype (Figure 14).

SMS320 was referred for developmental, speech, and motor delays, sleep disturbance, stereotypies, attention-seeking, and self-injurious behaviors. Her craniofacial and skeletal phenotype consists of brachydactyly, brachycephaly, midface hypoplasia, tented upper lip, and a broad, square face (Table 11). While originally referred for SMS, she carries a larger deletion of 2q37.3 that spans 3.02 Mb (Figure 14).

Subject 122 presented with developmental delay, mild facial dysmorphism, BDE and receptive language and social communication disorder. A detailed clinical description was published previously [107]. Case 122 carries a 3.21 Mb 2q37 deletion (Figure 14).

Subject 10780, who carries a 3.2 Mb deletion within 2q37 (Figure 14), presented with, developmental delay, short stature, facial dysmorphism, brachydactyly type E, and grand mal seizures (Table 11). A detailed clinical description has been published previously [71]

Table 11. Clinical features of cases with deletions or mutations involving 2q37.3.

PHENOTYPE	**SMS117** ***HDAC4*** **mutation**	**SMS272 del(2)(q37.3) including** ***HDAC4*** ***[96]***	**SMS320 del(2)(q37) including** ***HDAC4 [96]***	**122 del(2)(q37) including** ***HDAC4 [107]***	**10780 del(2)(q37) including** ***HDAC4*** ***[71]***	**2282 del(2)(q37.3) not including** ***HDAC4***
Sex	F	F	F	F	F	F
Age at evaluation	16 and 25 years	15 and 17 years	3 years	6 years	12 years	7 years
Developmental delay	+	+	+	+	+	+
Motor delay	+	-	+	+	+	+
Language impairment	+	+	+	+	+	+
Behavioral problems	Self-injurious behavior Pulling out of fingernails	Aggressive behavior Hand wringing Eye squint	Self-injurious behavior Head banging Skin picking Hand biting Hyperactivity	receptive language and social communication disorders	N/A	Autistic behavior Repetitive behaviors
Stereotypies		Self-hug, tics	Self-hug	N/A	N/A	N/A
Sleeping difficulties	+	-	+	N/A	N/A	N/A
Decreased sensitivity to pain	+	+	+	N/A	N/A	N/A
Hearing loss	+		-	-	-	-
Short stature	+/-	-	-	+	+	-
Seizures	essential tremor	-	-	-	+	+
Feeding difficulties	+	-	-	-	N/A	-
Obesity/overweight	+	+	+	-	-	+
Craniofacial manifestations						

Broad face	+	+	+	+	+	+
Upslanting eyes	+	+	+	-	-	
Brachycephaly	+	+	+			
Midface hypoplasia	+	+	+		+	+
Broad, upturned nose	+	+	+		+	-
Skeletal abnormalities						
Brachydactyly	+, type E	+, type E	-	+	+	-
Proximally placed 4th toe, shortened 4th metatarsal	+	+	-	+	+	-
OTHER	Subvalvar aortic stenosis; mitral stenosis; pacemaker Hypothyroidism Hirsuitism Spina bifida occulta Very friendly	Sinus arrhythmia Myopia Very friendly		2-3 toe syndactyly	Craniosynostosis	Hypothyroidism Precocious puberty

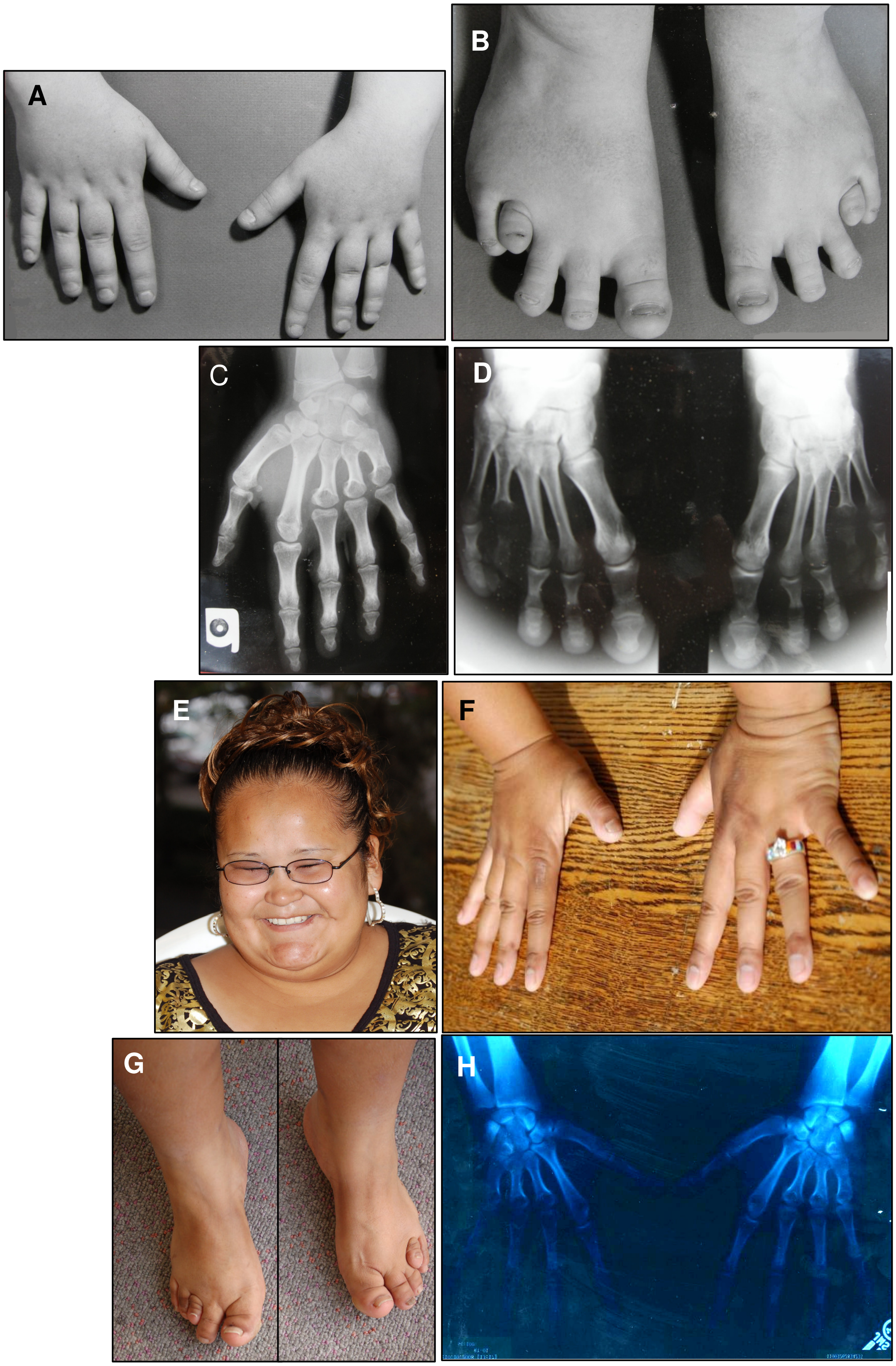

A
B
C
D
E
F
G
H

Figure 14. Skeletal anomalies observed in BDMR. A) Hands of case SMS117. B) Feet of SMS117, showing shortened 4th metatarsal and wide spacing between toes. C) Radiograph of left hand of SMS117 showing shortened 3rd, 4th, and 5th metacarpals. D) Radiograph of SMS117 feet showing proximally placed and shortened 4th metatarsal and bilaterally widely spaced 1st, 2nd, and 3rd toes. E). Photo of SMS272 at 17 years of age. F) Hands of SMS272. Note shortened 3rd and 4th fingers. G) Feet of SMS272 with shortened 3rd toes and proximally placed and shortened 4th toes. H) Radiograph documenting shortened 3rd and 4th metacarpals in the hands of SMS272.

Refinement of the BDMR critical region

Refining the critical region for deletion 2q37 syndrome has proven difficult because of poor genotype:phenotype correlation and a paucity of interstitial deletions that would help define minimal overlapping candidate regions. However, because of deletions identified within a small overlapping region within the 2q37.3 band [96, 107], we were able to narrow the critical region to ~200 kb region including *HDAC4* (Figure 14), with SMS272 critical to this analysis due to the very small deletion identified by aCGH. Further supporting the refinement of this region to ~200 kb is subject 2282 who does not have the BDE phenotype However this patient does have developmental delay, facial dysmorphism and autistic behaviors (Table 11), indicating that other aspects of 2q37 deletion syndrome are complex and like many deletion/mutation syndromes multiple genes may contribute to the full spectrum of the phenotype.

Two del(2)(q37.3) subjects were analyzed by MLPA to assess the specific location of the break points observed (Figure 15, Table 10). As shown in Figure 15 , the proximal breakpoints of the deletions are within the *HDAC4* gene. The breakpoint identified in case 122 is in the second intron of *HDAC4*, while the breakpoint for case 10780 is in intron 9 (Figure 15). Both deletions should result in haploinsufficiency for *HDAC4* as the promoter and initial coding exons are deleted in both cases (*HDAC4* is transcribed on the minus strand of DNA). These two subjects have a strikingly similar phenotype to that of SMS117 and to other BDMR cases, including intellectual disabilities, developmental delay, brachydactyly type E, and facial dysmorphism. Case 10780 also has short stature with obesity, further strengthening the case for mutations in *HDAC4* to contribute to the core phenotypes of BDMR.

Further supporting the notion that deletion of *HDAC4* is responsible for the brachydactyly type E phenotype, subject 2282 has a 2.68 Mb terminal deletion with a breakpoint that is distal of *HDAC4* (Figure 14). This individual presented with developmental delay, obesity, autistic behaviors and a history of seizures but had no evidence of metacarpal or metatarsal shortening. Facial dysmorphism included a broad face with bi-temporal narrowing and small nose with flat nasal bridge. There was no evidence of heart or other major organ defect. An additional case, subject 2232, was diagnosed with a terminal 2q37 deletion approximately 7 Mb in size (Figure 14). Unfortunately no clinical details were available and thus this case is not included in Table 11.

These molecular and clinical data supported mutation analysis of *HDAC4* as the logical next step toward identification of a single gene that, when mutated, results in BDMR. Further supporting sequence analysis of cases with the BDMR phenotype but without 2q37 deletion is the *Hdac4*$^{-/-}$ mouse [104], which has a significant skeletal phenotype [104]. Although important in electron transport, *NDUFA10* was eliminated as a probable cause for BDMR.

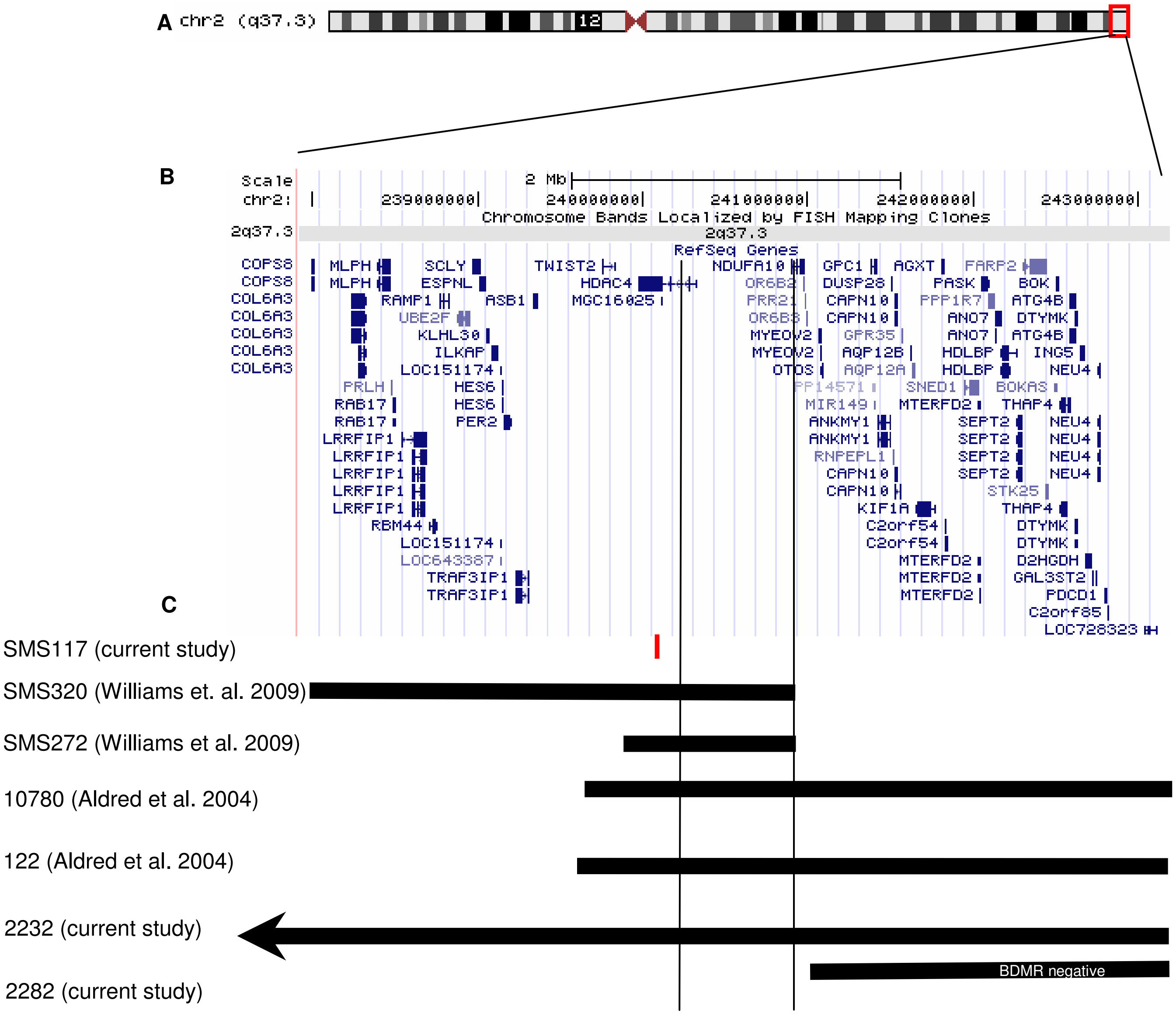

A
chr2 (q37.3)
12
B
Scale
chr2:
2 Mb
239000000|
240000000|
241000000|
242000000|
243000000|
Chromosome Bands Localized by FISH Mapping Clones
2q37.3
RefSeq Genes
COPS8
COL6A3
MLPH
SCLY
ESPNL
RAMP1
ASB1
UBE2F
KLHL30
ILKAP
LOC151174
PRLH
HES6
RAB17
PER2
LRRFIP1
RBM44
LOC643387
TRAF3IP1
TWIST2
HDAC4
MGC16025
NDUFA10
OR6B2
PRR21
OR6B3
MYEOV2
OTOS
GPC1
DUSP28
CAPN10
GPR35
AQP12B
AQP12A
PP14571
MIR149
ANKMY1
RNPEPL1
KIF1A
C2orf54
SNED1
MTERFD2
AGXT
FARP2
PASK
PPP1R7
ANO7
HDLBP
BOK
ATG4B
DTYMK
ING5
NEU4
BOKAS
THAP4
SEPT2
STK25
D2HGDH
GAL3ST2
PDCD1
C2orf85
LOC728323
C
SMS117 (current study)
SMS320 (Williams et. al. 2009)
SMS272 (Williams et al. 2009)
10780 (Aldred et al. 2004)
122 (Aldred et al. 2004)
2232 (current study)
2282 (current study)
BDMR negative

Figure 15. Delineation of the BDMR critical region. A) Ideogram of chromosome 2 band q37.3. B) Schematic representation of 2q37.3 region and RefSeq Genes included (www.genome.ucsc.edu, accessed Feb. 25, 2010). C) Horizontal bars indicate the regions of deletion in each of these key 2q37 deletion syndrome cases. The brachydactyly mental retardation syndrome critical region is indicated by the vertical bars. The SMS117 *HDAC4* point mutation is indicated by the small red vertical line.

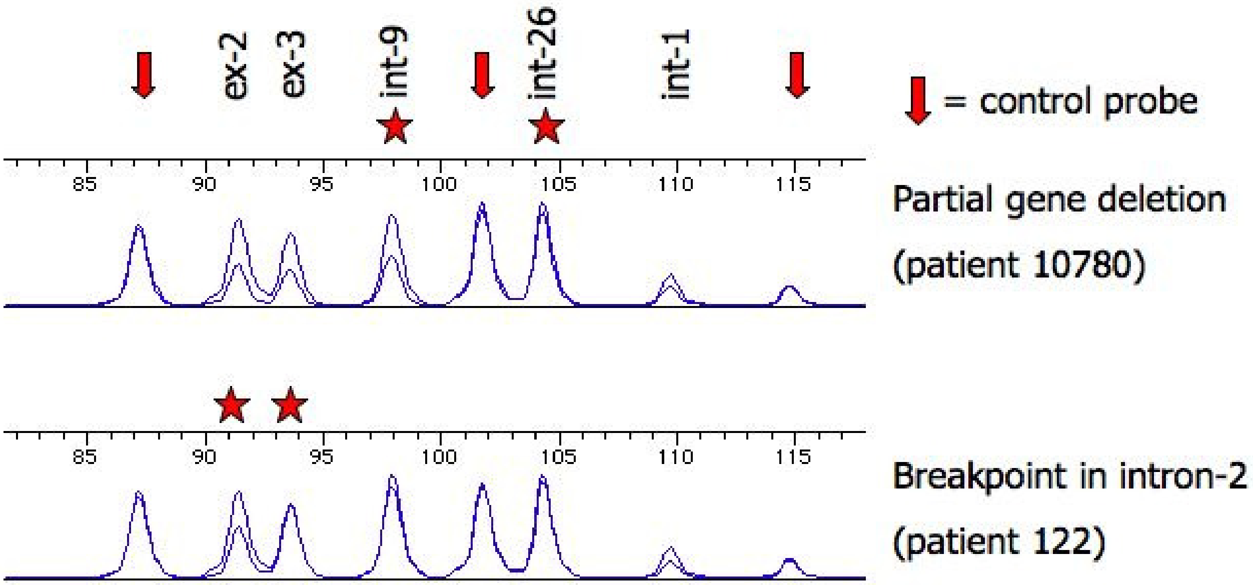


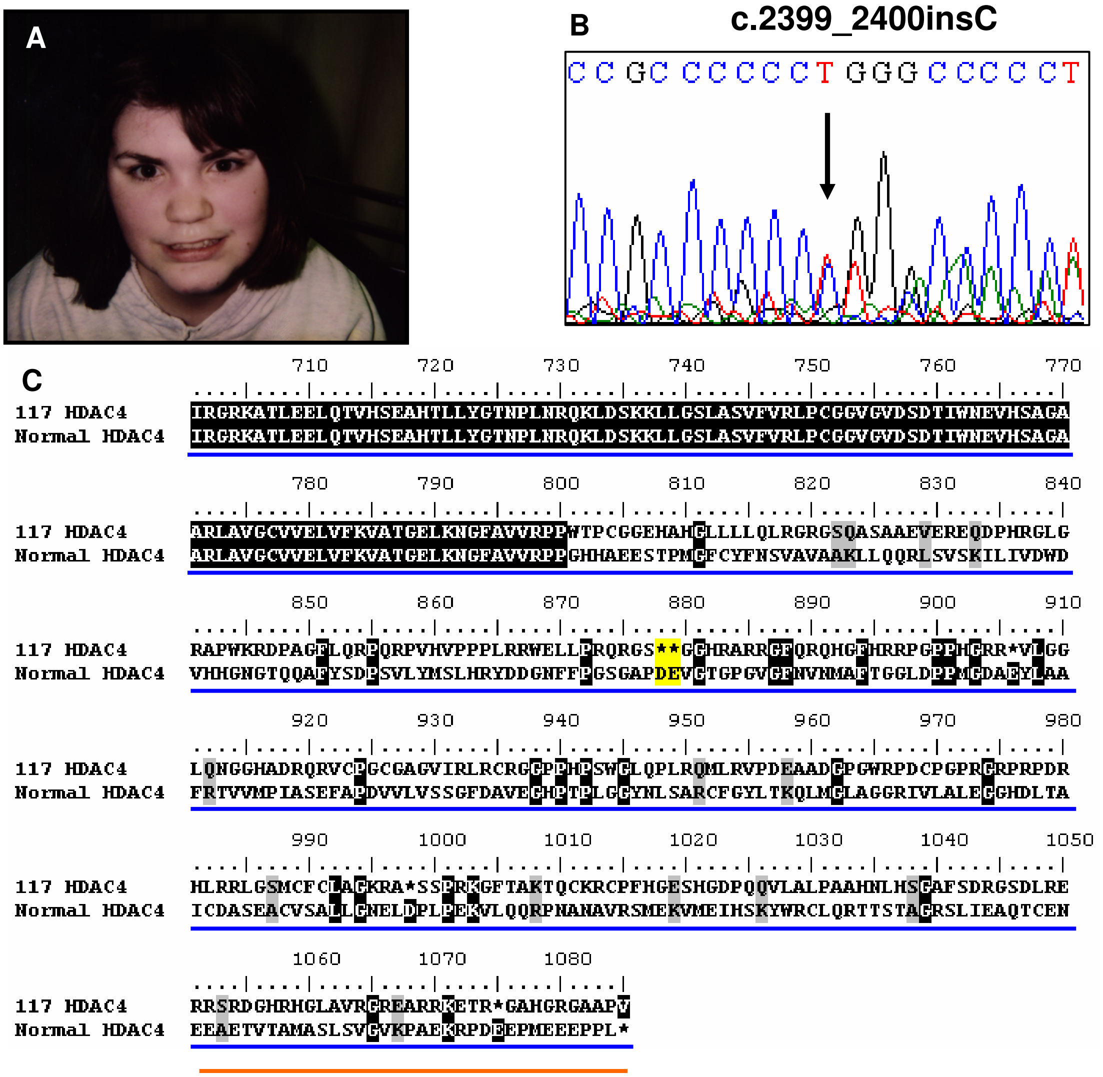
Figure 16. MLPA analysis of *HDAC4* deletion breakpoints in key 2q37.3 deletions. *Top:* Analysis of patient 10780 showed an *HDAC4* deletion extending at least through intron 9 but the breakpoint occurring prior to intron 26. *Bottom:* Patient 122 revealed a deletion including intron 1 and exon 2, with the breakpoint occurring between the probes for exon 2 and exon 3. Stars indicate regions between which the breakpoint occurs for each case. Control probes are indicated by the arrows and show 2 alleles of equal intensity for each.

SMS117 Clinical phenotype

SMS117, who is deletion 2q37 negative, is a female French-Canadian born at 43 weeks to nonconsanguineous parents (Figures 16, 17 Table 11). Birth was complicated by two induction attempts. Weight at birth was 2.85 kg (10-25%), and length was 51 cm (25-50%). The 1st year was marked with feeding difficulties. Subvalvular aortic stenosis was identified at 4 months. Psychomotor development was delayed, with walking at 18 months, and speaking in sentences at 6 years of age. The clinical phenotype worsened with age. She received a mitral valve replacement and permanent pacemaker at 5 years. At 13.5 years, she was referred to genetics where she presented with dysmorphic features, including midface hypoplasia, broad face and nose, brachycephaly, frontal bossing, a down turned lower lip, and upslanted eyes. At 25 years, physical exam revealed obesity (96 kg, >95%; 157cm, 10-25%). Neurologically, decreased deep tendon reflexes, and cerebral atrophy were also observed, as well as sensorineural hearing loss (>40 dB) for which she wears two hearing aids. Cerebral atrophy was identified by head CT, and spina bifida occulta was also observed. Decreased sensitivity to pain, onychotillomania, hyperactivity, and a decreased attention span were also observed, in addition to a hospitalization for hypersomnolence and precipitous loss of consciousness. Sleep abnormalities began at 3 years of age with several arousals throughout the night, which lasted until the age of 8; however at 24 years, she can sleep for up to 18 hours if not awakened. A complete absence of REM sleep as per 44 hours of EEG video telemetry was noted. Skeletal features include brachydactyly type E with proximally placed 3rd, 4th, and 5th fingers (shortened 3rd, 4th, and 5th metacarpals on X-ray) and bilateral proximally placed 4th toes and bilateral widely spaced 1st, 2nd, and 3rd toes (Figure 16). Total hand length was 14 cm (<3%) and total foot length was 22.5 cm (10-25%). There is a striking similarity between the morphology of the hands and feet of SMS117 (*HDAC4* mutation) and SMS272 (2q37 deletion) (Figure16), lending further support to the importance of *HDAC4* in this syndrome. Family history was unremarkable; however, the patient's father died at age 50 years of a pulmonary embolism, and no other abnormalities were reported.

Mutation analysis

All coding exons for *HDAC4,* in SMS117, were sequenced and assessed for mutation (PCR primers found in Table 8). When analyzing the nucleotide sequence for exon 19, an insertion of a single cytosine was observed (c.2399_2400insC; Figure 17). Because this insertion lies within a run of five cytosines, the exact location of the insertion cannot be determined, but this does not impact the resulting location of the premature stop codon (Gly801TrpfsX1) (Figure 17).



```
                   710       720       730       740       750       760       770
               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
117 HDAC4      IRGRKATLEELQTVHSEAHTLLYGTNPLNRQKLDSKKLLGSLASVFVRLPCGGVGVDSDTIWNEVHSAGA
Normal HDAC4   IRGRKATLEELQTVHSEAHTLLYGTNPLNRQKLDSKKLLGSLASVFVRLPCGGVGVDSDTIWNEVHSAGA

                   780       790       800       810       820       830       840
               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
117 HDAC4      ARLAVGCVVELVFKVATGELKNGFAVVRPPWTPCGGEHAHGLLLLQLRGRGSQASAAEVEREQDPHRGLG
Normal HDAC4   ARLAVGCVVELVFKVATGELKNGFAVVRPPGHHAEESTPMGFCYFNSVAVAAKLLQQRLSVSKILIVDWD

                   850       860       870       880       890       900       910
               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
117 HDAC4      RAPWKRDPAGFLQRPQRPVHVPPPLRRWELLPRQRGS**GGHRARRGFQRQHGFHRRPGPPHGRR*VLGG
Normal HDAC4   VHHGNGTQQAFYSDPSVLYMSLHRYDDGNFFPGSGAPDEVGTGPGVGFNVNMAFTGGLDPPMGDAEYLAA

                   920       930       940       950       960       970       980
               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
117 HDAC4      LQNGGHADRQRVCPGCGAGVIRLRCRGGPPHPSWGLQPLRQMLRVPDEAADGPGWRPDCPGPRGRPRPDR
Normal HDAC4   FRTVVMPIASEFAPDVVLVSSGFDAVEGHPTPLGGYNLSARCFGYLTKQLMGLAGGRIVLALEGGHDLTA

                   990      1000      1010      1020      1030      1040      1050
               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
117 HDAC4      HLRRLGSMCFCLACKRA*SSPRKGFTAKTQCKRCPFHGESHGDPQQVLALPAAHNLHSCAFSDRGSDLRE
Normal HDAC4   ICDASEACVSALLCNELDPLPEKVLQQRPNANAVRSMEKVMEIHSKYWRCLQRTTSTACRSLIEAQTCEN

                  1060      1070      1080
               ....|....|....|....|....|....|....|
117 HDAC4      RRSRDGHRHGLAVRGREARRKETR*GAHGRGAAPV
Normal HDAC4   EEAETVTAMASLSVGVKPAEKRPDEEPMEEEPPL*
```

Figure 17. SMS117 *HDAC4* mutation and impact on HDAC4 protein. A) SMS117 at 13 years. B) Electropherogram trace showing nucleotide insertion of a single cytosine at c.2399 in the *HDAC4* gene. C) Partial amino acid alignment of the normal HDAC4 protein with the altered protein likely produced in SMS117 due to frame shift mutation. Premature stop codon highlighted in yellow. The histone deacetylase domain resides at amino acids 655-1084 (Blue line), and the nuclear export signal is at amino acids 1051-1084 (Orange line). The mutation disrupts the HDAC domain and removes the nuclear export signal

Discussion

We describe five individuals with del(2)(q37.3) four of which include or interrupt the *HDAC4* gene, as well as one individual with a *HDAC4* insertion which results in a frame shift and truncated protein. Additional cases from the literature supporting this conclusion include patient 2 from Shrimpton et al (2004) with BDE, short stature and cognitive delay and a recent report of a patient with BDE, developmental delay, autism and seizures Felder et al, 2009. Although *HDAC4* was not specifically examined in these two cases, the breakpoints mapped in close proximity to *HDAC4* suggesting this gene was likely disrupted. Taken together, we conclude that haploinsufficiency of *HDAC4* results in brachydactyly mental retardation syndrome.

HDAC4

HDAC4 is a class II HDAC which acts as a co-repressor for DNA-binding transcription factors. Acting in concert with HDAC9 and HDAC3, HDAC4 acts to inhibit a variety of transcription factors including, MEF2C and RUNX2, and serves to deacetylate histones, both of which are essential for proper skeletal development. This deacetylation condenses chromatin making it inaccessible to factors that drive transcription. Additionally, class II HDACs are antagonists of cardiac hypertrophy needed for proper cardiac development [108]. Because of this molecular evidence it is logical to conclude that deletion or mutation of HDAC4 would lead to the core phenotypes seen in BDMR.

HDAC4 mutation

The impact of this insertion, found in SMS117, on the HDAC4 protein is significant. The premature stop codon interrupts the histone deacetylase domain potentially limiting or preventing all histone deacetylase activity from this *HDAC4* allele. This deacetylase activity is essential for protein function, as HDAC4 has the ability to deacetylate all four core histones [109]. Fischle et al. [110] showed that if amino acids 1002–1058 or 803-846 of HDAC4 (the most conserved domain of HDAC4 across species) are deleted, interaction with HDAC3 and all histone deacetylase activity are eliminated. As shown in Figure 17, the insertion of a cytosine at codon 801 causes a frame-shift that significantly reduces amino acid identity (5%) or similarity (5%) between the SMS117 mutant and wild type alleles. Taken together, these data provide evidence that the mutant HDAC4 protein is completely void of histone deacetylase activity and can no longer bind with HDAC3. Further, when mutated from histidine to alanine, a charged polar to nonpolar change at p.802 and p.803, Fischle et al. showed that binding to N-CoR was eliminated [110]. N-CoR is essential for HDAC3 recruitment and HDAC activity of the complex. The SMS117 mutation of *HDAC4* shows a p.H802T, and p.H803P change at these essential amino acids leading to charged polar to uncharged polar and charged polar to nonpolar changes, respectively, potentially having a great impact on the activity of this mutant protein (Figure 17).

Hdac4$^{-/-}$ and related mouse models

The *Hdac4*$^{-/-}$ mouse has a variety of significant phenotypes and was shown to have early onset chondrocyte hypertrophy and premature bone ossification [104]. These mice present with severe bone malformations and are much smaller than wild-type littermates due to premature ossification of the developing bone. A reflection of the importance of *HDAC4* in bone malformation in humans can be seen when studying the hands and feet of SMS117 and SMS272

(Figure 16). In addition, in the same manuscript by Vega et al. 2004, it was shown that HDAC4 also directly controls Runt-related transcription factor 2 (*Runx2*) activity. The *Hdac4*$^{-/-}$ mouse model directly mimics the phenotype seen with constitutive *Runx2* expression in chondrocytes [111, 112].

MEF2C activity which is necessary for proper chondrocyte hypertrophy and bone development is also regulated by HDAC4 [113, 114]. Mice deleted for the *Mef2c* gene have impaired hypertrophy, cartilage angiogenesis, ossification, and longitudinal bone growth, a converse phenotype to that of the *Hdac4*$^{-/-}$ mouse [115], consistent with HDAC4 acting as a negative co-repressor of *MEF2C*.

Clinical implications

In the analysis of twenty 2q37 deletion cases, Aldred et al. concluded that the region most likely associated with congenital heart defects in the BDMR syndrome included the *HDAC4* gene. In this report, we support this finding with the fact that SMS117 has a significant cardiac defect. The role of HDAC4 in cardiac development is further supported by data that implicate MEF2C, which is regulated by HDAC4, in cardiac development [116]. Additional studies and cases are required to confirm this hypothesis.

Based upon our findings reported here, we conclude that haploinsufficiency of the *HDAC4* gene results in BDMR. Data support a significant role for HDAC4 in normal skeletal development, specifically in metacarpal and metatarsal growth and craniofacial development, and in neuronal function, with a prominent role in behavior. Autism or autism spectrum disorder (ASD) has been reported in many cases with 2q37.3 deletions, and with behavioral and developmental findings consistent with the cases we report here [107, 117, 118, 119]. However, two of our older cases are reported to be "very friendly" but with poor communication skills, while the younger subjects have significant speech delay and poor communication with behavioral problems often prohibiting positive interactions. These findings suggest that reduced expression or function of HDAC4 is a contributor to the ASD phenotype observed in many cases with 2q37 deletion but that interpersonal skills may improve with age and with reduction of negative behaviors. Given that the genetic basis of autism is highly complex, it is not surprising that autistic behavior was also reported in case 2282 with a terminal 2q deletion that does not include *HDAC4* and in a published case with a much smaller subtelomeric deletion suggesting that multiple genes at 2q37 may be involved in this phenotype [120]. A more proximal locus, *CENTG2* at 2q37 (now renamed *AGAP1*), has also been suggested by two-point linkage analysis in autism families, although this was not supported by multipoint analysis [121].

Additionally, a recent publication by Klopocki et al. 2010 [122], has linked mutations in the parathyroid hormone related protein (*PTHLH*) to result in brachydactyly type E. Interestingly, Pthlh was identified as an inhibitor of Runx2 (regulated by Hdac4) in mice [123], and *Pthlh*-/- mice results in lethal short-limbed chondrodysplasia and speaks to the probability of common pathway involved in the development of metacarpals and metatarsals in the hands and feet in humans.

Given the significant phenotypic overlap with Smith-Magenis syndrome, BDMR and SMS should be considered together in the differential diagnosis. While type E brachydactyly is unique to BDMR, it is only penetrant in 50-60% of cases, so its absence does not remove the possibility of a positive diagnosis for BDMR. Also, other phenotypes within the syndrome also show variable expressivity, especially heart defects and behavioral problems. Given this variability, a genetic “two hit” model is the most likely applicable to neuropsychiatric disorder

[124] and given the range of phenotypes and variable expressivity seen dependent on deletion size this makes sense with regard to BDMR. The diagnostic challenges and the elucidation of core phenotypes from syndrome to syndrome make it important to note all molecular findings and detail specific phenotypic findings in all cases.

As such, we recommend molecular evaluation in cases with phenotypic findings consistent with BDMR and/or SMS for deletions involving chromosomes 2q37.3 and 17p11.2, respectively (karyotype, FISH, and/or aCGH). For those cases without 2q37.3 deletion and with type E brachydactyly, sequencing of *HDAC4* is appropriate. For those subjects without a deletion of 17p11.2 and without type E brachydactyly, *RAI1* sequence analysis should be undertaken, and if negative, followed by *HDAC4* sequencing. We anticipate additional mutations in the *HDAC4* gene will be identified lending further support to this gene playing an essential role in proper cognitive, bone, and cardiac development in humans.

Chapter 5

RAI1 transcriptionally regulates *CLOCK* and is a major contributor to circadian homeostasis

Introduction

The retinoic acid induced one gene is the gene that when mutated or deleted results in Smith-Magenis syndrome. However, little is known about the molecular function of RAI1. Given the phenotypic consequences of *RAI1* mutation or deletion, *RAI1* must be involved in pathways associated with development, behavior, neurological function, and circadian rhythm.

Mutations in transcription factors have long been associated with human disease, including ATRX syndrome, myeloid leukemia and autoimmune dysfunction [125]. Because transcription factors can regulate many genes, deciphering their global function can be difficult. With concern to *RAI1*, there is an additional layer of complexity because, other than bioinformatic analysis, very little is known about the true molecular function of RAI1.

Rai1, originally *called Gt1,* was identified in mouse P19 embryonal carcinoma cells which can be differentiated into neurons and glial cells when treated with retinoic acid [126], and three human isoforms are predicted by bioinformatic analysis. RAI1 contains a plant homeodomain (PHD), and a bipartite nuclear localization signal, which are commonly seen in transcription factors (Figure 18). PHD domains typically bind zinc and are involved in chromatin remodeling, whereas nuclear localization signals have the ability to recruit chaperone proteins, therefore facilitating transport into the nucleus [127]. Transcription factors will then either bind directly to nuclear DNA or other proteins which will form the initiation complex, to start of transcription via RNA polymerase II. However, transcription factors can also repress transcription acting to a negatively regulate transcription by blocking the enhancer binding regions.

Past studies have shown that Rai1 is expressed in all tested tissues and highly expressed in neural tissue [126]. We have previously shown that an RAI1 protein tagged with GFP (RAI1-GFP) localizes to the nucleus (Figure 18) (Burnes et al. in review). Complicating the story, the three predicted isoforms of RAI1 may serve different functions dependent on spatial and temporal expression.

Identification of genes targeted by transcription factors can be difficult. Before the advent of newer technologies, initial screens for gene targets of transcription factors could be a very labor intensive task [128]. Traditionally, the project would start with a chromatin immunoprecipitation using agarose beads bound to an antibody specific for the protein of interest. The protein would be pulled down with chromatin bound to the protein of interest, in complex with the antibody coated beads, and these chromatin fragments would be cloned into plasmids. Sequencing of these plasmids would follow. Hundreds of these individual clones would have to be sequenced making this process very time consuming [128]. Recently, chromatin immunoprecipitation coupled with microarray has become an important tool in deciphering the targets of transcription factors. Global ChIP-chip is a high throughput and effective tool for screening the regulatory regions of all known genes and has reduced the time needed to identify targets of transcription factors [129]. Through this set of studies we set out to identify genes regulated by RAI1.

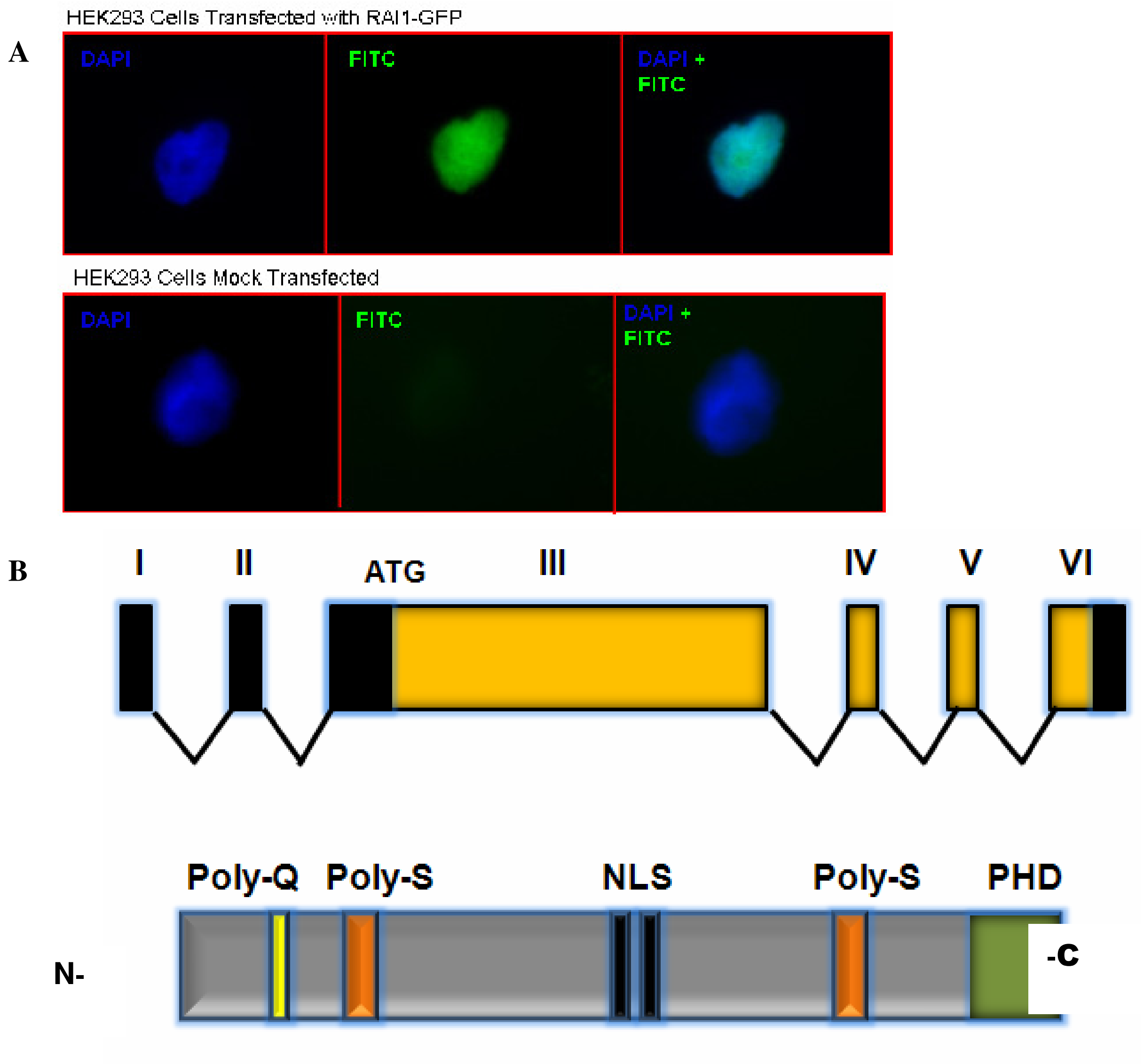

A
HEK293 Cells Transfected with RAI1-GFP
DAPI
FITC
DAPI +
FITC
HEK293 Cells Mock Transfected
DAPI
FITC
DAPI +
FITC
B
I
II
ATG
III
IV
V
VI
Poly-Q
Poly-S
NLS
Poly-S
PHD
N-
-C

Figure 18. RAI1 localization and genomic and protein structure of *RAI1*. A) HEK293 cells transfected with RAI1a-GFP. B) The top panel shows genomic structure of *RAI1* (isoform A) with 6 exons (4 coding: yellow and 2 noncoding: black). The bottom panel shows the protein structure of RAI1 including the polyglutamine (Poly-Q) and polyserine (Poly-S) tracts, the bipartite nuclear localization signals (NLS), and a C-terminal PHD (plant homeodomain). Created, in part, by Santhosh Girirajan.

Circadian Rhythm

Circadian rhythm

Mammalian circadian rhythm is an essential regulator of not only the sleep-wake cycle but also body temperature, feeding cycles, hormone secretion, drug and xenobiotic metabolism, glucose homeostasis, and cell-cycle progression [130]. The master regulator of the circadian clock resides in the suprachiasmatic nucleus located in the mammalian hypothalamus. The primary synchronizer of this clock is the 24 hour light dark cycle. Over this 24 hour light dark cycle, a feed back loop of gene transcription and subsequent degradation of gene products occur signaling the mentioned processes to react accordingly. Because of the complexity of this circadian loop, disturbances in the timing of gene expression can have a major impact on not only one's sleep pattern, but also ones disease susceptibility and behavioral stability. Genes involved in fatty acid metabolism, energy metabolism, and cholesterol biosynthesis seem to be expressed in a circadian manner [130] so one can anticipate that sleep disturbance could be tied to obesity and increased risk for disease. Interestingly, mice with a homozygous mutation in the gene *Clock* (one of the first and most important circadian genes identified) are obese and hyperphagic [131]. Further, altered circadian rhythm has been linked to bipolar disorder [132], and variants in *CLOCK* have been linked to sleep disorders in humans [133]. It has been shown through caregiver surveys that partial alleviation of the sleep disturbance in patients with SMS can result in behavioral improvement [134]

Materials and methods

Creation of Plasmids

RAI1^{Flag}: *RAI1* coding sequence was cloned into pEntr/D-TOPO using standard manufacturer protocols (Invitrogen, Carlsbad, CA). A 5' TOPO poly linker was added to cDNA

PCR forward primer (CACC) to ensure proper directional cloning. Following PCR 2 uL of PCR product was mixed with 1uL salt solution, 2 uL ddH20, and 1 uL pEntr/D-TOPO. The reaction was transformed into One Shot® Competent *E. coli* (Invitrogen, Carlsbad, CA) according to manufacturer's instructions. Cells were spread on LB plates containing 50 µg/ml kanamycin. After 18 hours colonies were picked and mixed with LB broth containing 50 ug/ml kanamycin and shaken for 18 hours at 37^0C. Plasmid DNA was isolated using Fermentus GeneJET^{TM} plasmid mini kit (Burlington, Ontario, Canada) according to standard manufacturer's instructions. Inserts were confirmed using *RAI1* cDNA primers and sequenced using standard Sanger techniques. Next, $RAI1^{pEntr/D\text{-}TOPO}$ was recombined with $pDest26^{Flag}$ [135]using the standard gateway protocol provided by Invitrogen(Carlsbad, CA) to create pDest26$RAI1^{Flag}$.

$CLOCK^{Luc}$: The *CLOCK* 1st intron element was PCR amplified using standard techniques (forward primer:5'-GGACCTTTGCAAGAGCCCAAG-3', reverse primer: 5'-GCAGAGCACAGAGGGCTTTTAGGCCGATGT-3') and cloned into StrataClone™ PCR Cloning vector using the standard protocol provide by the manufacture(Agilent Technologies, Santa Clara, CA) creating $CLOCK^{Strata}$ plasmid. The insert was confirmed by standard Sanger sequencing, and removed by restriction digest with *KpnI* and *SacI* enzymes (New England Biolabs, Ipswich, MA). The resulting product was run on a 2% agarose gel for 30 min and purified using Qiagen (Germantown, MD) Qiaquick gene extraction kit according to standard manufacturer's protocols. *Kpn1* and *Sac1* enzymes were then used to "prepare" the pGL3pro vector (Promega Corp, Madison, WI) and the cut out insert from $CLOCK^{Strata}$ was directionally ligated into the pGL3pro vector to create $CLOCK^{Luc}$ using standard T4 ligation protocols provided by the manufacturer (Promega Corp, Madison, WI).

Transfections

Human embryonic kidney (HEK293T) cells were maintained in 6-well dishes containing Dulbecco's modified Eagle's medium with 10% (v/v) FBS, 2 mM L-glutamine, and 100 µg/ml of penicillin/streptomycin (Invitrogen, Carlsbad, CA) at 37°C in a 5% CO_2 incubator. The cells were counted using trypan blue exclusion to ensure >90% viability. Transfections with pUC19, psvβ-Gal, $BDNF^{Luc}$ and $RAI1^{Flag}$ were performed by lipofection using Lipofectamine™ 2000 (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Briefly, ~4×10^5 cells were plated in 2.0 ml of growth medium without antibiotics 24 hours prior to transfection. A total of 4 ug of total plasmid DNA, using pUC19 plasmid as "filler" DNA, was diluted in 250 ul OptiMEM® Reduced Serum Medium (Invitrogen, Carlsbad, CA). Similarly, 10 ul of Lipofectamine™ 2000 was diluted in 250 µl of OptiMEM® Reduced Serum Medium, mixed well, and incubated for 5 minutes. After incubation, diluted plasmid DNA was mixed with diluted Lipofectamine™ 2000 to a total volume of 500 µl and incubated for 20 minutes. Plasmid:Lipofectamine complexes (500 µl) were added to each well and mixed by rocking. The cells were then incubated at 37°C in a 5% CO_2 incubator, for 24 hours before Trizol RNA extraction using standard protocols provided by Invitrogen (Carlsbad, CA).

Luciferase assay

Twenty-four hours after DNA transfections cells were washed with 2 mL DPBS (Invitrogen, Carlsbad, CA), and Tropix Glacto-Light™ (Applied Biosystems, Bedford, MA) standard protocol was used. Briefly, 250 uL of lysis solution was added to each well of the six well plate and scraped until all cells were detached. Lysates were collected and centrifuged at

12,000 RPM for 2 min to pellet cell debris. Next, of the resulting supernatant 50 uL was transferred to a 4 wells of a 96 well white luminometer plate. Two wells were treaded with 70 uL diluted Galacton® substrate (1:100, Galacton:Reaction buffer diluent) and incubated for 30 min. To the 2 wells that contained the diluted Galacton® substrate, 100 uL of Accelerator(-II) was added. To the 2 wells that did not contain the diluted Galacton® substrate, 100 uL of Steady-Glo® Luciferase substrate (Promega Corp, Madison, WI) was added. Each well was read using the Wallac *1420 VICTOR*2™ luminometer (PerkinElmer, Waltham, MA) on a maximum linear scale.

Calculation of relative luciferase activity

Relative luciferase activity, from each individual transfection, was calculated by dividing the average number of light units read from the wells containing the Steady-Glo® Luciferase substrate (Promega Corp, Madison, WI) by the average number of light units from the wells containing the Galacton® substrate and Accelerator(-II). The equation (ΔLuc/$\Delta\beta$-Gal=Relative luciferase activity) was used. Wells containing pUC19*, psvβ-Gal,* and *CLOCK*^{Luc} was used as baseline luciferase activity. Each experiment was performed independently no less than three times.

Statistical analysis

P-values were generated by averaging relative luciferase activity from each independent study and performing a two-tailed student's t-test. Standard deviations were generated using Excel.

Mouse tissue collection

Rai1+/- mice were obtained from the Jackson labs [42]. These mice were bred with wild-type C57Bl/6J females to obtain *Rai1+/-* and wild type pups. Genotyping was performed using PCR primers, reported in Bi et al 2005, specific for the *Rai1*-targeted allele and the wild type allele [42]. Mice euthanized in the day time hours were taken at ~8 hours Zeitgeber time and mice euthanized in the night time hours were taken at ~16 hours Zeitgeber time. WT C57Bl/6J and Rai1+/- mice were euthanized by CO_2 anesthesia and tissues were collected in accordance with standard protocols and frozen at -80°C.

Quantitative PCR

qPCR was performed as previously described [81].

Chromatin immunoprecipitation with microarray (ChIP-Chip)

HEK293t cells were transfected with *RAI1*^{Flag} plasmid as stated above by scaling the reaction up for a T75 culture dish. ChIP-Chip was processed with mouse IgG Dynabeads (Invitrogen, Carlsbad, CA) and monoclonal anti-Flag IgG antibody produced in mouse (Sigma Aldrich, St. Lewis, MO) according to manufacture's protocols using Nimblegen "ChIP sample preparation protocol v2" (Roche Nimblegen, Madison, WI). Briefly, after 24 hours, transfected cells were formaldehyde treated to cross-link protein complexes with nuclear chromatin and frozen at -80°C overnight. Nuclei were next isolated by cytoplasmic lysis and centrifugation. Isolated nuclei were then sonicated on ice with a Branson Sonifier, used at 25% power, to shear DNA:protein complexes to between ~200 and 800bp. Lysates were then centrifuged at 4,000

RPM at 4°C to pellet debris. Protein:chromatin products were then mixed with IgG Dynabeads, that had been washed and mixed with anti-Flag IgG antibody at 4°C for 24 hours, for 16 hours overnight. Using a DynaMag™-Spin magnet (Invitrogen, Carlsbad, CA) protein:chromatin products were isolated and nuclear extracts discarded. Chromatin;protein complexes were then washed 6 times according to manufacturers protocols and the crosslink was reversed by using heat treatment at 65°C for 16 hours. Following reverse crosslink chromatin products were whole genome amplified, as per recommendation by Nimblegen, by using WGA2-10RXN kit (Sigma-Aldrich, St. Lewis, MO).

Arrays were processed by Nimblegen using the Nimblegen HG18 RefSeq promoter array according to standard manufacturer's protocols. Control chromatin was processed as per manufacturers instructions using nuclear DNA isolated from HEK293t cells that were non-transfected and non-cross linked but were sonicated as stated above. Data were pre processed by Nimblegen according to standard manufacturer's protocols. $RAI1^{Flag}$ binding sites were generated by using genomic data points, Peak_Start and Peak_End, from Nimblegen pre-processed data files.

Candidate gene filtration

Ideally, regions of potential RAI1 binding would replicate between both ChIP-Chip experiments but this was not a defined criteria. RAI1 probable "hits" were defined by false discovery rate (FDR<.20), preprocessed by Nimblegen, reproducibility between arrays, known biological function, and Signal Map peak score (Roche Nimblegen, Madison, WI). FDR "is generated from the scaled $\log_2$-ratio data. NimbleScan detects peaks by searching for 4 or more probes whose signals are above the specified cutoff values, ranging from 90% to 15%, using a

500bp sliding window. The cutoff values are a percentage of a hypothetical maximum, which is the mean + 6[standard deviation]. The ratio data is then randomized 20 times to evaluate the probability of "false positives." Each peak is then assigned a FDR score based on the randomization (Roche Nimblegen, Madison, WI)." Peak scores were identified by loading preprocessed .GFF files. Peak scores are between 1-4 and only scores between 1-2 were considered. Peak scores are based off of an algorithm created by Nimblegen.

Results

Chromatin immunoprecipitation with microarray

Genes identified by ChIP-Chip:

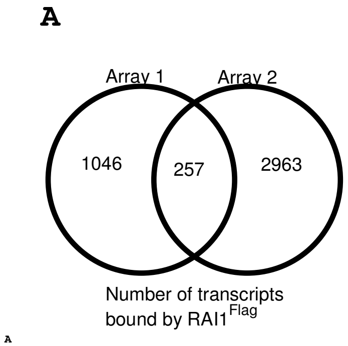
Two independent ChIP-Chips were performed by transfecting RAI1^{Flag} plasmid into HEK293 cells. HEK293 cells were used because of their ease of transfection and biological relevance as RAI1 is highly expressed in the human kidney [136].

By performing the two ChIP-chip experiments in HEK293 cells we were able to identify 257 transcripts which overlap between the arrays (Figure 19, Table 12). As suspected, we identified genes important in circadian rhythm, cognition, chromatin modification, development, and neurological stability were identified (Figure 19, Table 12). To refine our candidate gene list which would be further explored we used SignalMap software (provided by Roche Nimblegen), and known biological function to pick the best transcripts of interest. Peak scores range from 1-4. Peaks created in SignalMap which held a value of 1 or 2 were identified as best hits. A value of 1 or 2 correlates with a low false discovery rate and high log2 ratio of test/control indicating a high scoring peak. A ranked order of top genes can be seen in Table 13.

Table 12. List of gene regulatory regions which overlap between RAI1 ChIP-Chip arrays.

ABCB1
AGL
AKAP7
AKR1C1
AKR1C3
ALS2CR16
AMELY
AMY2B
ANXA1
APIN
APRT
ARHGEF7
ASAH2
ASB4
ATF6
ATP5E
AZI2
BAZ2B
BCHE
BMP5
C10orf61
C10orf68
C14orf10
C14orf150
C20orf103
C21orf34
C3orf58
C4BPB
C4orf7
C9orf102
C9orf150
CAB39L
CCDC46
CD69
CDH12
CDY2B
CENTD1
CEP76
CEPT1
CHL1
CHODL
CHRNA5
CLEC5A
CNTF
COG6
COMMD3
CRBN
CREM
CRYGS
CYLC2
CYorf15A
CYorf15B
DEFB112
DGKB
DIAPH3
DNAJC5B
DSPP
DTNA
ELOVL2
EYA1
FAM55D
FBXL7
FBXO4
FGB
FGF7
FGL2
FH
FLJ22662
FLJ32745
FLJ43706
FLJ44048
FLJ46120
FOXK2
FYTTD1
GABRG1
GALNT13
GLRA3
GPR52
HHLA2
HIST1H2AC
HIST1H2BL
HIST1H2BM
HIST1H3H
HIST1H4B
HIST1H4J
HLA-DPA1
HLA-DPB1
HS322B1A
HSFY1
HSFY2
HSN2
HTR2B
IFNA10
IL12RB2
IL15
IL17D
IMPG2
INVS
JMJD1A
KCNK10
KCNMB4
KCTD4
KIAA0391
KIAA1712
KLF14
KLHL4
KLHL9
KLRC4
KLRK1
KRT12
KRT24
KRTAP1-1
KRTAP1-3
KRTAP4-12
KRTAP6-3
LCE3B
LEF1
LOC129285
LOC401137
LOC642891
LOC646215
LOC81558
LPHN2
LRRN3
LRRTM4
LST-3TM12
LUM
MAS1L
MEF2A
MEP1A
MGAT4C
MGC33530
MGC34713
MGC35043
MGC35212
MKRN3
MMP13
MTRF1L
MUC7
MYL1
N/A
NAP1L3
OGN
OMD
OR10T2
OR11H4
OR13C3
OR13F1
OR1Q1
OR2B11
OR2J3
OR2L2
OR2L3
OR2M5
OR2T10
OR2T4
OR4A47
OR4C16
OR51A7
OR52A1
OR52N5
OR56B1
OR5B2
OR5BU1
OR5H14
OR5H6
OR5K3
OR5K4
OR5M3
OR5R1
OR5T2
OR5V1
OR5W2
OR6M1
OR8H1
OR8H2
OR8I2
OR8K3
OR9G1
OR9G9
OTUD6B
PALLD
PCDH11X
PCDH15
PCDHA1
PCDHA11
PCDHA3
PCDHB16
PCDHB8
PCDHGA12
PCDHGA9
PDE1A
PHIP
PHYHIPL
PIGF
PLCB4
PLRG1
PMCH
PNLIP
POU1F1
PRKACB
RAP1B
RGS21
RHOB
RP1-32F7.2
RP11-35N6.1
RP11-54H7.1
RP13-11B7.1
RREB1
RSRC1
SERPINI2
SGCG
SGCZ
SHOX2
SI
SLC12A1
SLC16A7
SLC25A32
SLC26A7
SLCO1B3
SLITRK6
SMC6L1
SNPH
SNRPN
SPAG6
SPANXA1
SPANXA2
SPANXE
SPATA18
TAF2
TAIP-2
TAS2R7
TAS2R8
TDRD4
TEX15
TFEC
TMEM77
TMPRSS11B
TNFSF18
TNFSF5IP1
TRPC3
TRPM8
TUSC1
TXNDC13
UBE3A
UFM1
UGT2A1
UNQ9356
UTS2D
UTY
VWA2
WDSOF1
XRN1
YIPF7
ZNF306
ZNF614

Full gene names are provided in Appendix E.


A
Array 1
Array 2
1046
257
2963
Number of transcripts bound by $RAI1^{Flag}$
A

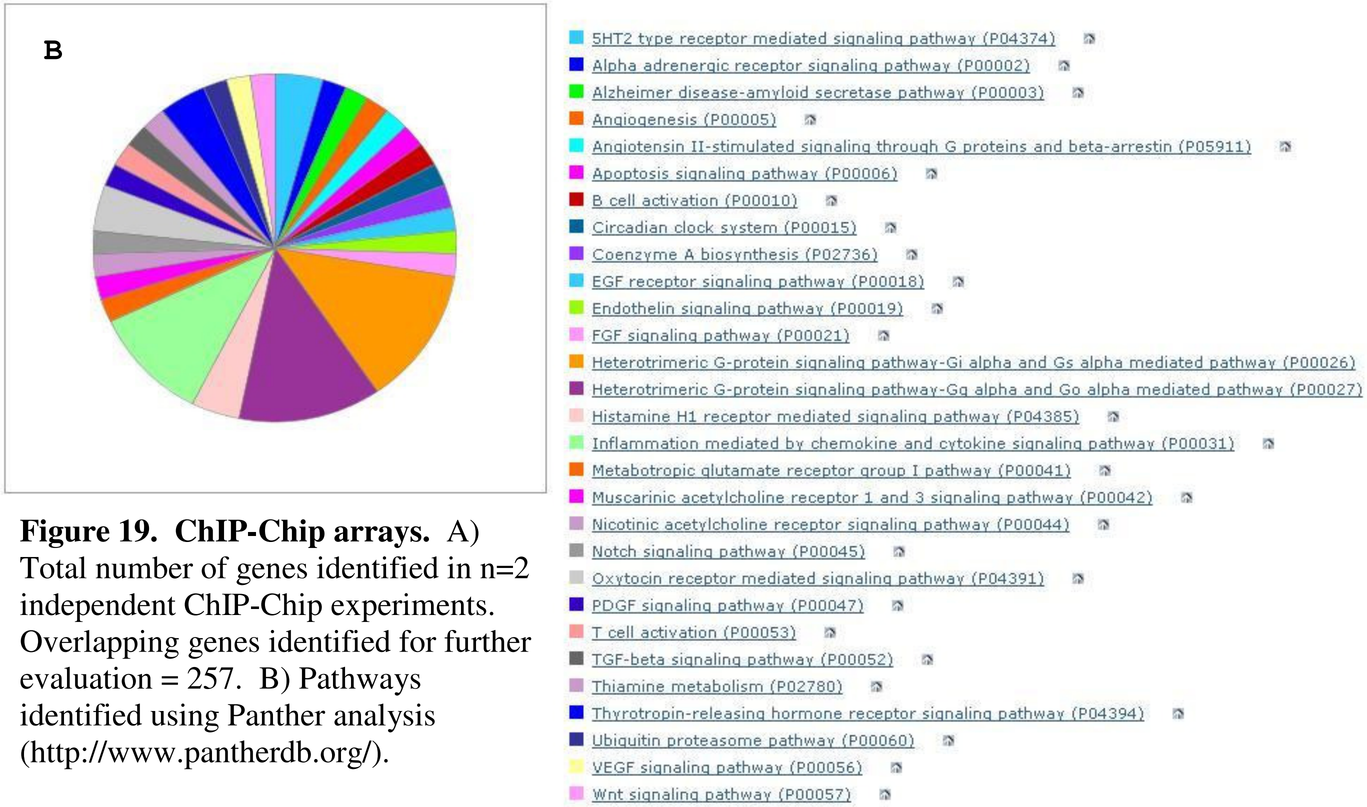


Figure 19. ChIP-Chip arrays. A) Total number of genes identified in n=2 independent ChIP-Chip experiments. Overlapping genes identified for further evaluation = 257. B) Pathways identified using Panther analysis (http://www.pantherdb.org/).

Table 13. Ranked order of top 10 gene regulatory regions identified in ChIP-Chip experiments. Disease relevance collected from Panther (www.pantherdb.org) and PubMed (http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed)

Gene Symbol	Gene Name	Disease Relevance	FDR
SNRPN	Small nuclear ribonucleoprotein polypeptide	Prader-Willi and Angelman syndromes	0.05
UBE3A	Ubiquitin protein ligase E3A	Angelman syndrome	0.10
BDNF	Brain-derived neurotrophic factor	Obesity, neurotransmission, metabolism	0.18
CLOCK	Circadian locomotor output cycles kaput	Sleep, circadian rhythm mood disorder, schizophrenia, insomnia, obesity	0.19
MEF2A	Myocyte enhancer factor 2A	Muscle development, neuronal differentiation, cell growth control, apoptosis	0.15
BMP5	Bone morphogenetic protein 5	Skeletal development, dendritic growth	0.01
TRPC3	Transient receptor potential cation channel, subfamily C, member 3	Alzheimer syndrome, BDNF function	0.14
BCHE	Butyrylcholinesterase	Neurotransmission, drug metabolism, skeletal muscle development, butyrylcholinesterase deficiency	0.15
CREM	cAMP responsive element modulator	Apoptosis, transcription, cell cycle regulation	0.09

Genes identified for further evaluation

CLOCK: The mammalian circadian *locomotor output cycles kaput* (*CLOCK*) was identified and cloned by King et al. [137] in 1997. *CLOCK* is thought to be the master regulator of the central clock of circadian rhythm. The protein product of the *CLOCK* gene is a basic-helix-loop-helix transcription factor and heterodimerizes with BMAL1 to activate transcription of a variety of downstream genes essential for proper modulation of circadian rhythm. The period genes (*PER1, PER2, PER3*), the cryptochrome genes (*CRY1, CRY2*), the nuclear receptor subfamily 1 group D member 2 (*NR1D2*/*Rev-Erbβ*), and RAR-related orphan receptor A genes (*RORA, RORB, RORC*). All of these genes have an E-box (consensus sequence: CANNTG) to which the CLOCK/BMAL1 complex binds.

SNRPN: The *small nuclear ribonucleoprotein polypeptide N* (*SNRPN*) resides in the Prader-Willi/Angelman syndromes region on chromosome 15q11.2. The protein encoded by this gene is one polypeptide of a small nuclear ribonucleoprotein complex and belongs to the snRNP SMB/SMN family. The protein may play a role in pre-mRNA processing. Although individual snRNPs are believed to recognize specific nucleic acid sequences through RNA-RNA base pairing, the specific role of *SNRPN* is unknown. The genomic location of *SNRPN* is on the opposite strand of the *SNURF* gene (*SNRPN* upstream reading frame) being transcribed on the sense strand. The 5' UTR of *SNRPN* has been identified as an imprinting center. The most common cause for Prader-Willi syndrome (PWS) is paternal microdeletion of *SNRPN* (del(15)(q11-q13)pat) resulting in ~70% of cases, while the second most common cause is maternal uniparental disomy (UPD) of chromosome 15 (15 (upd(15)mat)) [138]. Interestingly, there is strong overlap between the physical and behavioral phenotypes seen in PWS and SMS. This gene is being pursued by Anam Bashir (Elsea Lab)

UBE3A: *Ubiquitin protein ligase E3A* (*UBE3A*) is located on chromosome 15q11.2 and functions to target proteins for degradation. Ubiquitination is an important step in all cells as dysfunctional gene products must be removed from the cellular machinery by cellular structures such as proteasomes. *UBE3A* is ubiquitously expressed from both the paternal and maternal alleles except for the brain where only the maternal copy of the gene is expressed and the paternal copy is imprinted. Maternal deletion, mutation, or imprinting error of *UBE3A* results in Angelman syndrome (AS). The most common lesions leading to AS are a common large deletion of 15q11-q13 (~70%), paternal uniparental disomy (~1%), and an imprinting defect (~4%) [138]. Additionally, *UBE3A* mutation account for approximately 8% of cases [139]. Paternal UPD led investigators to the conclusion that AS involves a maternally expressed gene. Based on the finding of point mutations in patients with AS, *UBE3A* has been identified as the gene affected in AS. Interestingly, imprinted *UBE3A* expression is restricted to the brain only.

BDNF: *Brain-derived neurotrophic factor* (BDNF) is located on chromosome 11. It is primarily active in the hippocampus, cortex, basal forebrain, cerebellum and olfactory bulb. It is also found in the retina, central nervous system (CNS), motor neurons, kidneys, and the prostate. BDNF functions in neuronal survival, growth and differentiation of new neurons and synapses, as well as control of food intake. *BDNF* has recently been reported to regulate food intake and glucose homeostasis in genetically obese animal models [140]. The *Bdnf*+/- mouse exhibits hyperphagic behavior and dramatic obesity [141]. Haploinsufficiency for *Bdnf* in mice was found to cause increased food intake, early-onset obesity, hyperactivity, and cognitive impairment [141]. Previous studies have also shown that up regulation of *Bdnf* rescued synaptic plasticity, long-term memory deficits and behaviors of the mutant Huntington's disease mice [142]. *In vitro* overexpression of *Bdnf* increased axonal length and axonal branching in Rett syndrome, while preventing the decrease in dendritic length commonly seen with the disease [143]. This gene is being pursued by Sun Kim (Elsea lab)

RAI1 transcriptionally activates *CLOCK* via an intron 1 enhancer element

The downstream targets of CLOCK are well characterized. However, the transcriptional regulator of *CLOCK* has not been defined. Using peak data acquired from our duel ChIP-Chip assays we were able to identify a ~350 bp region where $\mathrm{RAI1}^{\mathrm{Flag}}$ possibly binds. This element resides in intron 1 (Chr4:56072089- 56072440). Primers designed to flank this region were used to amplify this piece of genomic DNA, which was first cloned into the StrataClone TA cloning vector and then was subcloned into the pGL3-Promoter vector ($\mathrm{CLOCK}^{\mathrm{Luc}}$) which contains an SV40 promoter that provides baseline luciferase activity.

When $\mathrm{RAI1}^{\mathrm{Flag}}$ is transfected with $\mathrm{CLOCK}^{\mathrm{Luc}}$ there is a greater than 2-fold increase in luciferase activity (Figure 19). Taken together, the ChIP-Chip and luciferase data suggest that RAI1 binds, directly or in a complex, to the 1st intron of *CLOCK,* enhancing its transcriptional activity *in vitro*. Given these findings, it was logical to move forward with *in vivo* studies utilizing the *Rai1+/-* mice.

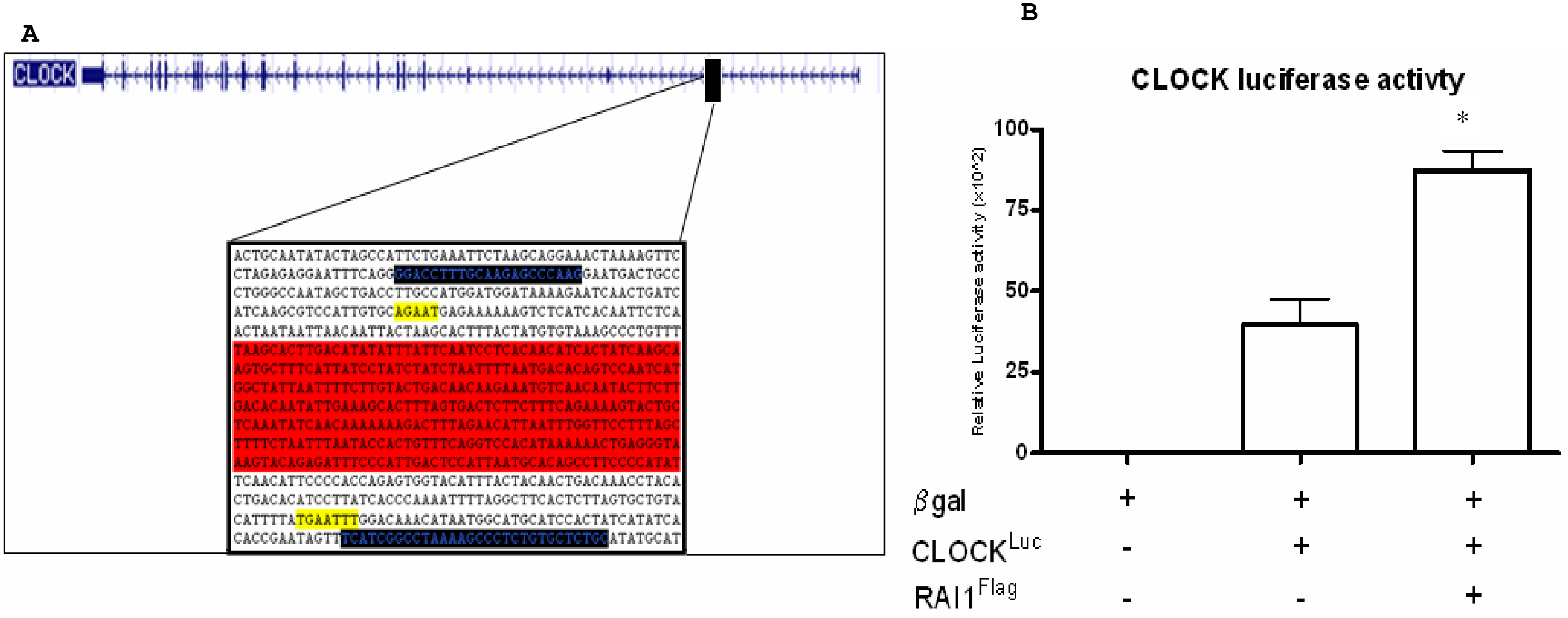


Figure 20. RAI1 impact on *CLOCK* transcription. A) Intron 1 *CLOCK* enhancer region used for luciferase assays. Shown in red: ~350bp region identified in Nimblegen array as most likely binding region of RAI1^{Flag}. Shown in yellow: Proposed RAI1 consensus binding sites as reported by Eri Kamura (Elsea Lab). Shown in black: PCR primers used to clone region (CLOCK^{Luc}) used for luciferase assays (724 bp). B) Luciferase activity of co-transfection of RAI1^{Flag} with CLOCK^{Luc} plasmids. N=3 studies with readings taken in duplicate. *, P=0.009.

Dysregulation of circadian genes in *Rai1+/-* mice

Rai1+/- mice are obese, hyperphagic, have decreased sensitivity to pain, low muscle strength, and an altered circadian rhythm [42, 45, 144]. Expression analysis of *Rai1*+/- kidney and whole brain shows a marked decrease in both *Rai1* and *Clock* (mouse orthologue to *CLOCK*) (Figure 20). Additionally, expression of *Rai1* and *Clock* in the mouse hypothalamus showed reduced expression (Figure 21). Further, we analyzed the expression of downstream targets of *Clock* including *Per2*, *Npas2*, and *Nr1d2* in the mouse hypothalamus, all of which had reduced expression as compared to WT C57Bl/6J littermates during the day time hours. These data show that there is circadian dysregulation of gene expression in the *Rai1+/-* mouse hypothalamus.

Next it was important to measure circadian gene expression in the mouse hypothalamus during the day time and night time hours to ensure that rhythmicity was not restored with regard to external cues such as feeding. Our results show that during the dark phase the expression of core circadian genes is dysregulated including *Per2* which when knocked out in mice results in a severe circadian phenotype [145] and *Nr1d2* which was identified as one of the top genes with reduced expression in a study of human cells knocked down for RAI1 [81] (Figure 22). Interestingly, *Clock* expression is upregulated in night time, suggesting a malfunction of the autoregulatory feedback loop.

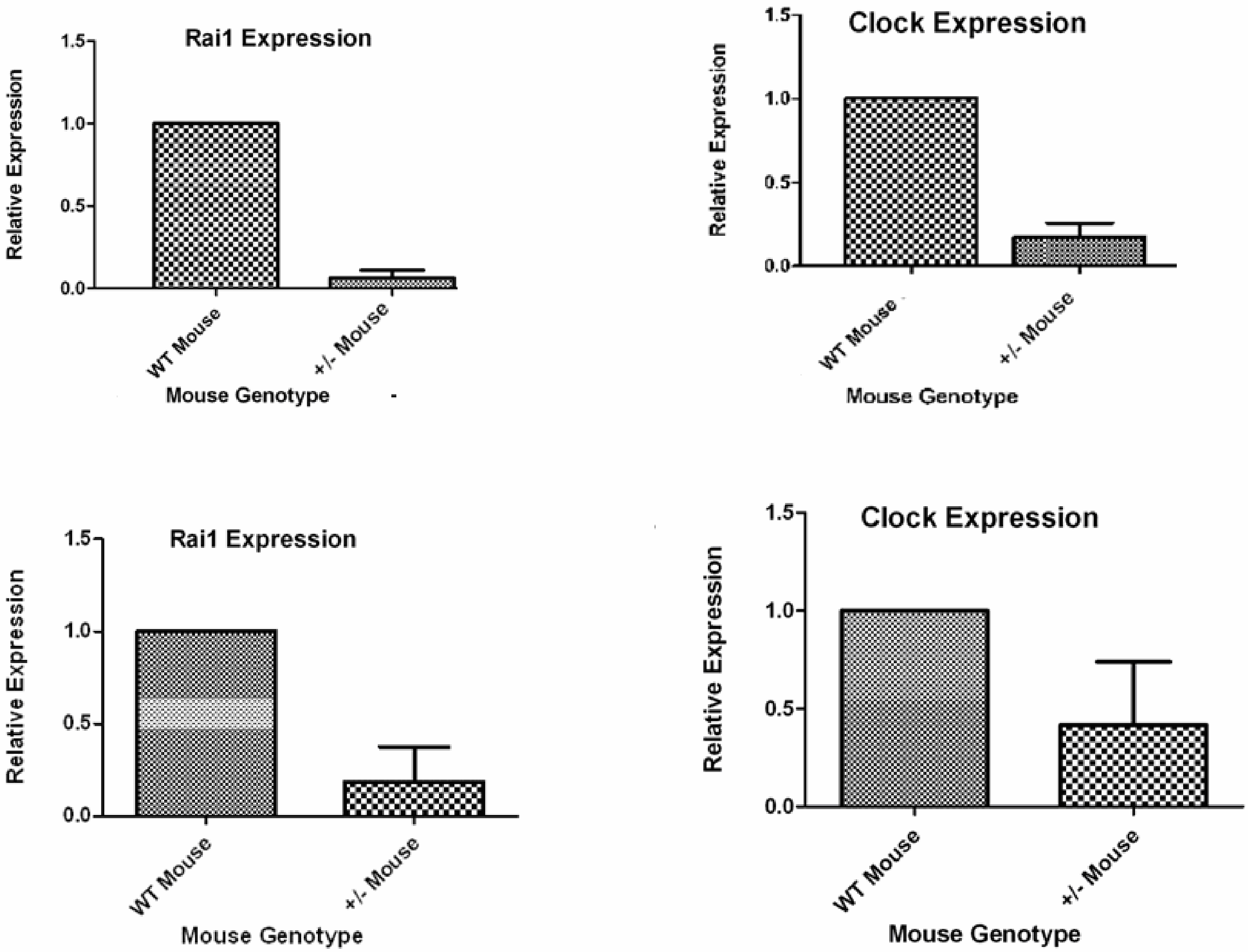


Figure 21. *Rai1* and *Clock* expression in mouse tissues. A-B) Mouse Brain. C-D) Mouse Kidney. N=3 mice per tissue. All mice sacrificed during the day time hours.

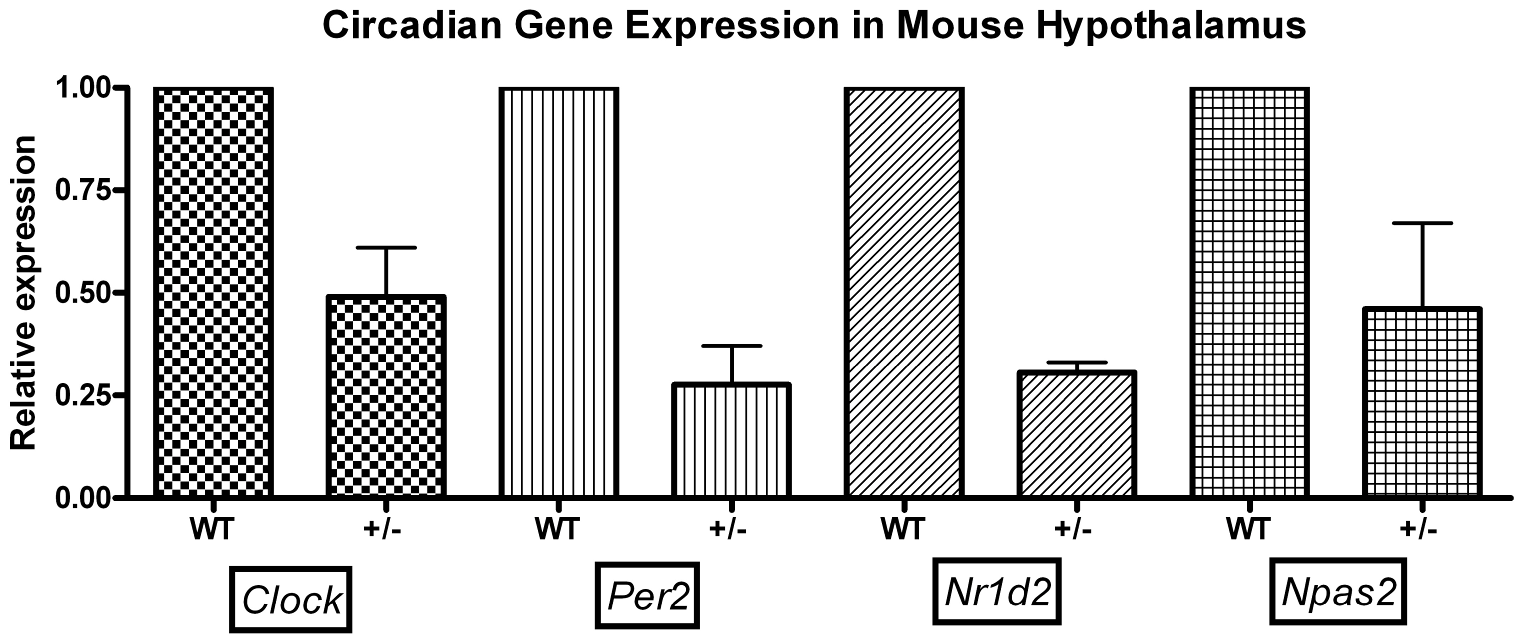


Figure 22. qPCR of circadian genes expressed in mouse hypothalamus. N=3 mice per tissue. All mice sacrificed during light phase.

Dysregulation of circadian genes in SMS fibroblasts

Human fibroblasts follow a 2 4hr rhythmicity *in vitro* just as they would in *vivo*. To assess the circadian cycling of *RAI1, CLOCK,* and *NR1D2* we took RNA samples from unaffected (GM637), deletion (SMS182), and mutation (SMS175) cell lines. Transcription of all of these genes is dysregulated when compared to unaffected fibroblasts. It is worth while to note that the hypothesis to date, with regard to the inverted circadian rhythm seen in SMS, is that an inverted secretion of melatonin causes sleep disturbance. The data presented here suggest that this inverted circadian rhythm is likely due to a dysregulation of circadian gene expression, and the inverted melatonin secretion seen in the majority of cases is likely a byproduct of said dysregulation. Further supporting our hypothesis is the fact that C57BL6 mice do not secrete melatonin [146].

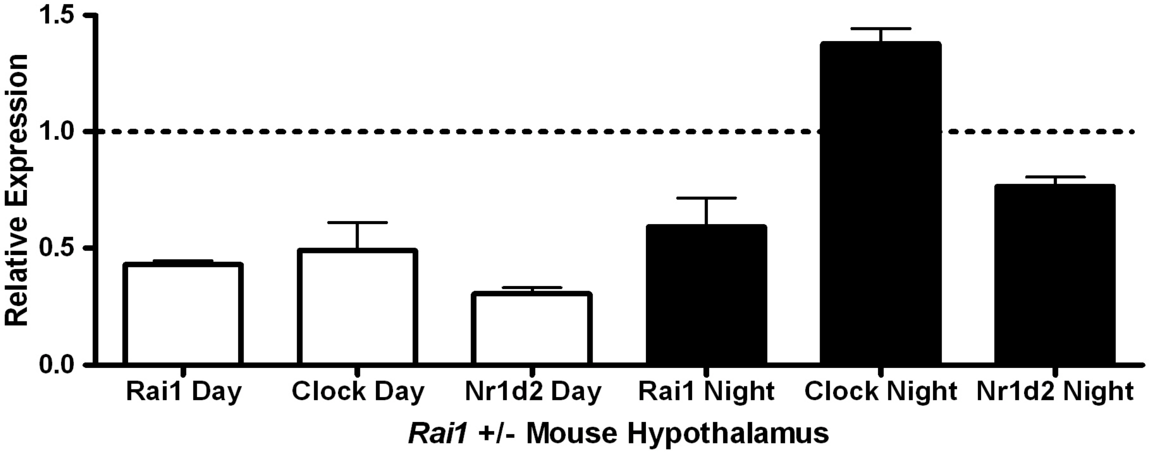


Figure 23. qPCR of core circadian gene expression in the mouse hypothalamus during day light (white boxes) and night time (black boxes) periods. Circadian rhythm remains disrupted during both light and dark phase hours relative to WT baseline at the same phase (1, dotted line). N=3 mice per tissue.

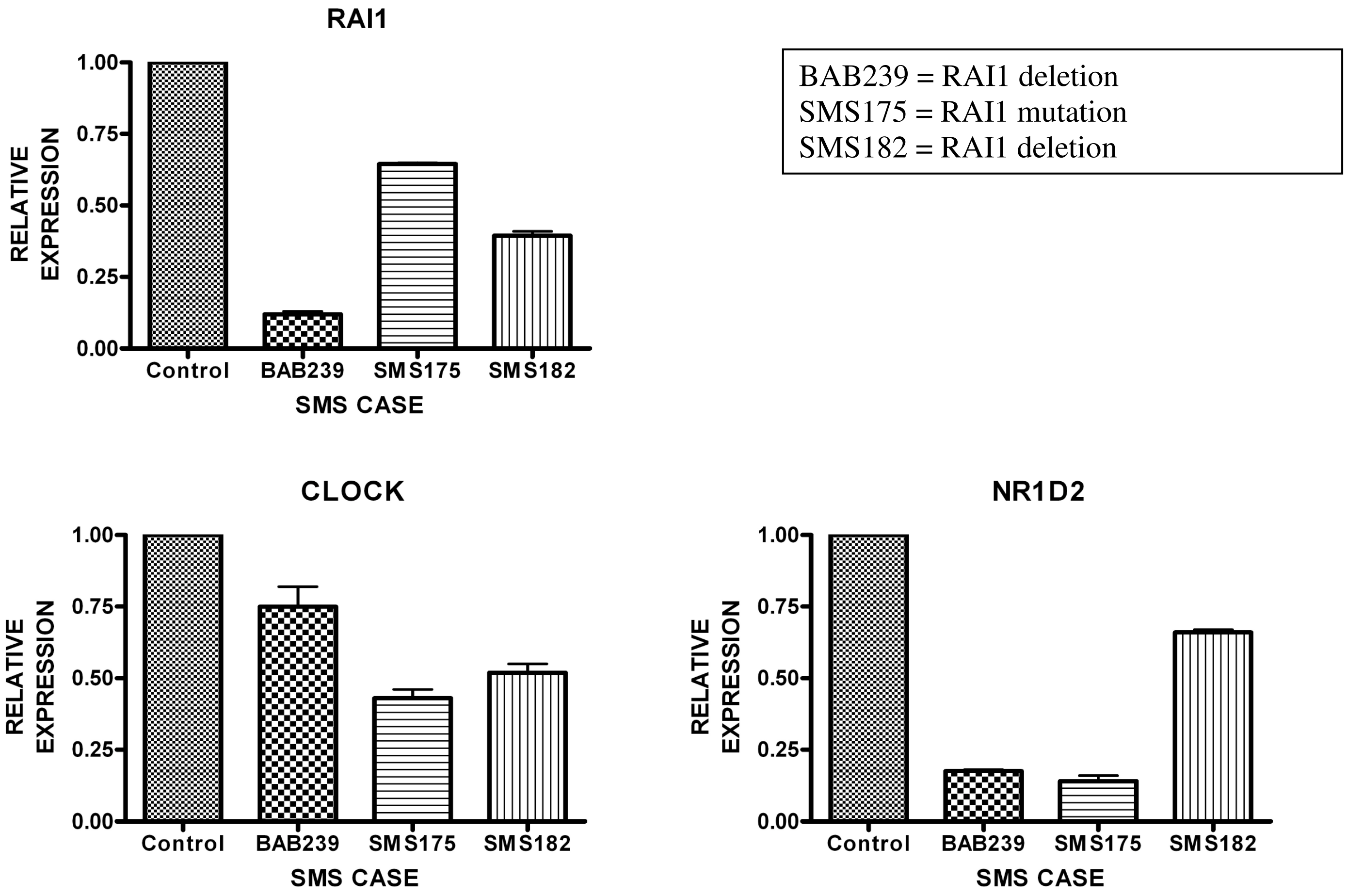


Figure24. Core circadian gene expression in unaffected and SMS fibroblasts. N=3 qPCR experiments per cell line performed in triplicate.

Discussion

The circadian rhythm defect seen in Smith-Magenis syndrome is thought to contribute to a more severe behavioral phenotype [14]. It is thought that if sleep rhythmicity is resolved, treatment of behavior becomes easier. This is why it is imperative to understand the molecular defects that result from *RAI1* mutation or deletion.

RAI1 is a proposed transcription factor containing a PHD domain and nuclear localization signal. However, it was previously unknown in what biochemical pathways RAI1 functions and what genes it may regulate. Identification of the molecular changes and subsequent affected pathways that result from *RAI1* haploinsufficiency holds the potential for chemical, habitual, or nutritional based interventions.

The *Clock/Clock* mutant mice were created my King et al. in 1996 [147]. By using a forward genetics approach, mutagenesis followed breeding and positional cloning, the authors identified a mutant mouse with an A to T splice site mutation which results in deletion of 51 amino acids of the protein. This mutation disrupts exons 16, 17, and 18 of the 24-exon gene product. This *Clock* mutant mouse was bred to generate the *Clock/Clock* mutant mouse which has an erratic and unpredictable circadian sleep cycle, with loss of periodicity, as compared to WT littermates indicating its importance in circadian homeostasis [137]. As a complementary approach to positional cloning, Antoch et al. 1997 [148] used *in vivo* complementation with BAC clones expressed in transgenic mice to identify the circadian *Clock* gene. A *Clock* transgene completely rescued both the long-period and the loss-of-rhythm phenotypes in *Clock* mutant mice. These over expressing *Clock* mice revealed a shortened period length lending to the fact that Clock is a dosage sensitive gene. Surprisingly, when DeBruyne et al. created the *Clock-/-*

complete knockout there was no loss of rhythmicity, indicating that multiple genes may play a critical role as “master regulators” or have compensatory effects [149].

As it is important to not only study the influence RAI1 has on *CLOCK* it is also necessary to analyze further downstream effects. The *period circadian protein homolog 2* (*PER2*) is a very important circadian transcript involved with the correct rhythmicity of circadian cycling. The *Per2* homozygous mutation mouse (*Per2*$^{m/m}$) was created by Zheng et al. in 1999 [150]. This mouse contains a deletion in the PAS domain of the mouse *Per2* gene. The PAS domain is highly conserved in circadian clock genes from a variety of species [150]. The *Per2*$^{m/m}$ mouse exhibits a variety of circadian phenotypes including a shorter circadian period in normal 12 hour light/12 hour dark conditions, followed by loss of circadian rhythmicity in constant darkness as measured by wheel running measurements [150]. Additionally, *Per1* mRNA levels were reduced in *Per2*$^{m/m}$ mice, indicating that *Per2* is upstream of *Per1* [150].

Further, Girirajan et al. 2009 [81] identified *NR1D2* in a *RAI1* knockdown study in HEK293t cells. This study included siRNA knock-down of *RAI1* followed by microarray analysis. *NR1D2* was one of the top genes down regulated in this study and adds to the validity of our findings.

Polymorphisms in the key players of circadian rhythm (*CLOCK, BMAL1, PER3*) have been implicated in a variety of mood disorders in humans [151, 152]. Additionally, single nucleotide polymorphisms (SNPs) in *CLOCK* have been associated with defects including, bipolar disorder [153], insomnia with bipolar disorder [154, 155], major depressive disorder [156] and obesity. Additionally *BMAl1* [157, 158], *PER3* [157] and *TIMLESS* [158] have all been associated with bipolar disorder. The behavioral phenotype seen in Smith-Magenis syndrome has also been described as manic in some cases lending to the idea that disruption of the circadian rhythm may not only complicate the behavioral phenotype but contribute directly to it.

Given the important role that *CLOCK* plays in circadian rhythm and its association with human disease, many studies have been performed to identify the function of CLOCK, but its transcriptional regulator(s) has remained vague. Here we present data that suggest that RAI1 is an important enhancer of *CLOCK* transcription. Using both *in vivo* and *in vitro* tools we were able to show that RAI1 regulates *CLOCK* via an enhancer region located in intron 1 which contains two proposed RAI1 consensus binding sites identified by Eri Kamura (Elsea lab) [159].

These data suggest that *RAI1* plays an important role in maintaining circadian rhythmicity and because circadian rhythms impact development, sleep, and behavior these findings should help to further delineate the pathways involved in syndromes wherein these functions are disrupted. Additionally, as better treatments are developed to target these pathways, we may one day be able to correct many of the sleep and behavioral phenotypes seen in Smith-Magenis syndrome.

Chapter 6

Summary

These studies have identified new loci associated with a Smith-Magenis-like phenotype. We have shown that SMS and SMS-like cohorts are not different with regard to the core phenotypes seen most commonly in SMS. We have also shown that alternate loci, that when deleted or duplicated, can result in a phenotype similar to SMS. Additionally, we have been able to make phenotypic connections between SMS and SMS-like syndromes which has already improved diagnosis of these disorders (Chapter 2).

The broad aCGH study allowed us to explore in detail the deletion 2q23 syndrome. Using overlapping deletions from our case, as well as new and reported cases, we were able to better define the phenotype of this syndrome, as well as narrow the critical region to one gene, *MBD5*. We believe that haploinsufficiency of *MBD5* contributes to the major phenotypes in deletion 2q23 syndrome, including microcephaly, intellectual disabilities, severe speech impairment, and seizures (Chapter 3). We believe that mutation screening of *MBD5* in individuals with a similar phenotype is an appropriate next step in confirming the weight of contribution that *MBD5* holds in this rare syndrome.

Additionally, using a similar approach as in Chapter 3 we were able to refine the critical region for Brachydactyly Mental Retardation syndrome (deletion 2q37 syndrome) and identified an insertion mutation of *HDAC4* in one case, SMS117. Mutation of *HDAC4* likely contributes to the major phenotypes seen in BDMR and identification of more mutation cases is likely in the future (Chapter 4).

Lastly, we were able to characterize, in part, the molecular function of *RAI1*, the gene that when mutated or deleted results in Smith-Magenis syndrome. There is a strong circadian phenotype associated with SMS wherein subjects have consistent night awakenings with frequent day time napping. We believe that *RAI1* plays a significant role in the central circadian pathway. Using both *in vivo* and *in vitro* methods, we have shown that RAI1 regulates *CLOCK* gene expression, which is a novel breakthrough, as the transcriptional regulators of *CLOCK* have not been defined (Chapter 5).

Taken together these studies have further defined and refined the molecular root for each syndrome. However, the question remains: "What specific pathways and genes are logical to explore in the future with regard to these syndromes and the impact of *RAI1* on human development and behavior?"

Development

In an exploratory bioinformatic study by Danielle Bartholomew (Elsea lab) *MEF* binding sites in the 5' untranslated region of *RAI1* were identified. As stated above *MEF2C* activity is necessary for proper chondrocyte hypertrophy and bone development. MEF2C is also inhibited by HDAC4. Additionally, the *bone morphogenic protein 5* gene (*BMP5*), a member of the transforming growth factor-beta superfamily of regulatory molecules, was identified in our ChIP-Chip studies lending to the idea that RAI1 regulates it on the transcriptional level. *BMP5* is necessary for proper skeletogenesis, and *BMP5* mutant mice have a variety of skeletal and soft tissue abnormalities [160]. It is also worthwhile to note that BMP5 has been shown to elicit dendritic growth *in vitro* [161], which may contribute to some of the neurological phenotypes seen in SMS. In these studies we have shown that mutation of *HDAC4* results in many of the features seen in SMS as well as distinct brachydactyly not seen in SMS, which is logical considering *HDAC4* may be upstream of *RAI1* (Figure 24). My belief is that MEF2C is a positive regulator of *RAI1* transcription which is modulated by HDAC4. I also believe that RAI1 may be a positive regulator of *BMP5* ultimately leading to proper skeletal development (Figure 24). However, it must be mentioned that HDAC4 is also a regulator of *RUNX2* which is known to be necessary for proper skeletogenesis suggesting multiple pathways involved. Additionally, as shown in Tables 12 and 13 RAI1 may regulate *MEF2A* suggesting a possible autoregulatory feedback loop with regard to RAI1 function in skeletogenesis.

Behavior and sleep

The behavioral phenotypes seen in SMS and SMS-like syndromes are significant. They include, but are not limited to, verbal outbursts, temper tantrums, self-abusive behavior, and stereotypies. These phenotypes all have a strong neurological element and are complicated by sleep disturbance. It has been shown that solving the sleep problems seen in SMS may relieve many of the behavioral phenotypes [14]. It is my proposition that *RAI1* plays a major role in circadian rhythm and that when RAI1 is mutated or deleted the resulting behavioral phenotypes are a direct result of this haploinsufficiency.

We have shown here that RAI1 is a positive regulator of *CLOCK*, the master regulator of central circadian gene expression. We used both animal models and cell based studies to show the disruption of the circadian cycle in SMS subjects. *CLOCK* has been shown to be associated with a variety of disorders including schizophrenia, bipolar disorder, mania and depression all of which have elements seen in the SMS phenotype. A schematic of the influence RAI1 has on the circadian cycle can been in seen in Figure 25.

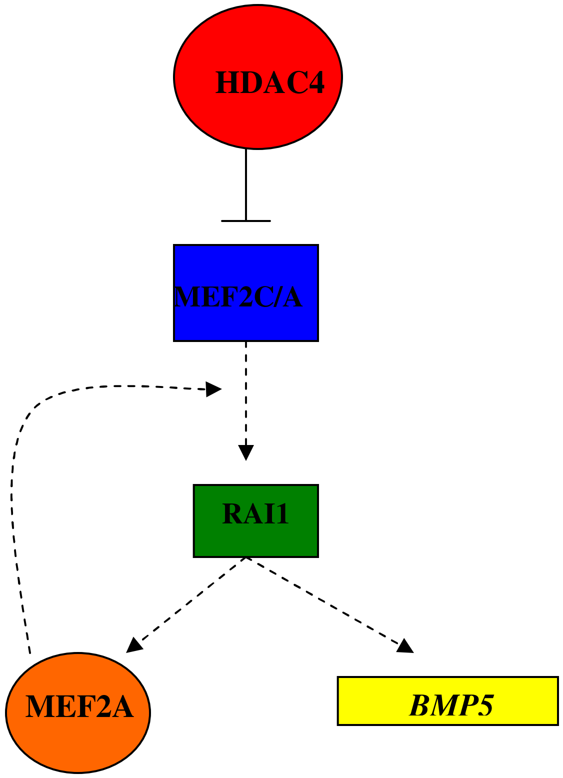


Figure 25. Proposed schematic of *RAI1* contribution to proper skeletal development. HDAC4 negatively regulates MEF2C on the protein level which then leads to reduced *RAI1* transcription leading to skeletal, neurological and behavioral phenotypes . Reduced *RAI1* transcription then leads to reduced *BMP5* and *MEF2A* transcription which may then lead to neurological and skeletal phenotypes as well as malfunction of the proposed feedback loop.

Theoretically, if *CLOCK* expression is reduced, downstream gene expression would also be reduced. This could have a variety of impacts; molecularly, this reduction of *CLOCK* and downstream genes would also result in dysregulation of the regulatory feedback loops. Clinically, this could result in an altered sleep cycle, mood disorder and even learning disability and developmental problems.

Future directions

With regard to the SMS-like study and resulting studies (Chapters 2-4), I would like to see more subjects ascertained and screened for *HDAC4* and *MBD5* mutations. I would also like to see molecular characterization of these mutations. For example, cDNA from subjects with mutations could be collected and cloned into a variety of vectors. A Flag vector already in use in the Elsea lab could be used to perform IP studies to identify presence or absence of specific binding partners. Also, stable transfection of these mutant alleles could help to identify weather they are functionally active and able to perform their transcriptional roles on downstream genes. The molecular follow up will allow further insight into the pathways involved in these syndromes and how mutation specifically affects pathways. I would also like to see a detailed panel of molecular tests, including aCGH, MLPA and gene sequencing, and phenotypic check lists developed such that the proper diagnosis for these individuals may be more easily acquired in the future, helping families, clinicians, and the subjects themselves.

With regard to the molecular function of RAI1, I would like to see a variety of studies performed. First and foremost I think that follow up with the remaining genes of interest from the ChIP-Chip study is imperative, specifically the top 10 genes identified in Table 13. Additionally, I would like to see a study performed that identifies binding partners of RAI1.

This can be easily performed utilizing the RAI1^{Flag} plasmid already available with IP and mass spec analysis. Alternatively, the RAI1 coding sequence could easily be cloned into a vector containing a HIS-tag and column chromatography could be utilized which may reduce some of the background associated with IP. If RAI1 binds to a well characterized protein, this will allow us further insight into the molecular functions of RAI1 and possibly lead us to proteins that regulate *RAI1* on the transcriptional level. Cloning of the RAI1 promoter would also help with this endeavor (being pursued by Dr. Deborah Zies). I would also like to see the *Rai1+/-* mice crossed with the *Clock-/-* mice which do not have a circadian phenotype. If the resulting mice do have a strong circadian phenotype we may be able to identify alternative circadian pathways that *RAI1* influences outside of *CLOCK* transcription. These studies will not only help SMS and SMS-like subjects but could have impact on the general public and those with psychiatric disorders.

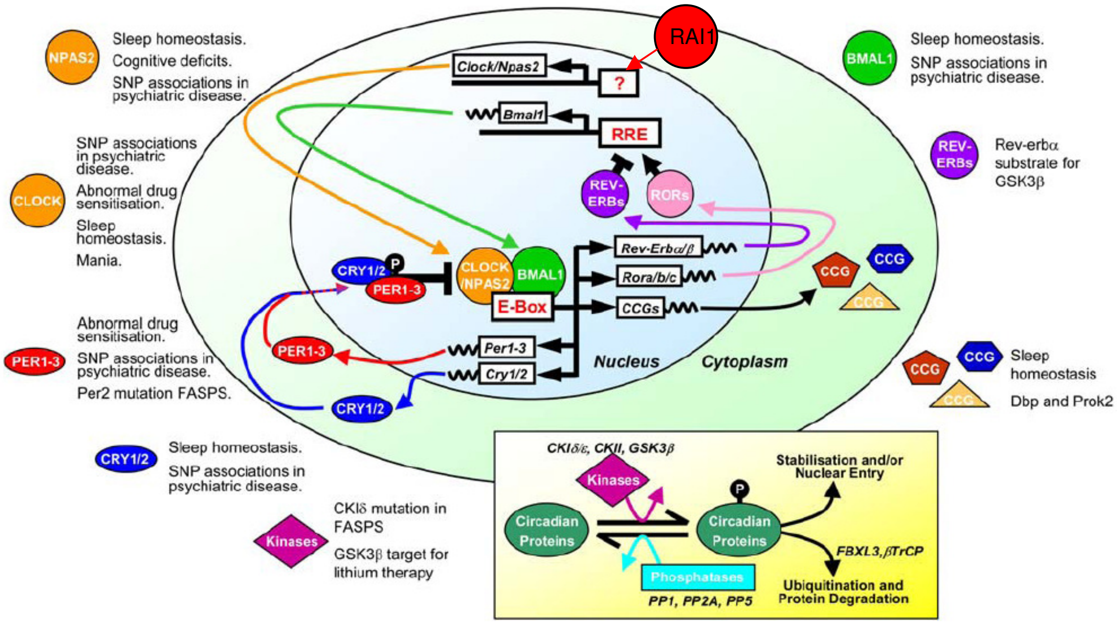

RAI1
NPAS2
Sleep homeostasis.
Cognitive deficits.
SNP associations in psychiatric disease.
BMAL1
Sleep homeostasis.
SNP associations in psychiatric disease.
Clock/Npas2
?
Bmal1
RRE
REV-ERBs
RORs
REV-ERBs
Rev-erbα substrate for GSK3β
CLOCK
SNP associations in psychiatric disease.
Abnormal drug sensitisation.
Sleep homeostasis.
Mania.
CRY1/2
P
PER1-3
CLOCK /NPAS2
BMAL1
E-Box
Rev-Erbα/β
Rora/b/c
CCGs
CCG
CCG
CCG
PER1-3
Per1-3
Cry1/2
Nucleus
Cytoplasm
CRY1/2
PER1-3
Abnormal drug sensitisation.
SNP associations in psychiatric disease.
Per2 mutation FASPS.
CCG
CCG
CCG
Sleep homeostasis
Dbp and Prok2
CRY1/2
Sleep homeostasis.
SNP associations in psychiatric disease.
Kinases
CKIδ mutation in FASPS
GSK3β target for lithium therapy
CKIδ/ε, CKII, GSK3β
Kinases
Circadian Proteins
P
Circadian Proteins
Stabilisation and/or Nuclear Entry
FBXL3,βTrCP
Phosphatases
PP1, PP2A, PP5
Ubiquitination and Protein Degradation

Figure 26. RAI1 regulates clock transcription and is a major element in the circadian cycle. Modified from Nolan et al. 2008.

References

1 Wagenstaller J, Spranger S, Lorenz-Depiereux B, Kazmierczak B, Nathrath M, Wahl D, Heye B, Glaser D, Liebscher V, Meitinger T, Strom TM. Copy-number variations measured by single-nucleotide-polymorphism oligonucleotide arrays in patients with mental retardation. *Am J Hum Genet* 2007;**81**(4):768-79.

2 Williams SR, Mullegama SV, Rosenfeld JA, Dagli AI, Hatchwell E, Allen WP, Williams CA, Elsea SH. Haploinsufficiency of MBD5 associated with a syndrome involving microcephaly, intellectual disabilities, severe speech impairment, and seizures. *Eur J Hum Genet*;**18**(4):436-41.

3 Smith AC, McGavran L, Robinson J, Waldstein G, Macfarlane J, Zonona J, Reiss J, Lahr M, Allen L, Magenis E. Interstitial deletion of (17)(p11.2p11.2) in nine patients. *Am J Med Genet* 1986;**24**(3):393-414.

4 Slager RE, Newton TL, Vlangos CN, Finucane B, Elsea SH. Mutations in RAI1 associated with Smith-Magenis syndrome. *Nat Genet* 2003;**33**(4):466-8.

5 Shinawi M, Cheung SW. The array CGH and its clinical applications. *Drug Discov Today* 2008.

6 Yatsenko SA, Treadwell-Deering D, Krull K, Lewis RA, Glaze D, Stankiewicz P, Lupski JR, Potocki L. Trisomy 17p10-p12 due to mosaic supernumerary marker chromosome: delineation of molecular breakpoints and clinical phenotype, and comparison to other proximal 17p segmental duplications. *Am J Med Genet A* 2005;**138**(2):175-80.

7 Zhang F, Potocki L, Sampson JB, Liu P, Sanchez-Valle A, Robbins-Furman P, Navarro AD, Wheeler PG, Spence JE, Brasington CK, Withers MA, Lupski JR. Identification of

uncommon recurrent Potocki-Lupski syndrome-associated duplications and the distribution of rearrangement types and mechanisms in PTLS. *Am J Hum Genet*;**86**(3):462-70.

8 Turner DJ, Miretti M, Rajan D, Fiegler H, Carter NP, Blayney ML, Beck S, Hurles ME. Germline rates of de novo meiotic deletions and duplications causing several genomic disorders. *Nat Genet* 2008;**40**(1):90-5.

9 Lee JA, Carvalho CM, Lupski JR. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell* 2007;**131**(7):1235-47.

10 Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. *Nat Rev Genet* 2006;**7**(2):85-97.

11 Lupski JR. Genomic rearrangements and sporadic disease. *Nat Genet* 2007;**39**(7 Suppl):S43-7.

12 Bi W, Yan J, Stankiewicz P, Park SS, Walz K, Boerkoel CF, Potocki L, Shaffer LG, Devriendt K, Nowaczyk MJ, Inoue K, Lupski JR. Genes in a refined Smith-Magenis syndrome critical deletion interval on chromosome 17p11.2 and the syntenic region of the mouse. *Genome Res* 2002;**12**(5):713-28.

13 Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill

X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. Global variation in copy number in the human genome. *Nature* 2006;**444**(7118):444-54.

14 Elsea SH, Girirajan S. Smith-Magenis syndrome. *Eur J Hum Genet* 2008;**16**(4):412-21.

15 Zori RT, Lupski JR, Heju Z, Greenberg F, Killian JM, Gray BA, Driscoll DJ, Patel PI, Zackowski JL. Clinical, cytogenetic, and molecular evidence for an infant with Smith-Magenis syndrome born from a mother having a mosaic 17p11.2p12 deletion. *Am J Med Genet* 1993;**47**(4):504-11.

16 Allanson JE, Greenberg F, Smith AC. The face of Smith-Magenis syndrome: a subjective and objective study. *J Med Genet* 1999;**36**(5):394-7.

17 Tomona N, Smith AC, Guadagnini JP, Hart TC. Craniofacial and dental phenotype of Smith-Magenis syndrome. *Am J Med Genet A* 2006;**140**(23):2556-61.

18 Edelman E, Girirajan S, Finucane B, Patel P, Lupski J, Smith A, Elsea S. Gender, genotype, and phenotype differences in Smith-Magenis syndrome: a meta-analysis of 105 cases. *Clin Genet* 2007;**71**(6):540-50.

19 Smith AC, Dykens E, Greenberg F. Sleep disturbance in Smith-Magenis syndrome (del 17 p11.2). *Am J Med Genet* 1998;**81**(2):186-91.

20 Gropman AL, Elsea S, Duncan WC, Jr., Smith AC. New developments in Smith-Magenis syndrome (del 17p11.2). *Curr Opin Neurol* 2007;**20**(2):125-34.

21 Di Cicco M, Padoan R, Felisati G, Dilani D, Moretti E, Guerneri S, Selicorni A. Otorhinolaringologic manifestation of Smith-Magenis syndrome. *Int J Pediatr Otorhinolaryngol* 2001;**59**(2):147-50.

22 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. Multi-

disciplinary clinical study of Smith-Magenis syndrome (deletion 17p11.2). *Am J Med Genet* 1996;**62**(3):247-54.

23 Liburd N, Ghosh M, Riazuddin S, Naz S, Khan S, Ahmed Z, Riazuddin S, Liang Y, Menon PS, Smith T, Smith AC, Chen KS, Lupski JR, Wilcox ER, Potocki L, Friedman TB. Novel mutations of MYO15A associated with profound deafness in consanguineous families and moderately severe hearing loss in a patient with Smith-Magenis syndrome. *Hum Genet* 2001;**109**(5):535-41.

24 Chen RM, Lupski JR, Greenberg F, Lewis RA. Ophthalmic manifestations of Smith-Magenis syndrome. *Ophthalmology* 1996;**103**(7):1084-91.

25 Finucane BM, Jaeger ER, Kurtz MB, Weinstein M, Scott CI, Jr. Eye abnormalities in the Smith-Magenis contiguous gene deletion syndrome. *Am J Med Genet* 1993;**45**(4):443-6.

26 Moeschler JB, Shevell M. Clinical genetic evaluation of the child with mental retardation or developmental delays. *Pediatrics* 2006;**117**(6):2304-16.

27 Madduri N, Peters SU, Voigt RG, Llorente AM, Lupski JR, Potocki L. Cognitive and adaptive behavior profiles in Smith-Magenis syndrome. *J Dev Behav Pediatr* 2006;**27**(3):188-92.

28 Gropman AL, Duncan WC, Smith AC. Neurologic and developmental features of the Smith-Magenis syndrome (del 17p11.2). *Pediatr Neurol* 2006;**34**(5):337-50.

29 Smith AC, Dykens E, Greenberg F. Behavioral phenotype of Smith-Magenis syndrome (del 17p11.2). *Am J Med Genet* 1998;**81**(2):179-85.

30 Boudreau EA, Johnson KP, Jackman AR, Blancato J, Huizing M, Bendavid C, Jones M, Chandrasekharappa SC, Lewy AJ, Smith AC, Magenis RE. Review of disrupted sleep

patterns in Smith-Magenis syndrome and normal melatonin secretion in a patient with an atypical interstitial 17p11.2 deletion. *Am J Med Genet A* 2009;**149A**(7):1382-91.

31 Dykens EM, Finucane BM, Gayley C. Brief report: cognitive and behavioral profiles in persons with Smith-Magenis syndrome. *J Autism Dev Disord* 1997;**27**(2):203-11.

32 Greenberg F, Guzzetta V, Montes de Oca-Luna R, Magenis RE, Smith AC, Richter SF, Kondo I, Dobyns WB, Patel PI, Lupski JR. Molecular analysis of the Smith-Magenis syndrome: a possible contiguous-gene syndrome associated with del(17)(p11.2). *Am J Hum Genet* 1991;**49**(6):1207-18.

33 Dykens EM, Smith AC. Distinctiveness and correlates of maladaptive behaviour in children and adolescents with Smith-Magenis syndrome. *J Intellect Disabil Res* 1998;**42 (Pt 6)**:481-9.

34 Girirajan S, Vlangos CN, Szomju BB, Edelman E, Trevors CD, Dupuis L, Nezarati M, Bunyan DJ, Elsea SH. Genotype-phenotype correlation in Smith-Magenis syndrome: evidence that multiple genes in 17p11.2 contribute to the clinical spectrum. *Genet Med* 2006;**8**(7):417-27.

35 Treadwell-Deering DE, Powell MP, Potocki L. Cognitive and behavioral characterization of the Potocki-Lupski syndrome (duplication 17p11.2). *J Dev Behav Pediatr*;**31**(2):137-43.

36 Potocki L, Bi W, Treadwell-Deering D, Carvalho CM, Eifert A, Friedman EM, Glaze D, Krull K, Lee JA, Lewis RA, Mendoza-Londono R, Robbins-Furman P, Shaw C, Shi X, Weissenberger G, Withers M, Yatsenko SA, Zackai EH, Stankiewicz P, Lupski JR. Characterization of Potocki-Lupski syndrome (dup(17)(p11.2p11.2)) and delineation of a

dosage-sensitive critical interval that can convey an autism phenotype. *Am J Hum Genet* 2007;**80**(4):633-49.

37 Potocki L, Chen KS, Park SS, Osterholm DE, Withers MA, Kimonis V, Summers AM, Meschino WS, Anyane-Yeboa K, Kashork CD, Shaffer LG, Lupski JR. Molecular mechanism for duplication 17p11.2- the homologous recombination reciprocal of the Smith-Magenis microdeletion. *Nat Genet* 2000;**24**(1):84-7.

38 Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Eswara P, Eyras E, Felsenfeld A, Fewell GA, Flicek P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S, Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigo R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, Karolchik D, Kasprzyk A, Kawai J, Keibler E, Kells C, Kent WJ, Kirby A, Kolbe DL, Korf I, Kucherlapati RS, Kulbokas EJ, Kulp D, Landers T, Leger JP, Leonard S, Letunic I, Levine R, Li J, Li M, Lloyd C, Lucas S, Ma B, Maglott DR, Mardis ER, Matthews L, Mauceli E, Mayer JH, McCarthy M, McCombie WR, McLaren S,

McLay K, McPherson JD, Meldrim J, Meredith B, Mesirov JP, Miller W, Miner TL, Mongin E, Montgomery KT, Morgan M, Mott R, Mullikin JC, Muzny DM, Nash WE, Nelson JO, Nhan MN, Nicol R, Ning Z, Nusbaum C, O'Connor MJ, Okazaki Y, Oliver K, Overton-Larty E, Pachter L, Parra G, Pepin KH, Peterson J, Pevzner P, Plumb R, Pohl CS, Poliakov A, Ponce TC, Ponting CP, Potter S, Quail M, Reymond A, Roe BA, Roskin KM, Rubin EM, Rust AG, Santos R, Sapojnikov V, Schultz B, Schultz J, Schwartz MS, Schwartz S, Scott C, Seaman S, Searle S, Sharpe T, Sheridan A, Shownkeen R, Sims S, Singer JB, Slater G, Smit A, Smith DR, Spencer B, Stabenau A, Stange-Thomann N, Sugnet C, Suyama M, Tesler G, Thompson J, Torrents D, Trevaskis E, Tromp J, Ucla C, Ureta-Vidal A, Vinson JP, Von Niederhausern AC, Wade CM, Wall M, Weber RJ, Weiss RB, Wendl MC, West AP, Wetterstrand K, Wheeler R, Whelan S, Wierzbowski J, Willey D, Williams S, Wilson RK, Winter E, Worley KC, Wyman D, Yang S, Yang SP, Zdobnov EM, Zody MC, Lander ES. Initial sequencing and comparative analysis of the mouse genome. *Nature* 2002;**420**(6915):520-62.

39 Walz K, Caratini-Rivera S, Bi W, Fonseca P, Mansouri DL, Lynch J, Vogel H, Noebels JL, Bradley A, Lupski JR. 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* 2003;**23**(10):3646-55.

40 Girirajan S, Williams SR, Garben JY, Nowak NJ, Hatchwell E, Elsea S. 17p11.2p12 triplication and del(17)q11.2q12 in a severely affected child with dup(17)p11.2p12 syndrome *Clin Genet* 2007;**71**:1-12, In Press.

41 Walz K, Spencer C, Kaasik K, Lee CC, Lupski JR, Paylor R. Behavioral characterization of mouse models for Smith-Magenis syndrome and dup(17)(p11.2p11.2). *Hum Mol Genet* 2004;**13**(4):367-78.

42 Bi W, Ohyama T, Nakamura H, Yan J, Visvanathan J, Justice MJ, Lupski JR. Inactivation of Rai1 in mice recapitulates phenotypes observed in chromosome engineered mouse models for Smith-Magenis syndrome. *Hum Mol Genet* 2005;**14**(8):983-95.

43 Girirajan S, Patel N, Slager RE, Tokarz ME, Bucan M, Wiley JL, Elsea SH. How much is too much? Phenotypic consequences of Rai1 overexpression in mice. *Eur J Hum Genet* 2008;**16**(8):941-54.

44 Girirajan S, Elsea SH. Abnormal maternal behavior, altered sociability, and impaired serotonin metabolism in Rai1-transgenic mice. *Mamm Genome* 2009;**20**(4):247-55.

45 Bi W, Yan J, Shi X, Yuva-Paylor LA, Antalffy BA, Goldman A, Yoo JW, Noebels JL, Armstrong DL, Paylor RE, Lupski JR. Rai1 deficiency in mice causes learning impairments and motor dysfunctions whereasRai1heterozygous mice display minimal behavioral phenotypes. *Hum Mol Genet* 2007.

46 Walz K, Paylor R, Yan J, Bi W, Lupski JR. Rai1 duplication causes physical and behavioral phenotypes in a mouse model of dup(17)(p11.2p11.2). *J Clin Invest* 2006;**116**(11):3035-41.

47 McLaren J, Bryson SE. Review of recent epidemiological studies of mental retardation: prevalence, associated disorders, and etiology. *Am J Ment Retard* 1987;**92**(3):243-54.

48 Roeleveld N, Zielhuis GA, Gabreels F. The prevalence of mental retardation: a critical review of recent literature. *Dev Med Child Neurol* 1997;**39**(2):125-32.

49 Flint J, Knight S. The use of telomere probes to investigate submicroscopic rearrangements associated with mental retardation. *Curr Opin Genet Dev* 2003;**13**(3):310-6.

50 Hunter AG. Outcome of the routine assessment of patients with mental retardation in a genetics clinic. *Am J Med Genet* 2000;**90**(1):60-8.

51 Knight SJ, Lese CM, Precht KS, Kuc J, Ning Y, Lucas S, Regan R, Brenan M, Nicod A, Lawrie NM, Cardy DL, Nguyen H, Hudson TJ, Riethman HC, Ledbetter DH, Flint J. An optimized set of human telomere clones for studying telomere integrity and architecture. *Am J Hum Genet* 2000;**67**(2):320-32.

52 Rauch A, Hoyer J, Guth S, Zweier C, Kraus C, Becker C, Zenker M, Huffmeier U, Thiel C, Ruschendorf F, Nurnberg P, Reis A, Trautmann U. Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. *Am J Med Genet A* 2006;**140**(19):2063-74.

53 Vlangos CN, Wilson M, Blancato J, Smith AC, Elsea SH. Diagnostic FISH probes for del(17)(p11.2p11.2) associated with Smith-Magenis syndrome should contain the RAI1 gene. *Am J Med Genet A* 2005;**132**(3):278-82.

54 Girirajan S, Elsas LJ, 2nd, Devriendt K, Elsea SH. RAI1 variations in Smith-Magenis syndrome patients without 17p11.2 deletions. *J Med Genet* 2005;**42**(11):820-8.

55 Shevell MI, Bejjani BA, Srour M, Rorem EA, Hall N, Shaffer LG. Array comparative genomic hybridization in global developmental delay. *Am J Med Genet B Neuropsychiatr Genet* 2008;**147B**(7):1101-8.

56 Slavotinek AM. Novel microdeletion syndromes detected by chromosome microarrays. *Hum Genet* 2008;**124**(1):1-17.

57 Nowak NJ, Miecznikowski J, Moore SR, Gaile D, Bobadilla D, Smith DD, Kernstine K, Forman SJ, Mhawech-Fauceglia P, Reid M, Stoler D, Loree T, Rigual N, Sullivan M, Weiss LM, Hicks D, Slovak ML. Challenges in array comparative genomic hybridization for the analysis of cancer samples. *Genet Med* 2007;**9**(9):585-95.

58 Nowak NJ, Snijders AM, Conroy JM, Albertson DG. The BAC resource: tools for array CGH and FISH. *Curr Protoc Hum Genet* 2005;**Chapter 4**:Unit 4 13.

59 Snijders AM, Nowak N, Segraves R, Blackwood S, Brown N, Conroy J, Hamilton G, Hindle AK, Huey B, Kimura K, Law S, Myambo K, Palmer J, Ylstra B, Yue JP, Gray JW, Jain AN, Pinkel D, Albertson DG. Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet* 2001;**29**(3):263-4.

60 Miliaras D, Conroy J, Pervana S, Meditskou S, McQuaid D, Nowak N. Karyotypic changes detected by comparative genomic hybridization in a stillborn infant with chorioangioma and liver hemangioma. *Birth Defects Res A Clin Mol Teratol* 2007;**79**(3):236-41.

61 Christian SL, Brune CW, Sudi J, Kumar RA, Liu S, Karamohamed S, Badner JA, Matsui S, Conroy J, McQuaid D, Gergel J, Hatchwell E, Gilliam TC, Gershon ES, Nowak NJ, Dobyns WB, Cook EH, Jr. Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder. *Biol Psychiatry* 2008;**63**(12):1111-7.

62 Eugster EA, Berry SA, Hirsch B. Mosaicism for deletion 1p36.33 in a patient with obesity and hyperphagia. *Am J Med Genet* 1997;**70**(4):409-12.

63 Reish O, Berry SA, Hirsch B. Partial monosomy of chromosome 1p36.3: characterization of the critical region and delineation of a syndrome. *Am J Med Genet* 1995;**59**(4):467-75.

64 Jaillard S, Dubourg C, Gerard-Blanluet M, Delahaye A, Pasquier L, Dupont C, Henry C, Tabet AC, Lucas J, Aboura A, David V, Benzacken B, Odent S, Pipiras E. 2q23.1 microdeletion identified by array-CGH: an emerging phenotype with Angelman-like features? *J Med Genet* 2008.

65 Elghezal H, Sendi HS, Monastiri K, Lapierre JM, Romdhane SI, Mougou S, Saad A. Large duplication 4q25-q34 with mild clinical effect. *Ann Genet* 2004;**47**(4):419-22.

66 Zahir F, Firth HV, Baross A, Delaney AD, Eydoux P, Gibson WT, Langlois S, Martin H, Willatt L, Marra MA, Friedman JM. Novel deletions of 14q11.2 associated with developmental delay, cognitive impairment and similar minor anomalies in three children. *J Med Genet* 2007;**44**(9):556-61.

67 Lupski JR, de Oca-Luna RM, Slaugenhaupt S, Pentao L, Guzzetta V, Trask BJ, Saucedo-Cardenas O, Barker DF, Killian JM, Garcia CA, Chakravarti A, Patel PI. DNA duplication associated with Charcot-Marie-Tooth disease type 1A. *Cell* 1991;**66**(2):219-32.

68 Patel PI, Franco B, Garcia C, Slaugenhaupt SA, Nakamura Y, Ledbetter DH, Chakravarti A, Lupski JR. Genetic mapping of autosomal dominant Charcot-Marie-Tooth disease in a large French-Acadian kindred: identification of new linked markers on chromosome 17. *Am J Hum Genet* 1990;**46**(4):801-9.

69 Kirov G, Grozeva D, Norton N, Ivanov D, Mantripragada KK, Holmans P, Craddock N, Owen MJ, O'Donovan MC. Support for the involvement of large cnvs in the pathogenesis of schizophrenia. *Hum Mol Genet* 2009.

70 de Vries BB, Pfundt R, Leisink M, Koolen DA, Vissers LE, Janssen IM, Reijmersdal S, Nillesen WM, Huys EH, Leeuw N, Smeets D, Sistermans EA, Feuth T, van Ravenswaaij-

Arts CM, van Kessel AG, Schoenmakers EF, Brunner HG, Veltman JA. Diagnostic genome profiling in mental retardation. *Am J Hum Genet* 2005;**77**(4):606-16.

71 Wilson LC, Leverton K, Oude Luttikhuis ME, Oley CA, Flint J, Wolstenholme J, Duckett DP, Barrow MA, Leonard JV, Read AP, et al. Brachydactyly and mental retardation: an Albright hereditary osteodystrophy-like syndrome localized to 2q37. *Am J Hum Genet* 1995;**56**(2):400-7.

72 Baynam G, Goldblatt J, Walpole I. Deletion of 8p23.1 with features of Cornelia de Lange syndrome and congenital diaphragmatic hernia and a review of deletions of 8p23.1 to 8pter? A further locus for Cornelia de Lange syndrome. *Am J Med Genet A* 2008;**146A**(12):1565-70.

73 Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J, Leotta A, Pai D, Zhang R, Lee YH, Hicks J, Spence SJ, Lee AT, Puura K, Lehtimaki T, Ledbetter D, Gregersen PK, Bregman J, Sutcliffe JS, Jobanputra V, Chung W, Warburton D, King MC, Skuse D, Geschwind DH, Gilliam TC, Ye K, Wigler M. Strong association of de novo copy number mutations with autism. *Science* 2007;**316**(5823):445-9.

74 Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, Shago M, Moessner R, Pinto D, Ren Y, Thiruvahindrapduram B, Fiebig A, Schreiber S, Friedman J, Ketelaars CE, Vos YJ, Ficicioglu C, Kirkpatrick S, Nicolson R, Sloman L, Summers A, Gibbons CA, Teebi A, Chitayat D, Weksberg R, Thompson A, Vardy C, Crosbie V, Luscombe S, Baatjes R, Zwaigenbaum L, Roberts W, Fernandez B, Szatmari P, Scherer SW. Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 2008;**82**(2):477-88.

75 Szatmari P, Paterson AD, Zwaigenbaum L, Roberts W, Brian J, Liu XQ, Vincent JB, Skaug JL, Thompson AP, Senman L, Feuk L, Qian C, Bryson SE, Jones MB, Marshall CR, Scherer SW, Vieland VJ, Bartlett C, Mangin LV, Goedken R, Segre A, Pericak-Vance MA, Cuccaro ML, Gilbert JR, Wright HH, Abramson RK, Betancur C, Bourgeron T, Gillberg C, Leboyer M, Buxbaum JD, Davis KL, Hollander E, Silverman JM, Hallmayer J, Lotspeich L, Sutcliffe JS, Haines JL, Folstein SE, Piven J, Wassink TH, Sheffield V, Geschwind DH, Bucan M, Brown WT, Cantor RM, Constantino JN, Gilliam TC, Herbert M, Lajonchere C, Ledbetter DH, Lese-Martin C, Miller J, Nelson S, Samango-Sprouse CA, Spence S, State M, Tanzi RE, Coon H, Dawson G, Devlin B, Estes A, Flodman P, Klei L, McMahon WM, Minshew N, Munson J, Korvatska E, Rodier PM, Schellenberg GD, Smith M, Spence MA, Stodgell C, Tepper PG, Wijsman EM, Yu CE, Roge B, Mantoulan C, Wittemeyer K, Poustka A, Felder B, Klauck SM, Schuster C, Poustka F, Bolte S, Feineis-Matthews S, Herbrecht E, Schmotzer G, Tsiantis J, Papanikolaou K, Maestrini E, Bacchelli E, Blasi F, Carone S, Toma C, Van Engeland H, de Jonge M, Kemner C, Koop F, Langemeijer M, Hijmans C, Staal WG, Baird G, Bolton PF, Rutter ML, Weisblatt E, Green J, Aldred C, Wilkinson JA, Pickles A, Le Couteur A, Berney T, McConachie H, Bailey AJ, Francis K, Honeyman G, Hutchinson A, Parr JR, Wallace S, Monaco AP, Barnby G, Kobayashi K, Lamb JA, Sousa I, Sykes N, Cook EH, Guter SJ, Leventhal BL, Salt J, Lord C, Corsello C, Hus V, Weeks DE, Volkmar F, Tauber M, Fombonne E, Shih A, Meyer KJ. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet* 2007;**39**(3):319-28.

76 Behjati F, Shafaghati Y, Firouzabadi SG, Kahrizi K, Bagherizadeh I, Najmabadi H, Bint S, Ogilvie C. M-banding characterization of a 16p11.2p13.1 tandem duplication in a child with autism, neurodevelopmental delay and dysmorphism. *Eur J Med Genet* 2008.

77 Martens MA, Wilson SJ, Reutens DC. Research Review: Williams syndrome: a critical review of the cognitive, behavioral, and neuroanatomical phenotype. *J Child Psychol Psychiatry* 2008;**49**(6):576-608.

78 Ohyama T, Groves AK. Expression of mouse Foxi class genes in early craniofacial development. *Dev Dyn* 2004;**231**(3):640-6.

79 Kumar RA, KaraMohamed S, Sudi J, Conrad DF, Brune C, Badner JA, Gilliam TC, Nowak NJ, Cook EH, Jr., Dobyns WB, Christian SL. Recurrent 16p11.2 microdeletions in autism. *Hum Mol Genet* 2008;**17**(4):628-38.

80 Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, Saemundsen E, Stefansson H, Ferreira MA, Green T, Platt OS, Ruderfer DM, Walsh CA, Altshuler D, Chakravarti A, Tanzi RE, Stefansson K, Santangelo SL, Gusella JF, Sklar P, Wu BL, Daly MJ. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 2008;**358**(7):667-75.

81 Girirajan S, Truong H, Blanchard C, Elsea S. A functional network module for Smith-Magenis syndrome. *Clin Genet* 2009.

82 Bijlsma EK, Aalfs CM, Sluitjer S, Oude Luttikhuis ME, Trembath RC, Hoovers JM, Hennekam RC. Familial cryptic translocation between chromosomes 2qter and 8qter: further delineation of the Albright hereditary osteodystrophy-like phenotype. *J Med Genet* 1999;**36**(8):604-9.

83 Falk RE, Casas KA. Chromosome 2q37 deletion: clinical and molecular aspects. *Am J Med Genet C Semin Med Genet* 2007;**145C**(4):357-71.

84 Schlesinger AE, Potocki L, Poznanski AK, Lupski JR. The hand in Smith-Magenis syndrome (deletion 17p11.2): evaluation by metacarpophalangeal pattern profile analysis. *Pediatr Radiol* 2003;**33**(3):173-6.

85 Jhas S, Ciura S, Belanger-Jasmin S, Dong Z, Llamosas E, Theriault FM, Joachim K, Tang Y, Liu L, Liu J, Stifani S. Hes6 inhibits astrocyte differentiation and promotes neurogenesis through different mechanisms. *J Neurosci* 2006;**26**(43):11061-71.

86 Kageyama R, Ohtsuka T, Kobayashi T. Roles of Hes genes in neural development. *Dev Growth Differ* 2008;**50 Suppl 1**:S97-103.

87 Hermey G, Plath N, Hubner CA, Kuhl D, Schaller HC, Hermans-Borgmeyer I. The three sorCS genes are differentially expressed and regulated by synaptic activity. *J Neurochem* 2004;**88**(6):1470-6.

88 Feldmann HM, Golozoubova V, Cannon B, Nedergaard J. UCP1 ablation induces obesity and abolishes diet-induced thermogenesis in mice exempt from thermal stress by living at thermoneutrality. *Cell Metab* 2009;**9**(2):203-9.

89 Kim-Han JS, Dugan LL. Mitochondrial uncoupling proteins in the central nervous system. *Antioxid Redox Signal* 2005;**7**(9-10):1173-81.

90 Jedele KB. The overlapping spectrum of rett and angelman syndromes: a clinical review. *Semin Pediatr Neurol* 2007;**14**(3):108-17.

91 Koolen DA, Vissers LE, Nillesen W, Smeets D, van Ravenswaaij CM, Sistermans EA, Veltman JA, de Vries BB. A novel microdeletion, del(2)(q22.3q23.3) in a mentally

retarded patient, detected by array-based comparative genomic hybridization. *Clin Genet* 2004;**65**(5):429-32.

92 Vissers LE, de Vries BB, Osoegawa K, Janssen IM, Feuth T, Choy CO, Straatman H, van der Vliet W, Huys EH, van Rijk A, Smeets D, van Ravenswaaij-Arts CM, Knoers NV, van der Burgt I, de Jong PJ, Brunner HG, van Kessel AG, Schoenmakers EF, Veltman JA. Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. *Am J Hum Genet* 2003;**73**(6):1261-70.

93 Ballif BC, Theisen A, Coppinger J, Gowans GC, Hersh JH, Madan-Khetarpal S, Schmidt KR, Tervo R, Escobar LF, Friedrich CA, McDonald M, Campbell L, Ming JE, Zackai EH, Bejjani BA, Shaffer LG. Expanding the clinical phenotype of the 3q29 microdeletion syndrome and characterization of the reciprocal microduplication. *Mol Cytogenet* 2008;**1**:8.

94 Ballif BC, Theisen A, McDonald-McGinn DM, Zackai EH, Hersh JH, Bejjani BA, Shaffer LG. Identification of a previously unrecognized microdeletion syndrome of 16q11.2q12.2. *Clin Genet* 2008;**74**(5):469-75.

95 Shaffer LG, McCaskill C, Han JY, Choo KH, Cutillo DM, Donnenfeld AE, Weiss L, Van Dyke DL. Molecular characterization of de novo secondary trisomy 13. *Am J Hum Genet* 1994;**55**(5):968-74.

96 Williams SR, Girirajan S, Tegay D, Nowak N, Hatchwell E, Elsea SH. Array comparative genomic hybridization of 52 subjects with a Smith-Magenis-like phenotype: identification of dosage-sensitive loci also associated with schizophrenia, autism, and developmental delay. *J Med Genet* 2009(In Press).

97 van Bon BW, Koolen DA, Brueton L, McMullan D, Lichtenbelt KD, Ades LC, Peters G, Gibson K, Moloney S, Novara F, Pramparo T, Dalla Bernardina B, Zoccante L, Balottin U, Piazza F, Pecile V, Gasparini P, Guerci V, Kets M, Pfundt R, de Brouwer AP, Veltman JA, de Leeuw N, Wilson M, Antony J, Reitano S, Luciano D, Fichera M, Romano C, Brunner HG, Zuffardi O, de Vries BB. The 2q23.1 microdeletion syndrome: clinical and behavioural phenotype. *Eur J Hum Genet*;**18**(2):163-70.

98 Marchese S DG, Garver K, Stadler M. De novo 46,XX,dup(2)(q21-q31) ascertained by prenatal cytogenetic analysis. *Am J Hum Genet* 1984;**36**:103.

99 Matos A, Nogueira A, Criado B, Pereira S, Castedo S, Montenegro N. Prenatal diagnosis of partial trisomy 2q. Case report. *Prenat Diagn* 1997;**17**(9):874-6.

100 Schumacher RE, Rocchini AP, Wilson GN. Partial trisomy 2q. *Clin Genet* 1983;**23**(3):191-4.

101 Farrell SA, Sajoo A, Maybury D, Speevak MD. Pure partial trisomy of 2q22-q23 secondary to a paternally inherited direct insertion: a rare duplication. *Clin Genet* 2003;**64**(3):255-7.

102 Nagase T, Kikuno R, Hattori A, Kondo Y, Okumura K, Ohara O. Prediction of the coding sequences of unidentified human genes. XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res* 2000;**7**(6):347-55.

103 Roloff TC, Ropers HH, Nuber UA. Comparative study of methyl-CpG-binding domain proteins. *BMC Genomics* 2003;**4**(1):1.

104 Vega RB, Matsuda K, Oh J, Barbosa AC, Yang X, Meadows E, McAnally J, Pomajzl C, Shelton JM, Richardson JA, Karsenty G, Olson EN. Histone deacetylase 4 controls chondrocyte hypertrophy during skeletogenesis. *Cell* 2004;**119**(4):555-66.

105 Chen B, Cepko CL. HDAC4 regulates neuronal survival in normal and diseased retinas. *Science* 2009;**323**(5911):256-9.

106 Aldred MA, Vijayakrishnan J, James V, Soubrier F, Gomez-Sanchez MA, Martensson G, Galie N, Manes A, Corris P, Simonneau G, Humbert M, Morrell NW, Trembath RC. BMPR2 gene rearrangements account for a significant proportion of mutations in familial and idiopathic pulmonary arterial hypertension. *Hum Mutat* 2006;**27**(2):212-3.

107 Aldred MA, Sanford RO, Thomas NS, Barrow MA, Wilson LC, Brueton LA, Bonaglia MC, Hennekam RC, Eng C, Dennis NR, Trembath RC. Molecular analysis of 20 patients with 2q37.3 monosomy: definition of minimum deletion intervals for key phenotypes. *J Med Genet* 2004;**41**(6):433-9.

108 Zhang CL, McKinsey TA, Chang S, Antos CL, Hill JA, Olson EN. Class II histone deacetylases act as signal-responsive repressors of cardiac hypertrophy. *Cell* 2002;**110**(4):479-88.

109 Grozinger CM, Hassig CA, Schreiber SL. Three proteins define a class of human histone deacetylases related to yeast Hda1p. *Proc Natl Acad Sci U S A* 1999;**96**(9):4868-73.

110 Fischle W, Dequiedt F, Hendzel MJ, Guenther MG, Lazar MA, Voelter W, Verdin E. Enzymatic activity associated with class II HDACs is dependent on a multiprotein complex containing HDAC3 and SMRT/N-CoR. *Mol Cell* 2002;**9**(1):45-57.

111 Takeda S, Bonnamy JP, Owen MJ, Ducy P, Karsenty G. Continuous expression of Cbfa1 in nonhypertrophic chondrocytes uncovers its ability to induce hypertrophic chondrocyte differentiation and partially rescues Cbfa1-deficient mice. *Genes Dev* 2001;**15**(4):467-81.

112 Ueta C, Iwamoto M, Kanatani N, Yoshida C, Liu Y, Enomoto-Iwamoto M, Ohmori T, Enomoto H, Nakata K, Takada K, Kurisu K, Komori T. Skeletal malformations caused by overexpression of Cbfa1 or its dominant negative form in chondrocytes. *J Cell Biol* 2001;**153**(1):87-100.

113 Black BL, Olson EN. Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins. *Annu Rev Cell Dev Biol* 1998;**14**:167-96.

114 McKinsey TA, Zhang CL, Olson EN. MEF2: a calcium-dependent regulator of cell division, differentiation and death. *Trends Biochem Sci* 2002;**27**(1):40-7.

115 Arnold MA, Kim Y, Czubryt MP, Phan D, McAnally J, Qi X, Shelton JM, Richardson JA, Bassel-Duby R, Olson EN. MEF2C transcription factor controls chondrocyte hypertrophy and bone development. *Dev Cell* 2007;**12**(3):377-89.

116 Munoz JP, Collao A, Chiong M, Maldonado C, Adasme T, Carrasco L, Ocaranza P, Bravo R, Gonzalez L, Diaz-Araya G, Hidalgo C, Lavandero S. The transcription factor MEF2C mediates cardiomyocyte hypertrophy induced by IGF-1 signaling. *Biochem Biophys Res Commun* 2009;**388**(1):155-60.

117 Smith M, Escamilla JR, Filipek P, Bocian ME, Modahl C, Flodman P, Spence MA. Molecular genetic delineation of 2q37.3 deletion in autism and osteodystrophy: report of a case and of new markers for deletion screening by PCR. *Cytogenet Cell Genet* 2001;**94**(1-2):15-22.

118 Reddy KS. Cytogenetic abnormalities and fragile-X syndrome in Autism Spectrum Disorder. *BMC Med Genet* 2005;**6**:3.

119 Felder B, Radlwimmer B, Benner A, Mincheva A, Todt G, Beyer KS, Schuster C, Bolte S, Schmotzer G, Klauck SM, Poustka F, Lichter P, Poustka A. FARP2, HDLBP and PASK are downregulated in a patient with autism and 2q37.3 deletion syndrome. *Am J Med Genet A* 2009;**149A**(5):952-9.

120 Wolff DJ, Clifton K, Karr C, Charles J. Pilot assessment of the subtelomeric regions of children with autism: detection of a 2q deletion. *Genet Med* 2002;**4**(1):10-4.

121 Wassink TH, Piven J, Vieland VJ, Jenkins L, Frantz R, Bartlett CW, Goedken R, Childress D, Spence MA, Smith M, Sheffield VC. Evaluation of the chromosome 2q37.3 gene CENTG2 as an autism susceptibility gene. *Am J Med Genet B Neuropsychiatr Genet* 2005;**136B**(1):36-44.

122 Klopocki E, Hennig BP, Dathe K, Koll R, de Ravel T, Baten E, Blom E, Gillerot Y, Weigel JF, Kruger G, Hiort O, Seemann P, Mundlos S. Deletion and Point Mutations of PTHLH Cause Brachydactyly Type E. *Am J Hum Genet*;**86**(3):434-9.

123 Guo J, Chung UI, Yang D, Karsenty G, Bringhurst FR, Kronenberg HM. PTH/PTHrP receptor delays chondrocyte hypertrophy via both Runx2-dependent and -independent pathways. *Dev Biol* 2006;**292**(1):116-28.

124 Girirajan S, Rosenfeld JA, Cooper GM, Antonacci F, Siswara P, Itsara A, Vives L, Walsh T, McCarthy SE, Baker C, Mefford HC, Kidd JM, Browning SR, Browning BL, Dickel DE, Levy DL, Ballif BC, Platky K, Farber DM, Gowans GC, Wetherbee JJ, Asamoah A, Weaver DD, Mark PR, Dickerson J, Garg BP, Ellingwood SA, Smith R, Banks VC, Smith W, McDonald MT, Hoo JJ, French BN, Hudson C, Johnson JP, Ozmore JR,

Moeschler JB, Surti U, Escobar LF, El-Khechen D, Gorski JL, Kussmann J, Salbert B, Lacassie Y, Biser A, McDonald-McGinn DM, Zackai EH, Deardorff MA, Shaikh TH, Haan E, Friend KL, Fichera M, Romano C, Gecz J, DeLisi LE, Sebat J, King MC, Shaffer LG, Eichler EE. A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat Genet*;**42**(3):203-9.

125 Capili AD, Schultz DC, Rauscher IF, Borden KL. Solution structure of the PHD domain from the KAP-1 corepressor: structural determinants for PHD, RING and LIM zinc-binding domains. *Embo J* 2001;**20**(1-2):165-77.

126 Imai Y, Suzuki Y, Matsui T, Tohyama M, Wanaka A, Takagi T. 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* 1995;**31**(1-2):1-9.

127 Bienz M. The PHD finger, a nuclear protein-interaction domain. *Trends Biochem Sci* 2006;**31**(1):35-40.

128 Nie L, Vazquez AE, Yamoah EN. Identification of Transcription Factor-DNA Interactions Using Chromatin Immunoprecipitation Assays. *Methods Mol Biol* 2009;**493**:311-22.

129 Kirmizis A, Farnham PJ. Genomic approaches that aid in the identification of transcription factor target genes. *Exp Biol Med (Maywood)* 2004;**229**(8):705-21.

130 Takahashi JS, Hong HK, Ko CH, McDearmon EL. The genetics of mammalian circadian order and disorder: implications for physiology and disease. *Nat Rev Genet* 2008;**9**(10):764-75.

131 Turek FW, Joshu C, Kohsaka A, Lin E, Ivanova G, McDearmon E, Laposky A, Losee-Olson S, Easton A, Jensen DR, Eckel RH, Takahashi JS, Bass J. Obesity and metabolic syndrome in circadian Clock mutant mice. *Science* 2005;**308**(5724):1043-5.

132 Salvatore P, Ghidini S, Zita G, De Panfilis C, Lambertino S, Maggini C, Baldessarini RJ. Circadian activity rhythm abnormalities in ill and recovered bipolar I disorder patients. *Bipolar Disord* 2008;**10**(2):256-65.

133 Mendlewicz J. Disruption of the circadian timing systems: molecular mechanisms in mood disorders. *CNS Drugs* 2009;**23 Suppl 2**:15-26.

134 De Leersnyder H. Inverted rhythm of melatonin secretion in Smith-Magenis syndrome: from symptoms to treatment. *Trends Endocrinol Metab* 2006;**17**(7):291-8.

135 Amr S, Heisey C, Zhang M, Xia XJ, Shows KH, Ajlouni K, Pandya A, Satin LS, El-Shanti H, Shiang R. A homozygous mutation in a novel zinc-finger protein, ERIS, is responsible for Wolfram syndrome 2. *Am J Hum Genet* 2007;**81**(4):673-83.

136 Seranski P, Heiss NS, Dhorne-Pollet S, Radelof U, Korn B, Hennig S, Backes E, Schmidt S, Wiemann S, Schwarz CE, Lehrach H, Poustka A. Transcription mapping in a medulloblastoma breakpoint interval and Smith-Magenis syndrome candidate region: identification of 53 transcriptional units and new candidate genes. *Genomics* 1999;**56**(1):1-11.

137 King DP, Zhao Y, Sangoram AM, Wilsbacher LD, Tanaka M, Antoch MP, Steeves TD, Vitaterna MH, Kornhauser JM, Lowrey PL, Turek FW, Takahashi JS. Positional cloning of the mouse circadian clock gene. *Cell* 1997;**89**(4):641-53.

138 Horsthemke B, Buiting K. Imprinting defects on human chromosome 15. *Cytogenet Genome Res* 2006;**113**(1-4):292-9.

139 Hitchins MP, Rickard S, Dhalla F, Fairbrother UL, de Vries BB, Winter R, Pembrey ME, Malcolm S. Investigation of UBE3A and MECP2 in Angelman syndrome (AS) and patients with features of AS. *Am J Med Genet A* 2004;**125A**(2):167-72.

140 Lebrun B, Bariohay B, Moyse E, Jean A. Brain-derived neurotrophic factor (BDNF) and food intake regulation: a minireview. *Auton Neurosci* 2006;**126-127**:30-8.

141 Gray J, Yeo GS, Cox JJ, Morton J, Adlam AL, Keogh JM, Yanovski JA, El Gharbawy A, Han JC, Tung YC, Hodges JR, Raymond FL, O'Rahilly S, Farooqi IS. Hyperphagia, severe obesity, impaired cognitive function, and hyperactivity associated with functional loss of one copy of the brain-derived neurotrophic factor (BDNF) gene. *Diabetes* 2006;**55**(12):3366-71.

142 Simmons DA, Rex CS, Palmer L, Pandyarajan V, Fedulov V, Gall CM, Lynch G. Up-regulating BDNF with an ampakine rescues synaptic plasticity and memory in Huntington's disease knockin mice. *Proc Natl Acad Sci U S A* 2009;**106**(12):4906-11.

143 Larimore JL, Chapleau CA, Kudo S, Theibert A, Percy AK, Pozzo-Miller L. Bdnf overexpression in hippocampal neurons prevents dendritic atrophy caused by Rett-associated MECP2 mutations. *Neurobiol Dis* 2009;**34**(2):199-211.

144 Girirajan S, Elsea SH. Distorted Mendelian transmission as a function of genetic background in Rai1-haploinsufficient mice. *Eur J Med Genet* 2009;**52**(4):224-8.

145 Zheng B, Albrecht U, Kaasik K, Sage M, Lu W, Vaishnav S, Li Q, Sun ZS, Eichele G, Bradley A, Lee CC. Nonredundant roles of the mPer1 and mPer2 genes in the mammalian circadian clock. *Cell* 2001;**105**(5):683-94.

146 Goto M, Oshima I, Tomita T, Ebihara S. Melatonin content of the pineal gland in different mouse strains. *J Pineal Res* 1989;**7**(2):195-204.

147 King DP, Takahashi JS. Forward genetic approaches to circadian clocks in mice. *Cold Spring Harb Symp Quant Biol* 1996;**61**:295-302.

148 Antoch MP, Song EJ, Chang AM, Vitaterna MH, Zhao Y, Wilsbacher LD, Sangoram AM, King DP, Pinto LH, Takahashi JS. Functional identification of the mouse circadian Clock gene by transgenic BAC rescue. *Cell* 1997;**89**(4):655-67.

149 Debruyne JP, Noton E, Lambert CM, Maywood ES, Weaver DR, Reppert SM. A clock shock: mouse CLOCK is not required for circadian oscillator function. *Neuron* 2006;**50**(3):465-77.

150 Zheng B, Larkin DW, Albrecht U, Sun ZS, Sage M, Eichele G, Lee CC, Bradley A. The mPer2 gene encodes a functional component of the mammalian circadian clock. *Nature* 1999;**400**(6740):169-73.

151 Hampp G, Ripperger JA, Houben T, Schmutz I, Blex C, Perreau-Lenz S, Brunk I, Spanagel R, Ahnert-Hilger G, Meijer JH, Albrecht U. Regulation of monoamine oxidase A by circadian-clock components implies clock influence on mood. *Curr Biol* 2008;**18**(9):678-83.

152 Albrecht U. Circadian clocks in mood-related behaviors. *Ann Med.*

153 Benedetti F, Serretti A, Colombo C, Barbini B, Lorenzi C, Campori E, Smeraldi E. Influence of CLOCK gene polymorphism on circadian mood fluctuation and illness recurrence in bipolar depression. *Am J Med Genet B Neuropsychiatr Genet* 2003;**123B**(1):23-6.

154 Serretti A, Cusin C, Benedetti F, Mandelli L, Pirovano A, Zanardi R, Colombo C, Smeraldi E. Insomnia improvement during antidepressant treatment and CLOCK gene polymorphism. *Am J Med Genet B Neuropsychiatr Genet* 2005;**137B**(1):36-9.

155 Serretti A, Benedetti F, Mandelli L, Lorenzi C, Pirovano A, Colombo C, Smeraldi E. Genetic dissection of psychopathological symptoms: insomnia in mood disorders and CLOCK gene polymorphism. *Am J Med Genet B Neuropsychiatr Genet* 2003;**121B**(1):35-8.

156 Pirovano A, Lorenzi C, Serretti A, Ploia C, Landoni S, Catalano M, Smeraldi E. Two new rare variants in the circadian "clock" gene may influence sleep pattern. *Genet Med* 2005;**7**(6):455-7.

157 Nievergelt CM, Kripke DF, Barrett TB, Burg E, Remick RA, Sadovnick AD, McElroy SL, Keck PE, Jr., Schork NJ, Kelsoe JR. Suggestive evidence for association of the circadian genes PERIOD3 and ARNTL with bipolar disorder. *Am J Med Genet B Neuropsychiatr Genet* 2006;**141B**(3):234-41.

158 Mansour HA, Wood J, Logue T, Chowdari KV, Dayal M, Kupfer DJ, Monk TH, Devlin B, Nimgaonkar VL. Association study of eight circadian genes with bipolar I disorder, schizoaffective disorder and schizophrenia. *Genes Brain Behav* 2006;**5**(2):150-7.

159 Kamura E. Exploring the Methylation Status of RAI1 and the RAI1 Consensus Binding Sequence. 2009.

160 Kingsley DM, Bland AE, Grubber JM, Marker PC, Russell LB, Copeland NG, Jenkins NA. The mouse short ear skeletal morphogenesis locus is associated with defects in a bone morphogenetic member of the TGF beta superfamily. *Cell* 1992;**71**(3):399-410.

161 Beck HN, Drahushuk K, Jacoby DB, Higgins D, Lein PJ. Bone morphogenetic protein-5 (BMP-5) promotes dendritic growth in cultured sympathetic neurons. *BMC Neurosci* 2001;**2**:12.

Appendix A

Overexpression of *RAI1* in HEK293t cells

As stated in Chapter 1, duplication of 17p11.2 results in dup17p11.2 syndrome (aka Potocki-Lupski syndrome). This syndrome is rarer than SMS, but because duplication of this genomic region, including *RAI1*, gives rise to a recognizable syndrome, it points to the fact that *RAI1* is a dosage-sensitive gene. Additionally, Girirajan et al. 2008, 2009 [43,[44] showed that *Rai1-Tg* mice have altered maternal behavior, altered sociability, and impaired serotonin metabolism [44]. Further, these mice have growth retardation, increased locomotor activity, and abnormal anxiety-related responses. *Rai1-Tg* mice also have an altered gait with short strides and long sways, impaired ability on a cage top hang test, decreased forelimb grip strength, and a dominant social behavior. Additional analyses of homozygous transgenic mice revealed a dosage-dependent exacerbation of the phenotype, including extreme growth retardation, severe neurological deficits, and increased hyperactivity [43].

These results laid the groundwork to examine the molecular changes that occur because of *RAI1* overexpression. We used transient over expression followed by gene expression microarray to explore the consequence of *RAI1* overexpression.

Materials and methods

Transfections

Transfections of HEK293t cells with RAI1^{Flag} were performed as described Chapter 5.

RNA isolation

RNA was isolated using the standard Trizol protocol provided by Invitrogen (Invitrogen, Carlsbad, CA)

Microarray

Microarray was performed as previously described [81] using the Affymetrix Human Genome, HG-U133A 2.0 Array chipset. Control RNA was isolated from non transfected HEK293t cells.

Results

Table A1. Top upregulated genes identified in HEK293t cells overexpressing *RAI1*.

***RAI1*/Control Signal Log Ratio**	**Gene Symbol**	**Gene Name**	**Pathway**	**Biological Process**
5.5	*TRPC1*	transient receptor potential cation channel, subfamily C, member 1	---	transport
5.5	*HNRNPU*	heterogeneous nuclear ribonucleoprotein U	mRNA_processing_Reactome	RNA processing
5.3	*MAT2A*	methionine adenosyltransferase II, alpha	---	S-adenosylmethionine biosynthetic process
5	*SRP72*	signal recognition particle 72kDa	---	protein amino acid phosphorylation /// SRP-dependent cotranslational protein targeting to membrane /// signal transduction
5	*SAR1A*	SAR1 gene homolog A	---	transport
5	*EZR*	ezrin	---	cytoskeletal anchoring /// regulation of cell shape
5	*PSAT1*	phosphoserine aminotransferase 1	---	L-serine biosynthetic process
4.9	*SELT*	selenoprotein T	---	selenocysteine incorporation /// cell redox homeostasis
4.8	*ASL*	argininosuccinate lyase	---	urea cycle
4.8	*TP53*	tumor protein p53	Apoptosis	response to tumor cell
4.8	*PRIM2*	primase, DNA, polypeptide 2 (58kDa)	DNA_replication_Reactome	DNA replication
4.7	*UBE2M*	ubiquitin-conjugating enzyme E2M	---	protein modification process ///regulation of protein metabolic process

4.7	*TMED2*	transmembrane emp24 domain trafficking protein 2	---	transport
4.7	*ENO1*	enolase 1, (alpha)	Glycolysis_and_Gluconeogenesis	negative regulation of transcription from RNA polymerase II promoter /// glycolysis
4.6	*MRPS12*	mitochondrial ribosomal protein S12	---	translation
4.6	*RNFT1*	ring finger protein, transmembrane 1	---	---
4.5	*DUSP3*	dual specificity phosphatase 3	---	protein amino acid dephosphorylation
4.5	*MAN1A1*	mannosidase, alpha, class 1A, member 1	---	protein amino acid glycosylation
4.5	*HIPK3*	homeodomain interacting protein kinase 3	---	transcription
4.5	*ADAM10*	ADAM metallopeptidase domain 10	Hypertrophy_model	in utero embryonic development
4.5	*ARFRP1*	ADP-ribosylation factor related protein 1	---	signal transduction /// small GTPase mediated signal transduction
4.4	*SLC29A1*	solute carrier family 29 member 1	---	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process
4.4	*HSPA8*	heat shock 70kDa protein 8	Circadian_Exercise	protein folding /// response to stress /// response to unfolded protein
4.3	*UTRN*	utrophin	---	muscle contraction /// muscle development
4.3	*PGK1*	phosphoglycerate kinase 1	Glycolysis_and_Gluconeogenesis	glycolysis /// phosphorylation
4.3	*GDAP1*	ganglioside-induced differentiation-associated protein 1	---	---

Table A2. Top downregulated genes identified in HEK293t cells overexpressing *RAI1*

***RAI1*/Control Signal Log Ratio**	**Gene Symbol**	**Gene Title**	**Pathway**	**Biological process term**
-5.5	*ANKRD36B*	ankyrin repeat domain 36B	---	cellular di-, tri-valent inorganic anion homeostasis
-4.5	*CDC42BPA*	CDC42 binding protein kinase alpha	---	protein amino acid phosphorylation
-4.2	*SPEN*	spen homolog, transcriptional regulator	---	transcription

-4.2	*SFRS11*	splicing factor, arginine/serine-rich 11	---	mRNA processing
-4.2	*SEMA6A*	sema domain, transmembrane domain, and cytoplasmic domain, 6A	---	apoptosis ///nervous system development /// axon guidance /// organ morphogenesis
-4.1	*NKTR*	natural killer-tumor recognition sequence	---	protein folding
-4.1	*SMA4*	Glucuronidase, beta pseudogene 1	---	skeletal development /// carbohydrate metabolic process /// nervous system development
-4	*SRRM2*	serine/arginine repetitive matrix 2	---	mRNA processing
-3.9	*RBM41*	RNA binding motif protein 41	---	---
-3.7	*MSH5*	mutS homolog 5 (E. coli) /// chromosome 6 open reading frame 26	Ovarian_Infertility_Genes	mismatch repair
-3.7	*GOLGA8A*	golgi autoantigen, golgin subfamily a, 8A	---	---
-3.7	*CEP350*	centrosomal protein 350kDa	---	protein targeting
-3.7	*ATHL1*	ATH1, acid trehalase-like 1	---	carbohydrate metabolic process
-3.7	*EBF2*	early B-cell factor 2	---	transcription
-3.7	*MCM3APAS*	minichromosome maintenance complex component 3 associated protein antisense	---	---
-3.7	*WNK1*	WNK lysine deficient protein kinase 1	---	protein amino acid phosphorylation
-3.6	*ATRX*	alpha thalassemia/mental retardation syndrome X-linked	---	DNA repair /// DNA repair /// DNA methylation

Discussion

In this study we have overexpressed *RAI1* in HEK293t cells to evaluate the molecular consequences of this action. Given that *RAI1* is a dosage sensitive gene these were important experiments to perform and examine pathways affected. We have identified gene which play critical roles in circadian rhythm, epigenetic regulation, DNA repair, RNA processing, transcription and a variety of other critical pathways. These studies will help to identify the molecular role *RAI1* plays in the cell and help to understand the consequence of abnormal *RAI1* dosage.

Appendix B

Clin Genet 2007: 72: 47–58
Printed in Singapore. All rights reserved

© *2007 The Authors*
Journal compilation © *2007 Blackwell Munksgaard*

CLINICAL GENETICS
doi: 10.1111/j.1399-0004.2007.00831.x

Short Report

17p11.2p12 triplication and del(17)q11.2q12 in a severely affected child with dup(17)p11.2p12 syndrome

Girirajan S, Williams SR, Garbern JY, Nowak N, Hatchwell E, Elsea SH. 17p11.2p12 triplication and del(17)q11.2q12 in a severely affected child with dup(17)p11.2p12 syndrome.
Clin Genet 2007: 72: 47–58. © Blackwell Munksgaard, 2007

Multiple congenital anomalies/mental retardation syndromes due to genomic rearrangements involving chromosome 17p11.2 include deletion resulting in Smith–Magenis syndrome and a reciprocal duplication of the same region resulting in the 17p11.2 duplication syndrome. We present the clinical and molecular analysis of an 8-year-old male with a dup(17p11.2p12) who was evaluated for unusual severity of the phenotype. Fluorescent *in situ* hybridization (FISH) analysis not only confirmed the 17p duplication but also identified an ~25% mosaicism for tetrasomy 17p11.2p12. Whole-genome array comparative genomic hybridization (aCGH) was performed to identify other genomic rearrangements possibly contributing to the severe phenotype and the unusual features in the patient. The 17p duplication was determined by FISH and aCGH to encompass ~7.5 Mb, from *COX10* to *KCNJ12*. An ~830 Kb deletion of 17q11.2q12, including exon 1 of an amiloride-sensitive cation channel neuronal gene, *ACCN1*, was also identified by aCGH; breakpoints of the deletion were confirmed by FISH. Sequencing the non-deleted allele of *ACCN1* did not show any mutations. Western analysis of human tissue-specific proteins revealed that *ACCN1* is expressed not only in the brain as previously reported but also in all tissues examined, including heart, liver, kidneys, and spleen. The large-sized 17p11.2p12 duplication, partial triplication of the same region, and the 17q11.2q12 deletion create a complex chromosome 17 rearrangement that has not been previously identified. This is the first case of triplication reported for this chromosome. Our study emphasizes the utility of whole-genome analysis for known cases with deletion/duplication syndromes with unusual or severe phenotypes.

S Girirajan[a], SR Williams[a], JY Garbern[b], N Nowak[c,d], E Hatchwell[e] and SH Elsea[a,f]

[a]Department of Human Genetics, Virginia Commonwealth University, Richmond, VA, USA, [b]Center for Molecular Medicine and Genetics, Wayne State University School of Medicine, Detroit, MI, USA, [c]Department of Biochemistry and Center of Excellence in Bioinformatics and Life Sciences, State University of New York at Buffalo, NY, USA, [d]Department of Cancer Prevention, Roswell Park Cancer Institute, Buffalo, NY, USA, [e]Department of Pathology, State University of New York at Stony Brook, NY, USA, and [f]Department of Pediatrics, Virginia Commonwealth University, Richmond, VA, USA

Key words: *ACCN1* – aCGH – cardiac anomalies – 17p11.2p12 triplication – 17q11.2q12 deletion – 17p11.2p12 duplication – FISH – mental retardation – tetrasomy 17p11.2p12

Corresponding author: Dr Sarah H. Elsea, PhD, FACMG, Departments of Pediatrics and Human Genetics, 1101 E Marshall St, 12-018 Sanger Hall, PO BOX 980441, Medical College of Virginia, Virginia Commonwealth University, Richmond, VA 23298, USA.
Tel.: +804 628 0987;
fax: +804 628 1609;
e-mail: selsea@vcu.edu

Received 26 January 2007, revised and accepted for publication 19 April 2007

Rearrangements due to homologous and non-homologous mechanisms within region-specific low copy repeats or segmental duplications cause deletions/duplications in disease-associated genomic regions (1). The genomic fragility of the short arm of chromosome 17 is due to extensive segmental duplications, especially in the Charcot–Marie–Tooth disease type 1A/Smith–Magenis syndrome (CMT1A/SMS) regions (1). Clinically recognized genomic disorders of proximal 17p include Charcot–Marie–Tooth disease type 1A caused by a duplication of the peripheral myelin protein 22 (*PMP22*) gene (2), hereditary neuropathy with liability to pressure palsies caused by a deletion of *PMP22* (3), and Smith–Magenis syndrome, associated with an interstitial deletion of 17p11.2 containing the retinoic acid induced 1 gene (*RAI1*) (4–10). Duplication of the SMS region and larger regions of the short arm of chromosome 17 has also been reported but less frequently (11–28).

Although less clearly defined, the clinical features of patients with duplication 17p consist of mild to severe mental retardation, developmental delay, failure to thrive, behavioral features, craniofacial anomalies, and short stature (16, 29). Craniofacial features include a triangular face, microcephaly, micrognathia, broad nasal bridge, high arched palate, and hypertelorism, while dental anomalies in the form of malocclusion of the teeth have also been reported (24). Limb abnormalities including flexion deformity of the fingers and club foot have also been described (12, 13, 15, 16, 18, 20, 21).

In this paper, we report the molecular and clinical evaluation of a child with severe phenotype carrying complex rearrangements of chromosome 17. A combination of fluorescent *in situ* hybridization (FISH) and array comparative genomic hybridization (CGH) was used to identify and confirm a 17p11.2p12 duplication, 17p11.2p12 mosaic triplication, and also 17q11.2q12 deletion in this patient. Both *de novo* rearrangements are seen on the same chromosome 17 homologue. While our study emphasizes the utility of interphase FISH for the identification of triplications; aCGH results indicate that identification of a single genomic rearrangement does not preclude the possibility of another genomic abnormality.

Materials and methods

Patient samples

Research samples were collected at Michigan State University and Wayne State University in accordance with Institutional Review Board approved protocols of Virginia Commonwealth University. Peripheral blood samples were collected, and metaphase chromosomes and DNA were prepared using standard cytogenetic methods.

Fluorescent *in situ* hybridization

In order to delineate the chromosome 17p duplication, probes were chosen from our previously published SMS contig map of bacterial artificial chromosome (BAC) and P1-derived artificial chromosome clones (30), and also from University of Southern California, Santa Cruz (UCSC) genome browser (May 2004 version). Probes to confirm the chromosome 17q deletion were obtained from the initial aCGH clones and verified using UCSC genome browser (March 2006 version). The FISH probes from centromeric to telomeric with their representative loci used to map the 17p11.2p12 rearrangement include: RP11-746M1 (D17S689), CTC-457L16 (D17S29), RP1-253P07 (*RAI1*), RP1-48J14 (*TNFRSF13B*), and RP11-764O6 (*PMP22*). The FISH probes used to map the 17q11.2q12 deletion (Table 1) include: RP11-329H16 (*CCL2*, *CCL7*, *CCL8*, and *CCL11)*, RP11-47G8, RP11-79K15, and RP11-797M22 (all *ACCN1* probes). FISH was performed as previously described (5). For every probe analyzed by FISH, ~100 interphase nuclei and metaphase spreads were observed. Parental samples were also evaluated to rule out any familial rearrangement and mosaicism.

Table 1. aCGH probes (from telomere to centromere) used to map the chromosome 17 duplication and deletion[a]

Probes	Loci
Chromosome 17p11.2p12	
RP11-661L20	*ELAC2*
RP11-210C3	*HS3ST3A*
RP11-668O24	*COX10*
RP11-78J16	*HS3ST3B1*
RP11-64B12	*FLJ45831*
RP11-465O5	*PMP22/TEKT3*
RP11-90G21	*FAM18B2*
RP11-459E6	*ADORA2B*
RP11-455I21	*NCOR1*
RP11-410D6	*PIGL*
RP11-404D6	*UBB*
RP11-803H14	*ZNF624*
RP11-793G13	*TNFRSF13B*
RP11-472C16	*COPS3/NT5M*
RP11-384M20	*TOM1L2/ATPAF2*
RP11-936C14	*DRG2/LLGL1*
RP11-80K17	*TRIM70/FBXW10*
RP11-977H13	*PRPSAP2*
RP11-160E2	*GRAP*
RP11-135L13	*EPN2/MAPK7*
RP11-1007K2	*AKAP10*
RP11-1065P14	*SPECC1*
RP11-746M1	*USP22*
RP11-64J19	*TMEM11/MAP2K3*
RP11-45G12	*KCNJ12*
RP11-730C17	*WSB1*
RP11-803B22	*KSR1*
Chromosome 17q11.2q12	
RP11-299F7	Exons 8–10 (*ACCN1*)
RP11-98K21	Exons 2–6 (*ACCN1*)
RP11-797M22	Intron 1 (*ACCN1*)
RP11-90A23	Intron 1 (*ACCN1*)
RP11-79K15	Intron 1 (*ACCN1*)
RP11-47G8	Intron 1 (*ACCN1*)
RP11-1137H3	Exon 1 (*ACCN1*)
RP11-329H16	*CCL2, CCL7, CCL8, CCL11*
RP11-521P1	*CCL1, CCL13*
RP11-42E9	*TMEM132E*

[a] A list of aCGH probes used to map the proximal 17p duplication and the 17q deletion along with their representative loci are shown. The shaded regions represent the extent of the 17p duplication and the deletion. Flanking unshaded regions represent genes/loci with normal copy number.

Genotyping

Four sequence tagged site markers localized to 17p11.2 (cen-D17S783, D17S805, D17S1794, and D17S900-tel) were used to determine the origin of the duplication. Polymerase chain reaction (PCR) for the markers was performed using fluorescently labeled reverse primers on the DNA samples from the patient and the parents. Genotyping was performed using an ABI Prism® 3100 Genetic Analyzer and analysis performed using GENESCAN analysis software (Applied Biosystems, Foster City, CA).

Whole-genome array-comparative genomic hybridization

The design of the tiling path BAC array was based around a pre-existing set of 6800 clones which had been assembled at the Roswell Park Cancer Institute (RPCI) to generate a first generation array used previously for cancer genomics (http://microarrays.roswellpark.org). This set of 6800 clones was used as a 'scaffold' around which an additional set of 13,000 clones were assembled to fill in the gaps, according to the following criteria: no gaps greater than 10 kb in the final design, overlap was only required if the gap was greater than 10 kb, and only RPCI-11 clones were used to fill in the gaps. A minimum number of BACs were used to fulfill the criteria and Perl scripts were written in order to access data from a full set of BAC end-pairs (www.genome.ucsc.edu).

Reference DNA, 500 ng from each parent, and 1 μg of the patient's (test) DNA was fluorescently labeled using the BioPrime aCGH Labeling System (GE HealthCare, Piscataway, NJ). Hybridization to the 19K CGH arrays was performed using a GeneMachine hybridization station (Genomic Solutions, Ann Arbor, MI) as described previously (31). Following post-hybridization washes, the slides were spun dry and scanned immediately at 5 μm resolution using a GenePix 4000B laser scanner (Axon Inc., Sunnyvale, CA). Image analysis was performed using the BlueFuse package from Cambridge BlueGnome (Cambridge, UK). aCGH clones used for mapping both the duplication (17p11.2p12) and deletion (17q11.2q12) are listed in Table 1.

ACCN1 sequence analysis

DNA from whole blood was isolated by the phenol–chloroform method. PCR to amplify coding exons of *ACCN1* was performed with 100 ng genomic DNA, 20 pmol each of forward and reverse primers, 10 mM Tris–HCl, 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM dNTPs, and 1 U *Taq* DNA polymerase in a total reaction volume of 25 μl. Standard PCR conditions were followed with annealing temperatures for all the primers ranging between 55°C and 64°C. PCR products (5 μl) were either purified by digesting with 2 U of shrimp alkaline phosphatase and 10 U of exonuclease I (USB Corp, Cleveland, OH) at 37°C, 15 min followed by 80°C, 15 min to remove excess primers or were gel purified using a Qiagen gel extraction kit (Qiagen Inc, Valencia, CA). Sequencing and analysis of electropherograms were performed as described previously (9). Primers used for amplifying the *ACCN1* gene are shown in Table 2.

Western analysis

Human tissue-specific protein medleys (20 or 40 μg with 0.5 mM DTT; Clontech, Mountain View, CA) were boiled, run on a NuPAGE™ 4–12% Bis–Tris gel using a XCell *SureLock*™ Mini-cell (Invitrogen, Carlsbad, CA), and transferred to 0.45 μm polyvinylidene fluoride membrane (Millipore Corporation, Bedford, MA), using XCell II™ Blot Module (Invitrogen). The blot was then blocked in phosphate-buffered saline with Tween 20 (PBST) (137 mM NaCl, 2.7 mM KCl, 8 mM

Table 2. Polymerase chain reaction primers used to amplify and sequence *ACCN1*

	Forward primer	Reverse primer
Exon 1	aaacactcacccctgtcagc	ccccactcaccctccatatt
Exon 1 (variant)	ttccactgctgcggact	tggtaacactgcaagaggcctgg
Exon 2	tttgctctttgatgccaaaa	cccatcctctgtccctctct
Exon 3	tttgccatgtaaggcaacct	tgcccaattttagggaacaa
Exon 4	gagccaggagttcaaaacca	cccaaacactggatgcttct
Exon 5	agctctggggtgcactgt	tcctctccatctccctctcc
Exon 6	gcaaaagtggtcccatttca	gcaggggctgtgtcttactc
Exon 7	gagtcaggcttcttggcttg	gctggggttctttccattct
Exon 8	ccttccaataagccatcagc	aaaacctcccatccccttg
Exon 9	gggcttgtgctacggtatgt	ctggcttctggcatgactct
Exon 10	aaggattccattcagcctca	caaggatgcgtcgtgttg

Na_2HPO_4, 2 mM KH_2PO_4 and 0.05% Tween-20) containing 10% non-fat dry milk overnight at 4°C. The blot was then incubated with rabbit anti-ACCN1 antibody (1:2000 dilution, Upstate/Chemicon, Temecula, CA) or anti-Lamin A antibody (1:20,000, Invitrogen) at 37°C for 1 h, washed 6×, 5 min each, with PBST containing 2% non-fat dry milk; incubated with horseradish peroxidase conjugated goat-anti-rabbit secondary antibody (1:20,000, Vector labs, Burlingame, CA) for 1 h at 37°C, washed 6×, 5 min each with PBST; and detected with Western Lightning™ Chemiluminescence Reagent Plus (PerkinElmer, Boston, MA).

Results

Clinical findings

SMS224 is an 8-year-old male with cognitive delay, facial dysmorphisms, severe behavioral abnormalities, weak extremities, sensorimotor neuropathy, and congenital heart disease (Fig. 1a). Initial evaluation by a high-resolution G-banding technique revealed a partial duplication of chromosome 17p.

The pre-natal period was complicated by maternal exposure to parvovirus B19 followed by delivery at 36 weeks. The post-natal period was complicated by neonatal unconjugated hyperbilirubinemia

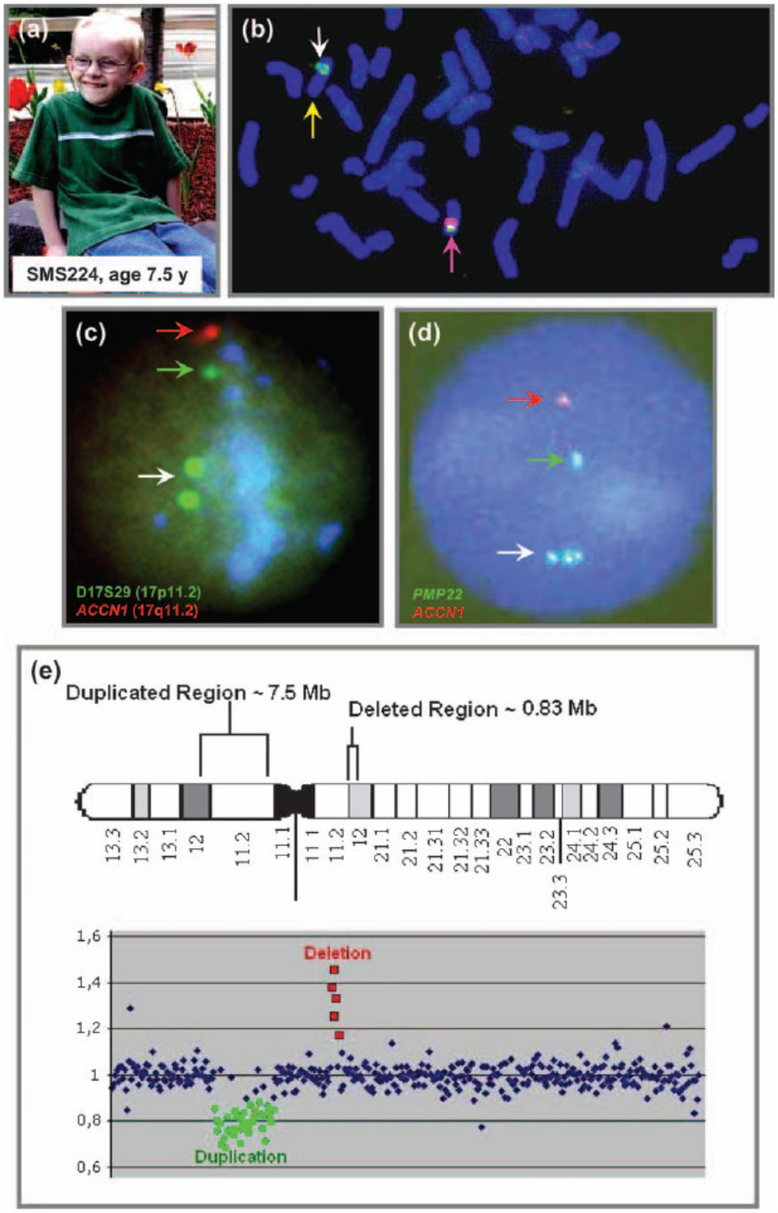

(a)
SMS224, age 7.5 y
(b)
(c)
D17S29 (17p11.2)
ACCN1 (17q11.2)
(d)
PMP22
ACCN1
(e)
Duplicated Region ~ 7.5 Mb
Deleted Region ~ 0.83 Mb
13.3
13.2
13.1
12
11.2
11.1
11.1
11.2
12
21.1
21.2
21.31
21.32
21.33
22
23.1
23.2
23.3
24.1
24.2
24.3
25.1
25.2
25.3
1,6
1,4
1,2
1
0,8
0,6
Deletion
Duplication

Fig. 1. Molecular analysis of SMS224: (**a**) Photo of SMS224 at the age of 7.5 years. (**b**) FISH analysis on a metaphase spread showing *ACCN1* (RP11-90A23) deletion (yellow arrow) and 17p11.2 (CTC-457L16/D17S29) duplication (white arrow) on the same homologue of chromosome 17. The normal chromosome 17 homologue is indicated (pink arrow). (**c**) FISH on an interphase nucleus shows a duplication (white arrow) and a normal copy (green arrow) for a 17p11.2 probe (CTC-457L16/D17S29). A single copy of *ACCN1* (RP11-90A23) is also indicated (red arrow). (**d**) FISH on an interphase nucleus shows a triplication for the *PMP22* (RP11-764O6) probe on one chromosome 17 (white arrow) and a single copy of *PMP22* on the normal chromosome 17; *ACCN1* probe (RP11-90A23) shows a single copy (red arrow). (**e**) Array CGH data for SMS224 showing gain of DNA indicated by a peak that spans 17(p11.2)(p12) and loss of DNA on 17(q11.2)(q12) represented by an ideogram of chromosome 17 (top).

and paresis of the right leg because of a thoracolumbar syrinx detected by magnetic resonance imaging (MRI). Syrinx decompression by shunt insertion was performed at about 1 year of age, after which his physical capacity and activity improved. Significant feeding difficulties and hypotonia were observed during early infancy with poor weight gain, and he was subsequently diagnosed with oropharyngeal dysphagia for which he was under nasogastric intubation until 6 months of age. A history of constipation prompted a rectal biopsy for Hirschsprung disease, which was normal. Because the subject developed recurrent, unexplained fevers and seemed to be anhidrotic, he was evaluated by a clinical geneticist who did not find evidence for anhidrotic ectodermal dysplasia. Developmentally, he has exhibited delays in physical and verbal milestones, as well as delays in social, emotional, and cognitive functioning. His first steps were at 18 month of age but have become independent for ambulation although he has a mild steppage gait. He currently has very little speech and has been communicating his needs by pointing or leading people to what he wants. He has normal dentition and normal tooth eruption.

Initially, a number of studies were performed including tests for urine and plasma amino acids, organic acids, peroxisomal panel, acylcarnitine panel, and immunological testing, all of which were normal. Subsequent analysis for fragile X and FISH studies for velocardiofacial syndrome were negative. Further clinical evaluations revealed a series of medical problems. An echocardiogram showed a number of congenital abnormalities including a bicuspid aortic valve, dilated aortic root and ascending aortic aneurysm, patent foramen ovale, and a ventriculoseptal defect for which he underwent a valve-sparing root and ascending aortic replacement surgeries. He also had mild mitral regurgitation and left ventricular hypertrophy with a shortening fraction of 36%. He also had several episodes per week of unexplained sinus tachycardia, with heart rates up to 190 beats/min, associated with syncope and cyanosis and that were investigated by cardiac catheterization. In addition to confirming the echocardiographic findings, he was found to have a narrow right coronary artery ostium. Episodic syncope was thought to be caused by poor flow into the right coronary artery, resulting in transient reduction in cardiac output, or by a Bezold–Jarisch reflex precipitated by obstruction to flow at the aortic valve. Syncopal spells resolved after replacement of the aortic root, aortic valve, and ascending aorta, and regrafting of the right coronary artery.

Frequent episodes of altered consciousness, compatible with non-convulsive seizures, were also seen in this child. Electroencephalogram studies documented multifocal epileptiform discharges on several occasions and almost continuous spike and slow wave activity during sleep. Occasionally, postictal periods were characterized by transient improvement in his ability to form coherent sentences. MRI of the brain showed patchy white matter abnormalities. 3-D time-of-flight magnetic resonance angiogram (MRA) of the circle of Willis showed no evidence of an arteriovenous malformation or aneurysm, although a slight asymmetry in the flow of the carotid arteries, the right being slightly smaller than the left, was documented. Magnetic resonance spectroscopy was also within normal limits.

Parents of SMS224 note that he had difficulty recognizing faces and following objects at 3–4 years of age. Ophthalmologic examination revealed severe hyperopia and some amblyopia for which refractive correction was given. External eye exams revealed no ptosis or proptosis, and no preferential fixation or nystagmus was noted. His eye movements were conjugate with intact corneal reflexes. The patient was sensitive to bright light and preferred dim light. Examination of the anterior segments showed no abnormalities; pupils were reactive to light with no relative afferent pupillary defect. Dilated fundus examination revealed a pink, healthy optic disc with no pallor or swelling. The history of photophobia was suspected for a cone dystrophy, although, his central vision was normal. A possible retinal dystrophy because of the deletion of *ACCN1*, implicated in light-induced damage to the retina, was considered; retinal fundus and electroretinogram studies showed no abnormalities.

Sleep studies revealed severe sleep disordered breathing with hypercapnia. He was reportedly very difficult to wake up as a baby but now has reversed and was reported as an extremely light sleeper. He also has obstructive and central apnea. SMS224 also had enlarged tonsils and adenoids for which he underwent surgery. Hearing exams revealed a moderately severe hearing loss. In addition, the patient was evaluated for malignant hyperthermia because of unexplained fevers and inappropriate sweating that is accompanied by muscle twitching, leukocytosis, and hyperglycemia after general anesthesia that usually resolved within 24 h. A malrotation of the left kidney was also identified by ultrasound.

Upon physical examination at the age of 7 years, he was unable to follow verbal cues and does not follow simple commands. He reached for objects and would accurately grasp them in a coarse fashion and in a purposeful manner. At the age of 7.5 years, his height was 120 cm (∼50th percentile), weight 22 kg ($<$5th percentile). His facial features showed a prominent forehead, downslanting palpebral fissures, mild mandibular hypoplasia, and a normal palate and uvula. He had dental caries and a notch in his left lower incisor superiorly. He had long fingers, hyperextensibility at the elbows and proximal interphalangeal joints, and flat arches of the feet. Cardiovascular examination was significant with a 3/6 harsh ejection systolic murmur with no diastolic murmur heard. His lungs were clear to auscultation bilaterally. Neurologically, he had some truncal hypotonia with increased tone symmetrically in both legs. Extremity examination showed metatarsus adductus bilaterally. He had delayed deep tendon reflexes, a withdrawal plantar response. He was non-verbal, ambulatory without assistance, and he had coarse motor control of his upper and lower extremities with a coarse gait and a slight foot slap when walking.

Peripheral nerve biopsy showed reduced packing density with expansion of the intervening endoneurial connective tissue consistent with moderately severe nerve fiber loss. On ultrastructural examination, many of the myelinated fibers showed retraction between axoplasm and myelin sheath; occasionally, rare fibers showed several concentrically arranged profiles of Schwann cell cytoplasm external to the myelinating Schwann cell, suggestive of early onion bulb formation. Gastrocnemius muscle biopsy showed moderately scattered atrophic fibers, relatively small diameter type 2 fibers, and mild fiber type grouping that is most prominent in the type 2 fiber population. Endomysial capillaries appeared dilated and prominent with thickening of the basal lamina. Electron microscopy showed well-formed myofibrils with no definite increase in myofiber glycogen, lipid, or mitochondrial content noted. Electromyography and nerve conduction studies ($\sim$15 m/sec) showed signs of demyelinating polyneuropathy, compatible with CMT1A.

There is no relevant family history of consanguinity, peripheral neuromuscular disease, or mental retardation. There is a history of miscarriages on the maternal side wherein the mother has had four miscarriages, and the grandmother has had seven. There is also a family history of cardiovascular manifestation with his sister having a dilated aorta and his maternal half-sister having a dilated aorta and long QT syndrome. In addition, there is a history on his paternal grandfather's side of individuals having undergone aortic replacement or rupture, including a cousin having an aortic rupture in his twenties. A possibility of Marfan syndrome in this family was assessed and subsequently ruled-out after a thorough clinical evaluation. A summary of clinical features in SMS224 compared with published data on patients with different sized chromosome 17 duplications is given in Table 3.

Molecular findings

Initial cytogenetic analysis of peripheral blood from SMS224 by karyotyping and standard GTG banding techniques showed a partial duplication of chromosome 17: 46,XY,dup(17)(p11.2p12). FISH studies confirmed the duplication and also identified a mosaic triplication or partial tetrasomy in $\sim$25% of the interphase nuclei. The duplication maps $\sim$7.5 Mb region within 17(p11.2p12) including the CMT1A and the SMS regions and encompassing more than 100 genes including (from telomere to centromere) *COX10, PMP22, ADORA2B, UBB, TNFRSF13B, RAI1, DRG2, PRPSAP2, EPN2, SPECC1, USP22* and *KCNJ12* (Table 1, Fig. 1b–d). Genotyping of patient and parental DNAs to determine the parent of origin for the 17p11.2 duplication revealed the presence of one paternal allele and both maternal alleles indicating a maternal inter-chromosomal recombination event. Further, aCGH was performed to identify any other genomic rearrangements that might contribute to the unusual features in this patient. Array CGH not only confirmed the breakpoints of the 17(p11.2p12) duplication but also identified a deletion containing exon 1 of the amiloride-sensitive cation channel neuronal 1 (*ACCN1*) gene and a cluster of genes coding for cytokine receptors (*CCL2, CCL7, CCL8,* and *CCL11*) on chromosome 17(q11.2q12) (Table 1, Fig. 1f). FISH was performed to confirm the extent of the deletion using locus specific probes; RP11-797M22, RP11-90A23, RP11-79K15, RP11-47G8, and RP11-329H16 (see Methods). The deletion was confirmed in all the cells examined by FISH. Parental chromosome analyses for both rearrangements on chromosome 17 using FISH and aCGH were negative.

Genotype–phenotype correlation

Comparison of clinical features of published cases with different sized chromosome 17 duplications to SMS224 shows that our patient has many features consistent with a large-sized duplication such as severe cognitive impairment, severe mental retardation, facial dysmorphism, cardiovascular, and renal defects (Table 3). To identify clinical features solely due to the 17p rearrangement, $\sim$40 clinical characteristics of reported cases with duplication 17p including dup17p11.2p11.2, dup17p11.2p12, dup17p11.2p10, and trisomy 17p were analyzed. While the 31 cases reported in the literature do show variability (summarized in Table 3), a common

Table 3. Clinical features of trisomy 17p, dup17p10p12, dup17p11.2p12, dup17p11.2p11.2, and SMS224[a]

Clinical features[b]	Trisomy 17p *n* = 12	Dup(17) (p10p12) *n* = 4	Dup(17) (p11.2p12) *n* = 7	Dup(17) (p11.2p11.2) *n* = 8	SMS224
Mental retardation	10/10	4/4	7/7	8/8	+
Pre- /post-natal growth retardation	12/12	3/4	7/7	3/3	+
Motor delay	12/12	4/4	7/7	2/3	+
Speech delay	2/2	4/4	5/5	2/4	+
Maternal history of spontaneous abortions	3/8	1/3	*n*	*n*	+
Hypotonia	6/10	4/4	2/3	2/4	+
Feeding difficulties	7/7	4/4	3/4	1/3	+
Hyperactivity/behavioral abnormalities	*n*	1/3	2/5	4/6	+
Seizures	2/12	2/4	1/6	3/7	+
Decreased nerve conduction	1/1	1/4	5/6	0/8	+
Delayed deep tendon reflexes	1/1	1/4	6/7	0/8	+
Craniofacial features					
Triangular face	*n*	3/3	2/4	1/2	+
Microcephaly	12/12	4/4	2/3	4/6	−
Micrognathia	11/12	3/4	1/4	2/3	−
Hypertelorism	8/12	2/4	3/3	1/3	−
Broad nasal bridge	11/12	4/4	4/5	2/4	−
Low-set/malformed ears	12/12	2/4	4/4	3/3	+
High arched palate	7/12	1/4	3/5	3/4	+
Cleft lip/palate	4/12	0/4	0/7	1/4	−
Ophthalmologic disorders					
Downslanting palpebral fissures	9/12	2/4	6/6	4/4	+
Ptosis	6/12	2/4	1/7	*n*	−
Myopia or hyperopia	*n*	2/4	2/7	1/3	+
Strabismus	*n*	0/4	1/7	1/2	−
Others eye anomalies	7/10	0/4	0/8	0/8	+
Short broad neck	7/12	*n*	1/7	*n*	−
Hearing loss	3/3	1/4	1/7	1/3	+
Cardiac anomalies					
Valvular defects	5/12	0/4	2/7	0/8	+
Dilated aorta/aneurysms/coarctation	1/12	0/4	0/7	0/8	+
Renal/urinary tract anomalies	3/12	1/4	1/7	2/8	+
Short stature	11/12	1/4	6/6	5/8	−
Decreased weight	11/12	1/4	4/4	3/5	+
Bone and limb defects					
Clinodactyly of fifth finger	7/12	*n*	4/7	2/8	−
Flexion contractures	10/12	1/4	3/7	0/8	−
Foot anomalies	5/12	2/4	6/7	0/8	−
Short sternum	7/12	*n*	1/6	*n*	*n*
Dental anomalies	*n*	1/4	4/7	5/8	−
Widely placed nipples	5/10	*n*	1/1	*n*	*n*
Single palmar crease	4/9	*n*	*n*	*n*	*n*
Malignant hyperthermia	*n*	*n*	*n*	*n*	+
Syringomyelia	*n*	*n*	*n*	*n*	+

[a] The numbers are based on published case reports of trisomy 17p [Bartsch-Sandhoff et al. (12); Feldman et al. (13); Jinno et al. 1982; Kulharya et al. 1998; Latta et al. 1974; Martsolf et al. (29); Mascarello et al. (15); Mikhail et al. 2006; Palutke et al. (11); Schrander-Stumpel et al. (17); Shabtai et al. 1979; Yamamoto et al. 1979], dup(17)(p10p12) [Docherty et al. (14); Shaw et al. (27); Stankiewicz et al. 2001; Yatsenko et al. (28)] dup(17)(p11.2p12), [Kozma et al. (18); Lupski et al. 1992; Magenis et al. (16); Pellegrino et al. (20); Potocki et al. 1999; Roa et al. (22); Upadhyaya et al. (19)], and dup(17)(p11.2p11.2) [Balarin et al. (23), Brown et al. (21), Potocki et al. (24), Schneider et al. (25)].

[b] Note that the denominators in each of the features are not equal to '*n*' numbers because of: incomplete or missing data and/or age not appropriate for assessment of that particular feature.

phenotype for dup17p syndrome did emerge. SMS224 has the clinical features common to these patients, such as characteristic craniofacial features, pre- and post-natal growth retardation, feeding difficulties, hyperactivity, cognitive impairment, speech delay, and hypotonia (Table 3). Further, the presence of delayed tendon reflexes, decreased nerve conduction, and peripheral neuromuscular disease is consistent with the involvement of *PMP22* (duplication of which causes CMT1A). However, limb abnormalities including flexion deformity of fingers, clinodactyly of fifth finger, and club foot that are commonly seen in patients with 17p duplications were not seen in our patient. Published cases of CMT1A homozygous for the 17p11.2 duplication have variable severity of clinical features suggesting that the severity and variability of the disease are not determined solely by the copy number of *PMP22* (32). Thus, the more severe manifestations and the unusual, previously unreported features in SMS224 (Table 3) can be because of the following: increased dosage of genes within the triplicated segment, involvement of genes such as *AKAP10*, *SPECC1*, *USP22*, *TMEM*, *MAP2K3*, and *KCNJ12* that are outside the SMS and CMT1A regions but within the chromosome 17p rearrangement breakpoints, or because of haploinsufficiency of genes on chromosome 17q11.2q12 (*ACCN1* and/or cytokine gene cluster).

Certain clinical manifestations seen in SMS224 such as severe cardiac manifestations, multiform epileptic seizures, thoracolumbar syrinx, and malignant hyperthermia are unusual and are not reported in patients with duplication 17p (Table 3). These features could be associated with the 17q11.2q12 deletion encompassing *ACCN1* and the cluster of cytokines or may be because of another, undetected, genetic abnormality.

Contribution of ACCN1 *towards severe phenotype in SMS224*

ACCN1 (also known as *BNAC1*, *BNC1*, *ASIC2*, or *MDEG*) maps to the long arm of chromosome 17 and is composed of 10 exons encoding a member of the degenerin/epithelial sodium channel (DEG/ENaC) superfamily (Fig. 2a) (33). Alternative splicing has been observed at this locus and two variants, encoding distinct isoforms differing in the first 184 amino acids, have also been identified (Fig. 2a). *ACCN1* is expressed during early embryogenesis and throughout post-natal life in the central and peripheral nervous systems of humans, specifically the brain (34). The members of this family are amiloride-sensitive sodium channels that contain intracellular N- and C-termini, 2 hydrophobic transmembrane regions (TM1 and TM2), and a large extracellular loop, which has many cysteine residues with conserved spacing (Fig. 2b) (35). ACCN1 is glycosylated at the two *N*-linked asparagine residues 365 and 392 during post-translational processing (Fig. 2b) (36, 37). The pre-transmembrane domain (pre-TM2) and the second transmembrane segment (TM2) form the amiloride-binding site and determine the selectivity of the channel (Fig. 2b) (38). Also, the conserved domain at the N-terminal end and the cysteine rich domain near the transmembrane domain 2 (TM2) are important for the basic function of the channel (39).

This class of acid-sensing channels also contributes to learning and memory as well as acidosis-induced injury, where they may act as mediators of stimuli that respond to the acidosis that accompanies stroke, epilepsy, and trauma under pathological conditions (40, 41). Comparison of clinical features in patients with NF1 with large 17q11.2 deletions encompassing *ACCN1* shows that variable mental retardation is the only common feature possibly contributed by *ACCN1* (42). To support this, Venturin et al. (43) performed a deletion gene

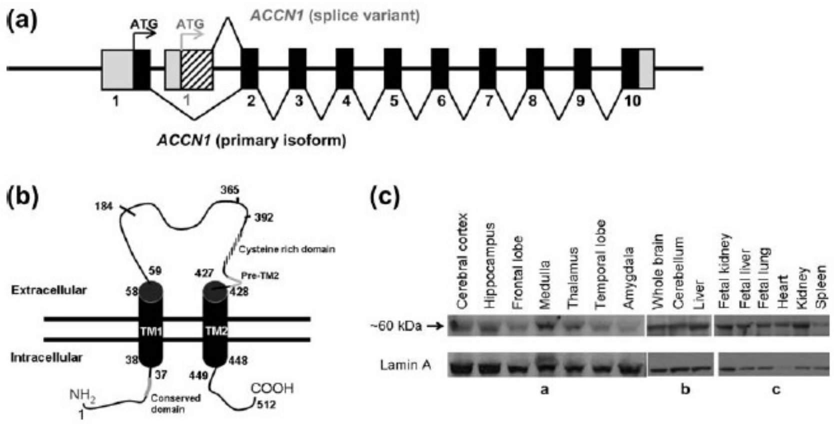


Fig. 2. Genomic structure, protein structure, and expression of ACCN1: (**a**) The genomic architecture of *ACCN1* with 10 exons (filled boxes) and 5′and 3′ UTRs (open boxes) is shown. Alternatively, spliced first exon (striped box) is also shown. (**b**) Schematic representation of ACCN1 showing intracellular N-terminus region, transmembrane domains (TM1 and TM2, black cylinders), a large extracellular loop, a pre-transmembrane domain (pre-TM2, gray segment), and a C-terminal intracellular tail (37). The amino acid residues forming the different components of ACCN1 are depicted. *N*-linked asparagine residues 365 and 392 are glycosylated during post-translational modification. (36) Other regions of importance are the conserved domain at the N-terminal end (gray segment) and the cysteine-rich domain (striped segment) near the TM2. (**c**) Western blot analysis of human tissue-specific protein medleys shows an ∼60 kDa band corresponding to ACCN1 (see Methods). Blots contain 40 μg (a &b) or 20 μg (c) of each protein lysate (see Methods).

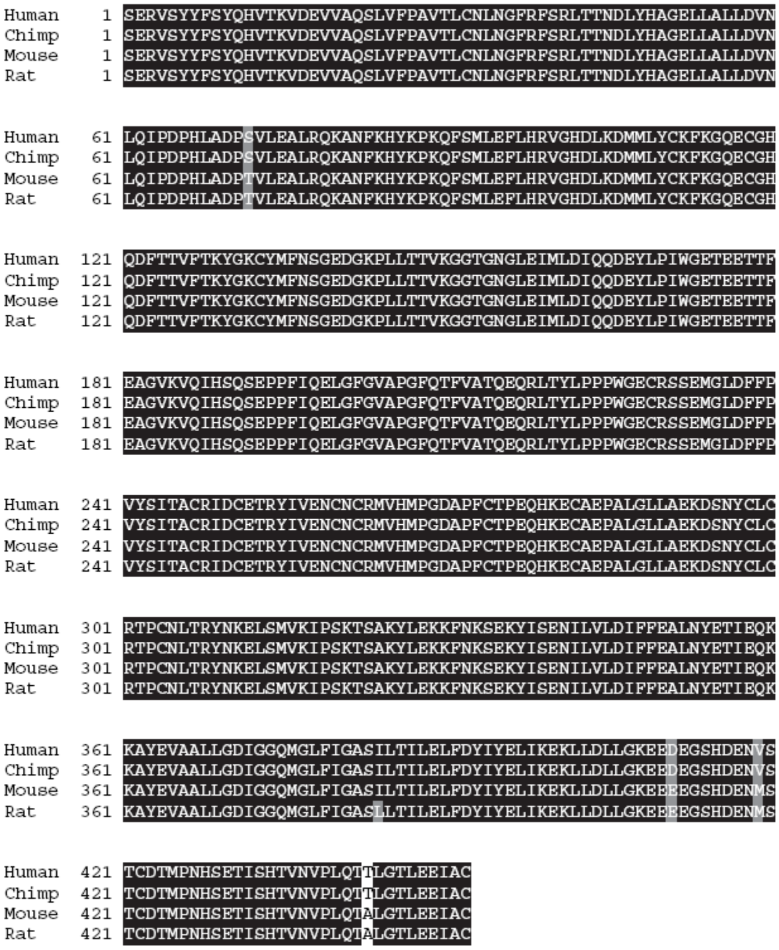
content analysis in patients with NF1 microdeletion and identified six genes including *ACCN1*, where haploinsufficiency may contribute to the mental retardation in these patients. Further, the *ACCN1* family of neuronal degenerins has been implicated in neurodegeneration in *Caenorhabditis elegans* and epilepsy in humans (34, 44). Adult *Accn1* knockout mice exhibit light-induced retinal degeneration as a result of uncontrolled rod photo transduction (34, 45).

To further understand the expression pattern of ACCN1 in human tissues, western analysis was performed using human tissue-specific protein medleys (Fig. 2c). Results indicate that the major isoform of ACCN1 (~60 kDa) is expressed in all of the major organs namely, fetal and adult brain (cerebral cortex, cerebellum, medulla, hippocampus, thalamus, amygdala, and frontal and temporal lobes), liver, heart, kidney, and spleen. *ACCN1* expression is not confined to the brain as previously reported (Fig. 2c) (34). Alignment of the ACCN1 protein sequences of human, chimp, mouse, and rat showed that human and mouse sequences are 99.2% identical and 99.4% similar; humans and rats have 99% identity and 99.2% similarity whereas humans and chimps share 100% identity (Fig. 3).

To further assess any possible role for *ACCN1* in the phenotype in SMS224, the entire *ACCN1* gene was sequenced including the coding exons and the splice variants (Table 2). Sequence analysis of the non-deleted allele of *ACCN1* and its splice variant showed no mutations.

Discussion

In this report, we describe a male patient with a multiple congenital abnormalities/mental retardation syndrome associated with unusual systemic manifestations and a severe phenotype. His initial cytogenetic findings showed a chromosome 17p11.2p12 duplication involving both the SMS and CMT1A regions. It is clear that other patients with duplications of the proximal 17p region have mild clinical abnormalities, but the severity seen in our patient prompted further genome-wide studies.



```
Human    1 SERVSYYFSYQHVTKVDEVVAQSLVFPAVTLCNLNGFRFSRLTTNDLYHAGELLALLDVN
Chimp    1 SERVSYYFSYQHVTKVDEVVAQSLVFPAVTLCNLNGFRFSRLTTNDLYHAGELLALLDVN
Mouse    1 SERVSYYFSYQHVTKVDEVVAQSLVFPAVTLCNLNGFRFSRLTTNDLYHAGELLALLDVN
Rat      1 SERVSYYFSYQHVTKVDEVVAQSLVFPAVTLCNLNGFRFSRLTTNDLYHAGELLALLDVN

Human   61 LQIPDPHLADPSVLEALRQKANFKHYKPKQFSMLEFLHRVGHDLKDMMLYCKFKGQECGH
Chimp   61 LQIPDPHLADPSVLEALRQKANFKHYKPKQFSMLEFLHRVGHDLKDMMLYCKFKGQECGH
Mouse   61 LQIPDPHLADPTVLEALRQKANFKHYKPKQFSMLEFLHRVGHDLKDMMLYCKFKGQECGH
Rat     61 LQIPDPHLADPTVLEALRQKANFKHYKPKQFSMLEFLHRVGHDLKDMMLYCKFKGQECGH

Human  121 QDFTTVFTKYGKCYMFNSGEDGKPLLTTVKGGTGNGLEIMLDIQQDEYLPIWGETEETTF
Chimp  121 QDFTTVFTKYGKCYMFNSGEDGKPLLTTVKGGTGNGLEIMLDIQQDEYLPIWGETEETTF
Mouse  121 QDFTTVFTKYGKCYMFNSGEDGKPLLTTVKGGTGNGLEIMLDIQQDEYLPIWGETEETTF
Rat    121 QDFTTVFTKYGKCYMFNSGEDGKPLLTTVKGGTGNGLEIMLDIQQDEYLPIWGETEETTF

Human  181 EAGVKVQIHSQSEPPFIQELGFGVAPGFQTFVATQEQRLTYLPPPWGECRSSEMGLDFFP
Chimp  181 EAGVKVQIHSQSEPPFIQELGFGVAPGFQTFVATQEQRLTYLPPPWGECRSSEMGLDFFP
Mouse  181 EAGVKVQIHSQSEPPFIQELGFGVAPGFQTFVATQEQRLTYLPPPWGECRSSEMGLDFFP
Rat    181 EAGVKVQIHSQSEPPFIQELGFGVAPGFQTFVATQEQRLTYLPPPWGECRSSEMGLDFFP

Human  241 VYSITACRIDCETRYIVENCNCRMVHMPGDAPFCTPEQHKECAEPALGLLAEKDSNYCLC
Chimp  241 VYSITACRIDCETRYIVENCNCRMVHMPGDAPFCTPEQHKECAEPALGLLAEKDSNYCLC
Mouse  241 VYSITACRIDCETRYIVENCNCRMVHMPGDAPFCTPEQHKECAEPALGLLAEKDSNYCLC
Rat    241 VYSITACRIDCETRYIVENCNCRMVHMPGDAPFCTPEQHKECAEPALGLLAEKDSNYCLC

Human  301 RTPCNLTRYNKELSMVKIPSKTSAKYLEKKFNKSEKYISENILVLDIFFEALNYETIEQK
Chimp  301 RTPCNLTRYNKELSMVKIPSKTSAKYLEKKFNKSEKYISENILVLDIFFEALNYETIEQK
Mouse  301 RTPCNLTRYNKELSMVKIPSKTSAKYLEKKFNKSEKYISENILVLDIFFEALNYETIEQK
Rat    301 RTPCNLTRYNKELSMVKIPSKTSAKYLEKKFNKSEKYISENILVLDIFFEALNYETIEQK

Human  361 KAYEVAALLGDIGGQMGLFIGASILTILELFDYIYELIKEKLLDLLGKEEDEGSHDENVS
Chimp  361 KAYEVAALLGDIGGQMGLFIGASILTILELFDYIYELIKEKLLDLLGKEEDEGSHDENVS
Mouse  361 KAYEVAALLGDIGGQMGLFIGASILTILELFDYIYELIKEKLLDLLGKEEEEGSHDENMS
Rat    361 KAYEVAALLGDIGGQMGLFIGASLLTILELFDYIYELIKEKLLDLLGKEEEEGSHDENMS

Human  421 TCDTMPNHSETISHTVNVPLQTTLGTLEEIAC
Chimp  421 TCDTMPNHSETISHTVNVPLQTTLGTLEEIAC
Mouse  421 TCDTMPNHSETISHTVNVPLQTALGTLEEIAC
Rat    421 TCDTMPNHSETISHTVNVPLQTALGTLEEIAC
```

Fig. 3. ACCN1 is highly conserved through species. Boxshade alignment of human ACCN1 (Q16515, NM_001094), chimp (NP_001085, uncurated sequences from UCSC genome browser), rat (Q62962, NM_001034014), and mouse Accn1 (Q925H0, NM_1034013) sequences are shown.

FISH experiments identified a mosaic tetrasomy of the same chromosomal region as the duplication. Array CGH analysis mapped the extent of this chromosome 17p rearrangement, spanning from *COX10* to *KCNJ12* and also found a 17q11.2q12 deletion encompassing the first exon of *ACCN1* and a cluster of cytokine genes. A detailed clinical evaluation of this patient identified a complex and severe phenotype, including several previously unreported systemic manifestations. Genotype–phenotype analysis suggests that the severity of the phenotype in this patient is either because of the tetrasomy of $\sim$100 genes on proximal 17p, some of which can be dosage sensitive, or because of the unusually large-sized duplication encompassing genes outside the SMS/CMT1A regions. The unusual features, hitherto not reported in patients with partial or complete duplication 17p, are possibly because of hemizygous deletion of the genes *ACCN1* and the cytokine cluster on 17(q11.2q12). SMS224 is legally blind and disruption of *ACCN1* may play a role in this aspect of the phenotype. Although an electroretinogram showed no indications of retinal degeneration in this patient, age-related progression of retinal degeneration cannot be ruled out. The cluster of cytokine genes, *CCL2, CCL7, CCL8*, and *CCL11,* is a family of secreted proteins involved in immunoregulatory and inflammatory processes. Further investigations needs to be carried out to decipher the role of this group of genes in contributing to the phenotype in our patient.

A detailed analysis of the family history from SMS224 documented the presence of certain manifestations that might have a hereditary basis suggesting a role for genes/loci other than that involved in the chromosome 17 rearrangement. There is a maternal history of miscarriages, suggesting the possibility of an inherited abnormality in the family even though parental karyotypes were normal. In addition, SMS224 was diagnosed with a dilated aortic root early in life and had aortic valve replacement surgeries. The sister of this patient was diagnosed with a dilated aortic root, and the half-sister from the mother's side also was diagnosed with a dilated aorta, long QT syndrome, scoliosis, extra teeth, and extreme flexibility. Also, the paternal side of the family has a history of aortic pathologies. While Marfan syndrome was ruled out in this family, the possibility of another locus elsewhere in the genome contributing to this connective tissue-associated phenotype requires further studies. Congenital bicuspid valve leading to aortic dilation has been reported earlier in sporadic and familial cases (46, 47). Thus, identification of more patients with similar hereditary defects might shed light in to the mode of inheritance and, possibly, help in the identification of susceptible loci.

We also report here for the first time a patient with tetrasomy 17p11.2p12. Triplications or partial tetrasomies for other chromosomal regions have been reported (Table 4) (48–62). Constitutional triplications of genomic regions have been associated with various disorders that include mental retardation syndromes, Pelizaeus–Merzbacher syndrome, Parkinsonism and recently, hereditary pancreatitis (63–66). The majority of chromosomal triplications reported so far are interstitial, with the exception of those cases reported by Devriendt et al. and Batanian et al. (48, 50) that have a triplication of the terminal chromosomal fragment.

Table 4. Summary of reported cases of intrachromosomal triplications

Clinical features	Triplicated segment	Reference
Multiple congenital anomalies associated with mental retardation, skeletal and craniofacial features, developmental delay	2q37	Rauch et al. (55)
	2q11.1q21	Wang et al. (61)
	3q25.3q29	Ounap et al. (54)
	5p14p15.33	Harrison et al. (52)
	7p21.3p22	Rivera et al. 1998
	8p22p23	Shieh et al. (58)
	9p13p22	Verheij et al. (60)
	9p22pter	Batanian et al. (48)
	10q26	Devriendt et al. (50)
	12p11.22p12.3	Eckel et al. (51)
	13q14	Brecevic et al. (49)
	13q22q23	Reddy and Logan 2001
	15q11q13	Roberts et al. (57)
	15q24q26	James et al. (53)
	22q11.2	Yobb et al. (62)
Hereditary pancreatitis	7q34 (*PRSS1*)	Marechal et al. (63)
Parkinsonism	Parkin gene (6q25.2q27)	Lucking et al. (64)
	A synuclein (4q21)	Singleton et al. (65)
Pelizaeus–Merzbacher disease	PLP (Xq22)	Wolf et al. (66)

The mechanism of duplication coupled with mosaic triplication in our patient could be because of an aberrant recombination during meiosis I followed by an intrachromosomal mitotic event resulting in mosaicism, similar to that described for other mosaic triplication cases (56). It is observed, in general, that tetrasomy of an autosomal segment causes a similar but more severe phenotype than trisomy of the same segment as demonstrated in tetrasomy 9p, 18p, and 15(q12-q13) (49). Thus, we can attribute the severity of the phenotype in our patient to increased copy number of dosage sensitive genes in 17p11.2p12 and unusual features to 17q11.2q12 deletion.

In conclusion, our study emphasizes the need for whole genome analysis in multiple congenital abnormalities/mental retardation disorders with unusual and severe phenotypes. Without the use of aCGH, this patient may never have been diagnosed with a 17q11.2q12 deletion possibly affecting his prognosis and treatment. Further, some duplications identified by cytogenetic analysis may, in fact, be triplications and therefore warrant further investigations by interphase FISH.

Acknowledgements

We thank our patient and his family for their interest and pursuit of research efforts. Drs Harold Dietz, Catherine Keegan, John Gerlach, and Michael Netzloff and Rebecca McFalda for helpful discussions, patient referral, and evaluation. This work was funded by a grant from the US National Institutes of Health (R01 HD38534) to S. H. E. and resources from Virginia Commonwealth University (to S. H. E.), and also the Roswell Park Cancer Center Support Grant CA16056 (to N. N.).

References

1. Lee JA, Lupski JR. Genomic rearrangements and gene copy-number alterations as a cause of nervous system disorders. Neuron 2006: 52: 103–121.
2. Lupski JR, de Oca-Luna RM, Slaugenhaupt S et al. DNA duplication associated with Charcot–Marie–Tooth disease type 1A. Cell 1991: 66: 219–232.
3. Chance PF, Alderson MK, Leppig KA et al. DNA deletion associated with hereditary neuropathy with liability to pressure palsies. Cell 1993: 72: 143–151.
4. Smith AC, McGavran L, Robinson J et al. Interstitial deletion of (17)(p11.2p11.2) in nine patients. Am J Med Genet 1986: 24: 393–414.
5. Vlangos CN, Yim DK, Elsea SH. 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 2003: 79: 134–141.
6. Vlangos CN, Wilson M, Blancato J et al. Diagnostic FISH probes for del(17)(p11.2p11.2) associated with Smith–Magenis syndrome should contain the RAI1 gene. Am J Med Genet A 2005: 132: 278–282.
7. Slager RE, Newton TL, Vlangos CN et al. Mutations in RAI1 associated with Smith–Magenis syndrome. Nat Genet 2003: 33: 466–468.
8. Bi W, Saifi GM, Shaw CJ et al. Mutations of RAI1, a PHD-containing protein, in nondeletion patients with Smith–Magenis syndrome. Hum Genet 2004: 115: 515–524.
9. Girirajan S, Elsas LJ 2nd, Devriendt K et al. RAI1 variations in Smith–Magenis syndrome patients without 17p11.2 deletions. J Med Genet 2005: 42: 820–828.
10. Girirajan S, Vlangos CN, Szomju BB et al. Genotype–phenotype correlation in Smith–Magenis syndrome: evidence that multiple genes in 17p11.2 contribute to the clinical spectrum. Genet Med 2006: 8: 417–427.
11. Palutke W, Chen H, Woolley P Jr et al. An extra small metacentric chromosome identified as a deleted chromosome no. 17. Clin Genet 1976: 9: 454–458.
12. Bartsch-Sandhoff M, Hieronimi G. Partial duplication of 17p. A new chromosomal syndrome. Hum Genet 1979: 49: 123–127.
13. Feldman GM, Baumer JG, Sparkes RS. Brief clinical report: the dup(17p) syndrome. Am J Med Genet 1982: 11: 299–304.
14. Docherty Z, Hulten MA, Honeyman MM. De novo tandem duplication 17p11 leads to cen. J Med Genet 1983: 20: 138–142.
15. Mascarello JT, Jones MC, Hoyme HE et al. Duplication (17p) in a child with an isodicentric (17p) chromosome. Am J Med Genet 1983: 14: 67–72.
16. Magenis RE, Brown MG, Allen L et al. De novo partial duplication of 17p [dup(17)(p12–p11.2)]: clinical report. Am J Med Genet 1986: 24: 415–420.
17. Schrander-Stumpel C, Schrander J, Fryns JP et al. Trisomy 17p due to a t(8;17) (p23;p11.2)pat translocation. Case report and review of the literature. Clin Genet 1990: 37: 148–152.
18. Kozma C, Meck JM, Loomis KJ et al. De novo duplication of 17p [dup(17)(p12 – p11.2)]: report of an additional case with confirmation of the cytogenetic, phenotypic, and developmental aspects. Am J Med Genet 1991: 41: 446–450.
19. Upadhyaya M, Roberts SH, Farnham J et al. Charcot–Marie–Tooth disease 1A (CMT1A) associated with a maternal duplication of chromosome 17p11.2–>12. Hum Genet 1993: 91: 392–394.
20. Pellegrino JE, Pellegrino L, Spinner NB et al. Developmental profile in a patient with monosomy 10q and dup(17p) associated with a peripheral neuropathy. Am J Med Genet 1996: 61: 377–381.
21. Brown A, Phelan MC, Patil S et al. Two patients with duplication of 17p11.2: the reciprocal of the Smith–Magenis syndrome deletion? Am J Med Genet 1996: 63: 373–377.
22. Roa BB, Greenberg F, Gunaratne P et al. Duplication of the PMP22 gene in 17p partial trisomy patients with Charcot–Marie–Tooth type-1 neuropathy. Hum Genet 1996: 97: 642–649.
23. Balarin MA, da Silva Lopes VL, Varella-Garcia M. A dup(17)(p11.2p11.2) detected by fluorescence in situ hybridization in a boy with Alport syndrome. Am J Med Genet 1999: 82: 183–186.
24. Potocki L, Chen KS, Park SS et al. Molecular mechanism for duplication 17p11.2 – the homologous recombination reciprocal of the Smith–Magenis microdeletion. Nat Genet 2000: 24: 84–87.
25. Schneider MC, Hughes CR, Forrester S et al. Mild phenotype due to tandem duplication of l7p11.2. Am J Med Genet 2000: 94: 296–299.
26. Stankiewicz P, Shaw CJ, Dapper JD et al. Genome architecture catalyzes nonrecurrent chromosomal rearrangements. Am J Hum Genet 2003: 72: 1101–1116.
27. Shaw CJ, Stankiewicz P, Christodoulou J et al. A girl with duplication 17p10-p12 associated with a dicentric chromosome. Am J Med Genet A 2004: 124: 173–178.
28. Yatsenko SA, Treadwell-Deering D, Krull K et al. Trisomy 17p10-p12 due to mosaic supernumerary marker chromosome: delineation of molecular breakpoints and clinical

phenotype, and comparison to other proximal 17p segmental duplications. Am J Med Genet A 2005: 138: 175–180.

29. Martsolf JT, Larson L, Jalal SM et al. Complete trisomy 17p a relatively new syndrome. Ann Genet 1988: 31: 172–174.
30. Lucas RE, Vlangos CN, Das P et al. 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 2001: 9: 892–902.
31. Cowell JK, Nowak NJ. High-resolution analysis of genetic events in cancer cells using bacterial artificial chromosome arrays and comparative genome hybridization. Adv Cancer Res 2003: 90: 91–125.
32. LeGuern E, Gouider R, Mabin D et al. Patients homozygous for the 17p11.2 duplication in Charcot-Marie-Tooth type 1A disease. Ann Neurol 1997: 41: 104–108.
33. Waldmann R, Voilley N, Mattei MG et al. The human degenerin MDEG, an amiloride-sensitive neuronal cation channel, is localized on chromosome 17q11.2-17q12 close to the microsatellite D17S798. Genomics 1996: 37: 269–270.
34. Waldmann R, Champigny G, Voilley N et al. The mammalian degenerin MDEG, an amiloride-sensitive cation channel activated by mutations causing neurodegeneration in Caenorhabditis elegans. J Biol Chem 1996: 271: 10433–10436.
35. Darboux I, Lingueglia E, Champigny G et al. dGNaC1, a gonad-specific amiloride-sensitive Na+ channel. J Biol Chem 1998: 273: 9424–9429.
36. Saugstad JA, Roberts JA, Dong J et al. Analysis of the membrane topology of the acid-sensing ion channel 2a. J Biol Chem 2004: 279: 55514–55519.
37. Snyder PM, McDonald FJ, Stokes JB et al. Membrane topology of the amiloride-sensitive epithelial sodium channel. J Biol Chem 1994: 269: 24379–24383.
38. Kellenberger S, Gautschi I, Schild L. A single point mutation in the pore region of the epithelial Na+ channel changes ion selectivity by modifying molecular sieving. Proc Natl Acad Sci U S A 1999: 96: 4170–4175.
39. Bianchi L, Driscoll M. Protons at the gate: DEG/ENaC ion channels help us feel and remember. Neuron 2002: 34: 337–340.
40. Wemmie JA, Chen J, Askwith CC et al. The acid-activated ion channel ASIC contributes to synaptic plasticity, learning, and memory. Neuron 2002: 34: 463–477.
41. Xiong ZG, Zhu XM, Chu XP et al. Neuroprotection in ischemia: blocking calcium-permeable acid-sensing ion channels. Cell 2004: 118: 687–698.
42. Riva P, Corrado L, Natacci F et al. NF1 microdeletion syndrome: refined FISH characterization of sporadic and familial deletions with locus-specific probes. Am J Hum Genet 2000: 66: 100–109.
43. Venturin M, Guarnieri P, Natacci F et al. Mental retardation and cardiovascular malformations in NF1 microdeleted patients point to candidate genes in 17q11.2. J Med Genet 2004: 41: 35–41.
44. Biagini G, Babinski K, Avoli M et al. Regional and subunit-specific downregulation of acid-sensing ion channels in the pilocarpine model of epilepsy. Neurobiol Dis 2001: 8: 45–58.
45. Ettaiche M, Guy N, Hofman P et al. Acid-sensing ion channel 2 is important for retinal function and protects against light-induced retinal degeneration. J Neurosci 2004: 24: 1005–1012.
46. Clementi M, Notari L, Borghi A et al. Familial congenital bicuspid aortic valve: a disorder of uncertain inheritance. Am J Med Genet 1996: 62: 336–338.
47. McKusick VA. Association of congenital bicuspid aortic valve and Erdheim's cystic medial necrosis. Lancet 1972: 1: 1026–1027.
48. Batanian JR, Chen X, Grange DK. Mosaic isodicentric chromosome 9 triplication (9p22-pter) and no deletion in an abnormal infant presenting with clinical features of trisomy 9: a new type of isodicentric chromosome formation. Am J Hum Genet 1994: 55: A98.
49. Brecevic L, Basaran S, Dutly F et al. Tandem triplication of chromosome 13q14 with inverted interstitial segment in a 4 year old girl. J Med Genet 2000: 37: 964–967.
50. Devriendt K, Matthijs G, Holvoet M et al. Triplication of distal chromosome 10q. J Med Genet 1999: 36: 242–245.
51. Eckel H, Wimmer R, Volleth M et al. Intrachromosomal triplication 12p11.22-p12.3 and gonadal mosaicism of partial tetrasomy 12p. Am J Med Genet A 2006: 140: 1219–1222.
52. Harrison KJ, Teshima IE, Silver MM et al. Partial tetrasomy with triplication of chromosome (5) (p14-p15.33) in a patient with severe multiple congenital anomalies. Am J Med Genet 1998: 79: 103–107.
53. James PA, Aftimos S, Oei P. Partial tetrasomy 15 due to a unique inverted triplication of chromosome15q24-q26. Am J Med Genet A 2004: 130: 208–210.
54. Ounap K, Ilus T, Bartsch O. A girl with inverted triplication of chromosome 3q25.3–> q29 and multiple congenital anomalies consistent with 3q duplication syndrome. Am J Med Genet A 2005: 134: 434–438.
55. Rauch A, Pfeiffer RA, Trautmann U. Deletion or triplication of the alpha 3 (VI) collagen gene in three patients with 2q37 chromosome aberrations and symptoms of collagen-related disorders. Clin Genet 1996: 49: 279–285.
56. Reddy KS, Logan JJ. Intrachromosomal triplications: molecular cytogenetic and clinical studies. Clin Genet 2000: 58: 134–141.
57. Roberts SE, Dennis NR, Browne CE et al. Characterisation of interstitial duplications and triplications of chromosome 15q11-q13. Hum Genet 2002: 110: 227–234.
58. Shieh JT, Hudgins L, Cherry AM et al. Triplication of 8p22-8p23 in a patient with features similar to Kabuki syndrome. Am J Med Genet A 2006: 140: 170–173.
59. Thomas PK, Marques W Jr, Davis MB et al. The phenotypic manifestations of chromosome 17p11.2 duplication. Brain 1997: 120 (Pt 3): 465–478.
60. Verheij JB, Bouman K, van Lingen RA et al. Tetrasomy 9p due to an intrachromosomal triplication of 9p13-p22. Am J Med Genet 1999: 86: 168–173.
61. Wang J, Reddy KS, Wang E et al. Intrachromosomal triplication of 2q11.2-q21 in a severely malformed infant: case report and review of triplications and their possible mechanism. Am J Med Genet 1999: 82: 312–317.
62. Yobb TM, Somerville MJ, Willatt L et al. Microduplication and triplication of 22q11.2: a highly variable syndrome. Am J Hum Genet 2005: 76: 865–876.
63. Le Marechal C, Masson E, Chen JM et al. Hereditary pancreatitis caused by triplication of the trypsinogen locus. Nat Genet 2006: 38: 1372–1374.
64. Lucking CB, Bonifati V, Periquet M et al. Pseudo-dominant inheritance and exon 2 triplication in a family with parkin gene mutations. Neurology 2001: 57: 924–927.
65. Singleton AB, Farrer M, Johnson J et al. Alpha-synuclein locus triplication causes Parkinson's disease. Science 2003: 302: 841.
66. Wolf NI, Sistermans EA, Cundall M et al. Three or more copies of the proteolipid protein gene PLP1 cause severe Pelizaeus–Merzbacher disease. Brain 2005: 128: 743–751.

Copyright of Clinical Genetics is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Appendix C

GENETIC TESTING
Volume 12, Number 1, 2008
© Mary Ann Liebert, Inc.
Pp. 67–74
DOI: 10.1089/gte.2007.0058

Diagnosing Smith–Magenis Syndrome and Duplication 17p11.2 Syndrome by *RAI1* Gene Copy Number Variation Using Quantitative Real-Time PCR

Hoa T. Truong,[1] Sara Solaymani-Kohal,[2] Kevin R. Baker,[3] Santhosh Girirajan,[4] Stephen R. Williams,[4]
Christopher N. Vlangos,[4] Ann C.M. Smith,[5,6] David J. Bunyan,[3] Paul E. Roffey,[1]
Christopher L. Blanchard,[1] and Sarah H. Elsea[4]

Smith–Magenis syndrome (SMS) and duplication 17p11.2 (dup17p11.2) syndrome are multiple congenital anomalies/mental retardation disorders resulting from either a deletion or duplication of the 17p11.2 region, respectively. The retinoic acid induced 1 (*RAI1*) gene is the causative gene for SMS and is included in the 17p11.2 region of dup17p11.2 syndrome. Currently SMS and dup17p11.2 syndrome are diagnosed using a combination of clinically recognized phenotypes and molecular cytogenetic analyses such as fluorescent *in situ* hybridization (FISH). However, these methods have proven to be highly expensive, time consuming, and dependent upon the low resolving capabilities of the assay. To address the need for improved diagnostic methods for SMS and dup17p11.2 syndrome, we designed a quantitative real-time PCR (Q-PCR) assay that measures *RAI1* copy number using the comparative C_t method, $\Delta\Delta C_t$. We tested our assay with samples blinded to their previous SMS or dup17p11.2 syndrome status. In all cases, we were able to determine *RAI1* copy number status and render a correct diagnosis accordingly. We validated these results by both FISH and multiplex ligation-dependent probe amplification (MLPA). We conclude that Q-PCR is an accurate, reproducible, low-cost, and reliable assay that can be employed for routine use in SMS and dup17p11.2 diagnosis.

Introduction

Smith–Magenis syndrome (SMS, OMIM #182290) is a multiple congenital anomalies/mental retardation disorder that includes distinctive facial features, cognitive and behavioral characteristics, self-injurious behaviors, and sleep disturbance (Smith *et al.*, 1986; Greenberg *et al.*, 1991; Greenberg *et al.*, 1996; Allanson *et al.*, 1999; Girirajan *et al.*, 2005). Nonallelic homologous recombination (NAHR) between SMS repeats (SMS-REPs) or other low copy repeats (LCRs) within the 17p11.2 region gives rise to SMS or the recombination reciprocal, duplication 17p11.2 (dup17p11.2) syndrome (Chen *et al.*, 1997; Potocki *et al.*, 2000; Bi W *et al.*, 2003). Individuals with dup17p11.2 syndrome also present with congenital anomalies, neurological and behavioral characteristics, as well as other clinical features that are distinct from that of SMS (Magenis RE *et al.*, 1986; Potocki *et al.*, 2007; Girirajan *et al.*, 2007b). The phenotypic spectrum of both SMS and dup17p11.2 syndrome is outlined in Table 1.

The majority of SMS patients carry an interstitial deletion of 17p11.2. Of these, 70% have a common deletion that spans ~3.5 Mb, while the remainder harbor larger or smaller deletions. The phenotypic manifestations observed in SMS are believed to be due to either the haploinsufficiency of the retinoic acid induced 1 gene (*RAI1*), which lies within this interstitial deletion region, or mutation of the *RAI1* gene (Slager *et al.*, 2003; Vlangos *et al.*, 2003; Bi *et al.*, 2004; Girirajan *et al.*, 2005). Although *RAI1* has not been implicated as the causative gene in dup17p11.2 syndrome, it is contained within the recently refined 1.3 Mb critical region for this syndrome (Potocki *et al.*, 2007).

The prevalence of SMS is approximated at 1:25,000, and the prevalence of dup17p11.2 syndrome should be equivalent to that of its recombination reciprocal (Greenberg *et al.*,

---

[1] Faculty of Science, Charles Sturt University, Wagga Wagga, NSW, Australia.
[2] University of the West of England, Bristol, United Kingdom.
[3] Wessex Regional Genetics Laboratory, Salisbury District Hospital, Salisbury, United Kingdom.
[4] Departments of Pediatrics and Human Genetics, Virginia Commonwealth University, Richmond, Virginia.
[5] Office of the Clinical Director, NHGRI, National Institutes of Health, Bethesda, Maryland.
[6] Georgetown University Medical Center, Washington, District of Columbia.

TABLE 1. CLINICAL MANIFESTATIONS OF SMITH–MAGENIS SYNDROME AND DUP17P11.2 SYNDROME

		SMS		
Craniofacial/skeletal	*Otolaryngologic*	*Neurological/behavioral*	*Ocular*	*Other*
Brachycephaly	Chronic ear infections	Variable mental retardation	Myopia	Cardiovascular anomalies
Midface hypoplasia	Hearing loss	Speech delay	Strabismus	Renal/urinary track abnormality
Prognathism	Hoarse, deep voice	Motor delay		Dental anomalies
Tented upper lip		Hypotonia		Hypercholesterolemia
Broad, square face		Seizures		
Synophrys		Sleep disturbance		
Brachydactyly		Self hugging		
Short stature		Attention seeking		
Scoliosis		Self-injurious behaviors		
		- Onychotillomania		
		- Polyembolokoilomania		
		- Head banging		

	Dup17p11.2		
Craniofacial/skeletal	*Neurological/behavioral*	*Ocular*	*Other*
Microcephaly	Variable mental retardation	Hypermetropia	EEG abnormalities
Triangular face	Speech delay		Low cholesterol
Pre- and postnatal growth delay	Motor delay		Cardiovascular anomalies
	Hypotonia		Oral-pharyngeal dysphagia
	Failure to thrive in infancy		Dental anomalies
	Sleep apnea		
	Hyperactivity		

1991; Potocki *et al.*, 2007). In the past decade, SMS diagnosis has relied upon a combination of clinical recognition of characteristic features, cytogenetic analysis, and molecular screening (Gropman *et al.*, 2006). The majority of cases of SMS have been diagnosed using G-banded chromosome analysis and/or fluorescent *in situ* hybridization (FISH). In patients where the SMS phenotype is present but there is no cytogenetically detectable deletion, the *RAI1* gene has been sequenced for mutations (Slager *et al.*, 2003; Bi *et al.*, 2004; Girirajan *et al.*, 2005). Routine G-banded chromosome analyses and/or FISH have typically been used to diagnose dup17p11.2 patients (Potocki *et al.*, 2007).

The incidence rate, however, is likely to be an underestimation due to the subtle phenotypes presented in both syndromes, the diversity in size of the deletions/duplications involved in perpetuating the syndrome, and also the reliability of the currently available diagnostic procedures (Struthers *et al.*, 2002; Potocki *et al.*, 2007). Many of the clinical and neurobehavioral features that are characteristic markers used in SMS diagnosis are subtle during infancy and early childhood but become increasingly pronounced with age, lending to a delayed diagnosis in many cases (Gropman *et al.*, 2006). Likewise, many of the features associated with dup17p11.2 syndrome escape recognition until later infancy or childhood (Potocki *et al.*, 2007).

Currently available cytogenetic methodologies that are routinely used in SMS and dup17p11.2 syndrome diagnoses have limitations. These include the low resolving power of G-banded karyotyping to visualize subtle deletions or duplications, the inability of FISH to detect deletions that may be smaller than the probes utilized, and the need for interphase rather than routine metaphase FISH to fully delineate duplicated regions. Further, from a practical diagnostic perspective, the cost, the time, and the labor intensity of these cytogenetic methodologies are considered to be quite high.

Together, all of these factors point to the need for an efficient, quick, and reliable diagnostic tool to aid in the identification of individuals with SMS or dup17p11.2 syndrome. This would allow a more timely diagnosis and will ensure that intervention measures for clinical manifestations are put into place as soon as possible. We have established a quantitative real-time PCR (Q-PCR) assay that is capable of measuring copy number variation in the 17p11.2 region, but more specifically, it measures the number of copies of the *RAI1* gene. We have validated this method by conducting a blinded study with samples that have a known deletion or duplication of *RAI1* and verified our results using FISH and multiplex ligation-dependent probe amplification (MLPA). Our approach is highly specific, fast, and reliable, making it an attractive method for routine use in SMS and dup17p11.2 diagnosis.

Materials and Methods

Subjects

To validate the Q-PCR assay for use in SMS and dup17p11.2 syndrome diagnosis, 17 SMS, 1 dup17p11.2, and 4 normal samples were blinded to the previously confirmed diagnoses of SMS or dup17p11.2 syndrome. Normal *RAI1* deletion and *RAI1* duplication control samples were also included in each run. Once the study was complete, all samples were decoded to reveal their copy number status, allowing for verification of the results obtained with FISH and MLPA. Additionally, the *RAI1* Q-PCR assay was tested on 42 samples of unrelated healthy individuals who were not previously tested for a deletion or duplication of the *RAI1* gene. Samples used in this study were collected in accordance to approved IRB protocols by either Virginia Commonwealth University (VCU3784) or the SMS Natural History Study (01-HG-0109) at the National Institutes of Health.

Genomic DNA

Patient and control genomic DNA was extracted by the phenol chloroform method from whole blood or from lymphoblastoid cell lines or saliva, using the QIAamp DNA Minikit (QIAGEN, GmbH, Hilden, Germany) or the Oragene DNA Purification Kit (DNA Genotek, Ottawa, Canada), respectively, and approximately 50 ng used per reaction. The concentration and purity of the genomic DNA preparations were measured at an absorbance of 260 and 280 nm. DNA was stored at −20°C until ready for use.

Fluorescence in situ *hybridization*

FISH was used to confirm previous clinical and cytogenetic diagnoses in all 18 patient samples used in the blinded study. The following BAC and PAC clones mapping to chromosome 17p11.2 were used for preparation of FISH probes: CTC-529I10, CTC-157E16, RP11-314M5, RP11-253P07 (representing the *RAI1* locus), and RP11-746M1. FISH was performed as previously described (Vlangos *et al.*, 2003).

Multiplex ligation-dependent probe amplification

The *RAI1* copy number status of the blinded samples was also tested using the P064B MLPA Mental Retardation 1 SALSA Kit (MRC-Holland, Amsterdam, The Netherlands), covering chromosome 17p11.2. Because probes specific for *RAI1* were not contained within this kit, additional probes were created. The design, location, sequence, and size of *RAI1* probes used in conjunction with the commercially available kit are outlined in Table II of Girirajan *et al.* (2007a). MLPA was performed as previously described (Girirajan *et al.*, 2007a).

MLPA data analysis

MLPA PCR products were separated on an ABI PRISM® 3100 Genetic Analyser and interpreted using Genotyper (v3.7) software (Applied Biosystems, Foster City, CA). *RAI1* copy number was determined by exporting peak heights into an Excel spreadsheet, which had been designed to assess the ratio of each test peak relative to all other peaks for the given individual. Further, ratios of test peaks to control peaks and ratios of control peaks to control peaks for each sample were compared to those of two normal individuals, which were included in each run (Bunyan *et al.*, 2004). The dosage quotient for a normal individual with two copies of *RAI1* is expected to be 1.0, 0.5 for a deletion (SMS), and 1.5 for a duplication (dup17p11.2 syndrome).

Quantitative real-time PCR

PCR was performed using an AB 7500 FAST Sequence Detection System (Applied Biosystems). Each reaction was performed in a total volume of 10 µL, containing 1× Taqman® Universal PCR Master Mix, No AmpErase® UNG, 1× Taqman Gene Expression Assay Mix (*RNase P*: P/N 431681 or *RAI1*: P/N Hs02570777_s1) at a final concentration of 900 nM of each primer, 250 nM of the probe, and 50 ng genomic DNA. All reactions were performed in triplicate and repeated at least once. PCR thermocycling conditions consisted of an initial polymerase activation and DNA denaturation step of 60°C for 15 min, followed by 40 cycles of 15 s at 95°C and 1 min at 60°C.

Q-PCR data analysis

Real-time PCR is a highly sensitive and specific assay that utilizes emission from fluorescent dyes (e.g., SYBR Green I) or fluorogenic probes (e.g., TaqMan Probes) as a direct measure of target amplification during the PCR reaction. Accumulated PCR products are measured when there is a noticeable increase in fluorescence above a set threshold level. The cycle at which this occurs is the threshold cycle (C_t). For quantification of gene copy number by real-time PCR, the C_t values of both an unknown sample and a calibrator sample are measured, and this difference represents *n*-fold copies of the target gene in the unknown sample, relative to the calibrator sample (Livak and Schmittgen, 2001).

Quantifying relative gene copies requires that the amplification efficiency of both reference and target genes are comparable. If amplification efficiencies are approximately equal, it is possible to employ the comparative C_t method, $\Delta\Delta C_t$, for determining relative gene copy (ABI PRISM User Bulletin #2). Efficiencies for *RNase P* (reference gene) and *RAI1* (target gene) were determined by constructing a standard curve over a 5 log dilution series and comparing the slopes of the resultant curves.

Background baseline and threshold levels for the amplification plots of *RAI1* and *RNase P* were automatically determined by the SDS Software 1.3.1 (Applied Biosystems). The threshold was set at 10 standard deviations above the mean baseline fluorescence within the logarithmic curve of the amplification plot. Using the relative quantification study of this program, *RAI1* copy number of the unknown sample was determined relative to the known copy number of a normal calibrator sample, using the expression: $2^{-(\Delta\Delta Ct \pm s)}$, where $\Delta\Delta C_t = [C_t$ *RAI1* (unknown sample) $- C_t$ *RNase P* (unknown sample)$] - [C_t$ *RAI1* (calibrator sample) $- C_t$ *RNase P* (calibrator sample)$]$. The standard deviation, *s*, is the difference of the standard deviations of the C_t values of *RAI1* and *RNase P*. An *RAI1* copy number of 1 is expected for normal diploid samples, 0.5 for samples with an *RAI1* deletion (SMS), and 1.5 for samples that have a duplication of *RAI1* (dup17p11.2 syndrome).

Results

In this study, we designed a Q-PCR assay to determine copy number of the *RAI1* gene, which lies within the critical deletion and duplication region for SMS and dup17p11.2 syndrome, respectively. The Taqman Gene Expression Assay specific to *RAI1* was chosen such that the primers and probe lie within the exon 6 boundary of this gene. *RNase P* was chosen as a reference for gene dosage because it is a single-copy gene and would be present as two copies in all individuals regardless of SMS or dup17p11.2 syndrome status.

The PCR amplification efficiency for the target gene and the reference gene was determined by constructing a standard curve over 5 log dilutions. The amplification efficiency based upon the slopes of the standard curves was 100% for *RAI1* and 97% for *RNase P*. Since the amplification efficiencies for both genes were approximately equivalent, it was possible to use the comparative C_t method ($\Delta\Delta C_t$) to determine the relative copy number of *RAI1*, as outlined in Materials and Methods.

Amplification plots for each triplicate showed near overlap, indicating that the assay was highly reproducible between replicates. However, in cases where laboratory error or issues with DNA quality led to insufficient amplification, samples were repeated. We noted that DNA samples that had been extracted by the phenol chloroform method had to be repeated more often than those extracted by commercial kits. Further, optimal amplification occurred using 40–80 ng DNA per reaction. Lower concentrations yielded unreliable amplification results.

We validated our *RAI1* Q-PCR assay using 17 SMS patients, 1 dup17p11.2 syndrome patient, and 4 normal individuals. FISH was used to confirm the diagnoses for these samples prior to use in Q-PCR. To prevent bias, we blinded the *RAI1* status of these samples. *RAI1* copy number of each sample was calculated by the $\Delta\Delta C_t$ ratio of $2^{-(\Delta\Delta Ct \pm s)}$. All samples were calibrated against a normal control, carrying two copies of *RAI1*. Once the study was complete, the samples were then decoded to allow for comparison of their expected genotype with the results obtained by Q-PCR. By Q-PCR, all 17 previously diagnosed SMS patients were deemed to carry a single copy of *RAI1*, the 4 normal individuals had two copies of *RAI1*, while the dup17p11.2 syndrome patient carried three copies of this gene (Fig. 1A).

The blinded samples were further subjected to MLPA to verify the results obtained by Q-PCR. Samples deleted for the *RAI1* gene exhibited half peak heights for all *RAI1* probes tested as compared to a normal control individual (Fig. 1C, bottom). Conversely, samples that carried a duplication of *RAI1* had peak heights that were 1.5 times that of a normal control individual for all *RAI1* probes (Fig. 1D, bottom). The *RA11* copy number status of the samples tested by MLPA confirmed the results we obtained by our Q-PCR assay as well as those by FISH (Fig. 1C, D, top).

We have established diagnostic ranges for detecting deletions and duplications of *RAI1* by Q-PCR based upon the samples that we have screened (Fig. 1B and Table 2). The diagnostic range for a *RAI1* deletion was 0.40–0.54 with a mean $\Delta\Delta C_t$ ratio of 0.47 ± 0.03. The $\Delta\Delta C_t$ ratio for a duplication of *RAI1* ranged between 1.53 and 1.78 with a mean of 1.66 ± 0.10. Samples from unrelated healthy individuals who were not previously screened for a deletion or duplication of the *RAI1* gene were used in the Q-PCR assay to determine a normal range. All 42 individuals were diploid for *RAI1* with a $\Delta\Delta C_t$ ratio that ranged from 0.72 to 1.28 and a mean of 0.98 ± 0.06.

Discussion

In recent years, Q-PCR has been employed to determine gene copy number for a variety of syndromes, including Charcot–Marie Tooth disease type 1A and hereditary neuropathy with liability to pressure palsy (Aarskog and Vedeler, 2000; Ruiz-Ponte *et al.*, 2000; Kim *et al.*, 2003; Thiel *et al.*, 2003), hereditary motor and sensory neuropathy I (Thiel *et al.*, 2003), Down syndrome (Hu *et al.*, 2004), Rubinstein–Taybi syndrome (Coupry *et al.*, 2004), and von Hippel Lindau syndrome (Hoebeeck *et al.*, 2005). Q-PCR offers several advantages in diagnostics, including speed, sensitivity, reproducibility, and low cost. It has the added advantage that there is no post-PCR manipulation, thereby reducing the possibility of introducing contamination.

In this study, we describe a Q-PCR assay that can be used to diagnose SMS and dup17p11.2 syndrome, based upon copy number variation of the *RAI1* gene. Since our Q-PCR assay is targeted specifically to this gene, it is especially useful in diagnosing atypical deletion cases in which SMS is suspected, but cytogenetic analysis was interpreted as normal and/or FISH did not include a probe specific to *RAI1* (Vlangos *et al.*, 2005). To evaluate our Q-PCR assay, a blinded study was designed using samples with normal, deleted, or duplicated copies of *RAI1* (Fig. 1A). The copy number status was accurately determined in 100% of the cases. Our Q-PCR findings were validated using both FISH and MLPA, illustrating the sensitivity and specificity of this assay.

Further, we established diagnostic ranges for each set of individuals who are normal, deleted, or duplicated for *RAI1*. The $\Delta\Delta C_t$ ratio for an *RAI1* deletion ranged from 0.40 to 0.54 with a mean of 0.47 ± 0.03, the $\Delta\Delta C_t$ ratio ranged between 1.53 and 1.78 with a mean of 1.66 ± 0.10 for *RAI1* duplications, and for a normal *RAI1* gene copy number this ratio was between 0.72 and 1.28 with a mean of 0.98 ± 0.06 (Table 2). Interestingly, we found that the Q-PCR results for SMS224, $\Delta\Delta C_t$ ratio $= 1.78$, were greater than our duplication control that had a $\Delta\Delta C_t$ ratio $= 1.53$ (Fig. 1A, B). Metaphase FISH and MLPA results indicated that this patient was duplicated for the *RAI1* gene (Fig. 1D).

Upon further investigation of SMS224 in our lab, it was determined by interphase FISH that this patient indeed had a duplication of 17p11.2 but was also 25% mosaic for tetrasomy 17p11.2p12 (Girirajan *et al.*, 2007b). Retrospectively, the mosaic tetrasomy for this region can possibly account for why our Q-PCR results were always greater than the expected $\Delta\Delta C_t$ ratio of 1.5 for a duplication of *RAI1*. It is, however, difficult to ascertain the extent of the effect that mosaicism lends to Q-PCR since we do not have any other mosaic deletion or duplication cases to evaluate at this time. But because of the limited number of duplication samples that we had available for Q-PCR testing, we have included the result for SMS224 within our *RAI1* Q-PCR duplication range. We expect that this range will fall in line with that of our control ($\Delta\Delta C_t$ ratio $= 1.53 \pm 0.11$) once more samples have been screened.

Accurate quantification by Q-PCR may be affected by a number of factors that warrant consideration. We carried out all reactions in a 96-well plate with a total reaction volume of 10 μL, which is half of what is recommended by the manufacturer. This reduction in reaction volume allowed for a marked decrease in the cost of reagents but did not appear to compromise our results, as reproducibility between replicates was consistent.

We noted, however, that amplification was affected by DNA quality whereby DNA that had been extracted using commercially available kits gave the most reliable results. Further, saliva provided high-quality DNA with added advantages of being a lower cost and far less invasive procedure compared to blood-based DNA collection. In cases where amplification was not sufficient because of poor DNA quality, new DNA samples were obtained and retested, yielding reproducible replicates. High-quality DNA is a critical factor for consistent amplification as well as contrib-

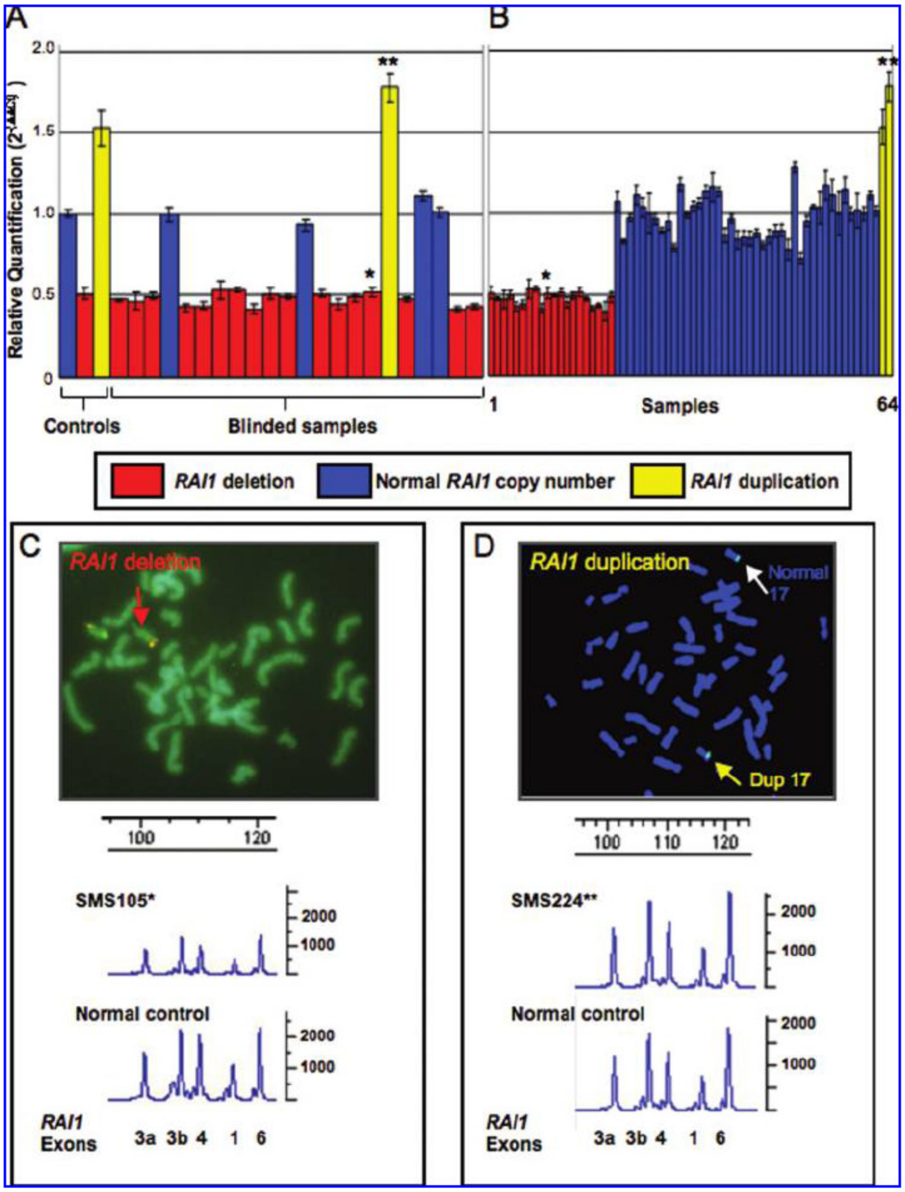


FIG. 1. Determining *RAI1* copy number variation by Q-PCR. (**A**) Mean results for the relative quantification, $2^{-(\Delta\Delta Ct)}$, and standard deviation of *RAI1* for 4 normal samples and 18 samples blinded to their previous diagnosis of either SMS or dup17p11.2 syndrome. Normal (blue), deletion (red), and duplication (yellow) controls were included in each run to ascertain the *RAI1* copy number status of the blinded sample. (**B**) All 64 samples assayed by Q-PCR provide nonoverlapping ranges for one (deletion), two (normal), or three (duplication) copies of *RAI1*. Q-PCR results were also validated by FISH and MLPA. Examples of FISH (top) and MLPA (bottom) results are shown for SMS (**C**) and dup17p11.2 syndrome (**D**). SMS105 is denoted with (*) and SMS224 with (**) in panels (**A**) and (**B**).

uting to a lower standard deviation between replicates (Thiel *et al.*, 2003).

Another important consideration when quantifying gene copy number is the use of an internal reference gene for normalization of sample-to-sample variation (Livak and Schmittgen, 2001). The choice of a suitable internal reference gene should be based upon whether the PCR amplification efficiencies of the target gene and the reference gene are similar. Amplification efficiencies closest to 100% ensure that accurate and reproducible results are obtained (ABI PRISM User Bulletin #2). Our internal reference gene, *RNase P*, had an amplification efficiency of 97% making it almost equivalent to the 100% amplification efficiency of *RAI1*. Thus *RNase P* was a suitable internal reference gene for our *RAI1* Q-PCR assay.

With the results that we have obtained with our Q-PCR assay outlined here, we propose that *RAI1* Q-PCR is a reliable method for diagnosis of SMS and dup17p11.2. Q-PCR offers several advantages over traditional cytogenetic diagnostics such as FISH and more recent methods such as

TABLE 2. NORMAL, DELETED, AND DUPLICATED $\Delta\Delta C_T$ RATIOS DETECTED BY *RAI1* Q-PCR

Samples	*Subjects*[a]	*Mean* ± *SD*[b]	*Range*
Normal	41	0.98 ± 0.06	0.72–1.28
Deletion	21	0.47 ± 0.03	0.40–0.54
Duplication	2	1.66 ± 0.10	1.53–1.78

[a] Number of different individuals tested, including control samples.
[b] Mean copy number ± standard deviation (SD).

MLPA. Of the three techniques used here, Q-PCR is the most cost effective and time conserving in terms of processing time and the acquisition of results, which are also important factors in patient diagnosis.

Based upon publicly available pricing from several diagnostic laboratories in the United States, we have found that the cost to the consumer ranges from US$500–800 for Q-PCR, FISH, or MLPA for diagnosing SMS, dup17p11.2, or other microdeletion/microduplication syndromes. However, the estimated cost per sample, excluding labor and assuming that the laboratory has the infrastructure required to carry out each method, is approximately US$150 for FISH, US$20 for MLPA, and US$9 for Q-PCR. It should be noted that this does not include the cost of including control samples or control probes that are required for relevant analyses in each of these methods, which would increase the costs stated here.

We estimate that the turnaround time, from the time that a sample is received until a diagnosis is rendered and a report generated, can be as little as 2 days for *RAI1* Q-PCR. The reaction set up and run time for Q-PCR can be completed in as little as 3 h. Comparatively, the turnaround time for FISH is approximately 7–10 days and for MLPA is 3–6 weeks. FISH is labor intensive and requires 30–35 h to complete. The turnaround time for FISH includes an additional time constraint of cell culturing for chromosome preparation that is unnecessary for Q-PCR or MLPA, as these two methods utilize genomic DNA. The long and variable turnaround time for MLPA is based on the fact that very few samples are sent for *RAI1* MLPA analysis; therefore, when samples are received they are placed in a batch system. Otherwise, MLPA for microdeletion/microduplication diagnosis can be performed within 24 h. Turnaround times for FISH and MLPA were based upon estimates given by national and international labs that currently perform these tests for SMS and dup17p11.2 syndrome.

In conclusion, we have developed a highly efficient, reliable, and cost-effective method of quantifying copy number aberrations of the *RAI1* gene. We have demonstrated that *RAI1* Q-PCR is capable of unequivocal diagnosis of SMS and dup17p11.2 syndrome and, further, offers several advantages over currently available techniques. We expect that as more diagnostic laboratories lean toward using methods that are economical but do not compromise on the accuracy of results, Q-PCR will be employed more routinely for screening gene copy number abnormalities.

Acknowledgments

The authors would like to acknowledge Sonia Benhamed (GeneDx, Gaithersburg, MD) for real-time PCR training. We would also like to thank Chris Miller (Applied Biosystems) for his expertise during the initial setup of the *RAI1* Q-PCR assay. Our gratitude extends to all SMS patients and families that provided samples, without whom our research would not be possible. This work was funded by a grant from the U.S. National Institutes of Health (R01 HD38534) to S.H.E. and resources from Virginia Commonwealth University to S.H.E. This research was also supported, in part, by the Intramural Research Program of the National Human Genome Research Institute, National Institutes of Health. Financial support to H.T.T. was provided by Charles Sturt University through a CSU postgraduate scholarship.

References

Aarskog NK, Vedeler CA (2000) Real-time quantitative polymerase chain reaction. A new method that detects both the peripheral myelin protein 22 duplication in Charcot-Marie-Tooth type 1A disease and the peripheral myelin protein 22 deletion in hereditary neuropathy with liability to pressure palsies. Hum Genet 107:494–498.

ABI PRISM User Bulletin #2. ABI Prism 7700 Sequence Detection System, December 11, 1997 (updated 10/2001). http://www.applied-biosystems.com.

Allanson JE, Greenberg F, Smith AC (1999) The face of Smith-Magenis syndrome: a subjective and objective study. J Med Genet 36:394–397.

Bi W, Park SS, Shaw CJ, Withers MA, Patel PI, Lupski JR (2003) Reciprocal crossovers and a positional preference for strand exchange in recombination events resulting in deletion or duplication of chromosome 17p11.2. Am J Hum Genet 73:1302–1315.

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.

Bunyan DJ, Eccles DM, Sillibourne J, Wilkins E, Thomas NS, Shea-Simonds J, Duncan PJ, Curtis CE, Robinson DO, Harvey JF, Cross NC (2004) Dosage analysis of cancer predisposition genes by multiplex ligation-dependent probe amplification. Br J cancer 91:1155–1159.

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.

Coupry I, Monnet L, Attia AA, Taine L, Lacombe D, Arveiler B (2004) Analysis of CBP (CREBBP) gene deletions in Rubinstein-Taybi syndrome patients using real-time quantitative PCR. Hum Mutat 23:278–284.

Girirajan S, Elsas LJ 2nd, Devriendt K, Elsea SH (2005) *RAI1* variations in Smith-Magenis syndrome patients without 17p11.2 deletions. J Med Genet 42:820–828.

Girirajan S, Mendoza-Londono R, Vlangos CN, Dupuis L, Nowak NJ, Bunyan DJ, Hatchwell E, Elsea SH (2007a) Smith-Magenis syndrome and moyamoya disease in a patient with del(17)(p11.2p13.1). Am J Med Genet A 143A:999–1008.

Girirajan S, Williams SR, Garben JY, Nowak NJ, Hatchwell E, Elsea S (2007b) 17p11.2p12 triplication and del(17)q11.2q12 in a severely affected child with dup(17)p11.2p12 syndrome. Clin Genet 72:47–58.

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.

Gropman AL, Duncan WC, Smith AC (2006) Neurologic and developmental features of the Smith-Magenis syndrome (del 17p11.2). Pediatr Neurol 34:337–350.

Hoebeeck J, van der Luijt R, Poppe B, de Smet E, Yigit N, Claes K, Zewald R, de Jong GJ, de Paepe A, Speleman F, Vandesompele J (2005) Rapid detection of VHL exon deletions using real-time quantitative PCR. Lab Invest 85:24–33.

Hu Y, Zheng M, Xu Z, Wang X, Cui H (2004) Quantitative real-time PCR technique for rapid prenatal diagnosis of Down syndrome. Prenat Diagn 24:704–707.

Kim SW, Lee KS, Jin HS, Lee TM, Koo SK, Lee YJ, Jung SC (2003) Rapid detection of duplication/deletion of the PMP22 gene in patients with Charcot-Marie-Tooth disease Type 1A and hereditary neuropathy with liability to pressure palsy by real-time quantitative PCR using SYBR Green I dye. J Korean Med Sci 18:727–732.

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402–408.

Magenis RE, Brown MG, Allen L, Reiss J (1986) *De novo* partial duplication of 17p [dup(17)(p12→p11.2)]: clinical report. Am J Med Genet 24:415–420.

Potocki L, Bi W, Treadwell-Deering D, Carvalho CM, Eifert A, Friedman EM, Glaze D, Krull K, Lee JA, Lewis RA, Mendoza-Londono R, Robbins-Furman P, Shaw C, Shi X, Weissenberger G, Withers M, Yatsenko SA, Zackai EH, Stankiewicz P, Lupski JR (2007) Characterization of Potocki-Lupski syndrome (dup(17)(p11.2p11.2)) and delineation of a dosage-sensitive critical interval that can convey an autism phenotype. Am J Hum Genet 80:633–649.

Potocki L, Chen KS, Park SS, Osterholm DE, Withers MA, Kimonis V, Summers AM, Meschino WS, Anyane-Yeboa K, Kashork CD, Shaffer LG, Lupski JR (2000) Molecular mechanism for duplication 17p11.2—the homologous recombination reciprocal of the Smith-Magenis microdeletion. Nat Genet 24:84–87.

Ruiz-Ponte C, Loidi L, Vega A, Carracedo A, Barros F (2000) Rapid real-time fluorescent PCR gene dosage test for the diagnosis of DNA duplications and deletions. Clin Chem 46: 1574–1582.

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 ACM, 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.

Struthers JL, Carson N, McGill M, Khalifa MM (2002) Molecular screening for Smith-Magenis syndrome among patients with mental retardation of unknown cause. J Med Genet 39:E59.

Thiel CT, Kraus C, Rauch A, Ekici AB, Rautenstrauss B, Reis A (2003) A new quantitative PCR multiplex assay for rapid analysis of chromosome 17p11.2-12 duplications and deletions leading to HMSN/HNPP. Eur J Hum Genet 11:170–178.

Vlangos CN, Wilson M, Blancato J, Smith AC, Elsea SH (2005) Diagnostic FISH probes for del(17)(p11.2p11.2) associated with Smith-Magenis syndrome should contain the *RAI1* gene. Am J Med Genet A 132:278–282.

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.

Address reprint requests to:
Sarah H. Elsea, Ph.D., FACMG
Department of Pediatrics and Human Genetics
Virginia Commonwealth University
1101 E. Marshall St., 12-018 Sanger Hall
P.O. Box 980441
Richmond, VA 23298

E-mail: selsea@vcu.edu

Appendix D

Mamm Genome (2008) 19:246–262
DOI 10.1007/s00335-008-9100-6

ORIGINAL CONTRIBUTIONS

Tom1l2 hypomorphic mice exhibit increased incidence of infections and tumors and abnormal immunologic response

Santhosh Girirajan · Paula M. Hauck · Stephen Williams · Christopher N. Vlangos · Barbara B. Szomju · Sara Solaymani-Kohal · Philip D. Mosier · Kimber L. White Jr. · Kathleen McCoy · Sarah H. Elsea

Received: 20 November 2007 / Accepted: 28 January 2008 / Published online: 15 March 2008
© Springer Science+Business Media, LLC 2008

Abstract Studies have shown that the TOM1 family of proteins, including TOM1 and TOM1L1, are actively involved in endosomal trafficking and function in the immune response. However, much less is known about the function of TOM1L2. To understand the biological importance of TOM1L2 and the potential significance of its cellular role, we created and evaluated *Tom1l2* gene-trapped mice with reduced *Tom1l2* expression. Mice hypomorphic for *Tom1l2* exhibited numerous infections and tumors compared to wild-type littermates. Associated with this increased risk for infection and tumor formation, apparently healthy *Tom1l2* hypomorphs also had splenomegaly, elevated B- and T-cell counts, and an impaired humoral response, although at a reduced penetrance. Furthermore, cellular localization studies showed that a Tom1l2-GFP fusion protein colocalizes with Golgi compartments, supporting the role of Tom1l2 in cellular trafficking, while molecular modeling and bioinformatic analysis of Tom1l2 illustrated a structural basis for a functional role in trafficking. These results indicate a role for *Tom1l2* in the immune response and possibly in tumor suppression.

Abbreviations

Tom1l2	Target of Myb-1-like 2
Tom1l1	Target of Myb-1-like 1
Tom1	Target of Myb-1

S. Girirajan and P. M. Hauck contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00335-008-9100-6) contains supplementary material, which is available to authorized users.

S. Girirajan · S. Williams · S. H. Elsea (✉)
Department of Human Genetics, Medical College of Virginia Campus, Virginia Commonwealth University, Richmond, VA 23298, USA
e-mail: selsea@vcu.edu

P. M. Hauck · C. N. Vlangos · B. B. Szomju ·
S. Solaymani-Kohal · S. H. Elsea
Department of Pediatrics, Medical College of Virginia Campus, Virginia Commonwealth University, Richmond, VA 23298, USA

Present Address:
P. M. Hauck
Molecular Oncology, McGuire Research Institute, Richmond, VA 23249, USA

Present Address:
C. N. Vlangos
Department of Pediatrics, University of Michigan, Ann Arbor, MI 48109, USA

P. D. Mosier
Department of Medicinal Chemistry and Center for the Study of Biological Complexity, Medical College of Virginia Campus, Virginia Commonwealth University, Richmond, VA 23298, USA

K. L. White Jr.
Department of Pharmacology and Toxicology, Medical College of Virginia Campus, Virginia Commonwealth University, Richmond, VA 23298, USA

K. McCoy
Department of Microbiology and Immunology, Medical College of Virginia Campus, Virginia Commonwealth University, Richmond, VA 23298, USA

Tollip	Toll interacting protein
SMS	Smith-Magenis syndrome
VHS	Vps27p/Hrs/Stam
GAT	GGA and Tom
GGA	Golgi localizing, Gamma-adaptin ear homology domain, ADP-ribosylation factor-binding protein
ARF1	ADP-ribosylation factor 1
SFK	Src family of protein kinases
IL-1R1	Interleukin-1 receptor type 1
SHIRPA	SmithKline Beecham Pharmaceuticals, Harwell MRC Mouse Genome Centre and Mammalian Genetics Unit, Imperial College School of Medicine at St Mary's, Royal London Hospital, St Bartholomew's, and the Royal London School of Medicine, Phenotype, Assessment

Introduction

Characterization of genes is of pivotal importance in understanding the biological processes that allow cells to function properly. This study focuses on *Tom1l2* (Target of Myb-1-Like 2), a 53.1-kb gene composed of 15 exons, encoding a 507-amino-acid protein ($\sim$55.7 kDa). *Tom1l2* belongs to a family of proteins involved in vesicular trafficking and endocytosis (Lohi and Lehto 1998). It is a homolog of the *Tom1* gene which was first discovered in a genetic screen to identify *Myb*-regulated genes (Burk et al. 1997). *Tom1* is a target of the proto-oncogene *v-Myb* when chicken myelomonocytic cells are transformed with avian myeloblastosis virus or avian leukemia virus E26 (Burk et al. 1997). The Tom1 family of proteins, including Tom1, Tom1l1 (alias Srcasm/Jerry), and Tom1l2, has a VHS (Vps27p/Hrs/Stam) domain, a GAT (GGA and Tom) domain, and clathrin-binding motifs. The VHS domain is involved in vesicular trafficking and endocytosis (Lohi and Lehto 1998), while the GAT domain of GGAs (Golgi localizing, Gamma-adaptin ear homology domain, ADP-ribosylation factor-binding protein) stabilizes membrane-bound ARF1 (ADP-ribosylation factor 1) in the GTP state (Takatsu et al. 2001). GGAs are monomeric clathrin adaptors (with clathrin boxes) that regulate transport and have been implicated in endosomal trafficking (Bonifacino 2004; Katoh et al. 2006). Clathrin-coated vesicles deliver cargo proteins from the plasma membrane and trans-Golgi network to the endosomes/lysosomes (Kirchhausen 2000).

Studies show that *Tom1* overexpression suppresses activation of transcription factors, including nuclear factor (NF)-κB and activator protein (AP-1), induced by either interleukin-1β or tumor necrosis factor-α, and that the VHS domain of Tom1 is indispensable for its suppressive activity (Yamakami and Yokosawa 2004). Tom1 is required for sorting of IL-1R1 (interleukin-1, receptor type 1) in late endosomes using Toll interacting protein (Tollip) as an endosomal adaptor, and Tom1 knockdown results in accumulation of IL-1R1 in late endosomes (Brissoni et al. 2006). Several studies have shown that Tom1 interacts with Tollip, ubiquitin, endofin, and clathrin (Akutsu et al. 2005; Katoh et al. 2004; Seet and Hong 2005; Seet et al. 2004; Shiba et al. 2004; Yamakami and Yokosawa 2004). Recently, human *TOM1* was implicated in both bipolar affective disorder (BPAD) and schizophrenia by a gene-based SNP haplotype mapping of a BPAD linkage region on chromosome 22q12.3 (Potash et al. 2008).

A partial functional redundancy appears to exist between members of the Tom1 family in regulating SFK (Src family of protein kinases) mitogenic signaling induced by growth factors (Franco et al. 2006). Tom1l1 and Tom1l2 act as adaptors for inhibition of mitogenesis (via disruption of SFK receptor complex formation), which is reversed by Myc expression or by SFK coexpression (Franco et al. 2006). In HEK293 cells, all human TOM1 members associated with SRC, with a stronger interaction for TOM1L1 and TOM1L2 (Franco et al. 2006). In addition to SFK, the tyrosine motifs at the C-terminal region of TOM1L1 mediate interactions with other signaling proteins such as GRB2, FYN, and the regulatory subunit of phosphoinositide 3-kinase, p85, in a phosphorylation-dependent manner (Puertollano 2005). Katoh et al. (2004) have shown that the GAT domains of TOM1 and TOM1L1 interact with both ubiquitin and TOLLIP. The VHS domain of TOM1 has been shown to interact with HRS (hepatocyte growth factor-regulated tyrosine kinase substrate), whereas a PTAP (Pro-Thr-Ala-Pro) motif, located between the VHS and GAT domain of TOM1L1, is responsible for binding to TSG101 (tumor susceptibility gene 101) (Puertollano 2005). Furthermore, TOM1L1 interacts with members of the multivesicular body-sorting machinery and may participate in the sorting of ubiquitinated proteins (Puertollano 2005). TOM1L1 and TOM1L2 differ from TOM1 in that they do not directly interact with endofin (Seet et al. 2004); however, like TOM1, they do interact with TOLLIP and clathrin (Katoh et al. 2006).

Human *TOM1L2* maps to chromosome 17p11.2, a region that is deleted in 90% of patients with Smith-Magenis syndrome (SMS) (Lucas et al. 2001). Murine *Tom1l2* lies in the syntenic region on chromosome 11. SMS is characterized by a constellation of approximately 30 clinical features, including mental retardation, sleep disturbance, speech and developmental delay, self-injurious and stereotypical behaviors, craniofacial and skeletal abnormalities, frequent infections, and hypotonia (Greenberg et al. 1991, 1996). The *RAI1* gene is responsible for most features of SMS (Slager et al. 2003; Smith et al.

1986), while other genes within the 17p11.2 deletion region contribute to the variable features and severity of the disorder (Girirajan et al. 2006), modifying the overall phenotype. The specific contribution of individual genes lying within the 17p11.2 deletion region to the SMS phenotype is not known.

The gene-trap approach has been a useful tool in generating mutant mouse (Gossler et al. 1989; Skarnes et al. 1995), Drosophila (Bellen et al. 1989; Bier et al. 1989), and zebrafish (Amsterdam et al. 1999; Talbot and Hopkins 2000) lines, allowing for functional studies to investigate the role of the trapped gene. To understand the functional consequences of *TOM1L2* deficiency and to decipher its potential contribution to the clinical features of SMS, a *Tom1l2* gene-trap mouse was created. Studies were performed to characterize the behavioral and immunologic phenotypes in the *Tom1l2* heterozygous and homozygous mice. *Tom1l2* cellular expression pattern, molecular modeling, and bioinformatic studies were also performed to obtain structural evidence for functional similarity between Tom1l2 and other Tom1 family members.

Materials and methods

Creation of *Tom1l2*-targeted mice

Publicly accessible databases from the BayGenomics consortium (http://www.baygenomics.ucsf.edu) were queried with *Tom1l2* EST sequence via BLAST analysis (Skarnes et al. 1992, 1995; Stanford et al. 2001; Stryke et al. 2003). BLAST results identified the XG909 embryonic stem (ES) cell line which has the gene-trapping vector (pGT1lxf) inserted between exons 10 and 11 in the mouse *Tom1l2* gene. XG909 ES cells were obtained and sent to the Transgenic Animal Model Core at The University of Michigan for injection into C57Bl/6 blastocysts. Chimeric males were mated to C57Bl/6J females to produce heterozygous (*Tom1l2+/−*) mice. The exact location of the insertion was determined by sequence analysis of a PCR product amplified using a vector-specific primer (5′ CGTAAGGAGAAAATACCGCATC 3′) and a *Tom1l2*-specific primer (5′ CAGAAAGCCACTGGGAAGAG 3′). Amplification of 50-100 ng of template DNA was performed using 1.25 U of *Taq* polymerase in a cocktail consisting of 0.4 µM of each primer, 0.2 mM dNTPs, and 2× PCR buffer (10× buffer contains 100 mM Tris-HCl, pH 8.3, 15 mM $MgCl_2$, 500 mM KCl, 0.01% gelatin). The samples were then subjected to an initial denaturation step at 94°C for 3 min, followed by 35 cycles of 94°C for 1 min, 60°C for 1 min, and 72°C for 1.5 min, followed by a final extension of 72°C for 10 min.

Genotyping of *Tom1l2* mice

DNA extraction Mouse tail (1–3 cm) was digested with 0.6 mg/ml proteinase K in TNES buffer (50 mM Tris, pH 7.5, 400 mM NaCl, and 100 mM EDTA, pH 8). DNA was then isolated using a standard phenol/chloroform procedure. Alternatively, DNA was extracted using the QIAamp DNA Minikit (Qiagen, Valencia, CA) as described in the Qiagen manual (tissue section).

Southern analysis Southern analysis was conducted by digesting 5 µg of DNA with 10 U/µg *Bgl*I (New England Biolabs, Beverly, MA) and 2.5 mM spermidine and the result was electrophoresed on 1% Tris-acetate (TAE; 0.04 M Tris-acetate, 1 mM EDTA) agarose gels. Gel preparation, DNA transfer to an Amersham Hybond-N+ nylon or Hybond XL membrane, crosslinking, probe-labeling and preassociation, prehybridization, hybridization, and washes were conducted as described previously (Lucas et al. 2001). The probe was created by PCR amplification of *Tom1l2* intron 10 sequence (5′ primer: CCCAAATGTGTCCACAGTTCTA, 3′ primer: GGGTGGAATGCACGACTAAA) followed by gel excision and band purification using the QiaQuick gel extraction kit (Qiagen) according to the manufacturer's instructions.

PCR genotyping The *Tom1l2*-specific primers (5′ primer: GGTGAAGCTAACAGCCATGGCC, 3′ primer: CCAGACAACAAATGCAGGTT) and insert-specific primers (5′ primer: GACGTCTCGTTGCTGCATAA, 3′ primer: GACCTGACCATGCAGAGGAT) were combined in each reaction. Amplification of 50–100 ng of template DNA was performed using the same cocktail and parameters as described above.

Tom1l2 expression analysis

Total RNA was isolated from 100 mg of brain, liver, or spleen tissues dissected from wild-type and *Tom1l2* gene-trapped mice using TRIzol reagent per the manufacturer's instructions. RNA was quantified using spectrophotometry, and 5 µg RNA was reverse transcribed by first-strand cDNA synthesis using SuperScript II RT (Invitrogen, Carlsbad, CA). Brain, liver, and spleen cDNAs from each mouse were PCR amplified using *Tom1l2* primers (5′ primer: CTTGGGGACAGAGAGTGTCAG, 3′ primer: CACAGGGCAAAGAGCGCATCGTC) and a standard curve was generated. An internal control PCR was performed with primers for the *Rai1* gene (5′ primer: TGTCAAGAACCTCGTGTCCA, 3′ primer: GGGAAACAGTCAAAAGCTGC), which is unaffected by the gene trap. Each PCR product (2 µl) was then loaded onto a 2% agarose gel and densitometry was performed using AlphaEaseFC software (Alpha Innotech, San Leandro, CA). A mean density for each tissue was calculated for *Tom1l2* relative to endogenous *Rai1*. Wild-type expression levels were then set to 1.0.

Animal husbandry

Mice were separated by sex and housed 2–4 per cage with access to a standard diet of Prolab® RMH 2000 chow (PMI Nutrition International, Inc., Brentwood, MO), and water *ad libitum*. The mouse-room was maintained on a 12-h/12-h light/dark cycle and all testing was done during the light phase. Ambient temperature (70°F) and humidity (40%) were maintained. Animals used in this study were cared for in accordance with the guidelines of the Institutional Animal Care and Use Committee of Virginia Commonwealth University and the *Guide for the Care and Use of Laboratory Animals* (National Research Council 1996). The mice were initially kept in wood chip bedding but were later switched to corncob bedding. Both bedding materials were sterilized prior to use.

Phenotypic assessment of *Tom1l2* mice

Assessments were conducted for 46 *Tom1l2* heterozygotes, 27 *Tom1l2* homozygotes, and 34 wild-type littermates. Mice of all genotypes were evaluated at 5, 10, 20, and 30 weeks. All animal studies were approved by the VCU IACUC committee. Physical assessment included whisker appearance and response, skin or fur abnormalities, condition of genital/rectal areas, nails and teeth, cage movement, and body measurements. The mice were weighed and body measurements, including distance between the bony landmarks of the craniofacial region and the limbs, were taken. In a fume hood, about 300 μl of isofluorane was used to anesthetize the mouse by inhalation. The following measurements (in mm) of the mouse were recorded: inner eye to inner eye, base of ear to base of ear, tip of nose to top of head, length of trunk from mandible, front limbs (toe to hip), and hind limbs (toe to thigh). The sizes of heterozygotes and homozygotes were compared to wild-type littermates.

General behavior assessments included wild running, freezing, sniffing, licking, rearing, defecation, urination, and movement around entire cage. *Gait* was assessed by footprint analysis (dipping each foot in ink and allowing the mouse to traverse across white paper under a tunnel). The average distance between adjacent paw prints and the average stride length were calculated for each mouse. In a *cage top-hang test*, the mouse was allowed to hold on to a cage-top which was then inverted onto the cage. The time (in sec) the mouse remained hanging was noted. Three trials were completed, and the mean was determined. All tests were modified from the behavioral battery described by Crawley and Paylor (1997).

Sensorimotor reflexes observed include *postural reflex* (ability to remain upright while the cage is shaking in all directions for 10 sec each), *righting reflex* (time to right itself after being turned onto its back), and the response to being picked up by the tail (the normal response is to raise the head, extend extremities and reach for the ground when lowered). *Whisker response* was measured by touching the whiskers of a freely moving mouse (normal mouse will stop moving its whiskers and will turn its head on the side of the whiskers that was touched). *Eye blink and ear twitch response* was elicited by touching the corner of the eye and the tip of the ear (pinna), respectively, with the tip of a cotton swab. *Sound orientation* (a brief sharp sound was made to the left and right sides of the mouse and response of the mouse, such as turning to the direction of the sound, flinching, startle response, or freezing, was noted) was analyzed as well. Mice were also assessed for pain sensitivity. In this assay, the mouse was placed on the hot plate set to 55°C. The time (in sec) for the mouse to jump, pick up a paw, paw lick, or paw shake is recorded (maximum duration is set at 30 sec). Hyperalgesia is indicated by an extreme sensitivity to the painful stimulus.

Statistical analyses

Statistical analyses were performed using GraphPad Prism version 4.02 for Windows (GraphPad Software, San Diego, CA). All data were analyzed using nonparametric unpaired *t* test and one-way (genotype) or two-way (genotype × age interval) analysis of variance (ANOVA). *Post hoc* comparisons were made using Bonferroni's multiple-comparison tests. Proportions of features in the genotypes were compared using Fisher's exact test. The level of significance was set at $P < 0.05$.

Histologic analysis

Spleens were isolated from *Tom1l2* mice and fixed in 4% paraformaldehyde/phosphate-buffered saline. Hematoxylin and eosin staining was performed at Michigan State University, Department of Pathology, according to standard procedures.

Hematologic evaluation

Blood was collected from the submandibular vein of mice of each genotype following standard protocols. Approximately 50 μl of blood were collected in a samplette micro whole-blood collector containing EDTA (Kendall Health Care, Mansfield, MA). The samples were kept at 4°C until analysis of hemoglobin, hematocrit, MCV (mean corpuscular volume), MCH (mean corpuscular hemoglobin), MCHC (mean corpuscular hemoglobin concentration), platelet, WBC (white blood cells), neutrophil, lymphocyte, monocyte, eosinophil, and basophil levels by the Comparative Pathology Lab (Davis, CA).

Flow cytometry

Monoclonal antibodies Cell-free culture supernatants from the following monoclonal antibody-producing B-cell hybridomas (American Type Culture Collection, Manassas, VA) were used for immunofluorescence staining: RA3-3A1 (anti-B220), GK1.5 (anti-CD4), and 53–6.72 (anti-CD8). Purified monoclonal anti-Mac-1 (M1/70), anti-Thy1.1 (OX-7), and fluoresceinated monoclonal anti-rat kappa light chain (MARK) antibodies were purchased from BD Biosciences Pharmingen (San Diego, CA). Biotinylated monoclonal anti-CD4 (GK1.5) antibody was purchased from eBioscience (San Diego, CA).

Immunofluorescence staining and flow cytometric analysis Single-cell suspensions of spleens, lymph nodes, and thymi were prepared, and cells were stained as previously described (Hartmann and McCoy 2004; Lewis et al. 1996). For one-color analysis, cells were incubated with monoclonal anti-B220, anti-CD4, anti-CD8, anti-Mac-1, or irrelevant anti-Thy-1.1 antibodies, followed by fluoresceinated MARK antibody. For two-color analysis, thymocytes were incubated with anti-CD8 or irrelevant anti-Thy1.1 antibodies followed by fluoresceinated MARK, and then the cells were incubated with biotinylated anti-CD4 antibody followed by phycoerythrin-conjugated streptavidin (Invitrogen, Carlsbad, CA). The fluorescence intensity of cells was measured with logarithmic amplification using a Becton-Dickinson FACScan (San Diego, CA) and analyzed by a G4 Macintosh computer. Data on 5,000 or 10,000 cells were collected, and forward-angle and side-scatter gates were set to exclude cell clumps and dead cells. To compare parameter measurements among the groups, values were analyzed by one-way ANOVA with a Dunnett's post-test. Significance was designated as $P < 0.05$.

Spleen IgM antibody-forming cell (AFC) response to sheep red blood cells

Using a modified hemolytic plaque assay of Jerne and Nordin (1963), the primary IgM response to sheep RBC was measured. Mice were sensitized with 7.5×10^7 sheep RBC (Lampire, Pipersville, PA) intravenously 4 days prior to sacrifice. On the day of sacrifice, spleen cell suspensions were prepared as previously described (Guo et al. 2001) and resuspended in RPMI 1640 (GibcoBRL, Grand Island, NY) medium with 10% fetal bovine serum (Hyclone, Logan, UT). Cell count was determined using a ZBII Coulter counter in the presence of ZAP-OGLOBIN II lytic reagent (Coulter Corporation, Miami, FL). An aliquot of cells was added to a test tube containing warm agar (46–48°C), guinea pig complement, and sheep RBC. The mixture was thoroughly mixed, plated in a petri dish, and covered with a microscope coverslip to incubate for 3 h at 37°C. Developed plaques were counted with a Bellco plaque viewer (Bellco Glass, Inc, Vineland, NJ). Experimental results are expressed as specific activity (AFC/10^6 spleen cells) and total spleen activity (AFC/spleen $\times 10^3$).

Anti-CD3 antibody-mediated T-cell proliferation

Proliferation of splenocytes in the presence of anti-CD3 antibody was performed as described by Guo et al. (2001). Briefly, single-spleen cell suspensions were prepared as above and resuspended in RPMI medium supplemented with 10% fetal bovine serum, sodium bicarbonate, HEPES, L-glutamine, gentamicin, and 2-mercaptoethanol. Splenocytes at 2×10^5/well were cultured in the microtiter wells coated with or without 1 μg/ml anti-CD3 antibody (PharMingen, San Jose, CA) at 37°C at 5% CO_2 and 95% humidity. Prior to harvest on day 3, cells were pulsed with ^3H-thymidine for 18–24 h. Incorporation of ^3H-thymidine into the proliferating cells was used as the endpoint of the assay, and the data were expressed as CPM/2×10^5 cells.

Natural killer cell activity

Activity of natural killer cells was assayed using the natural killer-sensitive target, $Na^{51}CrO_4$-labeled YAC-1 cells, as described by Guo et al. (2001). Briefly, the splenocytes at different dilutions were mixed with the target cells to obtain effector:target ratios of 200:1, 100:1, 50:1, 25:1, 12.5:1, and 6.25:1. Spontaneous release was determined by adding medium to replicate cultures containing the target cells. Maximum release was determined by adding 0.1% Triton X-100 to YAC-1 cells. Natural killer cell-specific lysis (%) of ^{51}Cr-labeled YAC-1 cells was used as the endpoint of the assay. Activity was calculated by the formula: (experimental release minus spontaneous release) / (maximum release minus spontaneous release) × 100%.

Bioinformatic analysis

Pairwise protein sequence alignments of TOM1l2 (Q6ZVM7), Tom1l2 (Q5SRX1), Tom1 (O88746), and Tom1l1 (Q923U0) were performed using the BioEdit sequence alignment editor, version 5.0.9 (Hall 1999), utilizing the ClustalW multiple alignment accessory application and the "full multiple alignment" setting. Following alignment similarity and identity was calculated using the BLOSUM62 matrix. The protein sequences were obtained from the Expert Protein Alignment System

(ExPASy) proteomics server of the Swiss Institute of Bioinformatics (SIB).

Creation of *Tom1l2*-GFP expressing cells

Construct preparation Tom1l2 was PCR amplified from mouse genomic DNA (5′ primer: ATATAATTTGCGGCCGCCCCCTTCACCatggagttcctgctggggaat) (3′ primer: ATATAAGGCGCGCCCACCCTTcagggcaaagagcgcatcgtc) and cloned into pBluescript. *Tom1l2* was directionally cloned into pENTR (Invitrogen, Carlsbad, CA) with *Not*1 and *Asc*1 (New England Biolabs, Beverly, MA) sites. *pENTR-Tom1l2* was then combined with pcDNA-DEST47 for an in-frame fusion with GFP, and an LRclonase reaction was performed according to the Gateway system instructions (Invitrogen). Digestions, ligation, transformations, and verification of all clones were conducted via standard molecular biology techniques.

Transient transfection and cellular localization of Tom1l2 COS-7 monkey kidney fibroblasts were transfected with pDEST47-Tom1l2 or a construct expressing GFP alone by Lipofectamine 2000 (Invitrogen). One day before transfection, cells were seeded on glass coverslips in DMEM medium at 2×10^5 cells/well in 6-well plates. *Tom1l2*-GFP plasmid DNA was complexed in serum-free media with Lipofectamine. After 20 min the complexes were added to COS-7 cells and cultured for 48 h. Transfected cells were stained with BODIPY TR C_5-ceramide (Molecular Probes, Eugene, OR) for visualization of the Golgi complex. Coverslips were mounted onto slides using Vectashield with DAPI (Vector Laboratories, Burlingame, CA) and examined under a Zeiss fluorescence microscope.

Molecular modeling

Homology modeling of the VHS and GAT domains of the mouse Tom1l2 protein was performed using the automodel feature of the MODELLER program (version 8.1) (Fiser and Sali 2003; Sali and Blundell 1993). An unambiguous primary protein sequence alignment of the VHS and GAT regions of Tom1 and Tom1l2 was achieved with ClustalX version 1.83 using the default parameter settings (Chenna et al. 2003). Three-dimensional coordinates for the homology models were based on the high-resolution X-ray crystal structures of the VHS (Misra et al. 2000) and GAT (Akutsu et al. 2005) domains of the human TOM1 protein. Subsequent structure optimization involved energy minimization of the protein domains using the Tripos Force Field in the SYBYL v7.3 (Tripos, Inc., St. Louis, MO) molecular modeling package. In the energy-minimization procedure, Gasteiger-Hückel charges were calculated and a distance-dependent dielectric constant ($D = 4.0$; cut-off = 8.0 Å) was employed. The stereochemical integrity of the modeled domains was assessed using the PROCHECK utility (Laskowski 1993) and the ProTable facility within SYBYL. Figures depicting the two domains were prepared using SYBYL. Figure 8A, C, and D were prepared using SYBYL. Figure 8B was prepared using SPOCK v1.0b170 (Christopher 1998). The molecular surface depicted in Fig. 8B was colored by electrostatic surface potential with saturating colors at 20 kT (blue) and −20 kT (red) and a salt concentration of 150 mM. The electrostatic charges used in the calculation of the electrostatic surface were also calculated using SPOCK.

Results

Generation of *Tom1l2* gene-trapped mice

To assess the function of the *Tom1l2* gene, we created *Tom1l2*-deficient mice by gene-trapping (Gossler et al. 1989; Skarnes et al. 1995). The XG909 cell line (BayGenomics Consortium) (Stryke et al. 2003) carries a gene-trapping vector inserted between exons 10 and 11 of *Tom1l2*. The XG909 cell line was created from 129/OlaHsd embryonic stem cells, expanded, and injected into C57Bl/6J blastocysts for creation of chimeric mice. Chimeric XG909 males were mated with C57Bl/6J females to produce an F_1 line. *Tom1l2* heterozygote males and females from the F_1 generation (selected by coat color and verified by Southern analysis) were bred to obtain an F_2 generation. *Tom1l2* mouse breeding generated heterozygous, homozygous, and wild-type pups in expected Mendelian ratios. To identify the exact location of the vector insertion, a DNA fragment was PCR amplified from a homozygous mouse (see Methods) using a 5′ primer specific to the gene-trapping vector and a 3′ primer specific to the endogenous *Tom1l2* gene. Sequencing of this product localized the insert to intron 10, 1978 bases downstream from the end of exon 10 (Fig. 1A). Further genotyping was conducted by PCR, and wild-type, heterozygous, and homozygous mice were easily distinguished (Fig. 1B).

To determine whether these gene-trapped mice had reduced levels of *Tom1l2* expression, a semiquantitative RT-PCR was performed using primers designed to amplify a region downstream of the gene-trap insertion. Homozygous *Tom1l2* mouse brain revealed an approximately 80% reduction in *Tom1l2* expression compared with that of sibling wild-type littermates (Fig. 1C). Similarly, an ~70% and ~50% decrease in *Tom1l2* expression was observed in liver and spleen, respectively, of homozygous targeted animals as compared with wild-type mice (Fig. 1C). These results, similar to other those of published reports of gene-trapped mice (Stanford et al. 2001; Voss et al. 1998), indicate that the presence of the gene-trapping

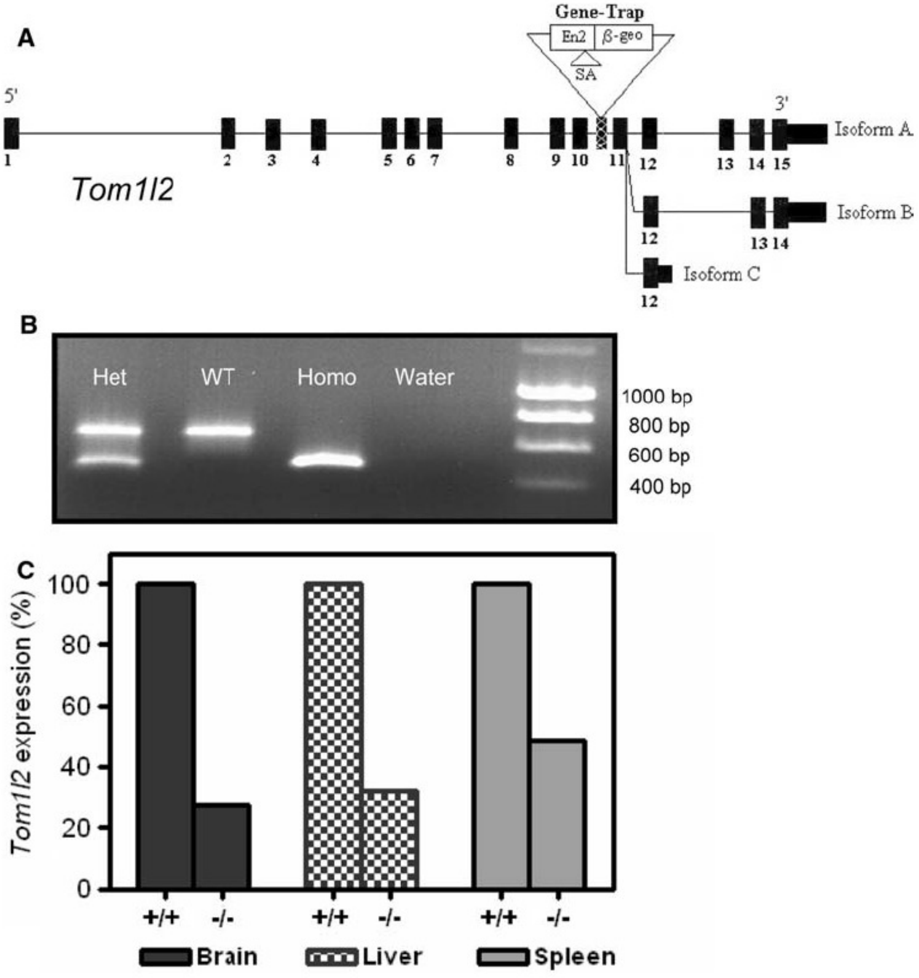


Fig. 1 *Tom1l2* gene-trapped mice. (**A**) *Tom1l2* genomic structure and gene trap. All isoforms (a, b, and c) of *Tom1l2* are shown. The gene trap is located in intron 10, as shown. The vector splice acceptor site (SA) is indicated by the arrow. Note that insertion of the *En2* gene and ß-geo interrupts all isoforms of *Tom1l2*. (**B**) *Tom1l2* PCR genotyping. The top band indicates the presence of the wild-type allele, while the appearance of the lower band indicates the presence of the gene-trapped allele. (**C**) RT-PCR of *Tom1l2* gene-trapped mice. Analysis of *Tom1l2* mRNA expression in brain, liver, and spleen of wild-type (+/+) and homozygous (−/−) *Tom1l2* gene-trapped mice indicates that the *Tom1l2−/−* mice exhibit reduced expression as compared to wild-type

vector in the position described above was sufficient to reduce *Tom1l2* levels significantly, resulting in *Tom1l2* hypomorphic mice. The term hypomorph, for the purposes of this study, is used to signify that mice, homozygous for the insertion, produce reduced levels of *Tom1l2* mRNA. A true knockout would have no detectable levels of expression. Heterozygous mice have one wild-type copy of *Tom1l2* and one copy of the gene with the insertion and should not be considered hypomorphs.

Anxiety-related behaviors in *Tom1l2* hypomorphs

We evaluated craniofacial, skeletal, neurologic, and behavioral aspects of the *Tom1l2* mice. Phenotypic assessments were based on a comprehensive set of protocols adapted from the SHIRPA, the Irwin battery, and the test battery described by Crawley and Paylor (1997; Irwin 1968; Rogers et al. 1997). Preliminary assessments were performed to delineate abnormal characteristics, including health, physical status, fecundity, and observable neurobehavioral features. Additional tests were employed to assess specific behavioral measures in all *Tom1l2* mice genotypes.

General behavioral evaluations showed significantly increased freezing response in the *Tom1l2* mice compared to wild-type mice (Fisher's exact test, $P < 0.05$; Table 1). Furthermore, a higher proportion of *Tom1l2* heterozygotes exhibited wild running (Table 1). To understand if the freezing behavior was related to fear or anxiety, open-field tests were performed. Heterozygous *Tom1l2* mice showed increased rearing in the open field compared with the wild-type animals, indicating fear- or anxiety-related behavior (Crawley and Paylor 1997). Physical assessments, including body weights, body lengths, and craniofacial measurements, were not significantly different among the genotypes (Supplementary Fig. 1 and data not shown). Other behavioral evaluations did not reveal any statistically significant effects due to reduced *Tom1l2* expression (Table 1).

Table 1 Phenoytpic assessment of *Tom1l2* hypomorphic mice

	Tom1l2+/+	*Tom1l2+/−*	*Tom1l2−/−*
Number of mice evaluated	34	46	27
Physical assessment and simple reflexes (% abnormal)			
Whisker appearance	11.7	8.6	14.8
Bald patches (%)	8.8	17.4	22.2*
Condition of teeth[a]	3	2.1	7.4
Righting reflex	0	0	0
Sound orientation	0	0	0
Whisker response	0	0	0
Eye blink/ear twitch	0	0	0
Pupil constriction/dilation	0	0	0
General observational measurement (%)			
Wild running	8.6	17.6*	3.7
Freezing	2.1	17.6*	14.8*
Sniffing	100	100	100
Licking	100	100	100
Rearing			
5 weeks	8.8	17.3	11.1
10 weeks	5.8	15.2*	7.4
20 weeks	8.8	15.2*	7.4
30 weeks	8.8	13*	0
Defecation	100	100	100
Urination	100	100	100
Stereotypies	0	0	0
Sensorimotor reflexes and strength (% abnormal)			
Postural reflex	0	0	0
Cage-top hang test (latency in sec ± SEM)			
Male (5 weeks)	42.8 ± 4.9	47.3 ± 3.5	42.4 ± 5.2
Female (5 weeks)	52.1 ± 3	56.4 ± 1.3	55.4 ± 3.8
Male (10 weeks)	26.7 ± 4.3	32.3 ± 5.1	34 ± 6.1
Female (10 weeks)	51.63 ± 4.8	48.9 ± 4	52.1 ± 2.9
Response to being picked up by tail	0	0	0
Hot-plate test (latency in sec ± SEM)			
Male	4.3 ± 0.03	3.7 ± 0.73	3.8 ± 0.36
Female	3.11 ± 0.39	3.46 ± 0.03	3.3 ± 1

* Significantly different from the wild-type ($P < 0.05$)

[a] Malocclusion of teeth

Tom1l2 gene-trapped mice exhibit increased incidence of infections and tumors

Physical assessments of *Tom1l2* homozygous, *Tom1l2* heterozygous, and wild-type littermates showed gross inflammatory lesions of the skin, including hair loss (bald patches), dermatitis, and skin ulcerations, in the *Tom1l2* mice compared with the wild-type mice (Tables 1 and 2). Homozygous mice showed more severe lesions than the heterozygotes, and wild-type mice rarely showed any abnormalities (Table 1). Initially, all mice were housed in woodchip bedding, which, with the benefit of retrospection, may have promoted the initiation and contributed to the severity of the infections observed. Eye manifestations observed in *Tom1l2* mice included conjunctivitis, corneal ulcers, exophthalmos, and ocular asymmetry (Table 2). Later, the mice were transferred to cages with corncob bedding and the infections were not only less severe but were less frequent in the *Tom1l2* mice. Infections did not occur in wild-type animals from the same cage. Serologic investigations, including swab culture from sites of skin and eye infections, revealed the presence of *Staphylococcus aureus*, a common skin pathogen. Other infectious pathology included bronchopneumonia (2% incidence) and murine mycoplasmosis (1%). Overall, the incidence of skin/eye infections in *Tom1l2* mice was 15–20%. Although individual mice did not develop the entire array of observed health problems, *Tom1l2* mice at least 9 months

Table 2 Pathologies identified during gross organ and phenotypic assessment

Tom1l2+/−	*Tom1l2−/−*
Splenomegaly	Splenomegaly
Skin ulcers	Skin ulcers
Eye infections	Eye infections
Kyphosis	Malocclusion of teeth
Exophthalmos	Bladder hemorrhage
Ocular assymetry	Hemorrhagic mass near spleen
Corneal ulcer	Kyphosis
Excess fat in the abdomen	Hydrocephaly
Hydrocephaly	Pleural effusion
Emphysema of lung	Tumors
Tumors	Skin
Skin	Small, cystic kidneys
Soft tissue mass in vulva	Liver polyps
Lachrymal gland enlargement	Large ovarian mass
Lung tumor	Mammary gland tumor
Ovarian cyst	Kidney tumor
Dark mass on prostate	Bony tumor in the hind limb
Kidney tumors	Lung tumor

old manifested an array of complications typically resulting in euthanasia, including hemorrhagic bladder, hemorrhagic masses adjacent to spleen and ovary, and thoracic hemorrhage. Furthermore, a variety of tumors were observed in ~20% of *Tom1l2* mice. Tumors were not limited to any particular organ and developed within the lung, ovary, kidney, liver, prostate, abdomen, thorax, and skin (Table 2). The types of tumor ranged from primary and benign adenomas (pulmonary/brochoalveolar, renal) to noninvasive or invasive neoplasms (basal cell carcinoma of the skin, liver, and renal adenocarcinoma, lymphoid hyperplasia, and tumors of the mammary gland, ovaries, and prostate) and metastatic neoplasia (renal adenocarcinoma with secondaries in the liver and lymph nodes). In addition, 3% of the *Tom1l2* mice developed excess fat deposition in the abdomen. None of these pathologies were observed in the wild-type littermates.

Splenomegaly in *Tom1l2* hypomorphs

Gross evaluation of spleens from homozygous mice older than 1 year revealed splenomegaly with weights from 1.5 to 2 times normal ($P < 0.01$). In one evaluation, four of six older *Tom1l2* mice had significantly larger, darker, abnormally shaped spleens. Homozygous mice showed an expanded splenic white pulp (Fig. 2A). Within the white pulp, both follicles (B-cell-enriched regions) and periarteriolar lymphoid sheath (T-cell-enriched regions) had increased cellularity. However, no germinal centers were present. There was no indication of erythropoiesis within the red pulp. Even apparently healthy homozygous mice from 12 to 16 weeks of age had significant splenomegaly (Fig. 2B). Body lengths and weights were assessed at 5, 10, 20 and 30 weeks of age and were not different among the genotypes (Supplementary Fig. 1). Therefore, differences in splenic weights are not due to differences in body size.

Because the white pulp is part of the immune system, we hypothesized that splenomegaly and the previously mentioned health issues were related. To further examine this possibility, a complete peripheral blood analysis was conducted on healthy mice of each genotype and sex (15-19 months old) (Fig. 3A). Although not statistically different, the total white blood cell (leukocyte) count in homozygous mice was decreased, with percentages of lymphocytes and neutrophils slightly lower than in wild-type mice, and the percentages of monocytes, eosinophils, and basophils were increased compared with those for wild-type mice (Fig. 3B). Homozygous mice had significantly lower platelet counts than wild-type mice (Fig. 3C), while the red blood cell (RBC) counts for all mice were normal (Fig. 3D). Other parameters concerning erythroid

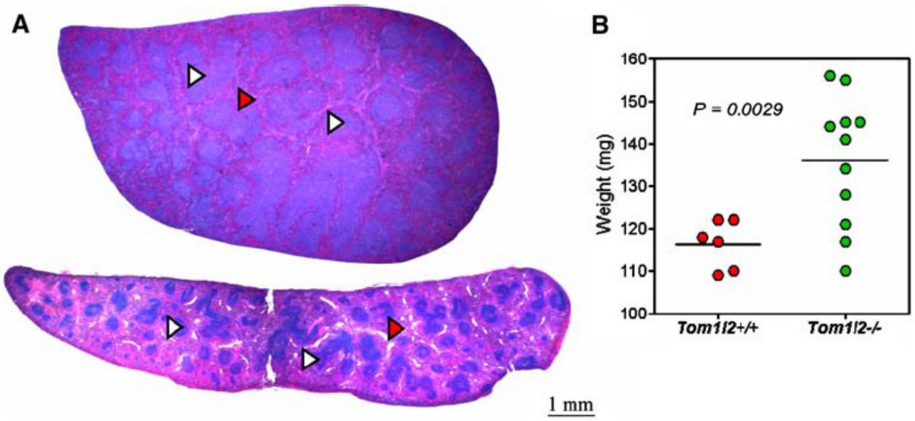


Fig. 2 Enlarged spleens in *Tom1l2* mice. (**A**) Histologic tissue sections of spleen from a *Tom1l2* hypomorph (homozygous for the insertion), illustrating splenomegaly and enlarged white pulp (upper), and wild-type mouse (lower). Red (pink arrowheads) and white (white arrowheads) pulp are indicated. Scale bar = 1 mm. (**B**) Spleen weights for *Tom1l2+/+* and *Tom1l2−/−* mice from 12 to 16 weeks of age. Mice from both sexes were evaluated

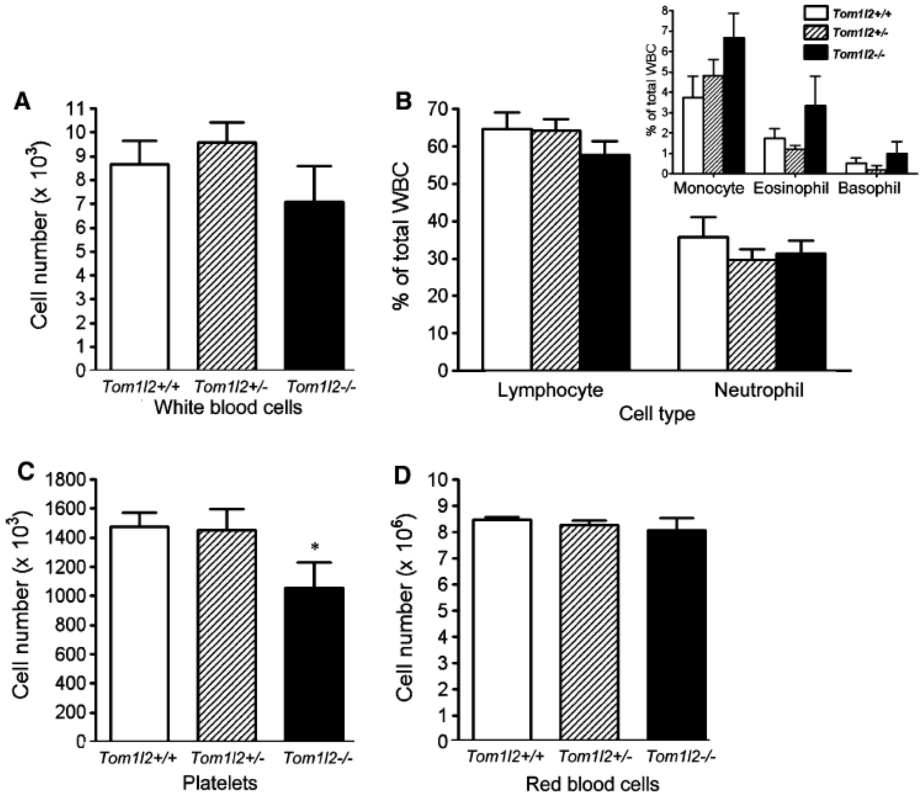


Fig. 3 Hematologic analyses of *Tom1l2* mice. (**A**) Total white blood cell count. (**B**) Percentages of lymphocytes and neutrophils. Inset: Percentages of monocytes, eosinophils, and basophils. (**C**) Platelet counts. * $P < 0.05$. (**D**) Red blood cell counts. Histogram represents *Tom1l2+/+* (white bars), *Tom1l2+/−* (hatched bars), and *Tom1l2−/−* (black bars) mice

cells were also evaluated. Hemoglobin, hematocrit, mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), and mean corpuscular hemoglobin concentration (MCHC) did not show any statistically significant differences among the genotypes (Supplementary Table 1). Thus, although homozygous mice had reduced platelet counts, they were not anemic. Overall, the decreased level of *Tom1l2* expression had no major effect on the erythroid system based on the examined parameters despite the darker color of the spleens in older mice.

Analyses of the cellular composition of major lymphoid organs, including spleen, lymph nodes, and thymus, were performed on mice that showed no signs of infection. For this purpose, young (7 weeks of age) and older (6 months of age) mice of each genotype were examined. A trend toward an increased number of nucleated cells (leukocytes) in the spleen was observed in homozygous mice for both age groups (Table 3). This increase in leukocytes likely explains the enlarged spleens in homozygous mice. By contrast, cellular counts for thymi for young mice were not different among the genotypes (Table 3). For evaluation of immune cell subpopulations, immunofluorescence staining and flow cytometric analysis were performed. Cells from the spleen, lymph nodes, and thymus were stained with monoclonal antibodies to detect CD4$^+$ (helper T-cell marker), CD8$^+$ (cytotoxic T marker), B220$^+$ (B-cell marker), and Mac-1$^+$ (macrophage marker) molecules. The various thymocyte subpopulations, including CD4 single-positive, CD8 single-positive, CD4-CD8 double-positive, and CD4-CD8 double-negative cells, were virtually identical in all groups of young mice (data not shown), suggesting that thymic maturation was normal.

Table 3 Number of nucleated cells (× 10^6) in spleen and thymus in *Tom1l2* mice

Age and genotype	Nucleated cells (× 10^6)	
	Spleen	Thymus
7 weeks		
Tom1l2+/+	91 ± 55	126 ± 50
Tom1l2+/−	80 ± 36	130 ± 82
Tom1l2−/−	114 ± 76	107 ± 60
6 months		
Tom1l2+/+	107 ± 23	ND
Tom1l2+/−	123 ± 54	ND
Tom1l2−/−	124 ± 32	ND

Spleens from 6-month-old hypomorphic mice showed a significant increase in the percentages of CD4$^+$ helper T cells ($P = 0.05$) and CD8$^+$ cytotoxic T cells ($P < 0.05$), while the percentage of macrophages was lower compared to the wild-type (Fig. 4B–D, Supplementary Table 2).

Although the difference was not statistically significant, B-cell levels tended to be elevated in spleens from both heterozygous and homozygous older mice (Fig. 4A, Supplementary Table 2). These results suggest that leukocyte subpopulations became skewed in older homozygous mice, while the relative proportion of the four cell types examined did not differ in the spleens of young mice (Supplementary Table 2). Lymph nodes of *Tom1l2* mice were not enlarged (data not shown) but had dramatically higher percentages of B cells and T cells in older mice and B cells in younger mice (Supplementary Table 2). The percentages of B cells, T cells, and macrophages in the lymph nodes from both young and old hypomorphic mice were highly variable compared to wild-type, which may suggest incomplete penetrance (Supplementary Table 2).

Reduced humoral immune response in *Tom1l2* hypomorphs

To assess immune competency of *Tom1l2* gene-trapped mice, we analyzed the ability of splenic cells to generate immune responses, which encompassed humoral immunity, pan-T-cell response, and innate immunity. First, the IgM antibody-forming cell (AFC) response to T-dependent antigen sheep RBC was examined to evaluate humoral immunity (Jerne and Nordin 1963) (Fig. 5A). In homozygous mice, the specific activity (AFC/10^6 spleen cells) was decreased by ~44% ($P < 0.05$, $t = 2.4$; $df = 8$) and the total spleen activity (AFC/spleen) was decreased by ~45% ($P < 0.05$, $t = 2.39$; $df = 8$) compared with the controls (Fig. 5A, B). A reduction in the AFC response has been implicated in decreased host resistance to infectious agents in mice (Luster et al. 1988). However, the cells required for this humoral response were not decreased in *Tom1l2* homozygous mice. Generalized immune responsiveness of T cells was measured by cell proliferation induced with anti-CD3 antibody that crosslinks the T-cell receptor complex regardless of antigen specificity. The proliferative response of T cells was not statistically different between homozygous and wild-type mice (Fig. 5C). This result is surprising considering that the homozygous mice had 1.6-fold more T cells. Because a corresponding increase in the CD3 assay was not seen despite the higher levels of T cells in the homozygous mice, these data

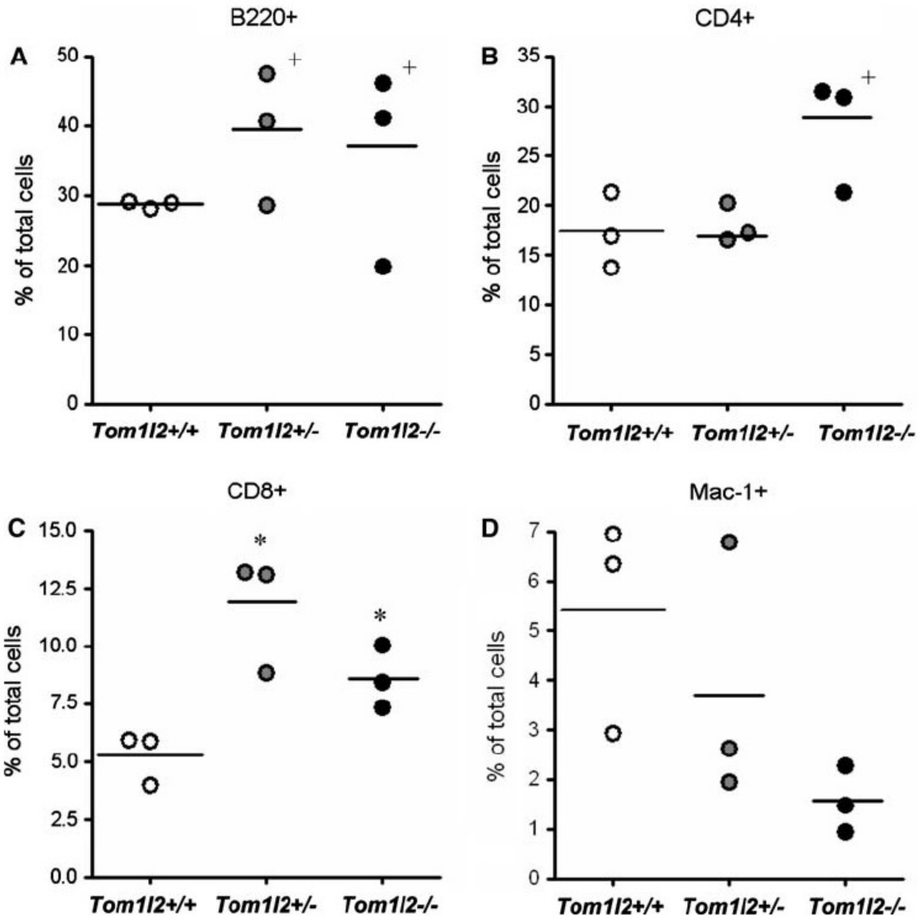


Fig. 4 Altered splenocyte subpopulations in 6-month-old *Tom1l2* hypomorphs. (**A**) $B220^+$ B cells. (**B**) $CD4^+$ helper T cells. (**C**) $CD8^+$ cytotoxic T cell. (**D**) Mac-1^+ macrophages. * $P < 0.05$. + = In those samples that were fully penetrant, the differences observed were significant; $P < 0.05$

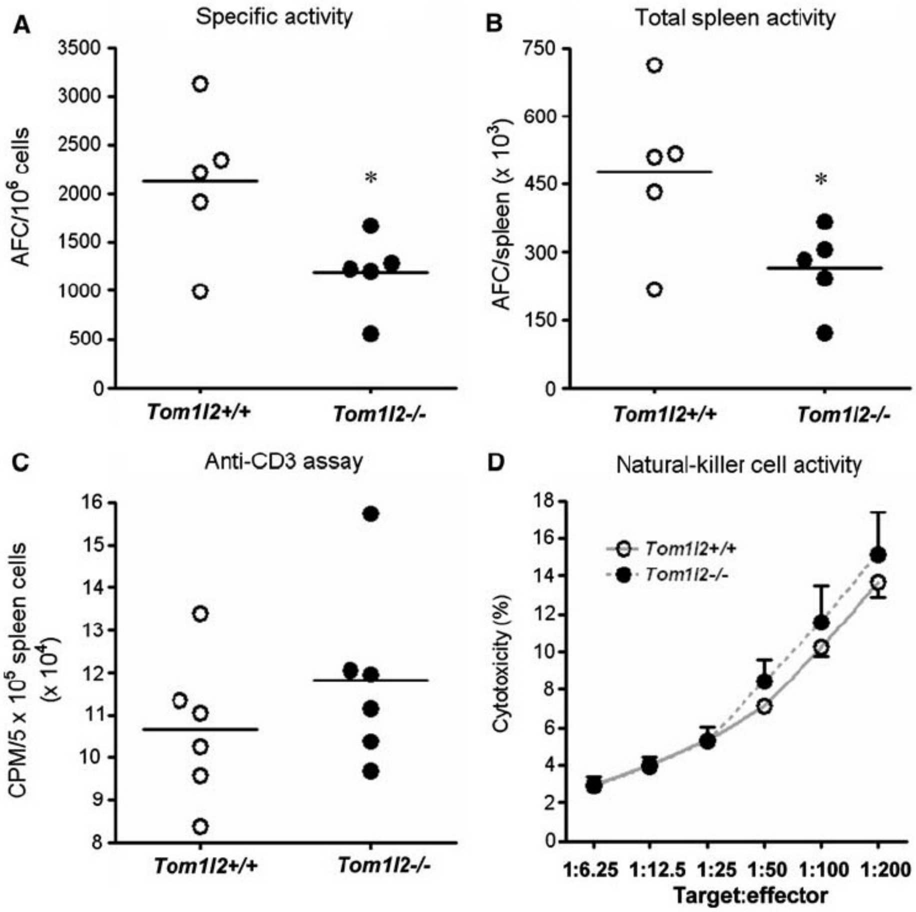


Fig. 5 Impaired adaptive immune responses by spleen cells from *Tom1l2* mice. IgM AFC response to the T-dependent sheep RBC antigen expressed as (**A**) specific splenic activity and (**B**) total splenic activity. (**C**) Pan-T-cell proliferative response to anti-CD3 antibody. (**D**) Natural killer cytotoxicity. * $P < 0.05$

confirm that there is a T-cell immune defect in the *Tom1l2* mice. Finally, innate immune responsiveness was evaluated by measuring natural killer cell cytotoxicity using ^{51}Cr-labeled YAC-1 cells as the target at various target:effector ratios (Fig. 5D). Although increased cytotoxicity was observed at higher ratios with *Tom1l2*-deficient cells, no statistically significant difference was observed (Fig. 5D).

Tom1l2 belongs to Tom1 family of proteins and Tom1l2 localizes to the trans-Golgi network

Bioinformatic studies were performed using the BioEdit pairwise sequence alignment editor. Tom1l2 has 60.8% amino acid identity (74.2% similarity) to Tom1 and 26.6% identity (47.3% similarity) to Tom1l1 (Fig. 6, Supplementary Table 3). The mouse Tom1l2 protein is 94.4% identical (97.8% similarity) to the human TOM1L2. Thus, a high level of similarity with conservation of functional domains exists among the Tom1 family of proteins, indicating a common role in cellular biochemical processes.

A GFP-*Tom1l2* fusion construct (pDEST47-*Tom1l2*) was transfected into COS-7 monkey kidney fibroblasts. Forty-eight hours after transfection, the Tom1l2-GFP fusion protein colocalized with Bodipy TR Ceramide, which stains the Golgi apparatus (Fig. 7). Thus, the Tom1l2-GFP fusion protein colocalizes with Golgi compartments and vesicles, providing supporting evidence that Tom1l2 is involved in trafficking and vesicular transport.

Molecular modeling was then performed based on protein homology using an unambiguous sequence alignment of the VHS and GAT domains of Tom1 and Tom1l2. Three-dimensional coordinates based on known high-resolution X-ray structures of VHS and GAT domains (Akutsu et al. 2005; Misra et al. 2000) show structural similarity of these functional domains between Tom1 and Tom1l2 (Fig. 8). While the GAT domains of TOM1 and TOM1L1 interact with both ubiquitin and TOLLIP, interaction of ubiquitin with TOM1L2 has not been shown (Katoh et al. 2004). Our molecular model suggests that Tom1l2 likely provides cognate sites, similar to other members of Tom1 family of proteins, for interactions with ubiquitin. Because a cellular trafficking role for Tom1 is

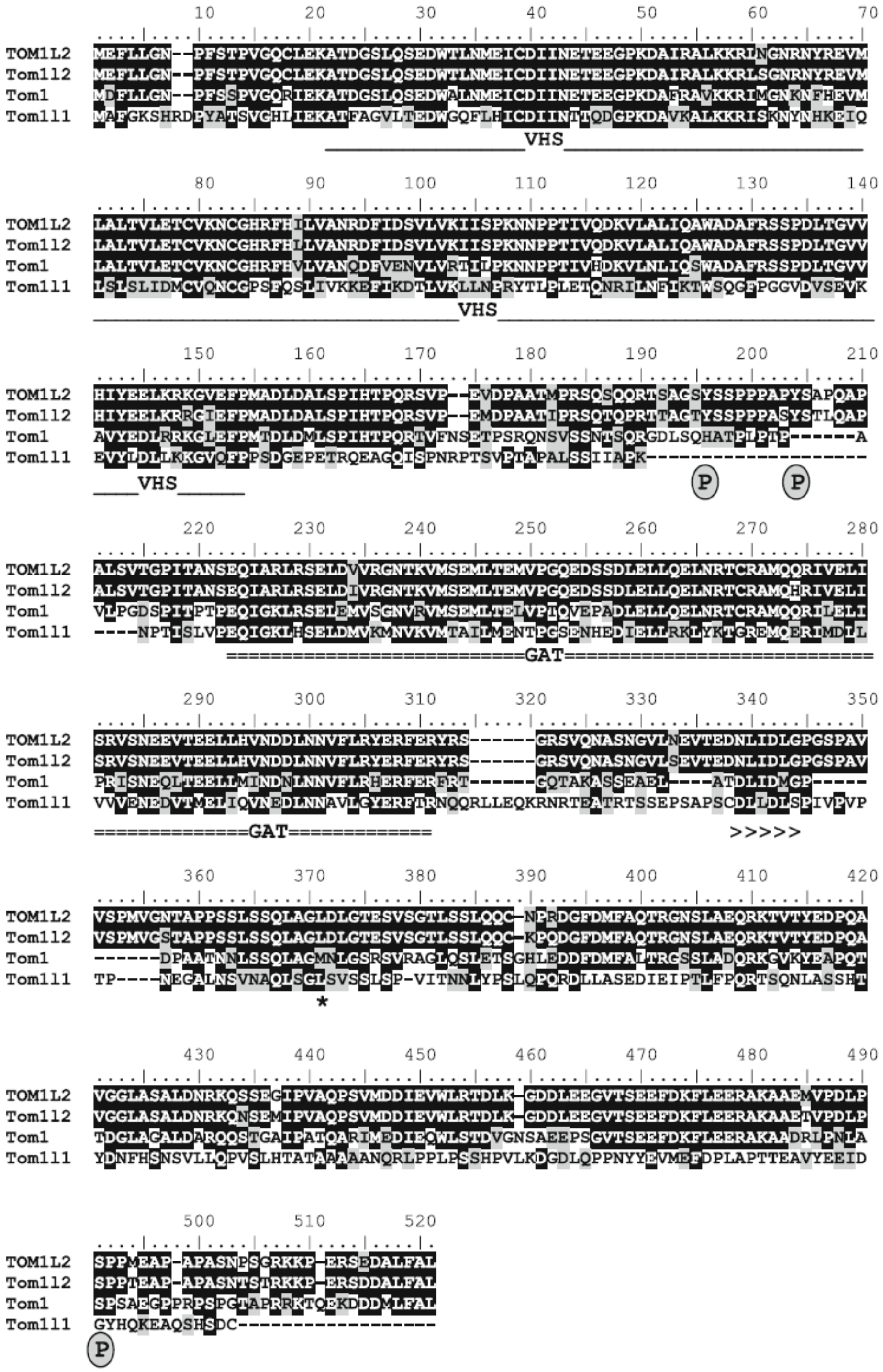


Fig. 6 Tom1l2 belongs to the Tom1 family of proteins. Human TOM1L2 and murine Tom1l2, Tom1, and Tom1l1 are shown aligned in decreasing homology. The VHS (solid line) and GAT (dashed line) domains are indicated. The clathrin box (indicated by >>>>>) is also shown. Potential tyrosine and serine phosphorylation sites are indicated by P. The gene-trap vector (indicated by *) is inserted into *Tom1l2* intron 10, as shown, and affects all splice variants

already established (Lohi and Lehto 1998), our homology model of Tom1l2 provides structural evidence for a role in trafficking for Tom1l2 in the cell as well and provides supporting evidence for Tom1l2 interactions with ubiquitin.

Discussion

Tom1l2 gene-trapped mice were created to determine the biological significance of this protein. The vector insertion site, located in intron 10, affects all known splice variants

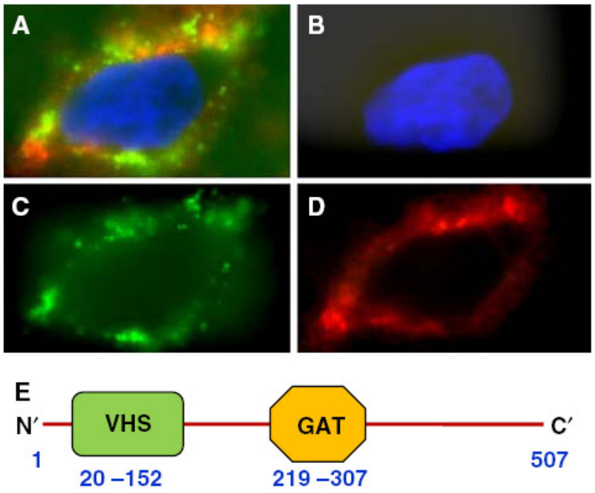


Fig. 7 Cellular localization of the Tom1l2 protein. Cells were imaged in live-cell format at 63×. (**A**) Tom1l2-GFP fusion protein expresses and localizes to Golgi compartment (*yellow*). (**B**) Nuclear staining (*blue*) with DAPI. (**C**) GFP (*green*) viewed under FITC filter. (**D**) Golgi apparatus stained with 2.5 μM fluorescent ceramide (*red*) and viewed under Texas Red filter. (**E**) Schematic of the Tom1l2 protein with location of the VHS and GAT domains, with corresponding amino acid numbers (*blue*)

of Tom1l2 (isoforms a, b, and c). In concert with many other reported studies on gene-trapped mouse models (Serafini et al. 1996; Skarnes et al. 1992; Yeo et al. 1997), these *Tom1l2* mice are not total knockouts but instead are hypomorphs that show ~30% residual *Tom1l2* expression. A concern for any gene-trapping model is that the trapped gene may still function and have significant levels of normal protein being produced due to leakiness of the gene-trapped vector. This is the result of the cell transcription machinery effectively "skipping" or "reading over" the vector insert in the intron, thus allowing production of some normally functioning protein. Although a consistent difference between the genotypes was observed, extensive phenotypic variability in the hypomorphs was revealed in most of the immunologic studies. This variability could be due to the genetic background, modifier genes altering the phenotypic consequence, or incomplete penetrance of the immune system phenotype, perhaps requiring a "trigger" to induce the immunologic effects. The penetrance of individual phenotypes, including splenomegaly, skin and eye infections, and tumors, varied between 15% and 20% in *Tom1l2* mice; thus, overall penetrance for any aspect of the global immune phenotype is approximately 50–60%.

The most prominent phenotype of the *Tom1l2* hypomorphs was the high frequency of infections and tumors found in these mice. To further understand this phenotype, we performed gross organ dissections, histologic analysis, and immunologic studies. Gross organ observations showed significant splenomegaly, further supported histologically by increased white pulp content. This was corroborated by an increase in $CD4^+$ and $CD8^+$ T cells and B cells in the spleen. While a corresponding increase in anti-CD3 proliferating T cells was expected in studies on humoral immune response, no significant difference was seen in the CD3 assay. The antibody-forming cell (AFC) assay is a sensitive indicator of the host's ability to mount an antibody response to a specific antigen (Luster et al. 1988). When T-dependent antigen sRBC are employed as a trigger, the IgM antibody response requires a coordinated interaction of T cells, B cells, and antigen-presenting cells such as macrophages and/or dendritic cells (Luster et al. 1988; Yang et al. 2003). Any alteration in antigen processing or presentation, cytokine production and release by T cells, and proliferation and/or differentiation of T cells and B cells could affect the final immunologic response (Yang et al. 2003). A decrease in splenic IgM AFC response was observed for *Tom1l2* homozygous mice compared to wild-type littermates. Taken together, these results suggest that in spite of increased T-cell population in the spleens, leading to enlarged spleens, the number of functional T cells is low, indicating a defect in the T-cell immune response.

Functional genomic analysis suggests that *TOM1L2* is likely involved in trafficking, which is a critical and important part of normal cellular function. Trafficking has been shown to be important for a variety of developmental disorders, including Chediak-Higashi syndrome, an autosomal recessive disease characterized by variable degrees of oculocutaneous albinism, recurrent infections, and a mild bleeding tendency due to mutations in *CHS1* (*LYST*), a protein involved in the regulation of lysosomal trafficking (Zarzour et al. 2005). Our cellular localization studies showed that Tom1l2 is localized in the trans-Golgi-network and in cytoplasmic vesicles, consistent with the presence of a VHS domain in the N-terminus of TOM1L2/Tom1l2 (Lohi and Lehto 1998). Our homology-based molecular model not only confirms the structural similarity of VHS and GAT domains of Tom1l2 to other Tom1 family trafficking proteins, but it also supports a similar functional role for Tom1l2.

A review of the literature related to *TOLLIP*, *TOM1*, and *TOM1L1*, including *in vitro* and *in vivo* studies, shows that the phenotypic features in *Tom1l2* hypomorphs could be due to disruption of several interacting proteins and pathways. It is noteworthy that TOLLIP, which interacts with the GAT domain of TOM1L2 (Katoh et al. 2006), also interacts with the IL-1R (Katoh et al. 2004), a master regulator of inflammation and innate immunity. Furthermore, TOM1, which is functionally and structurally similar to TOM1L1 and TOM1L2, also interacts with IL-1R (Brissoni et al. 2006). Interestingly, Tom1l1 interacts with Fyn (Seykora et al. 2002), and mice with mutations in *Fyn*

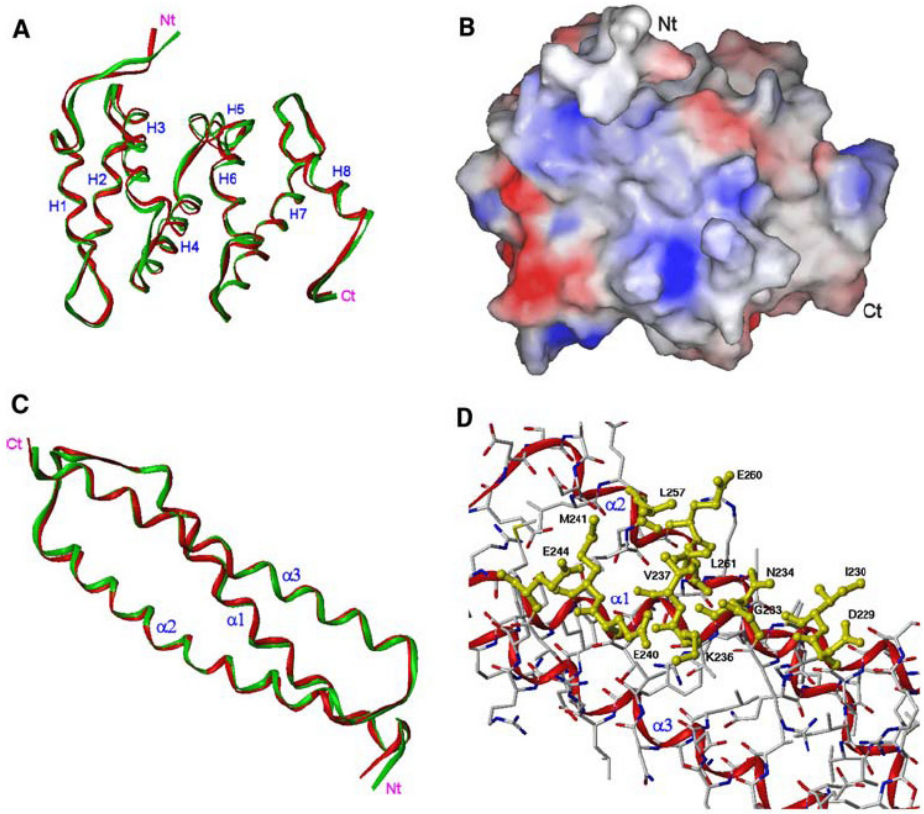


Fig. 8 Molecular modeling of Tom1l2 protein. (**A**) Ribbon diagrams depicting the secondary and tertiary structure of the VHS domains of TOM1 (green; residues Met1-Met153) crystal structure and Tom1l2 (red; residues Met1-Met153) homology model. Helix 1 through Helix 8 (H1-H8) and the N- and C-termini are labeled. The root-mean-square distance (RMSD) between the alpha carbon (C^{α}) atoms of TOM1 VHS and those in the corresponding Tom1l2 VHS model is 0.73 Å. (**B**) Molecular solvent-accessible surface of the Tom1l2 VHS domain. The surface is color-coded with respect to electrostatic surface potential, with acidic regions indicated in red and basic regions in blue. A pocket lined with basic residues can be seen, a feature also found in the VHS domains of other proteins (Misra et al. 2000). (**C**) Ribbon diagrams depicting the higher-order structure of the GAT domains of TOM1 (green; residues Glu215-Gly307) crystal structure and Tom1l2 (red; residues Glu219-Gly311) homology model. Helices $\alpha 1$ (a1) through $\alpha 3$ (a3) and the N- and C-termini are labeled. The RMSD between the C^{α} atoms of TOM1 GAT and those in the corresponding Tom1l2 GAT model is 0.46 Å. (**D**) Residues on helices $\alpha 1$ and $\alpha 2$ in the Tom1l2 GAT domain are proposed to be the interaction site of ubiqitin based on cognate residues in the ubiquitin-TOM1 GAT domain crystal structure (Akutsu et al. 2005)

have immune defects, with abnormalities in lymphocyte development and impaired brain function (Grant et al. 1992; Stein et al. 1992). Given these data, it is reasonable to hypothesize that a decrease in *Tom1l2* expression would result in impaired immune responses.

Tom1 is also a direct target of both oncogenic forms of the *v-Myb* gene (Burk et al. 1997). *Myb* is involved in cell differentiation processes, including hematopoeisis (Burk and Klempnauer 1999; Oh and Reddy 1999), which is intriguing since *Tom1l2* hypomorphic mice have greater percentages of T cells in their spleens. In addition, *c-Kit* (Hogg et al. 1997) and *Bcl-2* (Frampton et al. 1996), well known for their prosurvival role in cancer, are upregulated by *v-Myb* along with *Tom1* (Burk et al. 1997). *Tom1l1* downregulation resulted in greater than twofold increase in mitogenic response induced by platelet-derived growth factor (Franco et al. 2006), and TOM1L1 levels are decreased in cutaneous squamous cell carcinoma and associated precursor lesions (Li et al. 2005). Also, increased SFK activity is present in many human cancers and Tom1l1 acts as "rheostat" for activated SFKs that regulate cell proliferation (Aligayer et al. 2002; Biscardi et al. 2000; Park et al. 1993). Indeed, transgenic mice overexpressing a mutant *Tom1l1* exhibited a hyperproliferative skin phenotype (Li et al. 2007). Recently, expression profiling of osteosarcoma cell lines implicated *TOM1L2* as a candidate target gene for 17p amplicon-associated osteosarcomas (Atiye et al. 2005; van Dartel and Hulsebos 2004). These results are remarkable considering that *Tom1l2* hypomorphs have an increased incidence of tumors.

Understanding the role of *Tom1l2* in the immune response and its ultimate effect on immune function has broad implications in the study of infectious disease, the immune response, and tumor development. If its deficiency results in increased incidence of malignancies, then understanding its function is paramount. Furthermore, evaluating this gene will help provide much needed data to direct any future development of pharmacotherapies related to this protein and its known pathways. The role of TOM1L2 in the Smith-Magenis syndrome, in which 90% of affected individuals are hemizygous for *TOM1L2*, may be pursued further as a cause for the immune abnormalities observed in these patients. By identifying the pathways in which TOM1L2 functions, we will be able to better understand and more effectively investigate the role of *TOM1L2* in immune deficiencies and human cancers.

Acknowledgments The authors thank Dr. Mario Dance from the VCU Department of Animal Resources for helping with collection of blood samples from mice. They acknowledge Valerie Vinoverski for assistance with GFP studies and Constance B. Hartmann for technical assistance with immunofluorescence staining and flow cytometry. They also thank the Transgenic Animal Model Core at the University of Michigan, the Histopathology Core at Michigan State University, and the Comparative Pathology Lab at UC Davis for mouse tissue and serum studies. This work was supported in part by NIH R01HD38534 (SHE), NIH P50DA05275 (KM), NIH R01ES07199 (KM), the Jeffress Foundation (SHE), the A. D. Williams Trust Fund (KM & SHE), and by resources from Michigan State University and Virginia Commonwealth University.

References

Akutsu M, Kawasaki M, Katoh Y, Shiba T, Yamaguchi Y et al (2005) Structural basis for recognition of ubiquitinated cargo by Tom1-GAT domain. FEBS Lett 579:5385–5391

Aligayer H, Boyd DD, Heiss MM, Abdalla EK, Curley SA et al (2002) Activation of Src kinase in primary colorectal carcinoma: an indicator of poor clinical prognosis. Cancer 94:344–351

Amsterdam A, Burgess S, Golling G, Chen W, Sun Z et al (1999) A large-scale insertional mutagenesis screen in zebrafish. Genes Dev 13:2713–2724

Atiye J, Wolf M, Kaur S, Monni O, Bohling T et al (2005) Gene amplifications in osteosarcoma-CGH microarray analysis. Genes Chromosomes Cancer 42:158–163

Bellen HJ, O'Kane CJ, Wilson C, Grossniklaus U, Pearson RK et al (1989) P-element-mediated enhancer detection: a versatile method to study development in Drosophila. Genes Dev 3:1288–1300

Bier E, Vaessin H, Shepherd S, Lee K, McCall K et al (1989) Searching for pattern and mutation in the Drosophila genome with a P-lacZ vector. Genes Dev 3:1273–1287

Biscardi JS, Ishizawar RC, Silva CM, Parsons SJ (2000) Tyrosine kinase signalling in breast cancer: epidermal growth factor receptor and c-Src interactions in breast cancer. Breast Cancer Res 2:203–210

Bonifacino JS (2004) The GGA proteins: adaptors on the move. Nat Rev Mol Cell Biol 5:23–32

Brissoni B, Agostini L, Kropf M, Martinon F, Swoboda V et al (2006) Intracellular trafficking of interleukin-1 receptor I requires Tollip. Curr Biol 16:2265–2270

Burk O, Klempnauer KH (1999) Myb and Ets transcription factors cooperate at the myb-inducible promoter of the tom-1 gene. Biochim Biophys Acta 1446:243–252

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

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ et al (2003) Multiple sequence alignment with the clustal series of programs. Nucleic Acids Res 31:3497–3500

Christopher JA (1998) SPOCK. Texas A&M University, College Station, TX

Crawley JN, Paylor R (1997) A proposed test battery and constellations of specific behavioral paradigms to investigate the behavioral phenotypes of transgenic and knockout mice. Horm Behav 31:197–211

Fiser A, Sali A (2003) Modeller: generation and refinement of homology-based protein structure models. Methods Enzymol 374:461–491

Frampton J, Ramqvist T, Graf T (1996) v-Myb of E26 leukemia virus up-regulates bcl-2 and suppresses apoptosis in myeloid cells. Genes Dev 10:2720–2731

Franco M, Furstoss O, Simon V, Benistant C, Hong WJ et al (2006) The adaptor protein Tom1L1 is a negative regulator of Src mitogenic signaling induced by growth factors. Mol Cell Biol 26:1932–1947

Girirajan S, Vlangos CN, Szomju BB, Edelman E, Trevors CD et al (2006) Genotype-phenotype correlation in Smith-Magenis syndrome: evidence that multiple genes in 17p11.2 contribute to the clinical spectrum. Genet Med 8:417–427

Gossler A, Joyner AL, Rossant J, Skarnes WC (1989) Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. Science 244:463–465

Grant SG, O'Dell TJ, Karl KA, Stein PL, Soriano P et al (1992) Impaired long-term potentiation, spatial learning, and hippocampal development in fyn mutant mice. Science 258:1903–1910

Greenberg F, Guzzetta V, Montes de Oca-Luna R, Magenis RE, Smith AC et al (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 et al (1996) Multi-disciplinary clinical study of Smith-Magenis syndrome (deletion 17p11.2). Am J Med Genet 62:247–254

Guo TL, McCay JA, Karrow NA, Brown RD, Musgrove DL et al (2001) Immunotoxicity of sodium bromate in female B6C3F1 mice: a 28-day drinking water study. Drug Chem Toxicol 24:129–149

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Hartmann CB, McCoy KL (2004) Gallium arsenide exposure impairs processing of particulate antigen by macrophages: modification of the antigen reverses the functional defect. Life Sci 75:485–498

Hogg A, Schirm S, Nakagoshi H, Bartley P, Ishii S et al (1997) Inactivation of a c-Myb/estrogen receptor fusion protein in transformed primary cells leads to granulocyte/macrophage differentiation and down regulation of c-kit but not c-myc or cdc2. Oncogene 15:2885–2898

Irwin S (1968) Comprehensive observational assessment: Ia. A systematic, quantitative procedure for assessing the behavioral and physiologic state of the mouse. Psychopharmacologia 13:222–257

Jerne NK, Nordin AA (1963) Plaque formation in agar by single antibody-producing cells. Science 140:405

Katoh Y, Shiba Y, Mitsuhashi H, Yanagida Y, Takatsu H et al (2004) Tollip and Tom1 form a complex and recruit ubiquitin-conjugated proteins onto early endosomes. J Biol Chem 279:24435–24443

Katoh Y, Imakagura H, Futatsumori M, Nakayama K (2006) Recruitment of clathrin onto endosomes by the Tom1-Tollip complex. Biochem Biophys Res Commun 341:143–149

Kirchhausen T (2000) Clathrin. Annu Rev Biochem 69:699–727

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291

Lewis TA, Munson AE, McCoy KL (1996) Gallium arsenide selectively suppresses antigen processing by splenic macrophages for CD4+ T cell activation. J Pharmacol Exp Ther 278:1244–1251

Li W, Marshall C, Mei L, Dzubow L, Schmults C et al (2005) Srcasm modulates EGF and Src-kinase signaling in keratinocytes. J Biol Chem 280:6036–6046

Li W, Marshall C, Mei L, Gelfand J, Seykora JT (2007) Srcasm corrects Fyn-induced epidermal hyperplasia by kinase down-regulation. J Biol Chem 282:1161–1169

Lohi O, Lehto VP (1998) VHS domain marks a group of proteins involved in endocytosis and vesicular trafficking. FEBS Lett 440:255–257

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

Luster MI, Munson AE, Thomas PT, Holsapple MP, Fenters JD et al (1988) Development of a testing battery to assess chemical-induced immunotoxicity: National Toxicology Program's guidelines for immunotoxicity evaluation in mice. Fundam Appl Toxicol 10:2–19

Misra S, Beach BM, Hurley JH (2000) Structure of the VHS domain of human *Tom1* (target of myb-1): insights into interactions with proteins and membranes. Biochemistry 39:11282–11290

Oh IH, Reddy EP (1999) The myb gene family in cell growth, differentiation and apoptosis. Oncogene 18:3017–3033

Park J, Meisler AI, Cartwright CA (1993) c-Yes tyrosine kinase activity in human colon carcinoma. Oncogene 8:2627–2635

Potash JB, Buervenich S, Cox NJ, Zandi PP, Akula N et al (2008) Gene-based SNP mapping of a psychotic bipolar affective disorder linkage region on 22q12.3: Association with HMG2L1 and TOM1. Am J Med Genet B Neuropsychiatr Genet 147:59–67

Puertollano R (2005) Interactions of TOM1L1 with the multivesicular body sorting machinery. J Biol Chem 280:9258–9264

Rogers DC, Fisher EM, Brown SD, Peters J, Hunter AJ et al (1997) Behavioral and functional analysis of mouse phenotype: SHIRPA, a proposed protocol for comprehensive phenotype assessment. Mamm Genome 8:711–713

Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815

Seet LF, Hong W (2005) Endofin recruits clathrin to early endosomes via TOM1. J Cell Sci 118:575–587

Seet LF, Liu N, Hanson BJ, Hong W (2004) Endofin recruits TOM1 to endosomes. J Biol Chem 279:4670–4679

Serafini T, Colamarino SA, Leonardo ED, Wang H, Beddington R et al (1996) Netrin-1 is required for commissural axon guidance in the developing vertebrate nervous system. Cell 87:1001–1014

Seykora JT, Mei L, Dotto GP, Stein PL (2002) 'Srcasm: a novel Src activating and signaling molecule. J Biol Chem 277:2812–2822

Shiba Y, Katoh Y, Shiba T, Yoshino K, Takatsu H et al (2004) GAT (GGA and Tom1) domain responsible for ubiquitin binding and ubiquitination. J Biol Chem 279:7105–7111

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

Skarnes WC, Moss JE, Hurtley SM, Beddington RS (1995) Capturing genes encoding membrane and secreted proteins important for mouse development. Proc Natl Acad Sci U S A 92:6592–6596

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, McGavran L, Robinson J, Waldstein G, Macfarlane J et al (1986) Interstitial deletion of (17)(p11.2p11.2) in nine patients. Am J Med Genet 24:393–414

Stanford WL, Cohn JB, Cordes SP (2001) Gene-trap mutagenesis: past, present and beyond. Nat Rev Genet 2:756–768

Stein PL, Lee HM, Rich S, Soriano P (1992) pp59fyn mutant mice display differential signaling in thymocytes and peripheral T cells. Cell 70:41–750

Stryke D, Kawamoto M, Huang CC, Johns SJ, King LA et al (2003) BayGenomics: a resource of insertional mutations in mouse embryonic stem cells. Nucleic Acids Res 31:278–281

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

Talbot WS, Hopkins N (2000) Zebrafish mutations and functional analysis of the vertebrate genome. Genes Dev 14:755–762

van Dartel M, Hulsebos TJ (2004) Amplification and overexpression of genes in 17p11.2-p12 in osteosarcoma. Cancer Genet Cytogenet 153:77–80

Voss AK, Thomas T, Gruss P (1998) Efficiency assessment of the gene trap approach. Dev Dyn 212:171–180

Yamakami M, Yokosawa H (2004) Tom1 (target of Myb 1) is a novel negative regulator of interleukin-1- and tumor necrosis factor-induced signaling pathways. Biol Pharm Bull 27:564–566

Yang HM, Butterworth L, Munson AE, Meade BJ (2003) Respiratory exposure to diesel exhaust particles decreases the spleen IgM response to a T cell-dependent antigen in female B6C3F1 mice. Toxicol Sci 71:207–216

Yeo TT, Yang T, Massa SM, Zhang JS, Honkaniemi J et al (1997) Deficient LAR expression decreases basal forebrain cholinergic neuronal size and hippocampal cholinergic innervation. J Neurosci Res 47:348–360

Zarzour W, Kleta R, Frangoul H, Suwannarat P, Jeong A et al (2005) Two novel CHS1 (LYST) mutations: clinical correlations in an infant with Chediak-Higashi syndrome. Mol Genet Metab 85:125–132

Copyright of Mammalian Genome is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Appendix E

Overlapping genes identified in ChIP-Chip study (full names)

ATP-binding cassette, sub-family B (MDR/TAP), member 1;ABCB1
ATP-binding cassette, sub-family B (MDR/TAP), member 1;ABCB1
amylo-1, 6-glucosidase, 4-alpha-glucanotransferase (glycogen debranching enzyme, glycogen storage disease type III);AGL
amylo-1, 6-glucosidase, 4-alpha-glucanotransferase (glycogen debranching enzyme, glycogen storage disease type III);AGL
A kinase (PRKA) anchor protein 7;AKAP7
A kinase (PRKA) anchor protein 7;AKAP7
aldo-keto reductase family 1, member C1 (dihydrodiol dehydrogenase 1; 20-alpha (3-alpha)-hydroxysteroid dehydrogenase);AKR1C1
aldo-keto reductase family 1, member C3 (3-alpha hydroxysteroid dehydrogenase, type II);AKR1C3
aldo-keto reductase family 1, member C3 (3-alpha hydroxysteroid dehydrogenase, type II);AKR1C3
amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 16;ALS2CR16
amelogenin, Y-linked;AMELY
amelogenin, Y-linked;AMELY
amylase, alpha 2A; pancreatic;AMY2A
amylase, alpha 2B (pancreatic);AMY2B
annexin A1;ANXA1
annexin A1;ANXA1
APin protein;APIN
adenine phosphoribosyltransferase;APRT
adenine phosphoribosyltransferase;APRT
Rho guanine nucleotide exchange factor (GEF) 7;ARHGEF7
Rho guanine nucleotide exchange factor (GEF) 7;ARHGEF7
N-acylsphingosine amidohydrolase (non-lysosomal ceramidase) 2;ASAH2
N-acylsphingosine amidohydrolase (non-lysosomal ceramidase) 2;ASAH2
ankyrin repeat and SOCS box-containing 4;ASB4
activating transcription factor 6;ATF6
activating transcription factor 6;ATF6
ATP synthase, H+ transporting, mitochondrial F1 complex, epsilon subunit;ATP5E
ATP synthase, H+ transporting, mitochondrial F1 complex, epsilon subunit;ATP5E
5-azacytidine induced 2;AZI2
5-azacytidine induced 2;AZI2
bromodomain adjacent to zinc finger domain, 2B;BAZ2B
bromodomain adjacent to zinc finger domain, 2B;BAZ2B
butyrylcholinesterase;BCHE
butyrylcholinesterase;BCHE
bone morphogenetic protein 5;BMP5
bone morphogenetic protein 5;BMP5

chromosome 10 open reading frame 61;C10orf61
chromosome 10 open reading frame 61;C10orf61
chromosome 10 open reading frame 68;C10orf68
chromosome 10 open reading frame 68;C10orf68
chromosome 14 open reading frame 10;C14orf10
chromosome 14 open reading frame 10;C14orf10
chromosome 14 open reading frame 150;C14orf150
chromosome 16 open reading frame 73;C16orf73
chromosome 18 open reading frame 9;C18orf9
chromosome 20 open reading frame 103;C20orf103
chromosome 20 open reading frame 103;C20orf103
chromosome 21 open reading frame 34;C21orf34
chromosome 21 open reading frame 34;C21orf34
chromosome 3 open reading frame 58;C3orf58
complement component 4 binding protein, beta;C4BPB
complement component 4 binding protein, beta;C4BPB
chromosome 4 open reading frame 22;C4orf22
chromosome 4 open reading frame 7;C4orf7
chromosome 4 open reading frame 7;C4orf7
chromosome 5 open reading frame 36;C5orf36
chromosome 9 open reading frame 102;C9orf102
chromosome 9 open reading frame 150;C9orf150
chromosome 9 open reading frame 150;C9orf150
calcium binding protein 39-like;CAB39L
calcium binding protein 39-like;CAB39L
coiled-coil domain containing 46;CCDC46
CD69 antigen (p60, early T-cell activation antigen);CD69
CD69 molecule;CD69
cadherin 12, type 2 (N-cadherin 2);CDH12
cadherin 12, type 2 (N-cadherin 2);CDH12
chromodomain protein, Y-linked, 2B;CDY2B
centaurin, delta 1;CENTD1
centaurin, delta 1;CENTD1
centrosomal protein 76kDa;CEP76
choline/ethanolamine phosphotransferase 1;CEPT1
choline/ethanolamine phosphotransferase 1;CEPT1
complement factor H;CFH
cell adhesion molecule with homology to L1CAM (close homolog of L1);CHL1
cell adhesion molecule with homology to L1CAM (close homolog of L1);CHL1
chondrolectin;CHODL
chondrolectin;CHODL
cholinergic receptor, nicotinic, alpha 5;CHRNA5
cholinergic receptor, nicotinic, alpha polypeptide 5;CHRNA5
C-type lectin domain family 5, member A;CLEC5A
C-type lectin domain family 5, member A;CLEC5A
ciliary neurotrophic factor;CNTF

ciliary neurotrophic factor;CNTF
component of oligomeric golgi complex 6;COG6
component of oligomeric golgi complex 6;COG6
COMM domain containing 3;COMMD3
cereblon;CRBN
cereblon;CRBN
cAMP responsive element modulator;CREM
cAMP responsive element modulator;CREM
crystallin, gamma S;CRYGS
crystallin, gamma S;CRYGS
cylicin, basic protein of sperm head cytoskeleton 2;CYLC2
cylicin, basic protein of sperm head cytoskeleton 2;CYLC2
chromosome Y open reading frame 15A;CYorf15A
chromosome Y open reading frame 15B;CYorf15B
DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11 (CHL1-like helicase homolog, S. cerevisiae);DDX11
defensin, beta 112;DEFB112
defensin, beta 112;DEFB112
diacylglycerol kinase, beta 90kDa;DGKB
diacylglycerol kinase, beta 90kDa;DGKB
diaphanous homolog 3 (Drosophila);DIAPH3
diaphanous homolog 3 (Drosophila);DIAPH3
DnaJ (Hsp40) homolog, subfamily C, member 5 beta;DNAJC5B
DnaJ (Hsp40) homolog, subfamily C, member 5 beta;DNAJC5B
dentin sialophosphoprotein;DSPP
dentin sialophosphoprotein;DSPP
dystrobrevin, alpha;DTNA
dystrobrevin, alpha;DTNA
elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 2;ELOVL2
elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 2;ELOVL2
eyes absent homolog 1 (Drosophila);EYA1
eyes absent homolog 1 (Drosophila);EYA1
family with sequence similarity 55, member D;FAM55D
family with sequence similarity 55, member D;FAM55D
F-box and leucine-rich repeat protein 7;FBXL7
F-box and leucine-rich repeat protein 7;FBXL7
F-box protein 4;FBXO4
fibrinogen beta chain;FGB
fibrinogen beta chain;FGB
fibroblast growth factor 7 (keratinocyte growth factor);FGF7
fibrinogen-like 2;FGL2
fibrinogen-like 2;FGL2
fumarate hydratase;FH
fumarate hydratase;FH

hypothetical protein FLJ22662;FLJ22662
hypothetical protein FLJ32745;FLJ32745
FLJ44048 protein;FLJ44048
forkhead box K2;FOXK2
forkhead box K2;FOXK2
forty-two-three domain containing 1;FYTTD1
forty-two-three domain containing 1;FYTTD1
gamma-aminobutyric acid (GABA) A receptor, gamma 1;GABRG1
gamma-aminobutyric acid (GABA) A receptor, gamma 1;GABRG1
UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 13 (GalNAc-T13);GALNT13
UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 13 (GalNAc-T13);GALNT13
UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-like 2;GALNTL2
UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-like 2;GALNTL2
GDNF family receptor alpha like;GFRAL
glycine receptor, alpha 3;GLRA3
glycine receptor, alpha 3;GLRA3
G protein-coupled receptor 52;GPR52
G protein-coupled receptor 52;GPR52
HERV-H LTR-associating 2;HHLA2
HERV-H LTR-associating 2;HHLA2
histone 1, H2ac;HIST1H2AC
histone cluster 1, H2ac;HIST1H2AC
histone 1, H2bl;HIST1H2BL
histone cluster 1, H2bl;HIST1H2BL
histone 1, H2bm;HIST1H2BM
histone cluster 1, H2bm;HIST1H2BM
histone cluster 1, H3h;HIST1H3H
histone 1, H4b;HIST1H4B
histone cluster 1, H4b;HIST1H4B
histone cluster 1, H4j;HIST1H4J
major histocompatibility complex, class II, DP alpha 1;HLA-DPA1
major histocompatibility complex, class II, DP alpha 1;HLA-DPA1
major histocompatibility complex, class II, DP beta 1;HLA-DPB1
major histocompatibility complex, class II, DP beta 1;HLA-DPB1
hypothetical protein HS322B1A;HS322B1A
heat shock transcription factor, Y-linked 1;HSFY1
heat shock transcription factor, Y linked 2;HSFY2
hereditary sensory neuropathy, type II;HSN2
5-hydroxytryptamine (serotonin) receptor 2B;HTR2B
5-hydroxytryptamine (serotonin) receptor 2B;HTR2B
interferon, alpha 10;IFNA10
interleukin 12 receptor, beta 2;IL12RB2

interleukin 12 receptor, beta 2;IL12RB2
interleukin 15;IL15
interleukin 15;IL15
interleukin 17D;IL17D
interleukin 17D;IL17D
interphotoreceptor matrix proteoglycan 2;IMPG2
interphotoreceptor matrix proteoglycan 2;IMPG2
inversin;INVS
inversin;INVS
jumonji domain containing 1A;JMJD1A
jumonji domain containing 1A;JMJD1A
potassium channel, subfamily K, member 10;KCNK10
potassium channel, subfamily K, member 10;KCNK10
potassium large conductance calcium-activated channel, subfamily M, beta member 4;KCNMB4
potassium large conductance calcium-activated channel, subfamily M, beta member 4;KCNMB4
potassium channel tetramerisation domain containing 4;KCTD4
potassium channel tetramerisation domain containing 4;KCTD4
KIAA0391;KIAA0391
KIAA0391;KIAA0391
KIAA1712;KIAA1712
KIAA1712;KIAA1712
Kruppel-like factor 14;KLF14
Kruppel-like factor 14;KLF14
kelch-like 4 (Drosophila);KLHL4
kelch-like 4 (Drosophila);KLHL4
kelch-like 9 (Drosophila);KLHL9
kelch-like 9 (Drosophila);KLHL9
killer cell lectin-like receptor subfamily C, member 4;KLRC4
killer cell lectin-like receptor subfamily K, member 1;KLRK1
keratin 12 (Meesmann corneal dystrophy);KRT12
keratin 12 (Meesmann corneal dystrophy);KRT12
keratin 24;KRT24
keratin 24;KRT24
keratin associated protein 1-1;KRTAP1-1
keratin associated protein 1-1;KRTAP1-1
keratin associated protein 1-3;KRTAP1-3
keratin associated protein 1-3;KRTAP1-3
keratin associated protein 4-12;KRTAP4-12
keratin associated protein 4-12;KRTAP4-12
keratin associated protein 6-3;KRTAP6-3
late cornified envelope 3B;LCE3B
late cornified envelope 3B;LCE3B
low density lipoprotein receptor (familial hypercholesterolemia);LDLR
low density lipoprotein receptor (familial hypercholesterolemia);LDLR

lymphoid enhancer-binding factor 1;LEF1
smooth muscle myosin heavy chain 11 isoform SM1-like;LOC129285
C/EBP-induced protein;LOC81558
latrophilin 2;LPHN2
latrophilin 2;LPHN2
leucine rich repeat neuronal 3;LRRN3
leucine rich repeat neuronal 3;LRRN3
leucine rich repeat transmembrane neuronal 4;LRRTM4
leucine rich repeat transmembrane neuronal 4;LRRTM4
liver-specific organic anion transporter 3;LST-3
lumican;LUM
lumican;LUM
MAS1 oncogene-like;MAS1L
MAS1 oncogene-like;MAS1L
MADS box transcription enhancer factor 2, polypeptide A (myocyte enhancer factor 2A);MEF2A
MADS box transcription enhancer factor 2, polypeptide A (myocyte enhancer factor 2A);MEF2A
meprin A, alpha (PABA peptide hydrolase);MEP1A
meprin A, alpha (PABA peptide hydrolase);MEP1A
mannosyl (alpha-1,3-)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase, isozyme C (putative);MGAT4C
hypothetical protein MGC33530;MGC33530
hypothetical protein MGC34713;MGC34713
hypothetical protein MGC35043;MGC35043
hypothetical protein MGC35212;MGC35212
makorin, ring finger protein, 3;MKRN3
makorin, ring finger protein, 3;MKRN3
matrix metallopeptidase 13 (collagenase 3);MMP13
matrix metallopeptidase 13 (collagenase 3);MMP13
mitochondrial translational release factor 1-like;MTRF1L
mitochondrial translational release factor 1-like;MTRF1L
mucin 7, salivary;MUC7
mucin 7, secreted;MUC7
myosin, light chain 1, alkali; skeletal, fast;MYL1
myosin, light polypeptide 1, alkali; skeletal, fast;MYL1
nucleosome assembly protein 1-like 3;NAP1L3
nucleosome assembly protein 1-like 3;NAP1L3
odontogenic, ameloblast asssociated;ODAM
osteoglycin (osteoinductive factor, mimecan);OGN
osteoglycin (osteoinductive factor, mimecan);OGN
osteomodulin;OMD
osteomodulin;OMD
olfactory receptor, family 10, subfamily T, member 2;OR10T2
olfactory receptor, family 10, subfamily T, member 2;OR10T2
olfactory receptor, family 11, subfamily H, member 4;OR11H4

olfactory receptor, family 11, subfamily H, member 4;OR11H4
olfactory receptor, family 13, subfamily C, member 3;OR13C3
olfactory receptor, family 13, subfamily C, member 3;OR13C3
olfactory receptor, family 13, subfamily F, member 1;OR13F1
olfactory receptor, family 13, subfamily F, member 1;OR13F1
olfactory receptor, family 1, subfamily Q, member 1;OR1Q1
olfactory receptor, family 1, subfamily Q, member 1;OR1Q1
olfactory receptor, family 2, subfamily B, member 11;OR2B11
olfactory receptor, family 2, subfamily J, member 3;OR2J3
olfactory receptor, family 2, subfamily L, member 2;OR2L2
olfactory receptor, family 2, subfamily L, member 2;OR2L2
olfactory receptor, family 2, subfamily L, member 3;OR2L3
olfactory receptor, family 2, subfamily M, member 5;OR2M5
olfactory receptor, family 2, subfamily T, member 10;OR2T10
olfactory receptor, family 2, subfamily T, member 4;OR2T4
olfactory receptor, family 4, subfamily A, member 47;OR4A47
olfactory receptor, family 4, subfamily C, member 16;OR4C16
olfactory receptor, family 4, subfamily C, member 16;OR4C16
olfactory receptor, family 51, subfamily A, member 7;OR51A7
olfactory receptor, family 51, subfamily A, member 7;OR51A7
olfactory receptor, family 52, subfamily A, member 1;OR52A1
olfactory receptor, family 52, subfamily A, member 1;OR52A1
olfactory receptor, family 52, subfamily N, member 5;OR52N5
olfactory receptor, family 52, subfamily N, member 5;OR52N5
olfactory receptor, family 56, subfamily B, member 1;OR56B1
olfactory receptor, family 56, subfamily B, member 1;OR56B1
olfactory receptor, family 5, subfamily B, member 2;OR5B2
olfactory receptor, family 5, subfamily BU, member 1;OR5BU1
olfactory receptor, family 5, subfamily H, member 14;OR5H14
olfactory receptor, family 5, subfamily H, member 6;OR5H6
olfactory receptor, family 5, subfamily H, member 6;OR5H6
olfactory receptor, family 5, subfamily K, member 3;OR5K3
olfactory receptor, family 5, subfamily K, member 4;OR5K4
olfactory receptor, family 5, subfamily M, member 3;OR5M3
olfactory receptor, family 5, subfamily M, member 3;OR5M3
olfactory receptor, family 5, subfamily R, member 1;OR5R1
olfactory receptor, family 5, subfamily T, member 2;OR5T2
olfactory receptor, family 5, subfamily T, member 2;OR5T2
olfactory receptor, family 5, subfamily V, member 1;OR5V1
olfactory receptor, family 5, subfamily V, member 1;OR5V1
olfactory receptor, family 5, subfamily W, member 2;OR5W2
olfactory receptor, family 6, subfamily M, member 1;OR6M1
olfactory receptor, family 6, subfamily M, member 1;OR6M1
olfactory receptor, family 8, subfamily H, member 1;OR8H1
olfactory receptor, family 8, subfamily H, member 2;OR8H2
olfactory receptor, family 8, subfamily H, member 2;OR8H2

olfactory receptor, family 8, subfamily I, member 2;OR8I2
olfactory receptor, family 8, subfamily I, member 2;OR8I2
olfactory receptor, family 8, subfamily K, member 3;OR8K3
olfactory receptor, family 8, subfamily K, member 3;OR8K3
olfactory receptor, family 9, subfamily G, member 1;OR9G1
olfactory receptor, family 9, subfamily G, member 1;OR9G1
olfactory receptor, family 9, subfamily G, member 9;OR9G9
OTU domain containing 6B;OTUD6B
OTU domain containing 6B;OTUD6B
palladin, cytoskeletal associated protein;PALLD
protocadherin 11 X-linked;PCDH11X
protocadherin 11 Y-linked;PCDH11Y
protocadherin 15;PCDH15
protocadherin 15;PCDH15
protocadherin alpha 1;PCDHA1
protocadherin alpha 11;PCDHA11
protocadherin alpha 3;PCDHA3
protocadherin beta 16;PCDHB16
protocadherin beta 16;PCDHB16
protocadherin beta 8;PCDHB8
protocadherin beta 8;PCDHB8
protocadherin gamma subfamily A, 12;PCDHGA12
protocadherin gamma subfamily A, 9;PCDHGA9
phosphodiesterase 1A, calmodulin-dependent;PDE1A
phosphodiesterase 1A, calmodulin-dependent;PDE1A
pleckstrin homology domain interacting protein;PHIP
phytanoyl-CoA 2-hydroxylase interacting protein-like;PHYHIPL
phytanoyl-CoA hydroxylase interacting protein-like;PHYHIPL
phosphatidylinositol glycan anchor biosynthesis, class F;PIGF
phosphatidylinositol glycan, class F;PIGF
phospholipase C, beta 4;PLCB4
phospholipase C, beta 4;PLCB4
pleiotropic regulator 1 (PRL1 homolog, Arabidopsis);PLRG1
pleiotropic regulator 1 (PRL1 homolog, Arabidopsis);PLRG1
pro-melanin-concentrating hormone;PMCH
pancreatic lipase;PNLIP
pancreatic lipase;PNLIP
POU domain, class 1, transcription factor 1 (Pit1, growth hormone factor 1);POU1F1
POU domain, class 1, transcription factor 1 (Pit1, growth hormone factor 1);POU1F1
protein kinase, cAMP-dependent, catalytic, beta;PRKACB
protein kinase, cAMP-dependent, catalytic, beta;PRKACB
RAP1B, member of RAS oncogene family;RAP1B
RAP1B, member of RAS oncogene family;RAP1B
regulator of G-protein signalling 21;RGS21
ras homolog gene family, member B;RHOB
ras homolog gene family, member B;RHOB

ring finger protein 17;RNF17
ras responsive element binding protein 1;RREB1
ras responsive element binding protein 1;RREB1
arginine/serine-rich coiled-coil 1;RSRC1
arginine/serine-rich coiled-coil 1;RSRC1
serpin peptidase inhibitor, clade I (pancpin), member 2;SERPINI2
serpin peptidase inhibitor, clade I (pancpin), member 2;SERPINI2
sarcoglycan, gamma (35kDa dystrophin-associated glycoprotein);SGCG
sarcoglycan, gamma (35kDa dystrophin-associated glycoprotein);SGCG
sarcoglycan zeta;SGCZ
sarcoglycan zeta;SGCZ
short stature homeobox 2;SHOX2
short stature homeobox 2;SHOX2
sucrase-isomaltase (alpha-glucosidase);SI
sucrase-isomaltase (alpha-glucosidase);SI
silver homolog (mouse);SILV
silver homolog (mouse);SILV
solute carrier family 12 (sodium/potassium/chloride transporters), member 1;SLC12A1
solute carrier family 12 (sodium/potassium/chloride transporters), member 1;SLC12A1
solute carrier family 16 (monocarboxylic acid transporters), member 7;SLC16A7
solute carrier family 16, member 7 (monocarboxylic acid transporter 2);SLC16A7
solute carrier family 25, member 32;SLC25A32
solute carrier family 25, member 32;SLC25A32
solute carrier family 26, member 7;SLC26A7
solute carrier family 26, member 7;SLC26A7
solute carrier organic anion transporter family, member 1B3;SLCO1B3
solute carrier organic anion transporter family, member 1B3;SLCO1B3
SLIT and NTRK-like family, member 6;SLITRK6
SLIT and NTRK-like family, member 6;SLITRK6
structural maintenance of chromosomes 6;SMC6
SMC6 structural maintenance of chromosomes 6-like 1 (yeast);SMC6L1
syntaphilin;SNPH
syntaphilin;SNPH
small nuclear ribonucleoprotein polypeptide N;SNRPN
Sp6 transcription factor;SP6
Sp6 transcription factor;SP6
sperm associated antigen 6;SPAG6
sperm associated antigen 6;SPAG6
sperm protein associated with the nucleus, X-linked, family member A1;SPANXA1
SPANX family, member A2;SPANXA2
SPANX family, member E;SPANXE
spermatogenesis associated 18 homolog (rat);SPATA18
spermatogenesis associated 18 homolog (rat);SPATA18
TAF2 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 150kDa;TAF2
TAF2 RNA polymerase II, TATA box binding protein (TBP)-associated factor,

150kDa;TAF2
TGF-beta induced apotosis protein 2;TAIP-2
taste receptor, type 2, member 7;TAS2R7
taste receptor, type 2, member 7;TAS2R7
taste receptor, type 2, member 8;TAS2R8
taste receptor, type 2, member 8;TAS2R8
tudor domain containing 4;TDRD4
testis expressed sequence 15;TEX15
testis expressed sequence 15;TEX15
transcription factor EC;TFEC
transcription factor EC;TFEC
transmembrane protein 77;TMEM77
transmembrane protein 77;TMEM77
transmembrane protease, serine 11B;TMPRSS11B
transmembrane protease, serine 11B;TMPRSS11B
tumor necrosis factor (ligand) superfamily, member 18;TNFSF18
tumor necrosis factor (ligand) superfamily, member 18;TNFSF18
tumor necrosis factor superfamily, member 5-induced protein 1;TNFSF5IP1
tumor necrosis factor superfamily, member 5-induced protein 1;TNFSF5IP1
transient receptor potential cation channel, subfamily C, member 3;TRPC3
transient receptor potential cation channel, subfamily C, member 3;TRPC3
transient receptor potential cation channel, subfamily M, member 8;TRPM8
transient receptor potential cation channel, subfamily M, member 8;TRPM8
tumor suppressor candidate 1;TUSC1
tumor suppressor candidate 1;TUSC1
thioredoxin domain containing 13;TXNDC13
ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome);UBE3A
ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome);UBE3A
ubiquitin-fold modifier 1;UFM1
ubiquitin-fold modifier 1;UFM1
UDP glucuronosyltransferase 2 family, polypeptide A1;UGT2A1
UDP glucuronosyltransferase 2 family, polypeptide A1;UGT2A1
IVFI9356;UNQ9356
urotensin 2 domain containing;UTS2D
urotensin 2 domain containing;UTS2D
ubiquitously transcribed tetratricopeptide repeat gene, Y-linked;UTY
ubiquitously transcribed tetratricopeptide repeat gene, Y-linked;UTY
V-set and transmembrane domain containing 2;VSTM2
von Willebrand factor A domain containing 2;VWA2
WD repeat domain 89;WDR89
WD repeats and SOF1 domain containing;WDSOF1
WD repeats and SOF1 domain containing;WDSOF1
5'-3&apos; exoribonuclease 1;XRN1
5'-3' exoribonuclease 1;XRN1

Yip1 domain family, member 7;YIPF7
Yip1 domain family, member 7;YIPF7
zinc finger protein 306;ZNF306
zinc finger protein 306;ZNF306
zinc finger protein 614;ZNF614