Clemson University
TigerPrints

All Dissertations | Dissertations

12-2012

Analysis of Genotype, Phenotype, and Age Progression in Phelan-McDermid Syndrome

Sara Sarasua
Clemson University, sarasua@mindspring.com

Follow this and additional works at: https://tigerprints.clemson.edu/all_dissertations

 Part of the Genetics and Genomics Commons

Recommended Citation
Sarasua, Sara, "Analysis of Genotype, Phenotype, and Age Progression in Phelan-McDermid Syndrome" (2012). *All Dissertations*. 1032.
https://tigerprints.clemson.edu/all_dissertations/1032

This Dissertation is brought to you for free and open access by the Dissertations at TigerPrints. It has been accepted for inclusion in All Dissertations by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.

ANALYSIS OF GENOTYPE, PHENOTYPE, AND AGE PROGRESSION OF PHELAN-MCDERMID SYNDROME

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Genetics

---

by
Sara Moir Sarasua
December 2012

Accepted by:
Dr. Amy Lawton-Rauh, Committee Chair
Dr. Chin-Fu Chen
Dr. Leigh Anne Clark
Dr. Barbara R. DuPont
Dr. Alex Feltus

ABSTRACT

Phelan-McDermid syndrome is a developmental disability syndrome associated with deletions of the terminal end of one copy of chromosome 22q13. The observed chromosomal aberrations include simple terminal deletions, interstitial deletions, deletions and duplications, and duplications without deletions. All patients have some degree of developmental disability and many also have hypotonia, autism, minor dysmorphic features, and seizures. I performed an epidemiological and cytogenetic investigation to better understand the etiology of Phelan-McDermid syndrome and to provide information to patients and their families, clinicians, and researchers investigating developmental disabilities. Deletions vary widely in size, from 60 kb to more than 9 Mb, but almost all cases are missing one copy of the subtelomeric gene *SHANK3*, a candidate gene for neurological features. The results of this study established that larger deletions are associated with more severe disability establishing the rationale to investigate the role of additional genes or genomic regions for clinical features. Statistical association analyses identified specific genomic regions as associated with 22 clinical features. In particular, speech is highly correlated with deletion size indicating that speech-related genes or genomic elements located in genomic bands 22q13.2q13.31 may be critical in determining a patient's ability to communicate verbally. The use of protein interaction networks identified candidate genes within these narrowed genomic regions. Also, a longitudinal assessment of phenotypes observed among individuals aged 0.4 to 64 years established significant variation of phenotypes by age, such that future investigations need to take age into account when conducting genotype-phenotype studies. In particular, we find that behavioral difficulties subside and low

muscle tone becomes less prominent as children age, however seizures, autism, and some chronic diseases become more apparent in teens and adults.

DEDICATION

I dedicate this work to my family for their unwavering support.

ACKNOWLEDGMENTS

This work would not be possible without the dedicated efforts of the patients and families of patients with Phelan-McDermid syndrome.

I specifically acknowledge the support and guidance from Drs. Katy Phelan, Barbara DuPont, Luigi Boccuto, Curtis Rogers, Charles Schwartz, and Roger Stevenson. I thank my committee members Drs. Amy Lawton-Rauh, Chin-Fu Chen, Leigh Anne Clark, Barbara DuPont, and Alex Feltus for their helpful suggestions throughout my training.

Finally, I thank my late advisor, Dr. Julianne Collins, who set the standard to learn as much as possible about developmental disabilities and birth defects to help patients as much as possible.

TABLE OF CONTENTS

Page

Table of Contents (Continued)

Page

Table of Contents (Continued)

Page

LIST OF TABLES

List of Tables (Continued)

Table Page

LIST OF FIGURES

List of Figures (Continued)

Figure Page

PREFACE

In this work, I used statistical approaches to assess the potential associations of 22q13.3 aberrations and clinical features of Phelan-McDermid syndrome. Phelan-McDermid syndrome (PMS) is a developmental disability syndrome characterized by intellectual and developmental disability including severely delayed or absent speech, low muscle tone, minor dysmorphic features, autism spectrum disorders, seizures, and a number of other behavioral, medical, and physical features (Phelan and McDermid 2012). Presentation of Phelan-McDermid syndrome is associated with the deletion of the terminal end of one copy of chromosome 22q13. The two overarching questions I address in this work are 1) whether there are genes, in addition to *SHANK3*, that may contribute to the clinical features of the syndromes and 2) whether the clinical features of the syndrome change with patient age. The methods I used to study PMS and the improved knowledge of several potential clinical and molecular causes of this syndrome are applicable to the study of other syndromes. While no treatments are yet available, a better understanding of the clinical presentation and effect of deletion of different genes along the chromosome will assist patients, researchers, and clinicians in understanding and treating these patients.

This dissertation is divided into six chapters. In Chapter One, I review the literature regarding what is known about chromosome 22 and mechanisms that can lead to deletions or duplications. I also review what is known about Phelan-McDermid syndrome and the clinical features most commonly found. Because much of the Phelan-McDermid syndrome literature focuses on *SHANK3* as a candidate gene for some of the neurological features of the syndrome, I review what is known about this telomeric gene and its role in the brain, specifically its role in

the post synaptic density of excitatory neurons. I also review what is known about other genes and phenotypes associated with chromosome 22 in order to place this syndrome into context.

In Chapter Two, I present an analysis of deletion size and its relationship with clinical features which was published in the Journal of Medical Genetics (Sarasua and others 2011). The primary question addressed in this chapter was whether 22q13 genes or genomic regions, in addition to *SHANK3*, may be responsible for presence or severity of the clinical features in this syndrome. The specific hypothesis tested was that there is no difference in size of deletion between those with and without a given clinical feature. The alternative hypothesis was that there is a difference in deletion size between those with and without a given clinical feature. Deletion size was used as a proxy measure for the number of genes deleted, in addition to *SHANK3*. This work was based upon standardized clinical assessments of patients with simple terminal deletions of 22q13. I identified clinically relevant features significantly associated with deletion size and also reported the prevalence of 80 different phenotypes. One finding of particular significance was the association between better speech abilities in those with smaller deletions compared to larger deletions.

In Chapter Three, I expanded the genotype-phenotype study presented in Chapter Two to identify specific regions most significantly associated with each clinical feature. Protein interaction networks were used to identify candidate genes within these narrowed windows that may be most associated with the phenotypes identified. The research question being asked in this chapter was what specific genomic regions are associated with clinical features? I tested the null hypothesis that a given genomic region was not associated with a clinical feature. The alternative hypothesis was that a given genomic region was associated with a clinical feature.

In Chapter Four, I present a study on potential age-related and parent of origin related features in PMS. My primary research question was whether the syndrome changed with age or was affected by the parent of origin of the affected chromosome. My specific null hypothesis was that there was no difference in clinical features across age groups or parent of origin of the affected chromosome. My alternative hypothesis was that there was a difference in clinical features across age groups and parent of origin. To date, most studies have been based upon case reports, small case series, or studies of generally less than 40 individuals where no attempt was made to look at age-related differences in phenotypes. A consensus at the 2012 Phelan-McDermid syndrome Scientific Symposium held in July, 2012, in Orlando specifically identified the need for longitudinal studies of the syndrome to determine age-related changes. I assessed age-related changes both cross-sectionally and longitudinally among those patients who participated in multiple clinic visits. I also examined parent of origin effects on clinical features to look for potential imprinting whereby a gene or genomic region is epigenetically silenced depending on maternal or paternal inheritance. Further, I tested the effects of deletion size on clinical features. Finally, I confirmed the presence of the deletion in a buccal (cheek) specimen. This finding demonstrated that the cytogenetic rearrangements of PMS are not limited to the more commonly used whole blood specimens and are likely representative of cytogenetic rearrangements in other tissues, including other tissues of ectodermal origin such as brain.

In Chapter Five, I provide a summary of my findings and a look to the future.

As part of my graduate training, I participated in the 2008 Genetic Analysis Workshop and collaborated with statisticians and epidemiologists in the application of a novel method to adjust for population structure in a genome wide association study of rheumatoid arthritis. The

population genetics and statistical and programming skills I developed were helpful in my genetic epidemiology training, although this topic did not apply directly to my analysis of PMS. For this reason, I have included the paper I co-wrote in the appendix of this dissertation.

References

Phelan K and McDermid HE. 2012. The 22q13.3 deletion syndrome (Phelan-McDermid syndrome). Mol Syndromol 2(3-5):186-201.

Sarasua SM, Dwivedi A, Boccuto L, Rollins JD, Chen CF, Rogers RC, Phelan K, DuPont BR, Collins JS. 2011. Association between deletion size and important phenotypes expands the genomic region of interest in Phelan-McDermid syndrome (22q13 deletion syndrome). J Med Genet 48(11):761-6.

CHAPTER ONE

CHROMOSOMAL ANOMALIES AND CLINICAL FEATURES ASSOCIATED WITH CHROMOSOME 22Q13

Overview

This review describes and chronicles research into the chromosomal microdeletion syndrome Phelan-McDermid syndrome (PMS) [MIM 606232]. This syndrome, also known as 22q13.3 deletion syndrome, is associated with the loss, or occasional gain, of the terminal end of the long arm of chromosome 22. It is recognized as a syndrome, encompassing developmental delays including marked speech delay, intellectual disability, hypotonia, minor dysmorphic features and sometimes autism spectrum disorders (ASDs), epilepsy, and a wide range of other clinical features (Phelan and McDermid 2012). *SHANK3* has been identified as a candidate gene for neurological features. The finding that this gene is often associated with ASDs, and the substantial research interest into the causes of ASDs, has led to increased research into PMS in general and *SHANK3* in particular. The highly variable deletion sizes, along with the wide range in severity of the condition, suggest the need for epidemiologic and cytogenetic work to understand the genetic causes of clinical features observed in PMS. Further, little work has been done to understand the syndrome over the lifespan of the affected individuals. Age related changes in the clinical presentation of the syndrome may affect which features can be observed at different ages and may impact the design of genotype-phenotype studies as well as diagnoses.

In addition to describing what is known about PMS, this review also describes methods used to investigate the causes of a syndrome. This approach includes identifying a recurrent clustering of clinical features with an accompanying chromosomal rearrangement (deletion, duplication, translocation, or inversion) and identifying candidate genes within the affected genomic region. These investigations require evidence from case reports and clinical studies, as

well as incorporation of experimental information from *in vitro* and *in vivo* studies of model organisms. Statistical and bioinformatics tools can then be used to integrate extant knowledge and generate hypotheses for further investigative steps.

This review details what is known about chromosomal rearrangements and their causes, specifically what is known about chromosome 22, the history of PMS, what is known about *SHANK3*, and investigates how this information can be applied to broader phenotypes such as autism and intellectual disability, speech, hypotonia, and growth anomalies ---all prominent features of Phelan-McDermid syndrome.

Chromosome 22 and Chromosomal Rearrangements

The human genome is comprised of 22 pairs of autosomal chromosomes, the X and Y sex chromosomes, and the mitochondrial genome. Some anomalies in humans include whole chromosome duplication (for example, trisomy 21 also known as Down syndrome) or whole chromosome deletion (for example, monosomy X, also known as Turner Syndrome). Structural changes can include ring chromosomes, translocations whereby genomic material is exchanged between chromosomes, inversions whereby chromosomal material is not lost, but is inverted, and deletions and duplications of portions of a chromosome. As reviewed by Lupski and Stankiewicz, chromosomal deletions and duplications can impact phenotype by affecting gene dosage, gene interruption, gene fusion, position effects, unmasking recessive alleles, and by disruption of regulatory communication between alleles (Lupski and Stankiewicz 2005; Stankiewicz and Lupski 2002a). Chromosomal deletions and duplications can be found on all chromosomes.

Chromosome 22

Chromosome 22 is an acrocentric chromosome of approximately 51 Mb or 49.7 Mb in size according to the 2009 (hg19) build or the 2006 (NCBI36/hg18) build (International Human Genome Sequencing Consortium 2004), respectively. The 2006 human genome build will be used as reference in this review as it was used for the deletion breakpoint mapping in our data and many published studies refer to this build.

The PMS deletion region includes bands 22q13.2, 22q13.31, 22q13.32, and 22q13.33 (the terminus of 22q). 22q13.2, q13.31 and q13.33 are gene rich while band q13.32 is gene poor (Figure 1.1). Segmental duplications are scattered across this region, with more in the subtelomeric region and in 22q13.2. Several micro RNAs (miRNAs) are located in 22q13.2 and 22q13.31. MiRNAs are regulatory noncoding RNAs that bind to messenger RNA to regulate translation or degradation of the transcripts (Bartel 2004). These RNAs are typically about 22 nucleotides in size. Several miRNA target prediction algorithms have been developed to predict the targets of miRNAs including miRBase (Betel and others 2008; Griffiths-Jones and others 2006) and TargetScan (Friedman and others 2009). One miRNA located at 22q13.31, hsa-mir-1249, is predicted to target *SHANK3 (TargetScan)*. Other noncoding RNA, of unknown function, are also located on 22q13 (Figure 1.1). These RNAs, along with potential enhancer or regulatory elements, mean that associations between deletion regions and clinical features may be due to protein-coding or non-protein coding effects.

Copy Number Variants

Copy number variants (CNVs) are deleted or duplicated genomic segments and are common in the human genome. Distinguishing between benign and pathogenic CNVs remains a challenge. The Database of Genomic Variants is a compilation of CNVs observed in control

samples and is helpful as a tool to identify CNVs less likely to be pathogenic (Iafrate and others 2004; Redon and others 2006; Wong and others 2007; Zhang and others 2006). A study of 500 control individuals identified an average of 3 CNVs per genome with a median CNV size of 224 kb and 16% of CNVs were larger than > 1 Mb in size (Marshall and others 2008). Another genome-wide study found that 5-10% of individuals have CNVs larger than 500 kb and 1-2% of individuals have variants > 1 Mb (Itsara and others 2009). Individual CNVs > 100 kb are rare and it was found that the greater the gene density in a CNV and the greater the CNV size the smaller the frequency in the population (Itsara and others 2009). The authors found a 25-fold enrichment of CNVs between pairs of segmental duplications indicating that nonhomologous allelic recombination may be a mechanism underlying their origin. The most CNV-dense region on Chromosome 22 is 22q11.2 which contains segmental duplications and leads to DiGeorge Syndrome (Shaikh and others 2007). In comparison with 22q11, 22q13 is relatively poor in segmental duplications (Bailey and others 2002).

CNVs on chromosome 22 are common. In a study of 1654 pediatric patients tested because of developmental delay, ASD, seizures, dysmorphic features or congenital anomalies, 1298 benign chromosome 22 CNVs were observed (how these CNVs were determined to be benign was not provided). The same study identified 25 patients (1.5%) who had an abnormal chromosome 22 CNV (Yu and others 2011). Of the 25 with abnormal CNVs, 22 were in 22q11 region and three were in 22q13 (22q13.2qter, 22q13.31qter, 22q13.32qter).

CNVs may be associated with ASDs. In a study comparing 996 ASD patients and 1287 controls, the patients with ASD were found to have 19% more CNVs than controls and were 69% more likely to have a CNV > 500 Kb in size (Pinto and others 2010). In another study, similar numbers of CNVs were observed in 427 ASD patients as in 500 controls (Marshall and others

2008). The study identified some CNVs among patients which were not present in their controls or in the Database of Genomic Variants (Iafrate and others 2004) including two in 22q13.31 and one in 22q13.33.

Mechanisms Generating and Repairing Chromosomal Breakage

Two of the most common mechanisms for generating rearrangements are non-allelic homologous recombination (NAHR) and non-homologous end-joining (NHEJ) (Lupski and Stankiewicz 2005; Stankiewicz and Lupski 2002a).

Non-Allelic Homologous Recombination (NAHR). NAHR occurs in regions with segmental duplications/low-copy repeats (LCRs) (Lupski and Stankiewicz 2005; Stankiewicz and Lupski 2002b). These segmental duplications/LCRs are genomic segments of 10-400 Kb with > 95-97% similarity, which occur to varying degrees in all human chromosomes. During meiosis, misalignment of these highly similar, but non-homologous, regions allows for crossing over between sister chromatids. The result of NAHR is either duplication or deletion of the chromosomal region between the segmental duplications. Some of the best studied examples of NAHR occur on chromosome 17p11.2 where a recurrent 1.4 Mb deletion leads to hereditary neuropathy with liability to pressure palsies and the 1.4 Mb duplication leads to Charcot-Marie-Tooth disease type 1A (Lupski and Stankiewicz 2005). Smith-Magenis syndrome (SMS), associated with a recurrent 3.7 Mb deletion syndrome, is another syndrome associated with NAHR leading to a recurrent deletion (Elsea and Girirajan 2008). 22q11.2 deletion syndrome, also known as DiGeorge syndrome or velocardiofacial syndrome is likewise caused by NAHR leading to common deletions of 1.5 and 3 Mb (Stankiewicz and Lupski 2002a). The syndrome is characterized by developmental delay and speech delay along with cardiac and other anomalies (Kobrynski and Sullivan 2007). CNVs are especially common on 22q11 where numerous LCRs are

located (Yu and others 2011). Unlike 22q11, 22q13 has few LCRs, which explains the lack of common deletion sizes observed in PMS as shown in Figure 1.2. Thus, while PMS is a deletion syndrome like those described above, the mechanisms causing PMS are different.

Non-Homologous End Joining (NHEJ). In NHEJ, a double strand break is repaired by capturing the two broken ends, forming a synaptic complex to bridge the break. Overhangs are either filled in or removed causing gain or loss of genomic material, and then the two strands are ligated together (Weterings and Chen 2008).

Mechanisms of Chromosomal Deletion and Repair in PMS

The mechanisms by which 22q13 deletions were repaired have been studied in detail (Bonaglia and others 2011). Terminal deletions can be repaired by the synthesis of a new telomere using telomerase, by recombination, by telomere capture of another telomere leaving behind a derivative chromosome, and by circularization resulting in a ring chromosome. In their study, they sequenced the breakpoint regions for more than 40 individuals with PMS with terminal deletions, interstitial deletions, ring chromosomes, and derivative chromosomes (Bonaglia and others 2011). For terminal deletions, they found evidence of both telomere healing (Flint and others 1994) and telomere capture (Meltzer, Guan, Trent 1993). Frequently, but not always, they found repetitive elements near the breakpoints and suggest these regions may be more susceptible to double strand breaks (Hannes and others 2010; Zhao and others 2010). Telomere healing is detected by the presence of TTAGGG repeats at the breakpoint (Lamb and others 1993). In the case of telomere capture, frequently subtelomeric, in addition to telomeric, sequence is added (Lamb and others 1993; Meltzer, Guan, Trent 1993). In the Bonaglia study, they were able to identify the source of one captured telomere as being

homologous to Xp/Yp (Bonaglia and others 2011). In the cases of interstitial deletions, ring chromosomes, and translocations, they find most are consistent with NHEJ repair mechanisms.

History of 22q13 Deletion Syndrome

The history of Phelan-McDermid syndrome follows the development of improved cytogenetic testing methods and an evolving understanding of the role of cytogenetics in the pathophysiology of developmental delays. The first documentation of a chromosome 22 terminal deletion occurred in 1985 with a case report of an individual with intellectual disability and dysmorphic features who had a pericentric inversion of chromosome 22 resulting in deletion of the 22q12->qter (Watt and others 1985). This individual possessed a much larger deletion (22q12->qter) than is typically found (22q13.3) and was detected using chromosome staining techniques. The size of the deletion was not assessed. Over the next seven years, five more cases were reported in the literature of individuals with large 22q13 deletions and presenting with developmental delay, speech delay, hypotonia, and minor dysmorphic features (Herman, Greenberg, Ledbetter 1988; Kirshenbaum, Chmura, Rhone 1988; Narahara and others 1992; Phelan and others 1992; Romain and others 1990). In 1994 the first case series examined seven patients with 22q13.3 deletions and compared findings to previously reported cases. The cases were noted for developmental delay, hypotonia, severe delays in expressive speech, minor dysmorphic features, and normal or accelerated growth (Nesslinger and others 1994). The first time a 22q13.3 deletion patient had molecular characterization to identify the breakpoint and obtain deletion size was in 1997 when a patient was found to have a terminal deletion of 130 kb and who presented only with intellectual disability and speech delay (no

dysmorphic features) (Flint and others 1995; Wong and others 1997). Interestingly, this patient was identified from a cohort of 99 patients with intellectual disability of unknown cause for whom cytogenetic testing was performed. Based upon the screening study, the authors estimated that 6% of unexplained intellectual disability (those with normal routine karyotypes and no recognizable syndrome) possessed chromosomal anomalies. The finding of the individual with a small deletion of 130 kb indicated that the critical deletion region was the most distal portion. At the time, *SHANK3* had not been mapped and only *ACR* was mapped to the region.

A case report of a 22q13.3 deletion in an autistic patient expanded the phenotypes of concern to include the autism spectrum (Goizet and others 2000). This patient was identified when she was tested for DiGeorge syndrome (22q11.2 interstitial deletion) using FISH (fluorescence in-situ hybridization) and the control probe located distal to 22q13.3 was found to be deleted. The deletion size was not characterized. Deletion of the control probe for DiGeorge syndrome later became another common route for the diagnosis of 22q13.3 deletion syndrome.

Dr. Katy Phelan, a cytogeneticist who had identified some of these early cases, formed a Deletion 22q13 Support Group which met for the first time in 1998 and led to biannual meetings where patients and their families, clinicians, and researchers could meet to improve their understanding of this condition. This organization is now called the Phelan-McDermid Syndrome Foundation (www.22q13.org) which continues to support work on PMS. Researchers at the Greenwood Genetic Center developed standardized study protocols to assess medical history, obtain physical exams, and perform high resolution genotyping on a cohort of patients. Data from this effort are provided in chapters 2-4 and in the medical literature (Phelan,

Stapleton, Rogers 2010; Phelan and others 2001; Phelan, Brown, Rogers 2001; Phelan 2008; Rollins and others 2011; Sarasua and others 2011).

The first large series of cases was described in 2001 when information from 37 individuals was collected (Phelan and others 2001). The most common features of the syndrome were global developmental delay, hypotonia, absent or severely delayed speech, normal to accelerated growth, and minor dysmorphic features. Of the 37 cases, 29 (78%) had terminal deletions while 8 (22%) were the result of unbalanced translocations. The size of deletions was not provided. Following this study, a number of other investigations from across the globe have been conducted (Bonaglia and others 2011; Dhar and others 2010; Jeffries and others 2005; Koolen and others 2005; Lindquist and others 2005; Luciani and others 2003; Manning and others 2004; Philippe and others 2008; Wilson and others 2003).

While most deletions are simple terminal deletions, other chromosomal aberrations have been observed including ring 22 whereby genomic material is missing from the ends of the long and short arms of the chromosome which then fuse together (Battini and others 2004; De Mas and others 2002; Jeffries and others 2005; Luciani and others 2003; Manning and others 2004; Phelan and others 2001). While material may be missing from the short arm of this acrocentric chromosome, other than ribosomal RNA, no known functional material is located in this region and the clinical features are likely the result of missing 22qter material or mitotic instability of the ring chromosome (Guilherme and others 2011; Kosztolanyi 1987; Sigurdardottir and others 1999). Another presentations is as an unbalanced translocation with another chromosome, resulting in the loss of the distal portion of chromosome 22q13 (Bonaglia and others 2001; Luciani and others 2003; Manning and others 2004; Phelan and others 2001). Phenotypes observed in translocations may arise due to loss of genomic material from

22q13qter, gain of material from the other chromosome involved in the translocation, and changes in spatial location altering gene expression (Harewood and others 2010). Other presentations include inversions (Tagaya and others 2008), duplications due to translocations (Gajecka and others 2008; Jamsheer and others 2008), duplications and deletions (Koolen and others 2005; Lindquist and others 2005), mosaics (Bonaglia and others 2009; Jeffries and others 2005; Phelan, Brown, Rogers 2001; Phelan and others 2001), dicentric chromosomes (Babineau and others 2006), and even interstitial deletions (Wilson and others 2008).

A number of reviews of PMS have now been written (Bonaglia and others 2010; Cusmano-Ozog, Manning, Hoyme 2007; Havens and others 2004; Phelan, Stapleton, Rogers 2010; Phelan and McDermid 2012; Phelan 2008).

Testing Procedures

Currently, a chromosomal microarray is the recommended first tier test for patients with developmental disabilities or congenital anomalies (Miller and others 2010). Microarrays are able to detect small microdeletions and microduplications that may not be detectable on karyotypes. Array CHG became a more common detection method in the PMS literature after 2005 (Koolen and others 2005; Lindquist and others 2005). However, karyotypes and FISH are still useful for the identification of mosaicism and chromosomal rearrangements.

Most PMS cases are identified after postnatal testing reveals the 22q13 anomaly in a child or adult with developmental delays or dysmorphic features; however, prenatal testing has also identified 22q13 deletions. Several cases have been reported from amniocentesis following abnormal maternal serum screening (Koc and others 2009; Phelan, Brown, Rogers 2001), an amniocentesis performed due to malformations in a prior pregnancy (Chen and others

2005), or amniocentesis performed following abnormalities observed on prenatal ultrasounds (Koc and others 2009; Maitz and others 2008).

Parent of Origin

While almost all 22q13 deletions are *de novo* mutations, the affected chromosome is commonly paternal in origin. Estimates of paternal origin from several studies are as follows: 74% (Luciani and others 2003); 69% (Wilson and others 2003); 59% for ring chromosomes (Jeffries and others 2005); and 74% of terminal deletions, 100% for interstitial deletions, and 60% for ring chromosomes (Bonaglia and others 2011). In a study of 35 individuals with ring chromosomes, no differences in ages between parents transmitting or not transmitting an affected chromosome were observed (Jeffries and others 2005). Further, no differences in phenotypes were observed based on parent of origin implying a lack of imprinting (Jeffries and others 2005; Luciani and others 2003). One case of maternal uniparental disomy 22 has been reported without clinical features of PMS (Schinzel and others 1994). Assessment for possible imprinting is important as it has been found to be important in other syndromes such as Angelman and Prader-Willi. In these two syndromes, deletion or mutation of 15q11 results in distinct syndromes depending on the parent of origin (Buiting 2010). While parent of origin effects have not been identified in PMS, if present, they could impact analysis of genotype-phenotype correlations if not taken into account.

How Common Is Phelan-McDermid Syndrome?

The frequency of 22q13.3 deletions is unknown. Phelan and others estimated the incidence of PMS to be approximately 1 in 11,000 to 15,000 individuals (Phelan, Stapleton, Rogers 2010). They derived this estimate based upon two studies of ASD which found *SHANK3* deletions or mutations in 1% (Moessner and others 2007) and 1.4% (Durand and others 2007) of

cases. They multiplied these values with an estimated ASD prevalence of 1/150 to obtain their estimates. One could also make estimates based upon the frequency of observed 22q13.3 deletions in those with intellectual disability (ID) as almost all PMS cases have at least mild ID. In a study of 32,587 individuals with developmental delay, array CGH identified 59 (0.18%) as having a 22q13 deletion (Girirajan and others 2012). These 59 patients represented 2.55% of the 2,312 found to have a CNV in the large cohort of 32,587. No 22q13 deletions were found in 8,329 controls. In a study of 234 unexplained ID cases in China, 4 (1.7%) were found to have 22q13.3 deletions (Gong and others 2012). In their review of the literature, 0.24% of unexplained ID cases had 22q13 deletions using microarray studies (Gong and others 2012). Two other studies of ID found 22q13.3 deletions in 1 out of 95 cases (Hamdan and others 2011) and 1 out of 99 cases (Flint and others 1995). With estimates of the prevalence of ID of 1-3% (Leonard and Wen 2002; Roeleveld, Zielhuis, Gabreels 1997) and taking a range of 0.18-1.7% for 22q13.3 deletions in ID, one could estimate a range of 22q13.3 deletion frequencies of approximately 1 in 2,000 to 55,000. Clearly, the estimates vary widely and the actual rate of 22q13.3 deletions is unknown.

To better interpret the pathogenicity of 22q13.3 deletions, it is useful to determine the prevalence of these CNVs in unaffected populations. A study of 8,329 controls found no 22q13 deletions using array CGH (Girirajan and others 2012). The Database of Genomic Variants (DGV) (Iafrate and others 2004; Redon and others 2006; Wong and others 2007; Zhang and others 2006) is a compilation of CNVs observed in presumably healthy controls. The DGV contains several 22q13.33 deletions encompassing *SHANK3*. Moessner and others noted no CNVs encompassing *SHANK3* in 500 European controls or DGV HapMap samples using a high density 500k array (Moessner and others 2007). They reported that 10 22q13.33 deletions listed in the

DGV that were identified using BAC arrays were found to be false positive after use of quantitative PCR validation (data not shown) (Moessner and others 2007). Park and others conducted a high density genome-wide CNV study of 30 Asian individuals in the HapMap project and report a 24 Kb CNV disrupting *SHANK3* in one sample (Park and others 2010). Whether this finding has been validated is unknown. No other CNVs were observed in the DGV that include *SHANK3* exons. A study of ASD that examined more than 2500 controls identified 2 CNVs disrupting *SHANK3* (Glessner and others 2009). The DGV displays additional CNVs, primarily <100 Kb in size, scattered across 22q13.

The evidence from different studies can support or refute a 22q13.3 deletion being present in the normal population and estimates of the frequency of these deletions vary widely. Association between specific CNVs and specific clinical features is needed.

Interstitial Deletion

While the focus of investigation of PMS has been on terminal deletions of 22q13.33, and *SHANK3* in particular, several cases of interstitial deletions of 22q13 indicate that genomic regions centromeric to *SHANK3* may be involved. In the first case, a girl with a deletion of 22q13.1-q13.2 had developmental delay, including speech delay, hypotonia, and minor dysmorphic features (Fujita and others 2000). Two other patients with interstitial deletions centromeric to *SHANK3* had developmental delay and speech delay with deletions of 22q13.1-q13.2 (40.42-44.00 Mb) and 22q13 (41.22-45.37 Mb, 2006 Genome build), or approximately 4-5 Mb proximal to *SHANK3* (Wilson and others 2008).

Sporadic and Unique Phenotypic Features of PMS Case Reports

Most studies of PMS note neurologic, growth, and dysmorphic features of the syndrome. However, additional clinical features in PMS have been reported. These features

include central diabetes insipidus which resolved itself in a two year old girl with a 22q13.31 deletion (Barakat and others 2004). Two cases of autoimmune hepatitis requiring liver transplants have been reported. One case occurred in a 7 year old girl with a 1.535 Mb deletion of 22q13.3 (Tufano and others 2009) and another in a 4 year old girl with 22q13.31 deletion of 5.675 Mb (Bartsch and others 2010). Bartsch and others noted that the child showed developmental improvement after the transplant and suggest that chronic hepatic disease could contribute to developmental delay in some PMS patients (Bartsch and others 2010). This observation is intriguing as the bulk of work on developmental delay focuses on brain-expressed genes rather than other organ systems such as the liver. A case report described a case of an atypical teratoid/rhabdoid tumor in a girl with 22q13.3 deletion syndrome (Sathyamoorthi and others 2009). In a study of sudden infant death, one out of 27 cases examined had a 4.4 Mb deletion of 22q13.3 deletion and a 3 Mb duplication of chromosome 8q (Toruner and others 2009). Two brothers with identical 2.15 Mb 22q13.32q13.33 deletions had atypical bipolar disorder, in addition to intellectual disability and developmental delay of speech and language, and sleep disturbance (Verhoeven and others 2012). Two patients with intellectual disability and dysmorphic features with 22q13 deletions were found to have metachromatic leukodystrophy, caused by low ARSA levels (Bisgaard and others 2009). Their 22q13 deletions, in addition to *SHANK3* deletion, included deletion of *ARSA*, and their remaining *ARSA* allele was either a pseudo allele or had a pathogenic mutation. This deletion led to low ARSA activity which led to the recessive disease of metachromatic leukodystrophy (MLC). MLC presents with mental retardation, dysmorphic features and low ARSA activity (Bisgaard and others 2009). This is an example of a deletion unmasking a recessive allele. Interestingly, in the case of 22q13, little work has been done to sequence the remaining copy in patients with 22q13 deletions in

order to determine whether other phenotypes may be related to this phenomenon of a deletion unmasking a recessive allele.

PMS Clinical Features

PMS is one of a number of syndromes presenting with hypotonia, developmental delay, speech delay and/or autistic traits including Prader-Willi, Angelman, Williams, Smith-Magenis, Fragile X, Sotos, FG, trichorhinophalangeal and velocardiofacial syndromes, ASD, and cerebral palsy (Phelan 2008). A challenge in conducting genotype-phenotype studies is that many clinical features have multiple causes and variable severity and some features are present in only a fraction of all cases. The variability in phenotypes in PMS may be due to differences in deletion sizes and thus gene content deleted or could be due to incomplete penetrance, variable genetic background, or interacting factors. A listing of common features observed in PMS is provided in Table 1.1. A description of some of these presenting signs follows.

Developmental Delay

Developmental delay includes delayed attainment of typical milestones with gross or fine motor skills, language, social, and cognitive abilities that occur in childhood (Moeschler, Shevell, American Academy of Pediatrics Committee on Genetics 2006). Mild delays may resolve with age and may not be related to other life-long developmental disabilities including intellectual disability (ID). Intellectual disability cannot be diagnosed reliably until a child is at least 5 years of age and thus the term “developmental delay” is sometimes used in place of ID or a more specific diagnosis (Moeschler, Shevell, American Academy of Pediatrics Committee on Genetics 2006). Clinical evaluations of developmental delay or intellectual disability include clinical and family history, physical examination including dysmorphology and neurologic

examination, genetic testing, and targeted metabolic testing and brain imaging (Moeschler, Shevell, American Academy of Pediatrics Committee on Genetics 2006). Chromosomal microarrays (array CGH or SNP microarrays) are now the recommended first tier test for the diagnosis of developmental disabilities or congenital anomalies (Miller and others 2010). Developmental delay is observed in nearly all cases of PMS (Phelan, Stapleton, Rogers 2010).

Intellectual Disability

The American Association of Intellectual and Developmental Disabilities defines intellectual disability as characterized by limitations in intellectual functioning and adaptive behavior (www.aamr.org). It is measured with an IQ score with scores less than 70 indicating ID. The prevalence of ID is approximately 1-3% in the United States (Leonard and Wen 2002; Roeleveld, Zielhuis, Gabreels 1997). Approximately 10% of individuals with ID also have autism (Oeseburg and others 2011). Other common chronic conditions include epilepsy (22%), cerebral palsy (20%), anxiety disorder (17%), oppositional defiant disorder (12%), and Down syndrome (11%). Intellectual disability may be syndromic, appearing in conjunction with other clinical features, or non-syndromic, appearing in isolation. Intellectual disability and autism are highly interrelated (Schwartz and Neri 2012). ID and ASD frequently have similar implicated genes and share common pathways (Betancur, Sakurai, Buxbaum 2009; Betancur 2011; Kou and others 2012). More than 200 ID genes have been identified to date (Betancur, Sakurai, Buxbaum 2009; Kou and others 2012). ID genes are involved in synapses, glutamate signaling, cell adhesion, RHO pathway, synaptic vesicle trafficking and exocytosis, the ERK/MAP pathway, Zinc finger proteins, transcriptional regulation and chromatin remodeling (Kaufman, Ayub, Vincent 2010).

Autism Spectrum Disorders

Autism spectrum disorders (ASDs) are developmental disabilities characterized by deficits in communication and social interaction, and by repetitive or restricted interests and behaviors (American Psychiatric Association and American Psychiatric Association. Task Force on DSM-IV 2000). Subtypes of ASD include autistic disorder, childhood disintegrative disorder, Asperger syndrome, and pervasive developmental disorders, not otherwise specified. Approximately 1 in 88 children under age 8 years are estimated to have an autism spectrum disorder (Centers for Disease Control and Prevention 2012). Approximately 70% of individuals with autism also have ID (Schwartz and Neri 2012). PMS is just one of many genetic syndromes that include ASD (Cohen and others 2005). The prevalence of ASD in PMS is unknown, with estimates ranging from 0-94% (Sarasua and others 2011).

Speech and Language Delay

Absent or delayed speech is one of the hallmarks of PMS affecting more than 95% of individuals (Phelan 2008; Wilson and others 2003). Relatively little is known about the genetics of speech. *FOXP2*, a transcription factor, is the best studied gene associated with speech (Fisher and Scharff 2009; Konopka and others 2009; Spiteri and others 2007; Vernes and others 2007; Vernes and others 2008; Vernes and others 2011). *FOXP2* is associated with dyspraxia, which causes difficulties in the physical production of sound. No chromosome 22q13 targets were identified for FOXP2 (Konopka and others 2009; Spiteri and others 2007; Vernes and others 2007). Additionally, *CNTNAP2, CMIP, RIT2, ATP2C2*, and *SYT4* are potential speech related genes (Alarcon and others 2008; Bouquillon and others 2011; Newbury, Fisher, Monaco 2010). These other genes are associated with memory and learning related speech disorders. Speech abilities vary widely in PMS ranging from absent speech to being completely verbal. A range in speech

ability is also seen in those with intellectual disability and autism spectrum disorders (Flax and others 2010; Hu and Steinberg 2009; Hu and others 2009; Waga and others 2011).

In Rett syndrome, a neurodevelopmental disorder caused by mutations in *MECP2*, 45% of patients could speak no words, 55% had some words, and 14.5% could use two-word sentences. None had more than 40 words total (Uchino and others 2001). This level of speech is somewhat similar to that seen in PMS although the highest levels of speech in PMS include the use of thousands of words and being fully verbal (Sarasua and others 2011). In contrast, most typically developing children have at least 50 words and multi-word combinations by the time they reach two years of age (Rescorla 1989). In a study of two-year olds, 9.7% were language delayed as defined as having fewer than 50 words or no word combinations on the Language Development Survey (Rescorla and Alley 2001). Specific speech-language impairment, which is language impairment without cognitive delay, hearing impairment, autism, or other medical condition, is present in 5-8% of preschoolers (Newbury, Fisher, Monaco 2010). Language delay shows 70% heritability in twins showing the most delayed language (Dale and others 1998).

Hypotonia

Hypotonia, and neonatal hypotonia in particular, is another hallmark feature of PMS (Phelan 2008). Hypotonia is low muscle tone which is defined as resistance to stretch of the muscles as opposed to muscle strength (Hill 2005). While hypotonia is rare in newborns, it is a common presenting sign for systemic and nervous system diseases (Leyenaar, Camfield, Camfield 2005) and is one of the indications a newborn should be tested for PMS. In general, infants with hypotonia are more likely to have hypotonia due to a central nervous system cause (66%) than peripheral nerve cause (34%). The most common diagnoses of infants with hypotonia are hypoxic-ischemic encephalopathy (20%), intracranial hemorrhage (8%),

chromosomal abnormalities and syndromic disorders (22%) including Prader-Willi syndrome, Down syndrome, and other chromosomal abnormalities, brain malformations (10%), metabolic disorders (8%), and peripheral hypotonia caused by muscle disorders or motorneuron/nerve disorders (Laugel and others 2008; Richer, Shevell, Miller 2001). Central hypotonia is more likely to be accompanied by seizures, facial dysmorphisms, and cognitive delays than peripheral hypotonia (Harris 2008). Cognitive delays (>95%), seizures (>25%), and dysmorphic features (>25-50%) are frequent in PMS (Phelan 2008). In PMS, neonatal hypotonia may resolve and not be present at older ages (Phelan, Stapleton, Rogers 2010).

Dysmorphic Features

Dysmorphic features and congenital anomalies often accompany hypotonia or developmental delay and their presence contributes to the differential diagnosis of specific disorders (Greenwood Genetic Center 2011; Moeschler, Shevell, American Academy of Pediatrics Committee on Genetics 2006). In a standard physical examination for a clinical genetic workup, many features are examined. Several databases, including the London Dysmorphology database (Winter 2009) and Phenomizer (Kohler and others 2009), allow the clinician to input dysmorphic and other clinical features to assist with identifying candidate genes or syndromes. Background prevalence rates are not generally available for dysmorphic features. Standard terminology has been developed along with photographs to assist in describing physical features of the head and face (Allanson and others 2009), periorbital region (Hall and others 2009), the ear (Hunter and others 2009), and hands and feet (Biesecker and others 2009). During the physical evaluations of participants in the Greenwood Genetic Center PMS studies, standardized assessments by trained clinical geneticists were performed on a

consistent set of physical features to be comprehensive and standardized across all study participants. Table 1.1 lists some of the common dysmorphic features observed in PMS.

Growth and Stature

Growth and stature are some of the most commonly measured parameters in pediatrics and adult medicine and growth reference charts are available (Greenwood Genetic Center 2011; Kuczmarski and others 2002; Rollins, Collins, Holden 2010; World Health Organization 2006). Using growth charts, short stature is typically classified as being below the 5th percentile and tall stature as being above the 95th percentile for a given gender and age. Head circumference measurements are typically classified as macrocephaly with a head circumference >97th percentile or < 3rd percentile. Human genome wide association studies have identified a number of quantitative trait loci (QTL) for human height, but only one located in the 22q13.2q13.33 region (Sammalisto and others 2005). Sammalisto and others identified marker *D22S282*, located in the *MPPED1* gene at genomic position 42.1 Mb linked to human height. This genomic region also includes a region identified as a body mass QTL in rats (Rat Genome Database, (Rapp 2000) accessed on the UCSC genome browser. Several rat and mouse studies have identified QTL for growth in the 22q13 orthologous regions. Early studies of PMS suggest that growth tended to be normal or accelerated, but later systematic analysis identified that growth tends to be normal, but twice the expected number are either short or tall and 20% have macrocephaly (Rollins and others 2011).

Comparison with Other Deletion Syndromes

Deletion syndromes are associated with all chromosomes and among those tested with intellectual disability, 1p deletions were the most common with 1q, 2q, 4p, 5p, 6q, 18q, and 22q also being common (Heilstedt and others 2003a). Most *de novo* terminal deletions are paternal in origin, although 1p36 deletions are usually maternal in origin (Heilstedt and others 2003a). Candidate genes or genomic regions are identified consistently across genotype-phenotype studies by locating the smallest genomic region of overlap among affected individuals. These deletion syndromes, including 22q13 deletion syndrome, commonly have developmental delay accompanied by dysmorphic features. Several of the better known terminal deletion syndromes include 1p36 deletion syndrome, 5p deletion syndrome and 18q deletion syndrome, described below. In common with PMS, these deletion syndromes are characterized by developmental delay including speech delay and intellectual disability, hypotonia, and minor dysmorphic features and all have varying degrees of severity within each syndrome. None of the *SHANK* genes are locate on these chromosomes.

Deletion 1p36 syndrome is the most common subtelomeric deletion observed, estimated to occur in approximately 1 in 5000 births (Heilstedt and others 2003a) and are found in 0.5 to 1.2% of individuals with unexplained intellectual disability (Battaglia and others 2008). Many of the clinical features of 1p36 deletion syndrome are similar to PMS. Common clinical features include universal developmental delay, intellectual disability, and absent or delayed speech. Speech was absent in 75% and in 17% cases could speak individual words, and finally in 8% patients had 2 word phrases (Battaglia and others 2008). Additional features include hypotonia, delayed growth, heart defects, renal and genital abnormalities, minor facial dysmorphism, seizures, and deafness, among others. In contrast, deafness and delayed growth

are not commonly observed in PMS (Phelan and McDermid 2012). Terminal deletions, interstitial deletions and more complex rearrangements are found (Battaglia and others 2008; Shapira and others 1997). The *de novo* mutations tended to be from maternally inherited chromosomes (60% maternal inheritance) (Heilstedt and others 2003b) rather than paternally inherited chromosomes as is the case for PMS (Phelan and McDermid 2012). Larger deletions tended to be associated with a greater number of clinical phenotypes and genomic regions of approximately 2-4 Mb in size are associated with selected clinical features (Heilstedt and others 2003b). More recently, an assessment of the smallest region of overlap of five individuals with interstitial deletions has identified a genomic region of 200-800 kb and suggested several potential candidate genes (Rosenfeld and others 2010).

18q- syndrome includes deletions of the terminus of 18q21. Similar to Phelan-McDermid syndrome, the syndrome is characterized by developmental delay and intellectual disability, hypotonia, seizures, genital anomalies, and minor facial dysmorphic features but also includes growth deficiency (Feenstra and others 2007; Kline and others 1993). Larger deletions were correlated with more severe clinical features. The use of array CGH and examination of deletions for regions of overlap for specific features identified specific genomic regions for microcephaly, short stature, congenital aural atresia, cleft palate, and intellectual disability, among others (Feenstra and others 2007). These results were based on 23 cases of terminal deletions (deletion sizes ranged from 6.1 to 27.0 Mb) and six cases of interstitial deletions (deletion sizes ranged from 7.2 to 24.0 Mb). These deletions are much larger than those typically observed in PMS (generally < 9 Mb (Wilson and others 2003). No common deletion breakpoints were observed, also similar to the PMS deletion. A number of candidate genes have

been proposed (Cody and others 2009; Feenstra and others 2007; O'Donnell and others 2010; Overhauser and others 1994).

Another deletion syndrome, deletion of 5p, also known as cri du chat syndrome, is named for a characteristic high pitched cry (Mainardi and others 2001). In addition, patients commonly have microcephaly and growth delay, speech delay, intellectual disability, behavioral problems, and facial dysmorphism. Similar to PMS, *de novo* deletions are primarily paternal in origin (90%) (Mainardi and others 2001). Inherited deletions have also been reported (Fang and others 2008). Deletions are usually terminal, but also include interstitial deletions and deletions accompanied by 5p duplications (Zhang and others 2005). Deletion sizes are variable, range up to 37 Mb (Zhang and others 2005), much larger than those observed in PMS (Phelan and McDermid 2012). Deletion size is correlated with select clinical features and a genotype-phenotype map has been established based upon smallest common regions of overlap observed among cases (Mainardi and others 2001; Zhang and others 2005) although genetic background or other factors may also impact severity of clinical features (Fang and others 2008). By examining regions of common overlap, genomic regions have been associated with speech, severity of intellectual disability, and facial features (Gersh and others 1995; Overhauser and others 1994; Zhang and others 2005).

Challenges in Defining Phenotypes in Human Deletion Syndrome Studies

Challenges in understanding the genetic contribution to specific phenotypes include accurately defining phenotypes of study. Particularly during the early stages of investigations, phenotypes can be subjectively or inconsistently recorded. For instance, in studies of PMS,

manuscripts frequently classify phenotypes as present or absent (+/-) or combine categories of phenotype. Individuals with autism spectrum disorders may have extremely different phenotypes, but for research purposes are often combined blurring potentially important distinctions. When more detailed information is available, such as using the full set of variables obtained in the Autism Diagnostic Interview–Revised, investigators have been able to distinguish individuals with known genetic causes of autism from idiopathic autism (Bruining and others 2010) or distinguish subtypes based on differential gene expression (Hu and others 2009; Hu and Steinberg 2009). Some individuals may have no verbal speech and others may be mildly delayed but are often grouped into categories of "absent/delayed speech," again blurring important distinctions. Growth has typically been grouped as "normal to accelerated" (Nesslinger and others 1994) giving the impression that accelerated growth was a distinctive feature without separating out normal from accelerated. Manuscripts provide clinical descriptions of patients, but may report an inconsistent set of phenotypes such that the reader does not know if lack of reporting a feature means the patient did not exhibit the feature, whether that information was not assessed, or whether that feature was not of importance to the authors at the time of writing. These issues are of particular importance for rare diseases and during the early stages of identifying clinical syndromes where case reports and case series make up the bulk of the literature. These concerns have been voiced for other deletion syndromes (Stewart and Kleefstra 2007). As research into a syndrome matures, the use of standardized data collection instruments and larger sample sizes can improve accuracy of studies. Strengths of other PMS studies (see Chapters 2-4) includes the use of a standardized health history questionnaire with the same set of clinical features assessed across cases, trained

clinical geneticists to conduct standardized physical exams, and instruments that assess severity of phenotype.

Biological Basis of the Neurological Phenotypes of PMS and Identification of *SHANK3* as a Candidate Gene

The primary clinical features of concern (developmental delay, speech and language delay, intellectual disability, autism spectrum disorders, seizures) are brain-related phenotypes. Hypotonia, another prominent feature, is primarily related to central nervous system lesions, but can also be related to peripheral neurological causes (Laugel and others 2008; Richer, Shevell, Miller 2001). Since many of the known genes located on 22q13.3 are expressed in brain and nerve tissue and have neurologic functions, they could be potential candidate genes for the syndrome. Because *SHANK3* is the gene deleted or disrupted in all cases of PMS, other than the instances of interstitial deletion (Wilson and others 2008), it is considered to be the primary candidate gene for many of the neurological features of PMS (Wilson and others 2003). The discovery of a translocation that disrupted *SHANK3* gave additional evidence for *SHANK3* being the candidate gene for PMS (Bonaglia and others 2001). Further, this study looked at gene expression and found that *SHANK3* is highly expressed in brain tissue and that the full length transcripts were only found in cerebral cortex and cerebellum. Substantial experimental evidence in mice, described below, also supports a role for *Shank3* in neurologic phenotypes.

While most cases of PMS have large deletions encompassing the entire *SHANK3* gene, some cases have been observed that deleted only portions of the gene. These cases include those deleting the distal portion of *SHANK3*: a balanced translocation with chromosome 12 with

a breakpoint within exon 21 of *SHANK3* (Bonaglia and others 2001), a translocation with Xq21.33 deleting the last two exons of *SHANK3* (Misceo and others 2011), and three terminal deletions with breakpoints between exons 8 and 9 (Anderlid and others 2002; Bonaglia and others 2006). There are also reported cases where the proximal end of *SHANK3* is deleted with the distal portion intact (Delahaye and others 2009). The differences in breakpoint location may affect whether the affected copy is a loss of function mutation leading to haploinsufficiency or a gain of function mutation which could interfere with expression, localization, or interaction with binding partners.

SHANK3 is a scaffolding protein in the post-synaptic density of excitatory neurons. *SHANK3* is being investigated as a candidate gene for ASD (Betancur, Sakurai, Buxbaum 2009; Durand and others 2007; Grabrucker and others 2011) and to a lesser extent Alzheimer's disease, schizophrenia, and ID (Grabrucker and others 2011; Verpelli and Sala 2011). Therefore, the role of *SHANK3* at the synapse will be presented here.

Synapse Morphology and Function

Synapse development and connectivity are necessary for normal brain function (Gong and Lippa 2010; van Spronsen and Hoogenraad 2010). During growth, development and the learning process, new connections between synapses are made and old ones pruned (called synaptic plasticity). As shown in Figure 1.3, neurons communicate by synapse formation between the presynaptic axon and the postsynaptic dendrite. Electrical stimulation causes the release of neurotransmitters from the presynaptic axon that are received by receptors of the postsynaptic dendrite called the post synaptic density (PSD). Cell adhesion molecules such as cadherin, neurexin, and neuroligin maintain the synaptic connection. Actin, SHANK, and HOMER proteins form the scaffold of the PSD. Neurotransmitter receptors such as NMDA, AMPA, and

mGluR receive chemical signals. Alterations in synapse morphology and function may lead to ID, ASD, Alzheimer's disease, schizophrenia, Parkinson's disease, compulsive behavior, and addiction (Gong and Lippa 2010; van Spronsen and Hoogenraad 2010).

SHANK Family of Proteins

The three SHANK proteins, SHANK1 (chromosome 19q13.33), SHANK2 (chromosome 11q13.3), and SHANK3 (chromosome 22q13.33) are found in the post-synaptic density (Bockers and others 2004; Boeckers and others 2002). The SHANK family (SH3 and multiple ankyrin repeat domains) is also known as ProSAP (proline rich synapse associated protein) and all three have been associated with ID or ASD (Berkel and others 2010; Grabrucker and others 2011; Sato and others 2012; Verpelli and Sala 2011) although SHANK3 has been more extensively studied. The expression levels of different SHANK proteins vary by tissue type within the brain (Bockers and others 2004). SHANK3 is the only one of the three to be associated with a known deletion syndrome.

Structure of SHANK3

SHANK3 is a protein coding gene made up of 58,572 nucleotides with 23 exons producing a full length protein of 1,747 amino acids [UCSC genome browser]. In addition, two smaller isoforms are expressed. The two longest transcripts are only expressed in brain while the shorter transcript is expressed in brain, liver, heart, kidney, and placenta (Bonaglia and others 2001). The gene is located on 22q13.33 near the telomere. Figure 1.4 shows the location, splice variants, and expression of *SHANK3* as displayed on the UCSC genome browser. It is primarily expressed in brain, heart, and liver tissue. Within the brain, it is highly expressed in the post-synaptic density of excitatory neurons (Boeckers and others 2002).

The SHANK proteins are considered 'master scaffolding molecules' of the post synaptic density by cross-linking with other scaffolding proteins, binding the N-methyl-D-aspartic acid (NMDA), metabotropic glutamate receptors (mGluR), and α-amino-3-hydroxyl-5-methyl-4-isoazole-pronionic acid (AMPA) receptor complexes (Sheng and Kim 2000), and providing structural support for cell adhesion molecules (Betancur, Sakurai, Buxbaum 2009). See Figure 1.5 for a schematic of the proteins located in the post synaptic density.

SHANK proteins contain five protein-protein interaction domains: an ankyrin repeat domain, an Src homology 3 (SH3) domain, a PSD-95/discs large/zonula occludens-1 (PDZ) domain, proline rich regions, and a C-terminal sterile α-motif (SAM) domain. Each of these regions has specific protein binding partners (Boeckers and others 2002). The ankyrin repeat domain binds αFodrin which in turn binds to the actin cytoskeleton (Bockers and others 2001) as well as binds Sharpin which may cross-link multiple copies of SHANK3 (Lim and others 2001). The PDZ domain interacts with GKAP/SAPAP to bind to NMDA-receptors and cell adhesion molecules (Boeckers and others 1999; Naisbitt and others 1999). The PDZ may also bind to CIRL/CI1 and SSTR2 (Kreienkamp and others 2000; Tobaben, Sudhof, Stahl 2000; Zitzer and others 1999). Specific regions in the proline rich domain interact with Dyanamin-2 (Okamoto and others 2001), Homer (Hayashi and others 2009; Tu and others 1999), and Cortactin (Du and others 1998) which may bind to the mGluR (mGlutamate) receptors and actin cytoskeleton. The Homer and Shank proteins are thought to form a mesh-like scaffolding network in the PSD (Hayashi and others 2009). Finally, the SAM domain is thought to bind to the SAM domain of other SHANK3 proteins, forming a scaffold (Naisbitt and others 1999). Further, the 417aa C-terminus region of SHANK3, including the SAM domain, is required to localize SHANK3 to the

post synaptic density (Boeckers and others 2005). Therefore, individuals with deletions of the C-terminus may have gain-of-function mutations in addition to loss-of-function or haploinsufficiency. Figure 1.6 provides a schematic of SHANK3.

SHANK3 Expression

SHANK3 is predominantly expressed in the brain and expression appears to be regulated with tissue-specific methylation (Beri and others 2007; Ching and others 2005). Promoter associated CpG islands were highly methylated in peripheral blood where *SHANK3* expression is low and unmethylated in brain tissue where expression is high (Beri and others 2007; Ching and others 2005). In addition, chromatin conformational differences were also observed between tissue types (Beri and others 2007). Further analysis of *SHANK3* found tissue-specific methylation regulates five intragenic promoters within *SHANK3* and identified two novel transcripts (Maunakea and others 2010).

Most research into *SHANK3* has focused on brain expression, but SHANK3 has been found to be a scaffolding protein at the sarcolemma (Grubb and others 2011), in epithelial cells in the gut where it facilitates host-pathogen interfaces (Huett and others 2009), and in signal transduction in the immune system (Redecker, Bockmann, Bockers 2006).

Observed deletion breakpoints within the *SHANK3* gene may have varying effects from null mutations where the transcript is not translated, to truncating mutations that cause gain-of-function mutations affecting protein-protein binding or localization.

SHANK3 in Intellectual Disability

While most work related to *SHANK3* has examined ASD, *SHANK3* mutations have been observed in cases of intellectual disability (Hamdan and others 2011). In this study of 95 cases of nonsyndromic intellectual disability, one missense truncating mutation was found upstream of the PDZ domain and Homer and cortactin binding sites in SHANK3. While not examining *SHANK3* mutations *per se*, other investigators have found 22q13 deletions (including deletion of *SHANK3*) in 0.18% to 1% of unexplained ID (Flint and others 1995; Girirajan and others 2012; Gong and others 2012).

SHANK3 Mutations Observed in ASD in Humans

SHANK3 mutations, intragenic deletions, and 22q13.3 deletions have been observed in 0 – 4% of ASD cases (Boccuto and others 2012; Durand and others 2007; Gauthier and others 2010; Glessner and others 2009; Moessner and others 2007; Sykes and others 2009; Waga and others 2011). In a study of postmortem brain tissue, 28 miRNAs were differentially expressed in cases with ASD compared to controls (Abu-Elneel and others 2008). *SHANK3* and *Neurexin*, a synaptic cell adhesion molecule, are predicted targets for some of these miRNA. A linkage study found no linkage peaks on chromosome 22 (Stone and others 2004). In two studies specifically looking for associations with autism and speech delays, no chromosome 22 SNPs were identified (Cho and others 2011; Flax and others 2010). The evidence remains mixed on the role of *SHANK3* mutations or deletions as observed in human observational studies.

Mouse Models of *Shank3* Knockout, Knockdown, and Haploinsufficiency

Animal models are often used to estimate effects of genetic mutations in humans and have the advantage of being able to control the specific mutation under study and to examine phenotypic endpoints in a controlled, standardized manner. A particular challenge for studies of

human syndromes with intellectual or behavioral features of concern is to identify measurable and standardized phenotypes in the animal model that are analogous to the human feature. While there are no human equivalents of autism, intellectual disability, or other psychiatric diseases in mice, there is a substantial literature on the use of quantifiable intermediate traits in mice to study these conditions (Crawley 2004; Seong, Seasholtz, Burmeister 2002; Silverman and others 2010). Quantifiable intermediate traits in mice can be used to approximate the measurable intermediate traits in humans. In addition to anatomic/biochemical phenotypes, behavioral and cognitive phenotypes can be measured in mice. For instance, assays have been developed to measure learning and memory (e.g. various maze tests), motor function (e.g. hanging from a wire or balancing on a rotarod), and social interaction (e.g. sniffing or interacting with other mice) (Crawley 2004; Seong, Seasholtz, Burmeister 2002; Silverman and others 2010). Thus, mouse models can be used to assess genetic effects on quantifiable intermediate phenotypes with analogs in humans. In PMS, the clinical features of intellectual disability, ASD, and hypotonia have measurable intermediate traits of learning and memory, speech and communication, social interaction, muscle tone, and neurologic properties of synapse function.

Functional studies of *Shank3* in mouse models support the hypothesis that *Shank3* is critical for neurological function. Various mouse models have been developed that knock out select isoforms or create a partial knock-out to mimic the deletion syndrome in humans. *Shank3* is located in a syntenic region on mouse chromosome 15 at genomic position 89 Mb on the 103 Mb chromosome (Mouse genome assembly July 2007 NCBI37/mm9). Unlike human *SHANK3*, mouse *Shank3* is not proximal to a telomere.

In a mouse model most closely resembling the 22q13.33 deletion, Bangash and others created a heterozygous mutation whereby one copy of *Shank3* was missing the C-terminus of

the gene (*Shank3* (+/ΔC)(Bangash and others 2011). The deleted portion contains the region that binds Homer. This region of the protein also deletes the critical region for targeting Shank3 to the PSD (Boeckers and others 2005). Bangash and others find that the truncated protein interacts with the wild-type protein with a resulting reduction of more than 90% of Shank3 at the synapse. Polyubiquitinization and relocalization of the wild type Shank3 to the proteasomes was observed. In addition, the NR1 subunit of the NMDA receptor was polyubiquinated and down regulated. No effects were observed on Shank1, Shank2, GKAP, and AMPA glutamate receptors. Synapse morphology and number were also unaffected. Electrophysiological studies showed a reduced NMDAR response and reduced NMDAR-dependent long term potentiation. The mGluR-long term potentiation was increased. Behavioral studies found *Shank3*(+/ΔC) to have deficits in social interaction, increased response to amphetamine and NMDA antagonists, and reduced NR1 expression. Learning and memory functions were not affected. This model supports the hypothesis that haploinsufficiency could produce the phenotypes observed in humans. It also supports the hypothesis that a truncated Shank3 protein may act in a dominant negative manner and be the cause of deleterious phenotypes rather than haploinsufficiency.

In another mouse model, the two primary *Shank3* isoforms were completely knocked out and the third isoform was reduced (Peca and others 2011). The model was created by targeting *Shank3*α exons 4-7 (the ankyrin repeat domains) and *Shank3*β exons 13-16 (the PDZ domain). Both isoforms were eliminated and *Shank3*γ was reduced by 42%. Knockout mice exhibited self-injurious repetitive behaviors and deficits in social interaction (Peca and others 2011). Further, in the knock-out mouse brain tissue the molecular composition of the PSD was found to be altered, the morphology of the PSD was altered, and striatal postsynaptic function was reduced (Peca and others 2011). This model demonstrated the role of Shank3 in the

mouse, but did not address the question of whether elimination of only one copy of Shank3, as would happen in 22q13.3 deletion syndrome, would result in clinical features.

Whereas Peca and others used a complete knockout mouse model, haploinsufficiency was observed in another mouse model (Bozdagi and others 2010). In this case, *Shank3* deletions of exons 4-9 were created. The full length transcript was eliminated in the full knockouts and reduced in the heterozygotes. Detection of transcripts containing the C-terminus and exons 19 and 20 were reduced in the heterozygotes (but not quantified). The heterozygous mice demonstrated deficits in social interaction and communications and reduced synaptic function. This model supports the hypothesis that haploinsufficiency of Shank3 could result in clinically relevant phenotypes.

In another model of *Shank3* exons 4-9 complete knock out, the two largest and most common transcripts were eliminated, but shorter isoforms were observed (Wang and others 2011). Knockout mice had abnormal social behavior, communication, repetitive behaviors and deficits of learning and memory. Further, reduced levels of Homer1b/c, GKAP, GluA1 were found in the PSD. The authors concluded that the behavioral changes were similar to human ASD patients. The finding of changes in learning and behavior contrasts to the findings of Bangash and others who did not find changes in learning and memory in mice with C-terminal deletions of Shank3 (Bangash and others 2011). By using a complete knock out model, the authors demonstrate the role of Shank3 in the mouse, but could not test whether missing one copy of the gene, as in the case of PMS, would cause similar phenotypes. Further, the knockout did not eliminate the shorter transcripts which contain the C-terminus needed for localization of the protein.

Finally, in a study that used RNAi to knock down *Shank3* expression in cultured mouse neurons, expression of mGluR5 receptors was reduced and synapse morphology was abnormal (Verpelli and others 2011). The RNAi knocked down expression of the major isoforms by 70-80%. This study also demonstrates the importance of *Shank3* in mouse models and supports the hypothesis that *SHANK3* may be haploinsufficient in humans, although the knock downs reduced expression by more than the 50% expected with deletion of one allele.

In summary, functional work with mouse models demonstrated behavioral, biochemical, and synapse morphological changes with disruption of Shank3. Given the multiple intragenic promoters (Maunakea and others 2010; Wang and others 2011) along with tissue-specific expression (Beri and others 2007; Ching and others 2005) and multiple methods of measuring phenotypes, it becomes difficult to parse out Shank3-domain-specific genotype-phenotype correlations. It appears that mutations within Shank3 could work in a dose-dependent manner (leading to haploinsufficiency) or as a gain-of-function dominant manner. With many interacting proteins, there is a strong possibility of deficits of Shank3 being attenuated or exacerbated by changes in interacting partners. Nonetheless, it is clear that *Shank3* mutations can have measureable and deleterious effects in the mouse brain.

The evidence is mixed on whether *SHANK3* dosage changes alone are the critical change in PMS. In most cases of PMS, the entire *SHANK3* is missing, but typically so are additional genes which have largely been unstudied. In some instances of PMS, the *SHANK3* gene has an intragenic deletion or truncation which could affect protein regulation, localization, dimerization, and interactions with protein-interaction partners. The literature on *SHANK3* mutations in autism describe the combination of missense mutations along with 22q13.33 terminal deletions making it difficult to parse whether the implied effect is due to dosage or

aberrant protein production. The mouse model that was developed to directly address the question of the effect of a C-terminus deletion of Shank3 suggests that truncating mutations could be more deleterious than a single null allele because the mutant protein interacts with the wild type protein, leading to >90% reduced Shank3 at the PSD (Bangash and others 2011). The question whether *SHANK3* deletions have the same phenotype as *SHANK3* mutations in humans remains unanswered.

While the majority of PMS research efforts to date have focused on *SHANK3*, other 22q13.3 genes may contribute to the phenotype. Evidence to support this hypothesis comes from a report of two patients with PMS clinical features who have an interstitial, rather than terminal, deletion of 22q13.3 (Wilson and others 2008). In these patients, the breakpoints are >4 Mb proximal to *SHANK3 (SHANK3* is intact*).* This finding of cases with a similar clinical presentation, but not missing *SHANK3*, suggests that there may be additional genes or genomic factors on 22q13.3 responsible for clinical features (Wilson and others 2008).

Position Effects

Independent of *SHANK3*-specific deletions or mutations, a chromosomal deletion can have position effects on the remaining genes. Deletions and rearrangements can affect the presence of or relationship between genes and regulatory elements (Buchanan and Scherer 2008). Since most of the deletions are terminal deletions, telomere position effects may be of particular relevance.

In a terminal deletion, the repair of a broken chromosome results in the placing of a newly healed or captured heterochromatic telomere in close proximity to a genomic region not normally proximal to a telomere. Telomere position effects (TPE), whereby gene expression is inhibited in genes in close proximity to a telomere, has been observed in human cells (Baur and

others 2001). In HeLa cells, gene expression was markedly reduced in genes placed adjacent to a newly formed telomere and expression was also reduced with increased telomere size (Baur and others 2001). The distance over which TPE may work is estimated to be up to 100 kb (Kulkarni and others 2010). TPE has been found to affect replication timing in the affected region (Ofir and others 1999; Smith and Higgs 1999). Interestingly, this effect was studied in a lymphoblastoid cell line from a patient with a 130kb terminal deletion of 22q13 previously reported (Wong and others 1997). Gene expression of *ARSA*, located 54 kb from the breakpoint showed both alleles were expressed, although expression levels were not quantified (Ofir and others 1999) and appear qualitatively different. The authors concluded that TPE did not extend that distance or the effects were incomplete. No mention was found in the research record of any further work being done to investigate TPE on 22q13 genes in deletion patients or to determine if TPE could be different in brain tissue where gene expression differs. It remains plausible that TPE could reduce expression levels of genes located within 100kb of a breakpoint. Future research might explore the TPE phenomenon on cell lines established from individuals with larger deletions or in neurons obtained from induced pluripotent stem cells of 22q13 deletion patients.

Beyond PMS, What Is Known About Phenotypes Associated With 22q13?

To investigate what is already known about 22q13 associated phenotypes, the Catalog of Published Genome Wide Association Studies (GWAS) (Hindorff and others 2009a; Hindorff and others 2009b) was examined (Table 1.2 and Figure 1.7). This database is curated by staff of the National Human Genome Institute who regularly review published GWAS studies and

include findings when studies include at least 100,000 SNPs and achieve a level of statistical significance of 10^{-5} or more significant. An advantage to using this catalog is the comprehensive curation and weekly updating from the literature. A drawback is that it only includes SNPs that achieve a P-value of 10^{-5} or more significant and it is possible that important associations may have been found at a lower level of significance in smaller scale studies. The stringent criteria are used to weed out the expected high number of false positives given the large number of SNPs examined in modern SNP panels (100,000 to more than 1,000,000 SNPs assessed per individual). However, in studies with small sample sizes, a true association may not be able to achieve such a high level of significance and will be omitted from this curation. Table 1.2 summarizes known GWAS associations with 22q13.2-22q13.33 genomic locations. Associations found in 22q13.1 are beyond the largest observed deletion observed in PMS and are not included. Traits related to temperament, narcolepsy, multiple sclerosis, attention deficit/hyperactivity disorder appear most relevant to PMS. However, as more information is collected on PMS patients, particularly in studies related to deletions beyond *SHANK3*, these traits can be revisited. No GWAS studies found SNPs associated with intellectual disability, autism spectrum disorders, hypotonia, or speech/language disorders on 22q13.2q13.33.

GWAS identified a large number of associations with a wide variety of phenotypes on 22q13. In region 22q13.2 there were SNPs associated with Alzheimer disease and prostate cancer. In region 22q13.31 there were SNPs associated with liver disease and liver enzymes and response to methylphenidate in children with attention deficit/hyperactivity disorder. The ADHD related finding points to an SNP in an intergenic region between genes *TBC1D22A* and *FAM19A5* (Mick and others 2008). Band 22q13.32 had one association with pancreatic cancer. In region 22q13.33, associations were found for narcolepsy, hematologic phenotypes,

natriuretic peptide levels, and colitis. The narcolepsy finding was for a SNP located at base position between genes *CPT1B* and *CHKB* (Miyagawa and others 2008). Sleep disturbance has been reported in PMS (Sarasua and others 2011).The SNP is located 24kb proximal to the gene *IB2* and 95kb proximal to *SHANK3*. There were no associations in the Catalog when searching for "SHANK3". There were no hits when searching terms "hypotonia", "language", "speech", "stature", "dysmorphic OR dysmorphia", or "macrocephaly", "head circumference", "intellect", "IQ", or "intellectual". A search for "height" identified several studies, including one on 22q13.1, but outside the PMS deletion region (Estrada and others 2009). "Intelligence" found one study, but found no SNPs associated. "Mental retardation", "syndactyly", "vision", "acuity", "development" and "brain" found hits but nothing on chromosome 22. Interestingly, searching for "autism" found only 5 published studies. No chromosome 22 hits were found and only hits on chromosome 15 were identified. This is a surprising finding given the large number of genes that have been found to be associated with autism, including *SHANK3*. It may be that GWAS methods are of limited use for multifactorial conditions such as autism. They also require highly significant findings which may be difficult to achieve even in the face of a truly causal gene.

Mouse and Rat QTLs in the 22q13 Orthologous Regions

Again using the UCSC genome browser to facilitate review of relevant features, mouse and rat QTLs (Blake and others 2011; Eppig and others 2012; Finger and others 2011; Rapp 2000) in the orthologous regions were identified. At base position 43.4-43.6 there is a mouse QTL related to methamphetamine response and may be related to brain function (Blake and others 2011; Eppig and others 2012; Finger and others 2011; Palmer and others 2005). Reviewing the Rat Genome Database for human QTLs identified regions associated with high blood pressure, chronic obstructive pulmonary disease (Rapp 2000; Twigger and others 2007).

Online Mendelian Inheritance in Man (OMIM)

OMIM (www.omim.org) is a curated database of human genes and phenotypes focusing on Mendelian disorders. OMIM identified genes and their associated disorders include *CYB5R3* (methemoglobinemia) in 22q13.2; *UPK3A* (renal adysplasia), *FBLN1* (polydactyly), *ATXN10* (spinocerebellar ataxia), *PPARA* (lipid metabolism), and *TRMU* (liver failure, deafness) in 22q13.31; and *ALG12* (congenital disorders of glycosylation), *MLC1* (megalencephalic leukoencephalopathy with subcortical cysts), *TUBGCP6* (microcephaly with or without mental retardation), *SCO2* (cardioencephalomyopathy), *TYMP* (mitochondrial disorder), *CHKB* (muscular dystrophy), *ARSA* (metachromatic leukodystrophy), *SHANK3* (Phelan-McDermid syndrome, Schizophrenia), and *ACR* (male infertility) in 22q13.33. These genes are spread across the deletion region and are potential candidates for contributing to the PMS constellation of features, particularly given reports of kidney and liver problems, developmental delay, ataxia, and hypotonia in PMS (Phelan and McDermid 2012).

Genotype-Phenotype Study Methods

One of the goals of this work is to identify genes located in 22q13.2q13.33 that are deleted in addition to *SHANK3* that may contribute to the presentation of the syndrome. The typical method used to conduct genotype-phenotype studies in deletion syndromes is to identify the smallest region of overlap in a collection of cases which typically have deletions of varying sizes (Mefford and others 2012; Molin and others 2012; Talkowski and others 2011). In the case of PMS, the smallest region of overlap contains the three most terminal genes including *SHANK3*, *ACR*, and *RABL2B*. All cases have developmental delay, absent or delayed speech, and almost all have hypotonia. The severity of the condition varies as does the appearance of secondary medical and physical features. As has been found in other conditions (Talkowski and

others 2011) the gene commonly deleted or disrupted may not be responsible for all features of a syndrome. A contribution to the literature would be to delineate the features contributed by *SHANK3* deletion and the features contributed by other genes.

Summary

Genes located on 22q13 remain sparsely annotated. Efforts to determine genotype-phenotype correlations in PMS will aid in understanding the independent and contributory effects of *SHANK3* and 22q13 genes. Further, what is learned in studying PMS will help in understanding the etiology of other developmental disability syndromes caused by mutations or copy number variants. For instance, as array CGH is increasingly used, increasing numbers of microdeletions and microduplications are being detected and new syndromes are being recognized. Increased numbers of identified cases with phenotypes accompanied by refinements in breakpoint identification mean that more phenotype-specific candidate regions will be identified. Concurrently, as candidate genes are identified and confirmed in better understood syndromes with similar phenotype, this information will help in identifying candidate genes within narrowed candidate regions. Examples of these deletion syndromes without identified causal genes include microdeletion syndrome 15q24 (Magoulas and El-Hattab 2012), 6p22.3-p24.3 (Celestino-Soper and others 2012), and 2q22.1 (Mulatinho and others 2012). Progress in other deletion syndromes such as Angelman and Prader-Willi have identified the critical genes in the deletion regions (Buiting 2010). Examining genes known to be involved in other developmental disability syndromes and their interactions with 22q13 genes will be as

useful to PMS research as incorporating what we learn about PMS into research on other syndromes.

Table 1.1. Common Features of PMS (as provided in (Phelan, Stapleton, Rogers 2010)).

Proportion with feature	Feature
➢ 95%	Neonatal hypotonia Global developmental delay Absent or severely delayed speech Normal or accelerated growth
➢ 75%	Large, fleshy hands Dysplastic toenails Long eyelashes Decreased perception of pain Mouthing/chewing behaviors
➢ 50%	Dolichocephaly Poorly formed/large ears Wide brow Full/puffy cheeks Full/puffy eyelids Ptosis Deep-set eyes Flat midface Wide nasal bridge Bulbous nose Pointed chin Sacral dimple Decreased sweating
➢ 25%	Strabismus Renal problems Gastroesophageal reflux Epicanthal folds Long philtrum High-arched palate Malocclusion/wide-spaced teeth Toes 2-3 syndactyly Cardiac defects Hypothyroidism Lymphedema Precocious or delayed puberty

Table 1.2. Summary of Findings on Chromosome 22q13 from the Catalog of Published Genome Wide Association Studies.

Source	Region	Chr_pos hg 19	Chr_pos hg 18	Disease/Trait	Reported Gene(s)	Context	p-Value
(Man and others 2012)	22q13.2	41042091	39372036	Treatment response for severe sepsis	*MKL1*	Intergenic	6.00E-07
(Kim and others 2011a)	22q13.2	42218856	40548801	Alzheimer's disease biomarkers	*CCDC134*	intron	1.00E-06
(Eeles and others 2009)	22q13.2	43500212	41830155	Prostate cancer	*NR*	Intergenic	6.00E-29
(Schumacher and others 2011)	22q13.2	43518275	41848218	Prostate cancer	*BIK*	intron	6.00E-06
(Chambers and others 2011)	22q13.31	44324727	42656059	Liver enzyme levels (alanine transaminase)	*PNPLA3, SAMM50*	missense	1.00E-45
(Speliotes and others 2011)	22q13.31	44324727	42656059	Nonalcoholic fatty liver disease	*PNPLA3*	missense	4.00E-34
(Kim and others 2011b)	22q13.31	44325996	42657328	Metabolite levels	*PNPLA3*	intron	2.00E-18
(Kim and others 2011b)	22q13.31	44325996	42657328	Metabolite levels	*PNPLA3*	intron	2.00E-39
(Yuan and others 2008)	22q13.31	44332570	42663902	Liver enzyme levels	*PNPLA3, SAMM50*	intron	8.00E-16
(Greenwood and others 2012)	22q13.31	45961904	44340567	Temperament-related traits	*FBLN1*	intron	2.00E-08
(Dolmans and others 2011)	22q13.31	46421842	44800505	Dupuytren's disease	*RP11û398F12.1, WNT7B*	Intergenic	3.00E-33
(Yashin and others 2010)	22q13.31	47532396	45911059	Longevity	*TBC1D22A*	intron	1.00E-06
(Mick and others 2008)	22q13.31	48284514	46663177	Attention deficit hyperactivity disorder	*Intergenic*	Intergenic	3.00E-06
(Wu and others 2011)	22q13.32	48929569	47308232	Pancreatic cancer	*FAM19A5*	intron	1.00E-10
(Del Greco and others 2011)	22q13.33	50086373	48472376	Natriuretic peptide levels	*BRD1*	Intergenic	7.00E-06
(Agrawal and others 2011)	22q13.33	50350971	48736974	Cannabis dependence	*PIM3*	Intergenic	8.00E-06
(Anderson and others 2011)	22q13.33	50435480	48777606	Ulcerative colitis	*PIM3, IL17REL*	missense	2.00E-07
(Franke and others 2010)	22q13.33	50435480	48777606	Ulcerative colitis	*IL17REL*	missense	4.00E-08
(Kamatani and others 2010)	22q13.33	50966914	49313779	Hematological and biochemical traits	*NCAPH2,SCO2,TYMP, KLHDC7B*	intron	4.00E-08
(International Multiple Sclerosis Genetics Consortium and others 2011)	22q13.33	50971266	49318131	Multiple sclerosis	*ODF3B*	nearGene-5	2.00E-08
(Ganesh and others 2009)	22q13.33	50971752	49318617	Mean corpuscular volume	*ECGF1*	nearGene-5	1.00E-15

(Miyagawa and others 2008)	22q13.33	51017353	49364218	Narcolepsy	*CPT1B*	nearGene-3;nearGene-5	6.00E-08
(Yashin and others 2010)	22q13.33	51104680	49451545	Longevity	*AC000050.2*	Intergenic	1.00E-06

[1]Source: (Hindorff and others 2009a; Hindorff and others 2009b), database accessed May 28, 2012

Figure 1.1 Region of Deletions Observed in PMS (22q13.2q13.33). The UCSC genome browser (Kent and others 2002) was used for this graphic using the 2006 (NCBI36/hg18) genome build (International Human Genome Sequencing Consortium 2004).

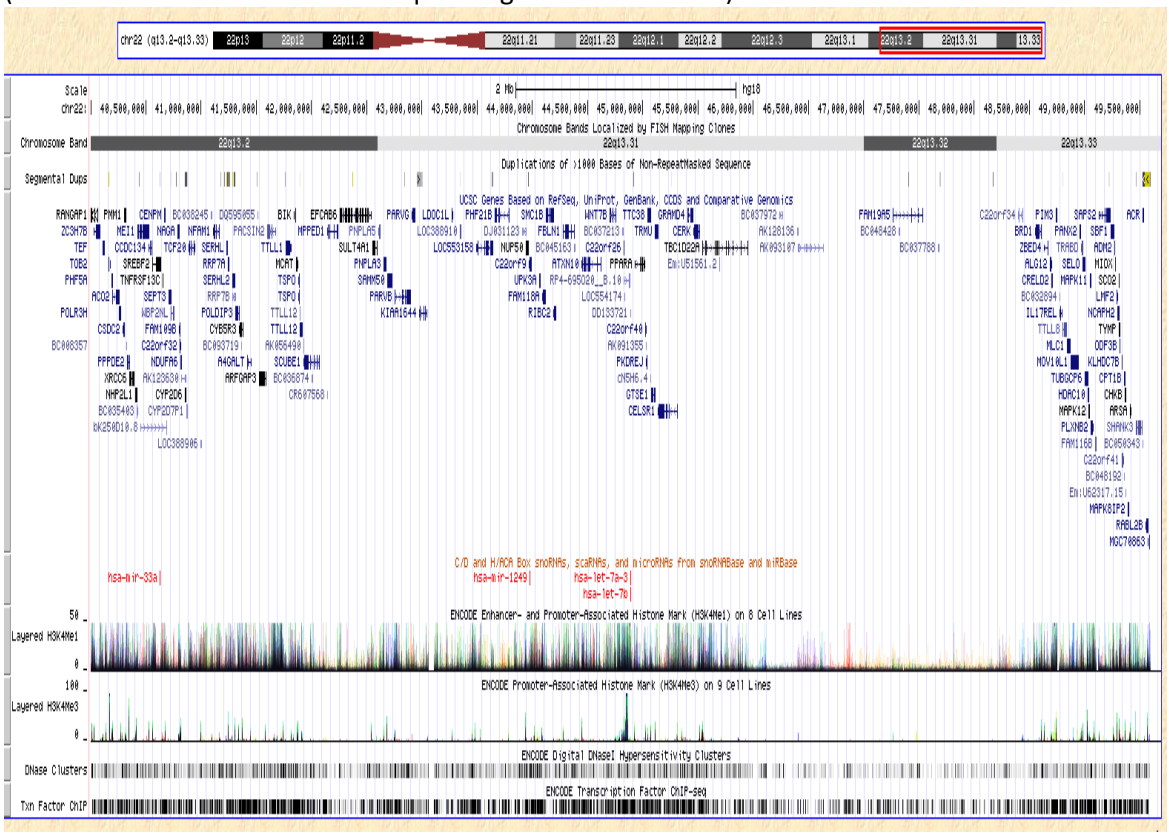


Figure 1.2 Deletions of 22q13.2q13.33 in Phelan-McDermid Syndrome. Terminal deletions are highly variable without common breakpoints in PMS. The UCSC genome browser (Kent and others 2002) was used for this graphic using the 2006 (NCBI36/hg18) genome build (International Human Genome Sequencing Consortium 2004). Reproduced from (Sarasua and others 2011) with permission from the BMJ Publishing Group Ltd.

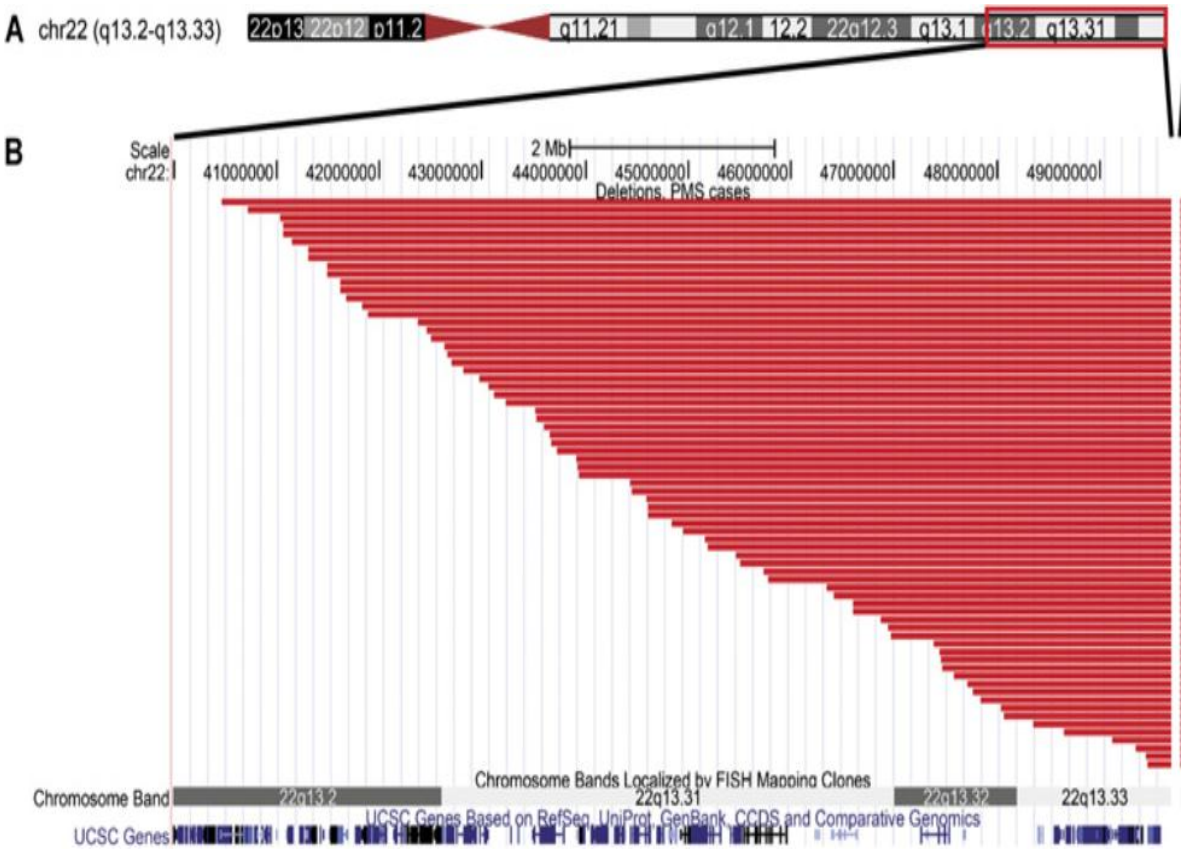


Figure 1.3. Schematic Depiction of the Synapse. Schematic shows key proteins including the cell adhesion molecules cadherin, neuroligin, and neurexin; signaling receptors AMPAR, NMDAR, and mGluR; and structural proteins in the post synaptic density Actin, Homer, and Shank. Figure provided with kind permission from Springer Science and Business Media. The figure originally appeared in Current Neurology and Neuroscience Reports, Synapse Pathology and Psychiatric Disease, Volume 10, 2010, page 208, Myrrhe van Spronsen and Casper C. Hoogenraad.

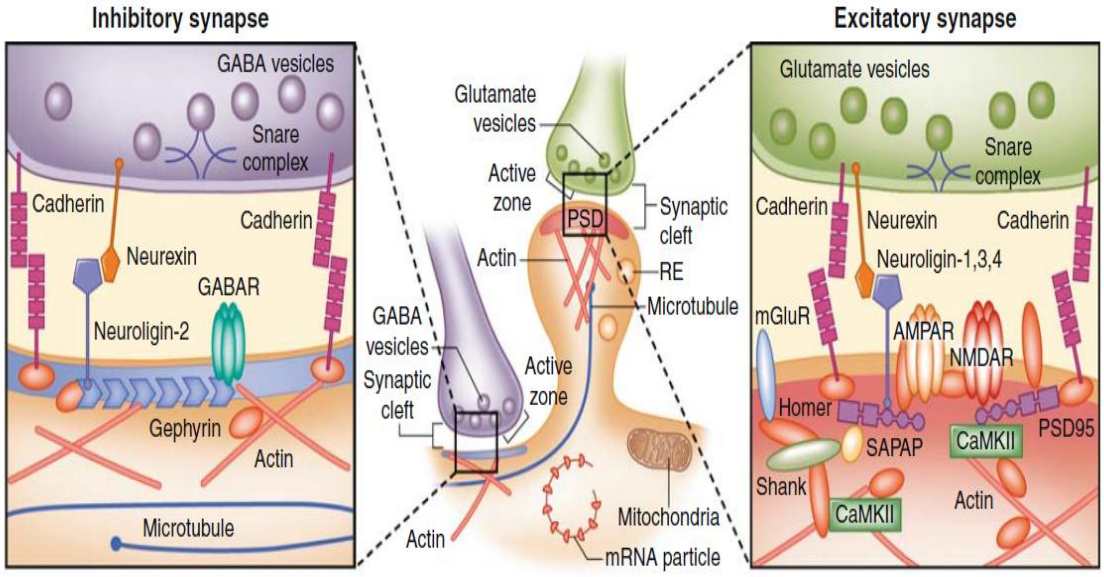


Figure 1.4. Location of *SHANK3* Splice Variants in 22q13.33, Proximity to the Telomere, Expression in Tissue Types, and Open Chromatin Regions from the UCSC Genome Browser (Kent and others 2002; Su and others 2004).

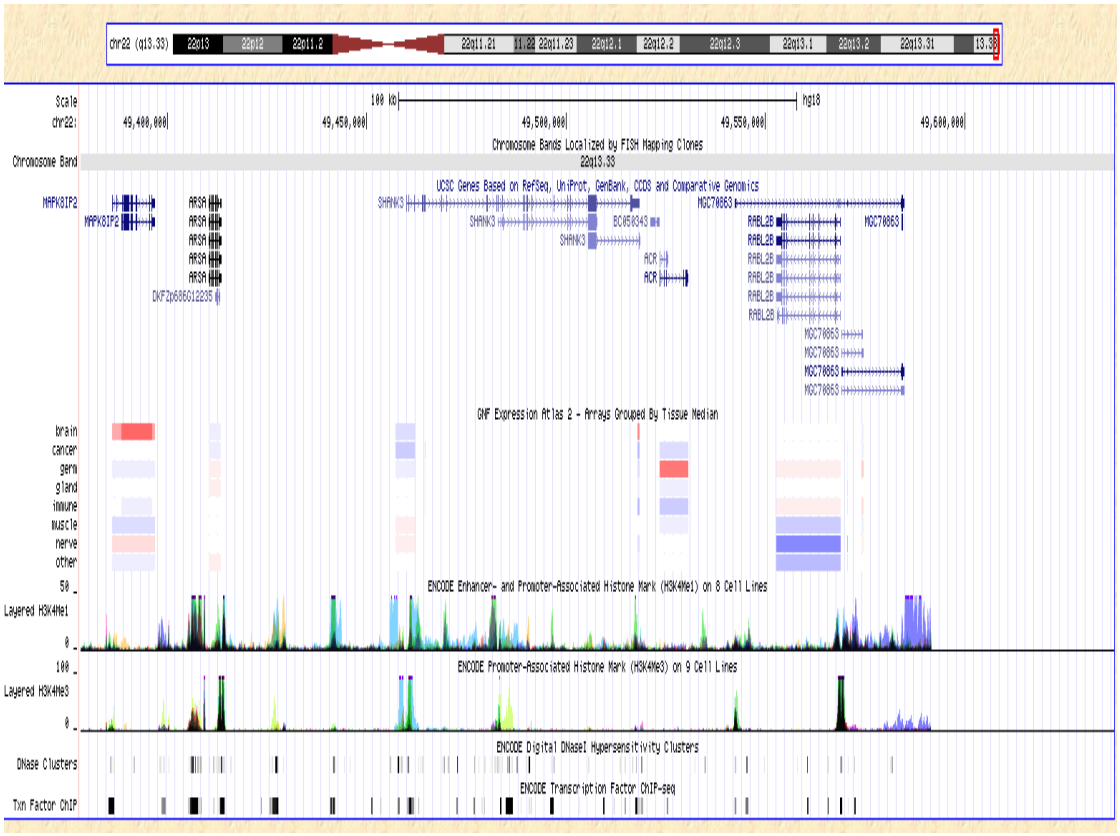


Figure 1.5. Location and Interaction of Proteins in the Post Synaptic Density. Reprinted from (Verpelli and Sala 2011) with permission from Elsevier.

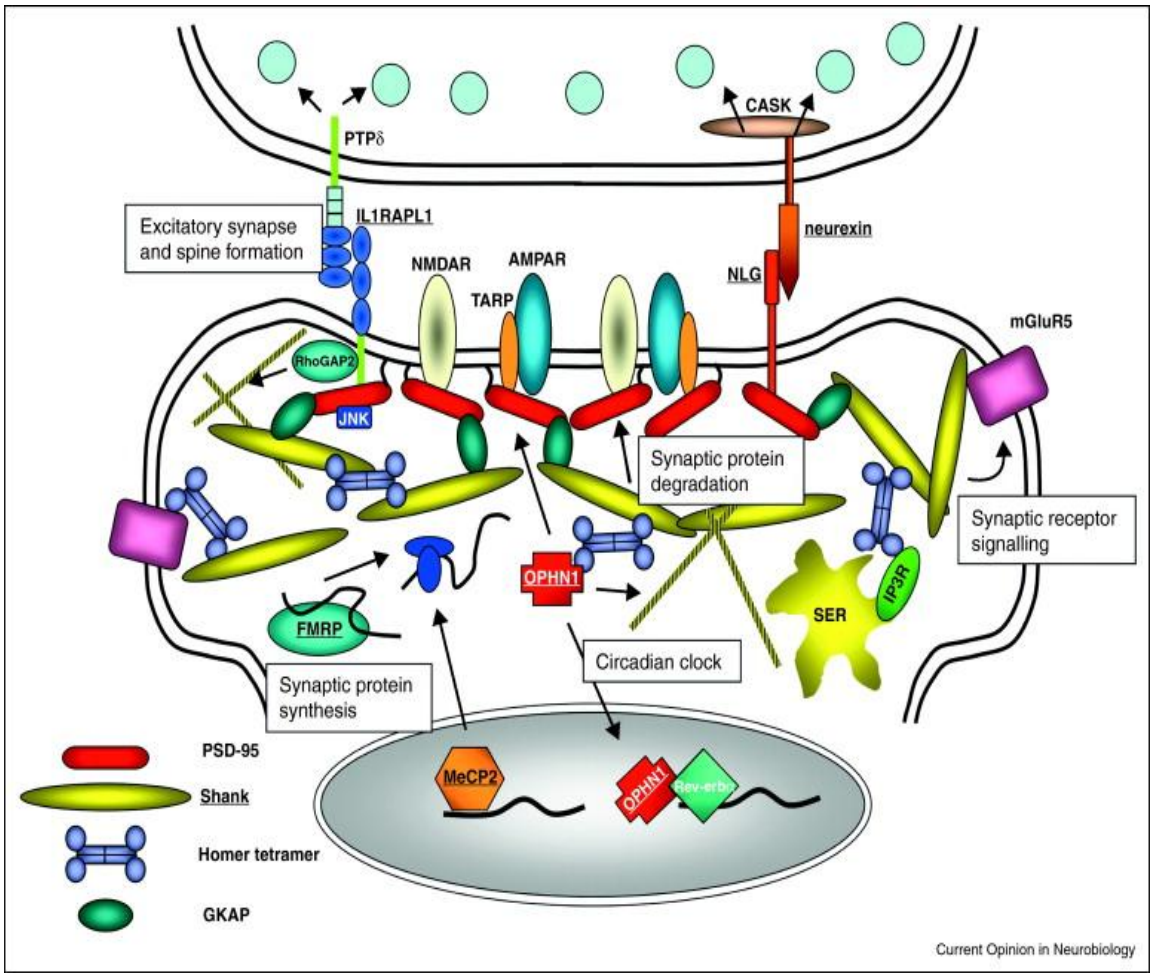


Figure 1.6. Schematic of SHANK3 and the PSD. Panel A provides a schematic of SHANK3 exons and conserved domains. Panel B provides a schematic of SHANK3 and biding partners in the post synaptic density. This figure is reprinted from (Phelan and McDermid 2012) with permission from S. Karger AG Basel. The figure was based upon (Kreienkamp 2008; Phelan and McDermid 2012) and is reprinted with kind permission from Springer Science and Business Media and Dr. Kreienkamp.

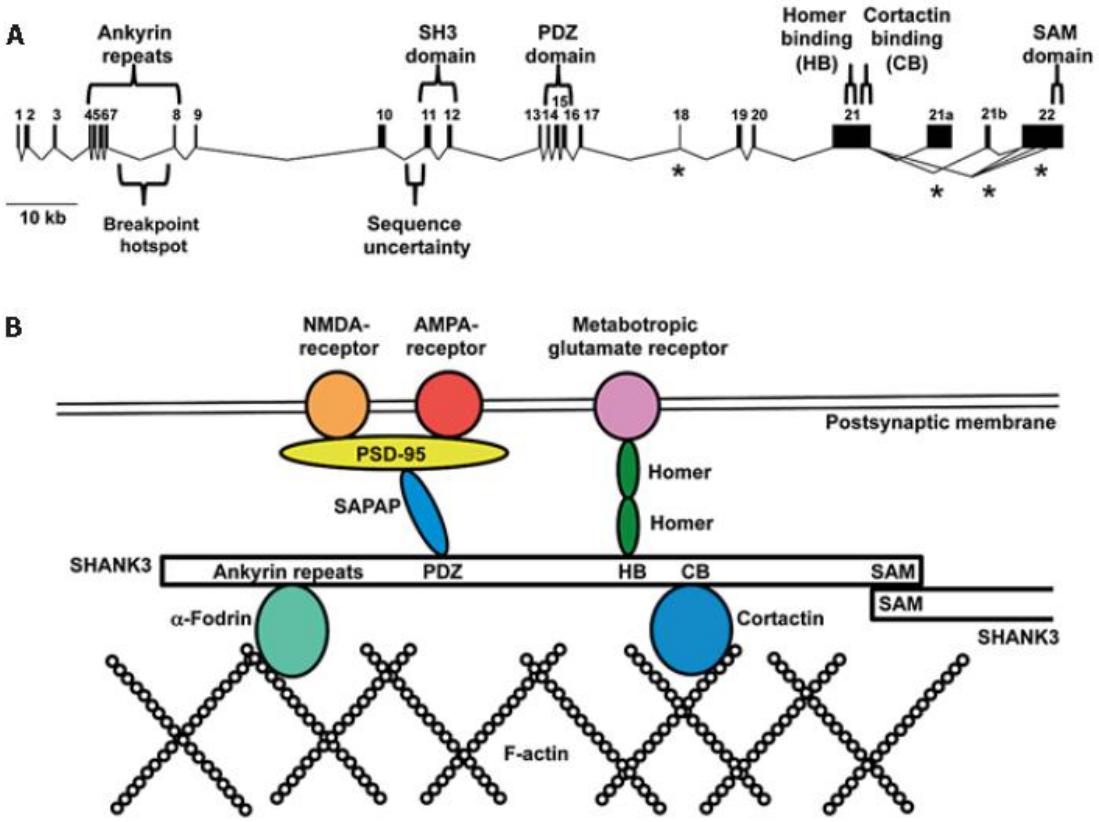
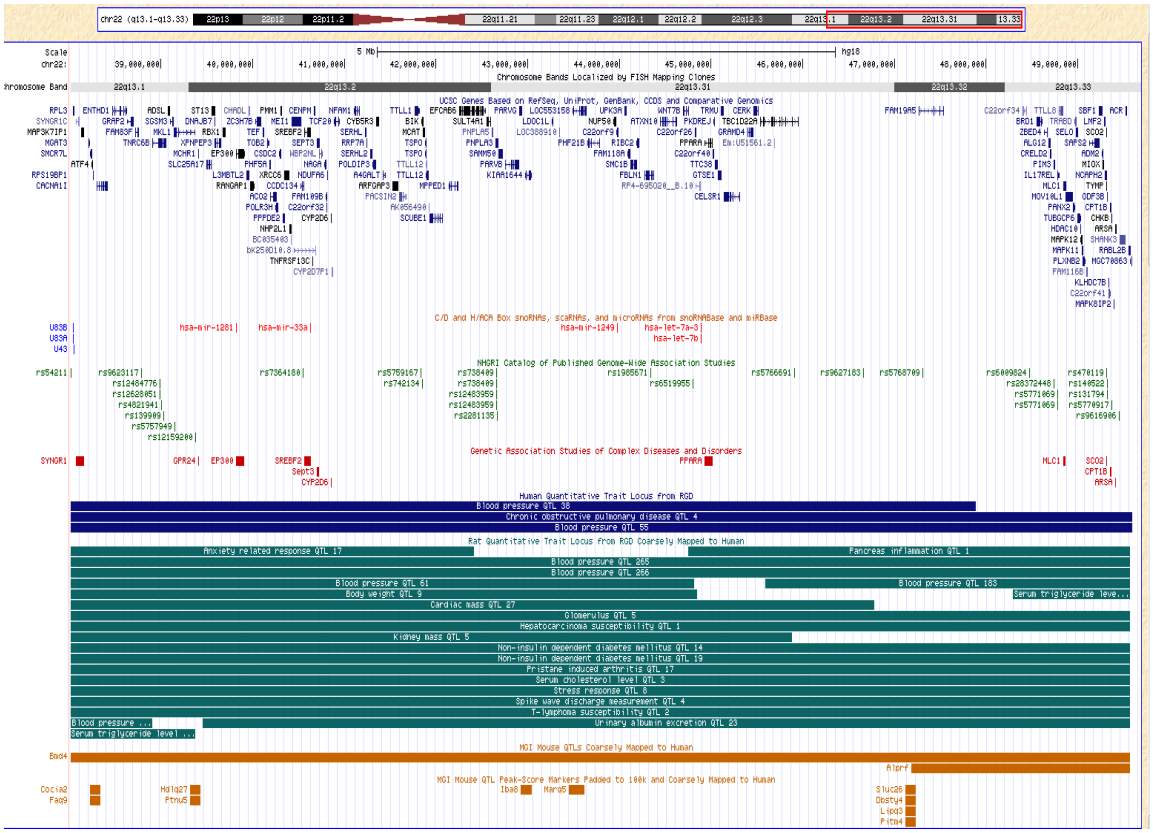


Figure 1.7. GWAS Findings in 22q13 from the UCSC Genome Browser and Catalog of Published Genome Wise Association Studies (Kent and others 2002; Hindorff and others 2009a; Hindorff and others 2009b).



References

Abu-Elneel K, Liu T, Gazzaniga FS, Nishimura Y, Wall DP, Geschwind DH, Lao K, Kosik KS. 2008. Heterogeneous dysregulation of microRNAs across the autism spectrum. Neurogenetics 9(3):153-61.

Agrawal A, Lynskey MT, Hinrichs A, Grucza R, Saccone SF, Krueger R, Neuman R, Howells W, Fisher S, Fox L, and others. 2011. A genome-wide association study of DSM-IV cannabis dependence. Addict Biol 16(3):514-8.

Alarcon M, Abrahams BS, Stone JL, Duvall JA, Perederiy JV, Bomar JM, Sebat J, Wigler M, Martin CL, Ledbetter DH, and others. 2008. Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. Am J Hum Genet 82(1):150-9.

Allanson JE, Cunniff C, Hoyme HE, McGaughran J, Muenke M, Neri G. 2009. Elements of morphology: Standard terminology for the head and face. American Journal of Medical Genetics Part A 149A(1):6-28.

American Psychiatric Association and American Psychiatric Association. Task Force on DSM-IV. 2000. Diagnostic and statistical manual of mental disorders: DSM-IV-TR. 4th , text revision ed. Washington, DC: American Psychiatric Association. 943 p.

Anderlid BM, Schoumans J, Anneren G, Tapia-Paez I, Dumanski J, Blennow E, Nordenskjold M. 2002. FISH-mapping of a 100-kb terminal 22q13 deletion. Hum Genet 110(5):439-43.

Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD, Lee JC, Goyette P, Imielinski M, Latiano A, and others. 2011. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. Nat Genet 43(3):246-52.

Babineau T, Wilson HL, Dawson AJ, Chodirker BN, Der Kaloustian VM, Demczuk S, McDermid HE. 2006. Unusual dicentric chromosome 22 associated with a 22q13 deletion. Am J Med Genet A 140(24):2819-23.

Bailey JA, Gu Z, Clark RA, Reinert K, Samonte RV, Schwartz S, Adams MD, Myers EW, Li PW, Eichler EE. 2002. Recent segmental duplications in the human genome. Science 297(5583):1003-7.

Bangash MA, Park JM, Melnikova T, Wang D, Jeon SK, Lee D, Syeda S, Kim J, Kouser M, Schwartz J, and others. 2011. Enhanced polyubiquitination of Shank3 and NMDA receptor in a mouse model of autism. Cell 145(5):758-72.

Barakat AJ, Pearl PL, Acosta MT, Runkle BP. 2004. 22q13 deletion syndrome with central diabetes insipidus: A previously unreported association. Clin Dysmorphol 13(3):191-4.

Bartel DP. 2004. MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell 116(2):281-97.

Bartsch O, Schneider E, Damatova N, Weis R, Tufano M, Iorio R, Ahmed A, Beyer V, Zechner U, Haaf T. 2010. Fulminant hepatic failure requiring liver transplantation in 22q13.3 deletion syndrome. Am J Med Genet A 152A(8):2099-102.

Battaglia A, Hoyme HE, Dallapiccola B, Zackai E, Hudgins L, McDonald-McGinn D, Bahi-Buisson N, Romano C, Williams CA, Brailey LL. 2008. Further delineation of deletion 1p36 syndrome in 60 patients: A recognizable phenotype and common cause of developmental delay and mental retardation. Pediatrics 121(2):404-10.

Battini R, Battaglia A, Bertini V, Cioni G, Parrini B, Rapalini E, Simi P, Tinelli F, Valetto A. 2004. Characterization of the phenotype and definition of the deletion in a new patient with ring chromosome 22. Am J Med Genet A 130A(2):196-9.

Baur JA, Zou Y, Shay JW, Wright WE. 2001. Telomere position effect in human cells. Science 292(5524):2075-7.

Beri S, Tonna N, Menozzi G, Bonaglia MC, Sala C, Giorda R. 2007. DNA methylation regulates tissue-specific expression of Shank3. J Neurochem 101(5):1380-91.

Berkel S, Marshall CR, Weiss B, Howe J, Roeth R, Moog U, Endris V, Roberts W, Szatmari P, Pinto D, and others. 2010. Mutations in the SHANK2 synaptic scaffolding gene in autism spectrum disorder and mental retardation. Nat Genet 42(6):489-91.

Betancur C. 2011. Etiological heterogeneity in autism spectrum disorders: More than 100 genetic and genomic disorders and still counting. Brain Res 1380:42-77.

Betancur C, Sakurai T, Buxbaum JD. 2009. The emerging role of synaptic cell-adhesion pathways in the pathogenesis of autism spectrum disorders. Trends Neurosci 32(7):402-1.

Betel D, Wilson M, Gabow A, Marks DS, Sander C. 2008. The microRNA.org resource: Targets and expression. Nucleic Acids Res 36(Database issue):D149-53.

Biesecker LG, Aase JM, Clercuzio C, Gurrieri F, Temple IK, Toriello H. 2009. Elements of morphology: Standard terminology for the hands and feet. American Journal of Medical Genetics Part A 149A(1):93-127.

Bisgaard AM, Kirchhoff M, Nielsen JE, Kibaek M, Lund A, Schwartz M, Christensen E. 2009. Chromosomal deletion unmasking a recessive disease: 22q13 deletion syndrome and metachromatic leukodystrophy. Clin Genet 75(2):175-9.

Blake JA, Bult CJ, Kadin JA, Richardson JE, Eppig JT, Mouse Genome Database Group. 2011. The mouse genome database (MGD): Premier model organism resource for mammalian genomics and genetics. Nucleic Acids Res 39(Database issue):D842-8.

Boccuto L, Lauri M, Sarasua SM, Skinner CD, Buccella D, Dwivedi A, Orteschi D, Collins JS, Zollino M, Visconti P, and others. 2012. Prevalence of SHANK3 variants in patients with different subtypes of autism spectrum disorders. Eur J Hum Genet (in press).

Bockers TM, Segger-Junius M, Iglauer P, Bockmann J, Gundelfinger ED, Kreutz MR, Richter D, Kindler S, Kreienkamp HJ. 2004. Differential expression and dendritic transcript localization of shank family members: Identification of a dendritic targeting element in the 3' untranslated region of Shank1 mRNA. Mol Cell Neurosci 26(1):182-90.

Bockers TM, Mameza MG, Kreutz MR, Bockmann J, Weise C, Buck F, Richter D, Gundelfinger ED, Kreienkamp HJ. 2001. Synaptic scaffolding proteins in rat brain. ankyrin repeats of the multidomain shank protein family interact with the cytoskeletal protein alpha-fodrin. J Biol Chem 276(43):40104-12.

Boeckers TM, Bockmann J, Kreutz MR, Gundelfinger ED. 2002. ProSAP/shank proteins - a family of higher order organizing molecules of the postsynaptic density with an emerging role in human neurological disease. J Neurochem 81(5):903-10.

Boeckers TM, Liedtke T, Spilker C, Dresbach T, Bockmann J, Kreutz MR, Gundelfinger ED. 2005. C-terminal synaptic targeting elements for postsynaptic density proteins ProSAP1/Shank2 and ProSAP2/Shank3. J Neurochem 92(3):519-24.

Boeckers TM, Winter C, Smalla KH, Kreutz MR, Bockmann J, Seidenbecher C, Garner CC, Gundelfinger ED. 1999. Proline-rich synapse-associated proteins ProSAP1 and ProSAP2 interact with synaptic proteins of the SAPAP/GKAP family. Biochem Biophys Res Commun 264(1):247-52.

Bonaglia MC, Giorda R, Ciccone R, Zuffardi O. 2010. Chromosome 22q13 rearrangements causing global developmental delay and autistic spectrum disorder. Knight SJL, editor. In: Genetics of mental retardation. Karger. 137 p.

Bonaglia MC, Giorda R, Borgatti R, Felisari G, Gagliardi C, Selicorni A, Zuffardi O. 2001. Disruption of the ProSAP2 gene in a t(12;22)(q24.1;q13.3) is associated with the 22q13.3 deletion syndrome. Am J Hum Genet 69(2):261-8.

Bonaglia MC, Giorda R, Mani E, Aceti G, Anderlid BM, Baroncini A, Pramparo T, Zuffardi O. 2006. Identification of a recurrent breakpoint within the SHANK3 gene in the 22q13.3 deletion syndrome. J Med Genet 43(10):822-8.

Bonaglia MC, Giorda R, Beri S, Bigoni S, Sensi A, Baroncini A, Capucci A, De Agostini C, Gwilliam R, Deloukas P, and others. 2009. Mosaic 22q13 deletions: Evidence for concurrent mosaic segmental isodisomy and gene conversion. Eur J Hum Genet 17(4):426-33.

Bonaglia MC, Giorda R, Beri S, De Agostini C, Novara F, Fichera M, Grillo L, Galesi O, Vetro A, Ciccone R, and others. 2011. Molecular mechanisms generating and stabilizing terminal 22q13 deletions in 44 subjects with Phelan/McDermid syndrome. PLoS Genet 7(7):e1002173.

Bouquillon S, Andrieux J, Landais E, Duban-Bedu B, Boidein F, Lenne B, Vallee L, Leal T, Doco-Fenzy M, Delobel B. 2011. A 5.3Mb deletion in chromosome 18q12.3 as the smallest region of overlap in two patients with expressive speech delay. Eur J Med Genet 54(2):194-7.

Bozdagi O, Sakurai T, Papapetrou D, Wang X, Dickstein DL, Takahashi N, Kajiwara Y, Yang M, Katz AM, Scattoni ML, and others. 2010. Haploinsufficiency of the autism-associated Shank3 gene leads to deficits in synaptic function, social interaction, and social communication. Mol Autism 1(1):1-15.

Bruining H, de Sonneville L, Swaab H, de Jonge M, Kas M, van Engeland H, Vorstman J. 2010. Dissecting the clinical heterogeneity of autism spectrum disorders through defined genotypes. PLoS One 5(5):e10887.

Buchanan JA and Scherer SW. 2008. Contemplating effects of genomic structural variation. Genet Med 10(9):639-47.

Buiting K. 2010. Prader-willi syndrome and angelman syndrome. Am J Med Genet C Semin Med Genet 154C(3):365-76.

Celestino-Soper PBS, Skinner C, Schroer R, Eng P, Shenai J, Nowaczyk MMJ, Terespolsky D, Cushing D, Patel GS, Immken LD. 2012. Deletions in chromosome 6p22. 3-p24. 3, including ATXN1, are associated with developmental delay and autism spectrum disorders. Molecular Cytogenetics 5(1):17.

Centers for Disease Control and Prevention. 2012. Prevalence of autism spectrum disorders -- autism and developmental disabilities monitoring network, 14 sites, united states, 2008. Morbidity and Mortality Weekly Report 61(SS-3):1-19.

Chambers JC, Zhang W, Sehmi J, Li X, Wass MN, Van der Harst P, Holm H, Sanna S, Kavousi M, Baumeister SE, and others. 2011. Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. Nat Genet 43(11):1131-8.

Chen CP, Hsu CY, Huang JK, Lee CC, Chen WL, Wang W. 2005. Prenatal diagnosis of partial trisomy 16q and distal 22q13 deletion associated with dolichocephaly and frontal bossing on second-trimester ultrasound. Prenat Diagn 25(10):964-6.

Ching TT, Maunakea AK, Jun P, Hong C, Zardo G, Pinkel D, Albertson DG, Fridlyand J, Mao JH, Shchors K, and others. 2005. Epigenome analyses using BAC microarrays identify evolutionary conservation of tissue-specific methylation of SHANK3. Nat Genet 37(6):645-51.

Cho SC, Yoo HJ, Park M, Cho IH, Kim BN, Kim JW, Shin MS, Park TW, Son JW, Chung US, and others. 2011. Genome-wide association scan of korean autism spectrum disorders with language delay: A preliminary study. Psychiatry Investig 8(1):61-6.

Cody JD, Heard PL, Crandall ALC, Carter EM, Li J, Hardies LJ, Lancaster J, Perry B, Stratton RF, Sebold C. 2009. Narrowing critical regions and determining penetrance for selected 18q-phenotypes. American Journal of Medical Genetics Part A 149(7):1421-30.

Cohen D, Pichard N, Tordjman S, Baumann C, Burglen L, Excoffier E, Lazar G, Mazet P, Pinquier C, Verloes A, and others. 2005. Specific genetic disorders and autism: Clinical contribution towards their identification. J Autism Dev Disord 35(1):103-16.

Crawley JN. 2004. Designing mouse behavioral tasks relevant to autistic-like behaviors. Ment Retard Dev Disabil Res Rev 10(4):248-5.

Cusmano-Ozog K, Manning MA, Hoyme HE. 2007. 22q13.3 deletion syndrome: A recognizable malformation syndrome associated with marked speech and language delay. Am J Med Genet C Semin Med Genet 145C(4):393-8.

Dale PS, Simonoff E, Bishop DV, Eley TC, Oliver B, Price TS, Purcell S, Stevenson J, Plomin R. 1998. Genetic influence on language delay in two-year-old children. Nat Neurosci 1(4):324-8.

De Mas P, Chassaing N, Chaix Y, Vincent MC, Julia S, Bourrouillou G, Calvas P, Bieth E. 2002. Molecular characterisation of a ring chromosome 22 in a patient with severe language delay: A contribution to the refinement of the subtelomeric 22q deletion syndrome. J Med Genet 39(4):e17.

Del Greco MF, Pattaro C, Luchner A, Pichler I, Winkler T, Hicks AA, Fuchsberger C, Franke A, Melville SA, Peters A, and others. 2011. Genome-wide association analysis and fine mapping of NT-proBNP level provide novel insight into the role of the MTHFR-CLCN6-NPPA-NPPB gene cluster. Hum Mol Genet 20(8):1660-71.

Delahaye A, Toutain A, Aboura A, Dupont C, Tabet AC, Benzacken B, Elion J, Verloes A, Pipiras E, Drunat S. 2009. Chromosome 22q13.3 deletion syndrome with a de novo interstitial 22q13.3 cryptic deletion disrupting SHANK3. Eur J Med Genet 52(5):328-32.

Dhar SU, del Gaudio D, German JR, Peters SU, Ou Z, Bader PI, Berg JS, Blazo M, Brown CW, Graham BH, and others. 2010. 22q13.3 deletion syndrome: Clinical and molecular analysis using array CGH. Am J Med Genet A 152A(3):573-81.

Dolmans GH, Werker PM, Hennies HC, Furniss D, Festen EA, Franke L, Becker K, van der Vlies P, Wolffenbuttel BH, Tinschert S, and others. 2011. Wnt signaling and dupuytren's disease. N Engl J Med 365(4):307-1.

Du YR, Weed SA, Xiong WC, Marshall TD, Parsons JT. 1998. Identification of a novel cortactin SH3 domain-binding protein and its localization to growth cones of cultured neurons. Mol Cell Biol 18(10):5838-51.

Durand CM, Betancur C, Boeckers TM, Bockmann J, Chaste P, Fauchereau F, Nygren G, Rastam M, Gillberg IC, Anckarsater H, and others. 2007. Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. Nat Genet 39(1):25-7.

Eeles RA, Kote-Jarai Z, Al Olama AA, Giles GG, Guy M, Severi G, Muir K, Hopper JL, Henderson BE, Haiman CA, and others. 2009. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. Nat Genet 41(10):1116-21.

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

Eppig JT, Blake JA, Bult CJ, Kadin JA, Richardson JE, Mouse Genome Database Group. 2012. The mouse genome database (MGD): Comprehensive resource for genetics and genomics of the laboratory mouse. Nucleic Acids Res 40(Database issue):D881-6.

Estrada K, Krawczak M, Schreiber S, van Duijn K, Stolk L, van Meurs JB, Liu F, Penninx BW, Smit JH, Vogelzangs N, and others. 2009. A genome-wide association study of northwestern europeans involves the C-type natriuretic peptide signaling pathway in the etiology of human height variation. Hum Mol Genet 18(18):3516-24.

Fang JS, Lee KF, Huang CT, Syu CL, Yang KJ, Wang LH, Liao DL, Chen CH. 2008. Cytogenetic and molecular characterization of a three-generation family with chromosome 5p terminal deletion. Clin Genet 73(6):585-90.

Feenstra I, Vissers LE, Orsel M, van Kessel AG, Brunner HG, Veltman JA, van Ravenswaaij-Arts CM. 2007. Genotype-phenotype mapping of chromosome 18q deletions by high-resolution array CGH: An update of the phenotypic map. Am J Med Genet A 143A(16):1858-67.

Finger JH, Smith CM, Hayamizu TF, McCright IJ, Eppig JT, Kadin JA, Richardson JE, Ringwald M. 2011. The mouse gene expression database (GXD): 2011 update. Nucleic Acids Res 39(Database issue):D835-41.

Fisher SE and Scharff C. 2009. FOXP2 as a molecular window into speech and language. Trends Genet 25(4):166-77.

Flax JF, Hare A, Azaro MA, Vieland VJ, Brzustowicz LM. 2010. Combined linkage and linkage disequilibrium analysis of a motor speech phenotype within families ascertained for autism risk loci. J Neurodev Disord 2(4):210-23.

Flint J, Wilkie AO, Buckle VJ, Winter RM, Holland AJ, McDermid HE. 1995. The detection of subtelomeric chromosomal rearrangements in idiopathic mental retardation. Nat Genet 9(2):132-40.

Flint J, Craddock CF, Villegas A, Bentley DP, Williams HJ, Galanello R, Cao A, Wood WG, Ayyub H, Higgs DR. 1994. Healing of broken human-chromosomes by the addition of telomeric repeats. Am J Hum Genet 55(3):505-12.

Franke A, Balschun T, Sina C, Ellinghaus D, Hasler R, Mayr G, Albrecht M, Wittig M, Buchert E, Nikolaus S, and others. 2010. Genome-wide association study for ulcerative colitis identifies risk loci at 7q22 and 22q13 (IL17REL). Nat Genet 42(4):292-4.

Friedman RC, Farh KK, Burge CB, Bartel DP. 2009. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19(1):92-105.

Fujita Y, Mochizuki D, Mori Y, Nakamoto N, Kobayashi M, Omi K, Kodama H, Yanagawa Y, Abe T, Tsuzuku T, and others. 2000. Girl with accelerated growth, hearing loss, inner ear anomalies, delayed myelination of the brain, and del(22)(q13.1q13.2). Am J Med Genet 92(3):195-9.

Gajecka M, Saadeh R, Mackay KL, Glotzbach CD, Spodar K, Chitayat D, Shaffer LG. 2008. Clinical and molecular cytogenetic characterization of four patients with unbalanced translocation der(1)t(1;22)(p36;q13). Am J Med Genet A 146A(21):2777-84.

Ganesh SK, Zakai NA, van Rooij FJ, Soranzo N, Smith AV, Nalls MA, Chen MH, Kottgen A, Glazer NL, Dehghan A, and others. 2009. Multiple loci influence erythrocyte phenotypes in the CHARGE consortium. Nat Genet 41(11):1191-8.

Gauthier J, Champagne N, Lafreniere RG, Xiong L, Spiegelman D, Brustein E, Lapointe M, Peng H, Cote M, Noreau A, and others. 2010. De novo mutations in the gene encoding the synaptic scaffolding protein SHANK3 in patients ascertained for schizophrenia. Proc Natl Acad Sci U S A 107(17):7863-8.

Gersh M, Goodart S, Pasztor L, Harris D, Weiss L, Overhauser J. 1995. Evidence for a distinct region causing a cat-like cry in patients with 5p deletions. Am J Hum Genet 56(6):1404.

Girirajan S, Rosenfeld JA, Coe BP, Parikh S, Friedman N, Goldstein A, Filipink RA, McConnell JS, Angle B, Meschino WS, and others. 2012. Phenotypic heterogeneity of genomic disorders and rare copy-number variants. N Engl J Med 367(14):1321-3.

Glessner JT, Wang K, Cai G, Korvatska O, Kim CE, Wood S, Zhang H, Estes A, Brune CW, Bradfield JP. 2009. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. Nature 459(7246):569-73.

Goizet C, Excoffier E, Taine L, Taupiac E, El Moneim AA, Arveiler B, Bouvard M, Lacombe D. 2000. Case with autistic syndrome and chromosome 22q13.3 deletion detected by FISH. Am J Med Genet 96(6):839-44.

Gong Y and Lippa CF. 2010. Review: Disruption of the postsynaptic density in alzheimer's disease and other neurodegenerative dementias. Am J Alzheimers Dis Other Demen 25(7):547-55.

Gong X, Jiang YW, Zhang X, An Y, Zhang J, Wu Y, Wang J, Sun Y, Liu Y, Gao X, and others. 2012. High proportion of 22q13 deletions and SHANK3 mutations in chinese patients with intellectual disability. PLoS One 7(4):e34739.

Grabrucker AM, Schmeisser MJ, Schoen M, Boeckers TM. 2011. Postsynaptic ProSAP/shank scaffolds in the cross-hair of synaptopathies. Trends Cell Biol 21(10):594-603.

Greenwood Genetic Center. 2011. Growth references. Rollins JD, Tribble LM, Collins JS, and others, editors. 3rd ed. Greenville, South Carolina: Greenwood Genetic Center.

Greenwood TA, Akiskal HS, Akiskal KK, Bipolar Genome Study, Kelsoe JR. 2012. Genome-wide association study of temperament in bipolar disorder reveals significant associations with three novel loci. Biol Psychiatry 72(4):303-10.

Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. 2006. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 34(Database issue):D140-4.

Grubb DR, Iliades P, Cooley N, Yu YL, Luo J, Filtz TM, Woodcock EA. 2011. Phospholipase Cbeta1b associates with a Shank3 complex at the cardiac sarcolemma. FASEB J 25(3):1040-7.

Guilherme RS, Meloni VF, Kim CA, Pellegrino R, Takeno SS, Spinner NB, Conlin LK, Christofolini DM, Kulikowski LD, Melaragno MI. 2011. Mechanisms of ring chromosome formation, ring instability and clinical consequences. BMC Med Genet 12:171.

Hall BD, Graham JM,Jr., Cassidy SB, Opitz JM. 2009. Elements of morphology: Standard terminology for the periorbital region. American Journal of Medical Genetics Part A 149A(1):29-3.

Hamdan FF, Gauthier J, Araki Y, Lin D, Yoshizawa Y, Higashi K, Park A-, Spiegelman D, Dobrzeniecka S, Piton A, and others. 2011. Excess of de novo deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability. Am J Hum Genet 88(3):306-1.

Hannes F, Van Houdt J, Quarrell OW, Poot M, Hochstenbach R, Fryns J, Vermeesch JR. 2010. Telomere healing following DNA polymerase arrest-induced breakages is likely the main mechanism generating chromosome 4p terminal deletions. Hum Mutat 31(12):1343-51.

Harewood L, Schutz F, Boyle S, Perry P, Delorenzi M, Bickmore WA, Reymond A. 2010. The effect of translocation-induced nuclear reorganization on gene expression. Genome Res 20(5):554-6.

Harris SR. 2008. Congenital hypotonia: Clinical and developmental assessment. Dev Med Child Neurol 50(12):889-92.

Havens JM, Visootsak J, Phelan MC, Graham JM,Jr. 2004. 22q13 deletion syndrome: An update and review for the primary pediatrician. Clin Pediatr (Phila) 43(1):43-5.

Hayashi MK, Tang C, Verpelli C, Narayanan R, Stearns MH, Xu RM, Li H, Sala C, Hayashi Y. 2009. The postsynaptic density proteins homer and shank form a polymeric network structure. Cell 137(1):159-71.

Heilstedt H, Ballif B, Howard L, Kashork C, Shaffer L. 2003a. Population data suggest that deletions of 1p36 are a relatively common chromosome abnormality. Clin Genet 64(4):310-6.

Heilstedt HA, Ballif BC, Howard LA, Lewis RA, Stal S, Kashork CD, Bacino CA, Shapira SK, Shaffer LG. 2003b. Physical map of 1p36, placement of breakpoints in monosomy 1p36, and clinical characterization of the syndrome. Am J Hum Genet 72(5):1200-12.

Herman GE, Greenberg F, Ledbetter DH. 1988. Multiple congenital anomaly mental-retardation (mca mr) syndrome with goldenhar complex due to a terminal del(22q). Am J Med Genet 29(4):909-15.

Hill A. 2005. Neonatal hypotonia. Maria BL, editor. In: Current management in child neurology. Third ed. BC Decker. 528 p.

Hindorff LA, MacArthur J, European Bioinformatics Institute, Wise A, Junkins HA, Hall PN, Klemm AK, Manolio TA. 2009a A catalog of published genome-wide association studies. <www.genome.gov/gwastudies>. Accessed 2012 02/06.

Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. 2009b. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A 106(23):9362-7.

Hu VW and Steinberg ME. 2009. Novel clustering of items from the autism diagnostic interview-revised to define phenotypes within autism spectrum disorders. Autism Res 2(2):67-7.

Hu VW, Sarachana T, Kim KS, Nguyen A, Kulkarni S, Steinberg ME, Luu T, Lai Y, Lee NH. 2009. Gene expression profiling differentiates autism case-controls and phenotypic variants of autism spectrum disorders: Evidence for circadian rhythm dysfunction in severe autism. Autism Res 2(2):78-97.

Huett A, Leong JM, Podolsky DK, Xavier RJ. 2009. The cytoskeletal scaffold Shank3 is recruited to pathogen-induced actin rearrangements. Exp Cell Res 315(12):2001-1.

Hunter A, Frias JL, Gillessen-Kaesbach G, Hughes H, Jones KL, Wilson L. 2009. Elements of morphology: Standard terminology for the ear. American Journal of Medical Genetics Part A 149A(1):40-6.

Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C. 2004. Detection of large-scale variation in the human genome. Nat Genet 36(9):949-51.

International Human Genome Sequencing Consortium. 2004. Finishing the euchromatic sequence of the human genome. Nature 431(7011):931-45.

International Multiple Sclerosis Genetics Consortium, Wellcome Trust Case Control Consortium 2, Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, Moutsianas L, Dilthey A, Su Z, and others. 2011. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 476(7359):214-9.

Itsara A, Cooper GM, Baker C, Girirajan S, Li J, Absher D, Krauss RM, Myers RM, Ridker PM, Chasman DI, and others. 2009. Population analysis of large copy number variants and hotspots of human genetic disease. Am J Hum Genet 84(2):148-61.

Jamsheer A, Smyk M, Wierzba J, Kolowska J, Wozniak A, Skolozdrzy J, Fischer M, Latos-Bielenska A. 2008. Subtle familial translocation t(11;22)(q24.2;q13.33) resulting in jacobsen syndrome and distal trisomy 22q13.3: Further details of genotype-phenotype maps. J Appl Genet 49(4):397-405.

Jeffries AR, Curran S, Elmslie F, Sharma A, Wenger S, Hummel M, Powell J. 2005. Molecular and phenotypic characterization of ring chromosome 22. Am J Med Genet A 137(2):139-47.

Kamatani Y, Matsuda K, Okada Y, Kubo M, Hosono N, Daigo Y, Nakamura Y, Kamatani N. 2010. Genome-wide association study of hematological and biochemical traits in a japanese population. Nat Genet 42(3):210-5.

Kaufman L, Ayub M, Vincent JB. 2010. The genetic basis of non-syndromic intellectual disability: A review. J Neurodev Disord 2(4):182-209.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. Genome Res 12(6):996-1006.

Kim S, Swaminathan S, Shen L, Risacher SL, Nho K, Foroud T, Shaw LM, Trojanowski JQ, Potkin SG, Huentelman MJ, and others. 2011a. Genome-wide association study of CSF biomarkers Abeta1-42, t-tau, and p-tau181p in the ADNI cohort. Neurology 76(1):69-7.

Kim YJ, Go MJ, Hu C, Hong CB, Kim YK, Lee JY, Hwang JY, Oh JH, Kim DJ, Kim NH, and others. 2011b. Large-scale genome-wide association studies in east asians identify new genetic loci influencing metabolic traits. Nat Genet 43(10):990-5.

Kirshenbaum G, Chmura M, Rhone DP. 1988. Long arm deletion of chromosome 22. J Med Genet 25(11):780.

Kline A, White M, Wapner R, Rojas K, Biesecker L, Kamholz J, Zackai E, Muenke M, Scott Jr C, Overhauser J. 1993. Molecular analysis of the 18q-syndrome--and correlation with phenotype. Am J Hum Genet 52(5):895.

Kobrynski LJ and Sullivan KE. 2007. Velocardiofacial syndrome, DiGeorge syndrome: The chromosome 22q11.2 deletion syndromes. Lancet 370(9596):1443-52.

Koc A, Arisoy O, Pala E, Erdem M, Kaymak AO, Erkal O, Karaoguz MY. 2009. Prenatal diagnosis of mosaic ring 22 duplication/deletion with terminal 22q13 deletion due to abnormal first trimester screening and choroid plexus cyst detected on ultrasound. J Obstet Gynaecol Res 35(5):978-82.

Kohler S, Schulz MH, Krawitz P, Bauer S, Dolken S, Ott CE, Mundlos C, Horn D, Mundlos S, Robinson PN. 2009. Clinical diagnostics in human genetics with semantic similarity searches in ontologies. Am J Hum Genet 85(4):457-64.

Konopka G, Bomar JM, Winden K, Coppola G, Jonsson ZO, Gao F, Peng S, Preuss TM, Wohlschlegel JA, Geschwind DH. 2009. Human-specific transcriptional regulation of CNS development genes by FOXP2. Nature 462(7270):213-7.

Koolen DA, Reardon W, Rosser EM, Lacombe D, Hurst JA, Law CJ, Bongers EM, van Ravenswaaij-Arts CM, Leisink MA, van Kessel AG, and others. 2005. Molecular characterisation of patients with subtelomeric 22q abnormalities using chromosome specific array-based comparative genomic hybridisation. Eur J Hum Genet 13(9):1019-24.

Kosztolanyi G. 1987. Does "ring syndrome" exist? an analysis of 207 case reports on patients with a ring autosome. Hum Genet 75(2):174-9.

Kou Y, Betancur C, Xu H, Buxbaum JD, Ma'ayan A. 2012. Network- and attribute-based classifiers can prioritize genes and pathways for autism spectrum disorders and intellectual disability. Am J Med Genet C Semin Med Genet 160C(2):130-42.

Kreienkamp HJ. 2008. Scaffolding proteins at the postsynaptic density: Shank as the architectural framework. Handb Exp Pharmacol (186)(186):365-80.

Kreienkamp HJ, Zitzer H, Gundelfinger ED, Richter D, Bockers TM. 2000. The calcium-independent receptor for alpha-latrotoxin from human and rodent brains interacts with members of the ProSAP/SSTRIP/shank family of multidomain proteins. J Biol Chem 275(42):32387-90.

Kuczmarski RJ, Ogden CL, Guo SS, Grummer-Strawn LM, Flegal KM, Mei Z, Wei R, Curtin LR, Roche AF, Johnson CL. 2002. 2000 CDC growth charts for the united states: Methods and development. Vital Health Stat 11 (246)(246):1-190.

Kulkarni A, Zschenker O, Reynolds G, Miller D, Murnane JP. 2010. Effect of telomere proximity on telomere position effect, chromosome healing, and sensitivity to DNA double-strand breaks in a human tumor cell line. Mol Cell Biol 30(3):578-89.

Lamb J, Harris PC, Wilkie AOM, Wood WG, Dauwerse JHG, Higgs DR. 1993. Denovo truncation of chromosome-16p and healing with (ttaggg)n in the alpha-thalassemia mental-retardation syndrome (atr-16). Am J Hum Genet 52(4):668-76.

Laugel V, Cossee M, Matis J, de Saint-Martin A, Echaniz-Laguna A, Mandel JL, Astruc D, Fischbach M, Messer J. 2008. Diagnostic approach to neonatal hypotonia: Retrospective study on 144 neonates. Eur J Pediatr 167(5):517-23.

Leonard H and Wen X. 2002. The epidemiology of mental retardation: Challenges and opportunities in the new millennium. Ment Retard Dev Disabil Res Rev 8(3):117-34.

Leyenaar J, Camfield P, Camfield C. 2005. A schematic approach to hypotonia in infancy. Paediatr Child Health 10(7):397-400.

Lim S, Sala C, Yoon J, Park S, Kuroda S, Sheng M, Kim E. 2001. Sharpin, a novel postsynaptic density protein that directly interacts with the shank family of proteins. Mol Cell Neurosci 17(2):385-97.

Lindquist SG, Kirchhoff M, Lundsteen C, Pedersen W, Erichsen G, Kristensen K, Lillquist K, Smedegaard HH, Skov L, Tommerup N, and others. 2005. Further delineation of the 22q13 deletion syndrome. Clin Dysmorphol 14(2):55-60.

Luciani JJ, de Mas P, Depetris D, Mignon-Ravix C, Bottani A, Prieur M, Jonveaux P, Philippe A, Bourrouillou G, de Martinville B, and others. 2003. Telomeric 22q13 deletions resulting from rings, simple deletions, and translocations: Cytogenetic, molecular, and clinical analyses of 32 new observations. J Med Genet 40(9):690-6.

Lupski JR and Stankiewicz P. 2005. Genomic disorders: Molecular mechanisms for rearrangements and conveyed phenotypes. PLoS Genet 1(6):e49.

Magoulas PL and El-Hattab AW. 2012. Chromosome 15q24 microdeletion syndrome. Orphanet Journal of Rare Diseases 7(1):2.

Mainardi PC, Perfumo C, Cali A, Coucourde G, Pastore G, Cavani S, Zara F, Overhauser J, Pierluigi M, Bricarelli FD. 2001. Clinical and molecular characterisation of 80 patients with 5p deletion: Genotype-phenotype correlation. J Med Genet 38(3):151-8.

Maitz S, Gentilin B, Colli AM, Rizzuti T, Brandolisio E, Vetro A, Zuffardi O, Guerneri S, Lalatta F. 2008. Expanding the phenotype of 22q13.3 deletion: Report of a case detected prenatally. Prenat Diagn 28(10):978-80.

Man M, Close SL, Shaw AD, Bernard GR, Douglas IS, Kaner RJ, Payen D, Vincent JL, Fossceco S, Janes JM, and others. 2012. Beyond single-marker analyses: Mining whole genome scans for insights into treatment responses in severe sepsis. Pharmacogenomics J .

Manning MA, Cassidy SB, Clericuzio C, Cherry AM, Schwartz S, Hudgins L, Enns GM, Hoyme HE. 2004. Terminal 22q deletion syndrome: A newly recognized cause of speech and language disability in the autism spectrum. Pediatrics 114(2):451-7.

Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, Shago M, Moessner R, Pinto D, Ren Y, and others. 2008. Structural variation of chromosomes in autism spectrum disorder. Am J Hum Genet 82(2):477-88.

Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, Johnson BE, Hong C, Nielsen C, Zhao Y, and others. 2010. Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 466(7303):253-7.

Mefford HC, Rosenfeld JA, Shur N, Slavotinek AM, Cox VA, Hennekam RC, Firth HV, Willatt L, Wheeler P, Morrow EM, and others. 2012. Further clinical and molecular delineation of the 15q24 microdeletion syndrome. J Med Genet 49(2):110-8.

Meltzer PS, Guan XY, Trent JM. 1993. Telomere capture stabilizes chromosome breakage. Nat Genet 4(3):252-5.

Mick E, Neale B, Middleton FA, McGough JJ, Faraone SV. 2008. Genome-wide association study of response to methylphenidate in 187 children with attention-deficit/hyperactivity disorder. Am J Med Genet B Neuropsychiatr Genet 147B(8):1412-8.

Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ, and others. 2010. Consensus statement: Chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. Am J Hum Genet 86(5):749-64.

Misceo D, Rodningen OK, Baroy T, Sorte H, Mellembakken JR, Stromme P, Fannemel M, Frengen E. 2011. A translocation between Xq21.33 and 22q13.33 causes an intragenic SHANK3 deletion in a woman with Phelan-McDermid syndrome and hypergonadotropic hypogonadism. Am J Med Genet A 155A(2):403-8.

Miyagawa T, Kawashima M, Nishida N, Ohashi J, Kimura R, Fujimoto A, Shimada M, Morishita S, Shigeta T, Lin L, and others. 2008. Variant between CPT1B and CHKB associated with susceptibility to narcolepsy. Nat Genet 40(11):1324-8.

Moeschler JB, Shevell M, American Academy of Pediatrics Committee on Genetics. 2006. Clinical genetic evaluation of the child with mental retardation or developmental delays. Pediatrics 117(6):2304-16.

Moessner R, Marshall CR, Sutcliffe JS, Skaug J, Pinto D, Vincent J, Zwaigenbaum L, Fernandez B, Roberts W, Szatmari P, and others. 2007. Contribution of SHANK3 mutations to autism spectrum disorder. Am J Hum Genet 81(6):1289-97.

Molin AM, Andrieux J, Koolen DA, Malan V, Carella M, Colleaux L, Cormier-Daire V, David A, de Leeuw N, Delobel B, and others. 2012. A novel microdeletion syndrome at 3q13.31 characterised by developmental delay, postnatal overgrowth, hypoplastic male genitals, and characteristic facial features. J Med Genet 49(2):104-9.

Mulatinho MV, de Carvalho Serao CL, Scalco F, Hardekopf D, Pekova S, Mrasek K, Liehr T, Weise A, Rao N, Llerena JC. 2012. Severe intellectual disability, omphalocele, hypospadia and high blood pressure associated to a deletion at 2q22. 1q22. 3: Case report. Molecular Cytogenetics 5(1):30.

Naisbitt S, Kim E, Tu JC, Xiao B, Sala C, Valtschanoff J, Weinberg RJ, Worley PF, Sheng M. 1999. Shank, a novel family of postsynaptic density proteins that binds to the NMDA receptor/PSD-95/GKAP complex and cortactin. Neuron 23(3):569-82.

Narahara K, Takahashi Y, Murakami M, Tsuji K, Yokoyama Y, Murakami R, Ninomiya S, Seino Y. 1992. Terminal 22q deletion associated with a partial deficiency of arylsulphatase A. J Med Genet 29(6):432-3.

Nesslinger NJ, Gorski JL, Kurczynski TW, Shapira SK, Siegel-Bartelt J, Dumanski JP, Cullen RF,Jr, French BN, McDermid HE. 1994. Clinical, cytogenetic, and molecular characterization of seven patients with deletions of chromosome 22q13.3. Am J Hum Genet 54(3):464-72.

Newbury DF, Fisher SE, Monaco AP. 2010. Recent advances in the genetics of language impairment. Genome Med 2(1):6.

O'Donnell L, Soileau B, Heard P, Carter E, Sebold C, Gelfond J, Hale DE, Cody JD. 2010. Genetic determinants of autism in individuals with deletions of 18q. Hum Genet 128(2).

Oeseburg B, Dijkstra GJ, Groothoff JW, Reijneveld SA, Jansen DE. 2011. Prevalence of chronic health conditions in children with intellectual disability: A systematic literature review. Intellect Dev Disabil 49(2):59-85.

Ofir R, Wong AC, McDermid HE, Skorecki KL, Selig S. 1999. Position effect of human telomeric repeats on replication timing. Proc Natl Acad Sci U S A 96(20):11434-9.

Okamoto PM, Gamby C, Wells D, Fallon J, Vallee RB. 2001. Dynamin isoform-specific interaction with the shank/ProSAP scaffolding proteins of the postsynaptic density and actin cytoskeleton. J Biol Chem 276(51):48458-65.

Overhauser J, Huang X, Gersh M, Wilson W, McMahon J, Bengtsson U, Rojas K, Meyer M, Wasmuth JJ. 1994. Molecular and phenotypic mapping of the short arm of chromosome 5: Sublocalization of the critical region for the cri-du-chat syndrome. Hum Mol Genet 3(2):247-52.

Palmer AA, Verbitsky M, Suresh R, Kamens HM, Reed CL, Li N, Burkhart-Kasch S, McKinnon CS, Belknap JK, Gilliam TC, and others. 2005. Gene expression differences in mice divergently selected for methamphetamine sensitivity. Mamm Genome 16(5):291-305.

Park H, Kim JI, Ju YS, Gokcumen O, Mills RE, Kim S, Lee S, Suh D, Hong D, Kang HP. 2010. Discovery of common asian copy number variants using integrated high-resolution array CGH and massively parallel DNA sequencing. Nat Genet 42(5):400-5.

Peca J, Feliciano C, Ting JT, Wang W, Wells MF, Venkatraman TN, Lascola CD, Fu Z, Feng G. 2011. Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. Nature 472:437-42.

Phelan MC, Stapleton GA, Rogers RC. 2010. Deletion 22q13 syndrome: Phelan-McDermid syndrome. Cassidy SB and Allanson JE, editors. In: Management of genetic syndromes. Third Edition ed. Hoboken, New Jersey: Wiley-Blackwell. 285 p.

Phelan K and McDermid HE. 2012. The 22q13.3 deletion syndrome (Phelan-McDermid syndrome). Mol Syndromol 2(3-5):186-201.

Phelan MC. 2008. Deletion 22q13.3 syndrome. Orphanet J Rare Dis 3:14-9.

Phelan MC, Brown EF, Rogers RC. 2001. Prenatal diagnosis of mosaicism for deletion 22q13.3. Prenat Diagn 21(12):1100.

Phelan MC, Thomas GR, Saul RA, Rogers RC, Taylor HA, Wenger DA, McDermid HE. 1992. Cytogenetic, biochemical, and molecular analyses of a 22q13 deletion. Am J Med Genet 43(5):872-6.

Phelan MC, Rogers RC, Saul RA, Stapleton GA, Sweet K, McDermid H, Shaw SR, Claytor J, Willis J, Kelly DP. 2001. 22q13 deletion syndrome. Am J Med Genet 101(2):91-9.

Philippe A, Boddaert N, Vaivre-Douret L, Robel L, Danon-Boileau L, Malan V, de Blois MC, Heron D, Colleaux L, Golse B, and others. 2008. Neurobehavioral profile and brain imaging study of the 22q13.3 deletion syndrome in childhood. Pediatrics 122(2):e376-82.

Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, and others. 2010. Functional impact of global rare copy number variation in autism spectrum disorders. Nature 466(7304):368-72.

Rapp JP. 2000. Genetic analysis of inherited hypertension in the rat. Physiol Rev 80(1):135-72.

Redecker P, Bockmann J, Bockers TM. 2006. Expression of postsynaptic density proteins of the ProSAP/shank family in the thymus. Histochem Cell Biol 126(6):679-85.

Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, and others. 2006. Global variation in copy number in the human genome. Nature 444(7118):444-5.

Rescorla L. 1989. The language development survey: A screening tool for delayed language in toddlers. J Speech Hear Disord 54(4):587-99.

Rescorla L and Alley A. 2001. Validation of the language development survey (LDS): A parent report tool for identifying language delay in toddlers. J Speech Lang Hear Res 44(2):434-45.

Richer LP, Shevell MI, Miller SP. 2001. Diagnostic profile of neonatal hypotonia: An 11-year study. Pediatr Neurol 25(1):32-7.

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

Rollins JD, Sarasua SM, Phelan MC, DuPont BR, Rogers RC, Collins JS. 2011. Growth in Phelan-McDermid syndrome. Am J Med Genet (155):2324-6.

Rollins JD, Collins JS, Holden KR. 2010. United states head circumference growth reference charts: Birth to 21 years. J Pediatr 156(6):907,13, 913.e1-2.

Romain DR, Goldsmith J, Cairney H, Columbanogreen LM, Smythe RH, Parfitt RG. 1990. Partial monosomy for chromosome 22 in a patient with del(22)(pter-]Q13 1=q13 33-]qter). J Med Genet 27(9):588-9.

Rosenfeld JA, Crolla JA, Tomkins S, Bader P, Morrow B, Gorski J, Troxell R, Forster-Gibson C, Cilliers D, Hislop RG. 2010. Refinement of causative genes in monosomy 1p36 through clinical and molecular cytogenetic characterization of small interstitial deletions. American Journal of Medical Genetics Part A 152(8):1951-9.

Sammalisto S, Hiekkalinna T, Suviolahti E, Sood K, Metzidis A, Pajukanta P, Lilja HE, Soro-Paavonen A, Taskinen MR, Tuomi T, and others. 2005. A male-specific quantitative trait locus on 1p21 controlling human stature. J Med Genet 42(12):932-9.

Sarasua SM, Dwivedi A, Boccuto L, Rollins JD, Chen CF, Rogers RC, Phelan K, DuPont BR, Collins JS. 2011. Association between deletion size and important phenotypes expands the genomic region of interest in Phelan-McDermid syndrome (22q13 deletion syndrome). J Med Genet 48(11):761-6.

Sathyamoorthi S, Morales J, Bermudez J, McBride L, Luquette M, McGoey R, Oates N, Hales S, Biegel JA, Lacassie Y. 2009. Array analysis and molecular studies of INI1 in an infant with deletion 22q13 (Phelan-McDermid syndrome) and atypical teratoid/rhabdoid tumor. Am J Med Genet A 149A(5):1067-9.

Sato D, Lionel AC, Leblond CS, Prasad A, Pinto D, Walker S, O'Connor I, Russell C, Drmic IE, Hamdan FF, and others. 2012. SHANK1 deletions in males with autism spectrum disorder. Am J Hum Genet 90(5):879-87.

Schinzel AA, Basaran S, Bernasconi F, Karaman B, Yuksel-Apak M, Robinson WP. 1994. Maternal uniparental disomy 22 has no impact on the phenotype. Am J Hum Genet 54(1):21-4.

Schumacher FR, Berndt SI, Siddiq A, Jacobs KB, Wang Z, Lindstrom S, Stevens VL, Chen C, Mondul AM, Travis RC, and others. 2011. Genome-wide association study identifies new prostate cancer susceptibility loci. Hum Mol Genet 20(19):3867-75.

Schwartz CE and Neri G. 2012. Autism and intellectual disability: Two sides of the same coin. Am J Med Genet C Semin Med Genet 160C(2):89-90.

Seong E, Seasholtz AF, Burmeister M. 2002. Mouse models for psychiatric disorders. Trends in Genetics 18(12):643-50.

Shaikh TH, O'Connor RJ, Pierpont ME, McGrath J, Hacker AM, Nimmakayalu M, Geiger E, Emanuel BS, Saitta SC. 2007. Low copy repeats mediate distal chromosome 22q11.2 deletions: Sequence analysis predicts breakpoint mechanisms. Genome Res 17(4):482-91.

Shapira SK, McCaskill C, Northrup H, Spikes AS, Elder F, Sutton VR, Korenberg JR, Greenberg F, Shaffer LG. 1997. Chromosome 1p36 deletions: The clinical phenotype and molecular characterization of a common newly delineated syndrome. Am J Hum Genet 61(3):642.

Sheng M and Kim E. 2000. The shank family of scaffold proteins. J Cell Sci 113 (Pt 11)(Pt 11):1851-6.

Sigurdardottir S, Goodman BK, Rutberg J, Thomas GH, Jabs EW, Geraghty MT. 1999. Clinical, cytogenetic, and fluorescence in situ hybridization findings in two cases of "complete ring" syndrome. Am J Med Genet 87(5):384-90.

Silverman JL, Yang M, Lord C, Crawley JN. 2010. Behavioural phenotyping assays for mouse models of autism. Nature Reviews Neuroscience 11(7):490-502.

Smith ZE and Higgs DR. 1999. The pattern of replication at a human telomeric region (16p13.3): Its relationship to chromosome structure and gene expression. Hum Mol Genet 8(8):1373-86.

Speliotes EK, Yerges-Armstrong LM, Wu J, Hernaez R, Kim LJ, Palmer CD, Gudnason V, Eiriksdottir G, Garcia ME, Launer LJ, and others. 2011. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. PLoS Genet 7(3):e1001324.

Spiteri E, Konopka G, Coppola G, Bomar J, Oldham M, Ou J, Vernes SC, Fisher SE, Ren B, Geschwind DH. 2007. Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. Am J Hum Genet 81(6):1144-57.

Stankiewicz P and Lupski JR. 2002a. Molecular-evolutionary mechanisms for genomic disorders. Curr Opin Genet Dev 12(3):312-9.

Stankiewicz P and Lupski JR. 2002b. Genome architecture, rearrangements and genomic disorders. Trends Genet 18(2):74-82.

Stewart DR and Kleefstra T. 2007. The chromosome 9q subtelomere deletion syndrome. Am J Med Genet C Semin Med Genet 145C(4):383-92.

Stone JL, Merriman B, Cantor RM, Yonan AL, Gilliam TC, Geschwind DH, Nelson SF. 2004. Evidence for sex-specific risk alleles in autism spectrum disorder. Am J Hum Genet 75(6):1117-23.

Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, Zhang J, Soden R, Hayakawa M, Kreiman G, and others. 2004. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A 101(16):6062-7.

Sykes NH, Toma C, Wilson N, Volpi EV, Sousa I, Pagnamenta AT, Tancredi R, Battaglia A, Maestrini E, Bailey AJ, and others. 2009. Copy number variation and association analysis of SHANK3 as a candidate gene for autism in the IMGSAC collection. Eur J Hum Genet 17(10):1347-53.

Tagaya M, Mizuno S, Hayakawa M, Yokotsuka T, Shimizu S, Fujimaki H. 2008. Recombination of a maternal pericentric inversion results in 22q13 deletion syndrome. Clin Dysmorphol 17(1):19-21.

Talkowski ME, Mullegama SV, Rosenfeld JA, van Bon BW, Shen Y, Repnikova EA, Gastier-Foster J, Thrush DL, Kathiresan S, Ruderfer DM, and others. 2011. Assessment of 2q23.1

microdeletion syndrome implicates MBD5 as a single causal locus of intellectual disability, epilepsy, and autism spectrum disorder. Am J Hum Genet 89(4):551-63.

Tobaben S, Sudhof TC, Stahl B. 2000. The G protein-coupled receptor CL1 interacts directly with proteins of the shank family. J Biol Chem 275(46):36204-10.

Toruner GA, Kurvathi R, Sugalski R, Shulman L, Twersky S, Pearson PG, Tozzi R, Schwalb MN, Wallerstein R. 2009. Copy number variations in three children with sudden infant death. Clin Genet 76(1):63-8.

Tu JC, Xiao B, Naisbitt S, Yuan JP, Petralia RS, Brakeman P, Doan A, Aakalu VK, Lanahan AA, Sheng M, and others. 1999. Coupling of mGluR/homer and PSD-95 complexes by the shank family of postsynaptic density proteins. Neuron 23(3):583-92.

Tufano M, Della Corte C, Cirillo F, Spagnuolo MI, Candusso M, Melis D, Torre G, Iorio R. 2009. Fulminant autoimmune hepatitis in a girl with 22q13 deletion syndrome: A previously unreported association. Eur J Pediatr 168(2):225-7.

Twigger SN, Shimoyama M, Bromberg S, Kwitek AE, Jacob HJ, RGD Team. 2007. The rat genome database, update 2007--easing the path from disease to data and back again. Nucleic Acids Res 35(Database issue):D658-62.

Uchino J, Suzuki M, Hoshino K, Nomura Y, Segawa M. 2001. Development of language in rett syndrome. Brain Dev 23 Suppl 1:S233-5.

van Spronsen M and Hoogenraad CC. 2010. Synapse pathology in psychiatric and neurologic disease. Curr Neurol Neurosci Rep 10(3):207-14.

Verhoeven WM, Egger JI, Willemsen MH, de Leijer GJ, Kleefstra T. 2012. Phelan-McDermid syndrome in two adult brothers: Atypical bipolar disorder as its psychopathological phenotype? Neuropsychiatr Dis Treat 8:175-9.

Vernes SC, Spiteri E, Nicod J, Groszer M, Taylor JM, Davies KE, Geschwind DH, Fisher SE. 2007. High-throughput analysis of promoter occupancy reveals direct neural targets of FOXP2, a gene mutated in speech and language disorders. Am J Hum Genet 81(6):1232-50.

Vernes SC, Newbury DF, Abrahams BS, Winchester L, Nicod J, Groszer M, Alarcon M, Oliver PL, Davies KE, Geschwind DH, and others. 2008. A functional genetic link between distinct developmental language disorders. N Engl J Med 359(22):2337-45.

Vernes SC, Oliver PL, Spiteri E, Lockstone HE, Puliyadi R, Taylor JM, Ho J, Mombereau C, Brewer A, Lowy E, and others. 2011. Foxp2 regulates gene networks implicated in neurite outgrowth in the developing brain. PLoS Genet 7(7):e1002145.

Verpelli C and Sala C. 2011. Molecular and synaptic defects in intellectual disability syndromes. Curr Opin Neurobiol 22(3):530-6.

Verpelli C, Dvoretskova E, Vicidomini C, Rossi F, Chiappalone M, Schoen M, Di Stefano B, Mantegazza R, Broccoli V, Bockers TM, and others. 2011. Importance of Shank3 protein in regulating metabotropic glutamate receptor 5 (mGluR5) expression and signaling at synapses. J Biol Chem 286(40):34839-50.

Waga C, Okamoto N, Ondo Y, Fukumura-Kato R, Goto Y, Kohsaka S, Uchino S. 2011. Novel variants of the SHANK3 gene in japanese autistic patients with severe delayed speech development. Psychiatr Genet 21(4):208-11.

Wang X, McCoy PA, Rodriguiz RM, Pan Y, Je HS, Roberts AC, Kim CJ, Berrios J, Colvin JS, Bousquet-Moore D, and others. 2011. Synaptic dysfunction and abnormal behaviors in mice lacking major isoforms of Shank3. Hum Mol Genet 20(15):3093-108.

Watt JL, Olson IA, Johnston AW, Ross HS, Couzin DA, Stephen GS. 1985. A familial pericentric inversion of chromosome 22 with a recombinant subject illustrating a 'pure' partial monosomy syndrome. J Med Genet 22(4):283-7.

Weterings E and Chen DJ. 2008. The endless tale of non-homologous end-joining. Cell Res 18(1):114-2.

Wilson HL, Wong AC, Shaw SR, Tse WY, Stapleton GA, Phelan MC, Hu S, Marshall J, McDermid HE. 2003. Molecular characterisation of the 22q13 deletion syndrome supports the role of haploinsufficiency of SHANK3/PROSAP2 in the major neurological symptoms. J Med Genet 40(8):575-84.

Wilson HL, Crolla JA, Walker D, Artifoni L, Dallapiccola B, Takano T, Vasudevan P, Huang S, Maloney V, Yobb T, and others. 2008. Interstitial 22q13 deletions: Genes other than SHANK3 have major effects on cognitive and language development. Eur J Hum Genet 16(11):1301-10.

Winter RM. 2009 LONDON DYSMORPHOLOGY DATABASE: Diagnostic and research applications. <http://eurogene.open.ac.uk/content/london-dysmorphology-database-diagnostic-and-research-applications>.

Wong AC, Ning Y, Flint J, Clark K, Dumanski JP, Ledbetter DH, McDermid HE. 1997. Molecular characterization of a 130-kb terminal microdeletion at 22q in a child with mild mental retardation. Am J Hum Genet 60(1):113-20.

Wong KK, deLeeuw RJ, Dosanjh NS, Kimm LR, Cheng Z, Horsman DE, MacAulay C, Ng RT, Brown CJ, Eichler EE, and others. 2007. A comprehensive analysis of common copy-number variations in the human genome. Am J Hum Genet 80(1):91-104.

World Health Organization. 2006. WHO child growth standards: Methods and development: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age. Geneva: World Health Organization.

Wu C, Miao X, Huang L, Che X, Jiang G, Yu D, Yang X, Cao G, Hu Z, Zhou Y, and others. 2011. Genome-wide association study identifies five loci associated with susceptibility to pancreatic cancer in chinese populations. Nat Genet 44(1):62-6.

Yashin AI, Wu D, Arbeev KG, Ukraintseva SV. 2010. Joint influence of small-effect genetic variants on human longevity. Aging (Albany NY) 2(9):612-20.

Yu S, Graf WD, Ramalingam A, Brawner SJ, Joyce JM, Fiedler S, Zhou XG, Liu HY. 2011. Identification of copy number variants on human chromosome 22 in patients with a variety of clinical findings. Cytogenet Genome Res 134(4):260-8.

Yuan X, Waterworth D, Perry JR, Lim N, Song K, Chambers JC, Zhang W, Vollenweider P, Stirnadel H, Johnson T, and others. 2008. Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes. Am J Hum Genet 83(4):520-8.

Zhang X, Snijders A, Segraves R, Zhang X, Niebuhr A, Albertson D, Yang H, Gray J, Niebuhr E, Bolund L. 2005. High-resolution mapping of genotype-phenotype relationships in cri du chat syndrome using array comparative genomic hybridization. The American Journal of Human Genetics 76(2):312-26.

Zhang J, Feuk L, Duggan GE, Khaja R, Scherer SW. 2006. Development of bioinformatics resources for display and analysis of copy number and other structural variants in the human genome. Cytogenet Genome Res 115(3-4):205-14.

Zhao J, Bacolla A, Wang G, Vasquez KM. 2010. Non-B DNA structure-induced genetic instability and evolution. Cellular and Molecular Life Sciences 67(1):43-62.

Zitzer H, Honck HH, Bachner D, Richter D, Kreienkamp HJ. 1999. Somatostatin receptor interacting protein defines a novel family of multidomain proteins present in human and rodent brain. J Biol Chem 274(46):32997-3001.

CHAPTER TWO

ASSOCIATION BETWEEN DELETION SIZE AND IMPORTANT PHENOTYPES EXPANDS THE GENOMIC REGION OF INTEREST IN PHELAN-MCDERMID SYNDROME (22Q13 DELETION SYNDROME)

[This paper was published as: Sarasua SM, Dwivedi A, Boccuto L, Rollins JD, Chen CF, Rogers RC, Phelan K, DuPont BR, Collins JS. 2011. Association between deletion size and important phenotypes expands the genomic region of interest in Phelan-McDermid syndrome (22q13 deletion syndrome). J Med Genet 48(11):761-6. Reprinted with permission from BMJ Publishing Group Ltd.]

Abstract

Background : The clinical features of Phelan-McDermid syndrome (also known as 22q13 deletion syndrome) are highly variable and include hypotonia, speech and other developmental delays, autistic traits, and mildly dysmorphic features. Patient deletion sizes are also highly variable, prompting this genotype-phenotype association study. **Methods:** Terminal deletion breakpoints were identified for 71 individuals in a patient cohort using a custom-designed high-resolution oligonucleotide array comparative genomic hybridization platform with a resolution of 100 bp. **Results:** Patient deletion sizes were highly variable, ranging from 0.22 to 9.22 Mb, and no common breakpoint was observed. *SHANK3*, the major candidate gene for the neurologic features of the syndrome, was deleted in all cases. Sixteen features (neonatal hypotonia, neonatal hyporeflexia, neonatal feeding problems, speech/language delay, delayed age at crawling, delayed age at walking, severity of developmental delay, male genital anomalies, dysplastic toenails, large or fleshy hands, macrocephaly, tall stature, facial asymmetry, full brow, atypical reflexes, and dolichocephaly) were found to be significantly associated with larger deletion sizes, suggesting the role of additional genes or regulatory regions proximal to *SHANK3*. Individuals with autism spectrum disorders (ASDs) were found to

have smaller deletion sizes (median deletion size of 3.39 Mb) than those without ASDs (median deletion size 6.03 Mb, *P*=0.0144). This may reflect the difficulty in diagnosing ASDs in individuals with severe developmental delay. **Conclusions:**
This genotype-phenotype analysis explains some of the phenotypic variability in the syndrome and identifies new genomic regions with a high likelihood for causing important developmental phenotypes such as speech delay.

Introduction

Phelan-McDermid syndrome (PMS [MIM 606232]), also known as 22q13 deletion syndrome, is a rare syndrome characterized by developmental delay, absent or impaired speech, neonatal hypotonia, autistic traits, and mild dysmorphic features.[1-8] Affected individuals have deletions ranging in size from 100 kb to over 9 Mb.[5] Because PMS is considered to be underdiagnosed, the true prevalence is unknown.[1] In an evaluation of more than 11,000 cases with developmental disabilities, 22q was the second most frequent subtelomeric rearrangement, identified in 0.2% of those evaluated.[9] Simple terminal deletions account for approximately 75% of PMS cases.[10] The other cases have been comprised of translocations in the 22q13 region,[2,11-13] ring chromosome 22,[1,2,14-16] and mosaics.[17,18]

Until recently, diagnostic technologies relied upon cytogenetic banding and fluorescence in-situ hybridization (FISH) or bacterial artificial chromosome comparative genomic hybridization (CGH), which are not always able to detect smaller deletions or to accurately

measure deletion size or breakpoints. Oligo array CGH allows for a much higher resolution of the chromosomal breakpoints.

Most of the published work suggests that the loss of one copy of *SHANK3* (SH3 and multiple ankyrin repeat domains 3, also referred to as *ProSAP2* or proline rich synapse associated protein 2*)* is responsible for the neurological features of the PMS phenotype, since *SHANK3* maps within the region of common deletion observed in patients.[5,6] A patient with t(12;22)(q24.1;q13.3), which disrupted *SHANK3,* gave evidence implicating this gene as a major candidate for the neurological features of the syndrome.[11] Further, a *de novo* interstitial deletion disrupting only *SHANK3* was observed in an individual with developmental delay, speech delay, and mildly dysmorphic features including ptosis, epicanthal folds, and cupped ears.[19] The authors concluded that haploinsufficiency of *SHANK3* alone, and not genes telomeric to it, was responsible for PMS.

SHANK3 is predominantly expressed in brain tissue and the expression appears to be regulated by tissue-specific methylation.[20,21] SHANK3, a structural protein, is considered to be critical in the assembly, maintenance, and plasticity of the post-synaptic density (PSD) at excitatory synapses in the brain.[22] SHANK3, along with other PSD components, including cell adhesion molecules such as neurexins[23-26] and neuroligins,[27,28] as well as scaffolding proteins, has been found to be associated with autism spectrum disorders (ASDs).[29-31] However, a recent finding of two individuals with PMS phenotypes and interstitial deletions outside of the *SHANK3* region has expanded the search for additional genetic causes of the 22q13 deletion phenotype.[32]

Another gene located near *SHANK3,* and deleted in most cases of PMS, is *IB2. IB2 (islet brain 2)*, also known as *MAPK8IP2 (mitogen-activated protein kinase 8 interacting protein 2)*, is 70kb proximal to *SHANK3* and performs critical neurologic functions.[33] Giza et al. demonstrated that the IB2 protein is located in the PSD and throughout the brain and is important in synaptic transmission and neural morphology. A full knock-out mouse model of IB2 demonstrated reduced cognitive ability, learning and social interaction.[33]

This analysis, the first genotype-phenotype comparison to use high-resolution deletion breakpoint mapping on a large sample size, addressed the hypothesis that additional genes or regions of chromosome 22q13 besides *SHANK3* contribute to the PMS phenotype.

Subjects and Methods

Subjects

Study subjects were previously diagnosed with PMS and most attended one or more PMS Family Support Conferences held in 2001, 2004, 2006, and 2008 in Greenville, South Carolina and in 2006 in Melbourne, Australia. The majority of blood specimens were collected at the 2006 and 2008 conferences in Greenville, South Carolina. The study was approved by the Institutional Review Board of Self Regional Healthcare (Greenwood, South Carolina), and all participants' parents or guardians provided signed informed consent forms. All participants have a terminal deletion encompassing the *SHANK3* gene. Individuals with known or self-reported chromosomal anomalies other than terminal deletions were not included in the

analysis in order to focus the study on 22q13 deletion effects. Some participants in the present study may have participated in other studies published elsewhere.

Information on physical features was obtained from physical examinations conducted by experienced clinical geneticists following standardized assessment checklists or, for two patients, abstracted from a medical record. Stature and head circumference were measured at the Family Conferences and the remaining physical features were evaluated based upon clinical judgment. These physician-confirmed features are listed in Supplemental Table 2.1. Medical history was obtained from standardized medical history questionnaires administered during an in-person interview at the Family Conferences or completed independently by parents and mailed or emailed to the investigators. When available, this parent-provided information was supplemented with information obtained from the physical examinations or from the abstracted medical record. These features are presented in Supplemental Table 2.2. For 72% of participants, more than one record source was available (physical examination, medical history questionnaire, or evaluation of the same individual across several years). In instances where a discrepancy between records was identified, physician-provided answers were used in place of parent-provided information. In cases where health information differed between different years of participation for an individual, positive responses were used such that the information represents “ever” reporting a feature. This information was entered in a Microsoft Access (Redmond, WA) database and checked for accuracy.

Genetic analysis

Genetic deletions were measured using a custom 4x44K 60-mer oligo array designed to cover chromosome 22q12.3-qter by Oxford Gene Technology (Oxford, UK). In brief, genomic DNA was isolated and subsequently purified using the Zymo DNA Clean & Concentrator™ kit (ZymoResearch, Orange, California) according to manufacturer's instructions. Reference DNA used in the comparative hybridization was obtained from Promega (Madison, Wisconsin). DNA concentration and purity were determined with a ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, Delaware). Two µg DNA from patients' peripheral blood and reference samples were digested at 37ºC for 2 hours with 5 U of *RsaI* and *AluI* (Promega). After heat inactivation of the enzymes, the samples were labeled with either Cy3- or Cy5-dUTP using the Agilent Genomic DNA labeling kit PLUS (Santa Clara, California). Samples were purified with YM-30 Microcon filters from Millipore (Bedford, Massachusetts). Hybridization and washes were conducted employing Oxford Gene Technology's (OGT) CytoSure™ Chromosome 22q specific array protocol (Oxford, UK). Arrays were scanned with the GenePix 4000B scanner (Molecular Devices, Sunnyvale, California). Array feature extraction was performed with GenePix Pro 6.1. Copy number/data analysis was performed with OGT's CytoSure™/Oligome viewer software package. Deletion breakpoints are accurate to 100 bp resolution and the array CGH genomic coordinates were established according to the 2006 human genome build 18 (GRCh 36/NCBI build 36.1).[34] Deletion sizes were plotted on the genome browser using the University of California at Santa Cruz Genome Browser.[35]

Statistical analysis

Statistical analyses were performed using SAS version 9.2 (SAS Institute Inc., Cary, NC) to examine the association between deletion size and clinical features. Because the distribution of deletion sizes was non-normal, non-parametric statistical tests including the Spearman rank correlation coefficient and two-sided Wilcoxon rank-sum test were used when the dependent variable was deletion size. In cases where expected cell sizes were small (<5), exact methods were used. Linear regression was used to examine the effect of the deletion size on continuous outcome measures such as speech and developmental features.

Results

Phenotype and deletion breakpoint data from 71 individuals were used in this analysis. Physical examination information was available for 54 (76%) and parent-provided medical history was available for 61 (86%) individuals. A total of 84 clinical phenotypic features were assessed. The age of participants ranged from 0.4 to 40 years, with a mean of 7.6 years (standard deviation of 2.5 years). The cohort was composed of 42 females and 29 males for a female to male ratio of 1.45:1. Median deletion size was similar for males (6.03 Mb) and females (5.24 Mb, *P*=0.2678). Deletion sizes ranged from 0.22 to 9.22 Mb (see Figure 2.1), with a mean of 5.08 Mb, a median of 5.25 Mb and a standard deviation of 2.56 Mb. *SHANK3* was deleted in all individuals and *IB2* was deleted in all but two individuals. Individuals with physical examination data tended to have smaller deletions (mean = 4.7 Mb) than those with only medical history information (mean = 6.5 Mb, *P*=0.0102), although the full range of deletion sizes

were observed for both groups. The most common features observed in this cohort include developmental delay, expressive speech/lanaguage delay, long eyelashes, increased pain tolerance, dysplastic toenails, and hypotonia.

Sixteen of 84 characteristics were found to be significantly associated with larger deletion sizes (Tables 2.1 and 2.2; a list of all assessed features is provided in Supplemental Tables 1 and 2). These characteristics tended to be related to physical features including: dolichocephaly, facial asymmetry, full (prominent) brow, large or fleshy hands, dysplastic toenails, tall stature, macrocephaly, atypical reflexes, and male genital anomalies; neonatal features: hypotonia, feeding problems, and hyporeflexia; and developmental delays: speech delay, developmental delay, later age to crawl, and later age to walk. Among those with atypical reflexes, median deletion sizes were similar among those with strong reflexes (n=7, median deletion size 6.46 Mb, range of 1.85-8.96 Mb) and those with weak reflexes (n=9, median deletion size 6.57, range of 1.98-8.20). Both were significantly larger than those with typical reflexes (Median deletion size 4.19 Mb, range 0.34-8.62). Behaviorial features were not associated with increased deletion size. However, two features, ASD and aggressive behavior, were associated with smaller median deletion sizes.

Prenatal and neonatal features

Features present in the neonatal period, including hypotonia, feeding problems, and hyporeflexia, were all associated with larger deletion sizes (Table 2.2). No association was found between deletion size and low birth weight (reported for 23% of the sample) or preterm birth (reported for 25% of the sample; Supplemental Table 2).

Autism Spectrum Disorders

Autism spectrum disorders were reported in 26% of the patients over age three years (Table 2.2). The median deletion size for those with ASDs was smaller (3.39 Mb, range 0.22 to 7.19 Mb) than the median deletion size for those without ASDs (6.03 Mb, range 0.57 to 9.22, *P*=0.0144). Similar to ASDs, aggressive behavior was also associated with smaller median deletion size (Table 2.2).

Developmental Delay

All patients had some degree of developmental delay. Severity of developmental delay, as rated by parents on an ordinal scale from mild, moderate, severe, and profound, was significantly and positively associated with deletion size when using linear regression (*P*=0.009). Similarly, later age to crawl (*P*=0.0032) and later age to walk (*P*=<0.0001) were significantly associated with larger deletion sizes.

Speech and Language Delay

Speech was absent or delayed for 100% of individuals. On the questionnaires administered in 2004 and 2006, parents were asked whether speech was absent or severely impaired and to note how many words the patient used. In the questionnaire administered in 2008, parents were asked whether speech was absent or severely impaired, how many words were used in a sentence, and to provide additional comments describing speech. Among those over three years of age for whom speech information was provided (n=50), half reported no speech. Another 28% had 40 or fewer words and did not report speaking in phrases or sentences. The final 22% reported having sentences or phrases, talking as the primary means of

communication, or more than 40 words and were coded as having "sentences" (Table 2.1). Median deletion size was higher for those with absent speech (6.72 Mb) and smaller for those with sentences (3.27 Mb). When deletion size was examined in detail, none of the 22 individuals with deletion sizes greater than 5.3 Mb was reported to speak in sentences whereas 39% of the 28 individuals with deletion sizes smaller than 5.3 Mb use sentences (*P*=0.001). Deletion size showed a significant negative correlation with the number of words spoken (*P*=0.0102).

Growth

Growth was found to be non-linearly associated with deletion size (Table 2.1); those with normal stature had the smallest median deletion size at 4.80 Mb, while those with tall stature (>95th percentile) had a median deletion size of 6.18 Mb, and those with short stature (<5th percentile) had the largest median deletion size at 8.06 Mb. Macrocephaly was associated with increased deletion size (median deletion size 6.99 Mb) whereas those with microcephaly had similar deletion sizes compared to those with normocephaly (median deletion size of 3.32 Mb compared to 3.34 Mb, respectively).

Discussion

This genotype-phenotype study is the largest to date to include high-resolution deletion breakpoint genotype information along with clinical features to identify associations between features of Phelan-McDermid syndrome and deletion sizes. Prior smaller studies found mixed indications of association between deletion size and the severity of the phenotype.[5] In

particular, our findings are consistent with previous reports of correlations between increased deletion size and the severity of selected phenotypes.[6,15,36] A case series of eight patients, which included neuroimaging, found that individuals with small deletions (~0.15 Mb) had decreased brain anomalies when compared to individuals with larger deletions.[7] In another study of 12 subjects with oligo array CGH breakpoint data, patients with large hands tended to have larger deletions,[36] in agreement with our findings. In a study of 30 individuals with 22q13 deletions manifesting as ring chromosomes, Jeffries et al. found increased deletion size positively correlated with dysmorphic features related to ears, toenails, and philtrum as well as the developmental features of increased severity of developmental delay and speech delay.[15] Wilson et al. found a positive correlation between increased deletion size and developmental delay, hypotonia, head circumference, ear infections, pointed chin, dental anomalies, and several measures of independent behavior and daily living abilities.[6] Our study also found an association of increased deletion size with neonatal hypotonia, head circumference, and facial features (Tables 2.1-2.2).

Supporting our finding of larger deletions being associated with more severe phenotypes are reports of three individuals with interstitial deletions overlapping the larger deletion regions in our patients.[32,37] Wilson et al. report two cases with an intact *SHANK3* gene, but presenting with speech delay (two words each and no sentences), macrocephaly, tall stature, hypotonia, delayed walking, and developmental delay.[32] Fujita et al. described an 18 month old Japanese girl with a del(22)(q13.1q13.2) with hypotonia, psychomotor delays, minor dysmorphic features (including dolichocephaly and full brow), and hearing loss due to inner ear

anomalies.[37] These cases suggest there are clinically important genes located proximal to *SHANK3*.

Speech and Language

Most prior studies reporting on the effect of deletion size on speech and language delay do not distinguish between absent speech and delayed or impaired speech. The present study found genomic differences by severity of language delay, particularly the dramatic difference in ability to speak in sentences being present in nearly 40% of those with smaller deletions but absent in those with larger deletions. Previous findings of severe speech impairment among two cases with interstitial deletions [32] support the presence of genes affecting speech in the regions proximal to *SHANK3*. None of the genes in the deletion region was identified as being a transcriptional target of FOXP2, a transcription factor known to be associated with speech.[38] Future work is critically needed to distinguish degrees of speech impairment with the genomic regions deleted.

Autism Spectrum Disorders

In this study, individuals reported by parents to have an autism spectrum disorder had smaller deletions than those without an ASD, but all are missing a copy of *SHANK3*, a gene implicated in autism.[29-31] Approximately 26% of those over age 3 were reported to have an ASD; all were reported to have some degree of developmental delay. The prevalence of ASD in the larger population of PMS patients is unknown, with published reports ranging from 0 to 94% (0/8 (0%),[7] 1/6 (17%),[39] 6/11 (55%),[40] 3/5 (60%),[36] 12/27 (44%) or 23/27 (85%) depending on the definition,[15] and 17/18 (94%) [5]). Complicating a comparison are the

different diagnostic and reporting criteria used, as well as ages and degree of developmental delay of patients assessed. Our analysis of autism is subject to several limitations. We relied upon parent report of an autism or autism spectrum disorder diagnosis rather than a standardized instrument such as the Autism Diagnostic Interview, revised, [41] which would have improved the validity of this assessment. We found that those with larger deletions tended to be more severely developmentally delayed. It is possible that more severely impaired individuals may have been less likely to have been assessed for autism, that a diagnosis of PMS may have "replaced" or precluded an autism diagnosis, that more severe physical and intellectual disabilities may obscure autistic features, or that the assessments may be difficult to administer to or be inappropriate for the more severely impaired.

Growth

One of the commonly associated phenotypes of PMS is "normal to accelerated growth."[4,5] A recent analysis of the same cohort noted that both tall (>95th percentile) and short stature (<5th percentile), as well as microcephaly (<3rd percentile) and macrocephaly (>97th percentile) are common in individuals with 22q13 deletion.[42] Additionally, the present analysis provides support for the presence of distinct deletion regions associated with short stature, tall stature, and macrocephaly.

Limitations

There are several limitations to the present analyses. A large number of phenotypes were examined with statistical tests and, as these analyses were considered exploratory, were not corrected for multiple testing. Given that 84 phenotypes were examined and using a *P*-

value cut-off of <0.05, one might expect four to be identified due to chance, whereas 18 phenotypes were found to be statistically significant (16 associated with larger deletion sizes and two with smaller deletion sizes). Additionally, some phenotypes may be correlated with each other. Further, high-resolution karyotype information was not available for all participants. Individuals with known or self-reported chromosomal anomalies other than simple deletions were removed from the analysis. It is still possible that some individuals in the present cohort include patients with r(22), translocations, or other anomalies. Nonetheless, prior studies of r(22) noted similar phenotypes to 22q13 deletion,[2,15,40] indicating that the ring structure does not alter the phenotype. The proportion of our study group reporting chromosomal anomalies (37/108 or 34%) is similar to those reported elsewhere.[1,10] Thus, it is unlikely that our results are significantly confounded by the presence of unaccounted-for chromosomal rearrangements. It should be noted that the use of chromosomal microarrays to detect copy number variants is now recommended as the primary test, over G-banding or FISH, for individuals with developmental delay.[43] Finally, analysis of physical features was restricted to those who had medical records or physical examinations at the family conferences. This restriction was used to most accurately characterize phenotypes. However, it was found that individuals with physical examinations tended to have smaller deletion sizes. It is possible that individuals with larger deletions, and possibly a more severe phenotype, were less likely to travel to participate in the family conferences and thus those included in the present analysis group may represent a somewhat less severely affected population. However, the full range of deletion sizes was observed in this group.

Summary

This study implicates genomic regions proximal to *SHANK3* in language, movement, developmental delay, and some dysmorphic features in individuals with Phelan-McDermid syndrome. This is the first study of PMS to distinguish different degrees of speech delay with deletion size. These findings are critical as they provide further regions of interest to help elucidate clinical implications for affected individuals. Current investigations of *SHANK3*, *IB2*, and other telomeric genes should be supplemented to determine the independent and additive impacts of the additional loss of genes, micro RNAs, or regulatory elements in this region of chromosome 22q13.

Acknowledgments The authors gratefully acknowledge the assistance of Kristi Atkinson, Kaitlyn O'Gorman, Alyssa Sattler, Keri Sprouse, and Taylor Sartain for data entry and verification. We also appreciate the review and helpful comments provided on the manuscript by Drs. Charles E. Schwartz and Roger E. Stevenson.

Funding Support for this study was provided by the Greenwood Genetic Center, the Phelan-McDermid Syndrome Foundation (SMS), the Genetics Endowment of South Carolina, the South Carolina Birth Defects Foundation, and the South Carolina Department of Disabilities and Special Needs.

Table 2.1. Physical Examination Phenotypes Showing Significant Differences in 22q13 Deletion Sizes. [a]

Phenotype	Observed	N	%	Median Deletion Size, Mb	Range of Deletion sizes, MB	*P*-value[b]
Dysplastic toenails	+	40	75%	5.89	0.22-8.96	0.0210
	-	13	25%	2.72	0.34-8.66	
Full (prominent) brow	+	31	58%	5.78	1.62-8.96	0.0287
	-	22	42%	3.18	0.22-8.66	
Large or fleshy hands	+	29	55%	6.03	1.65-8.96	0.0030
	-	24	45%	2.91	0.22-8.66	
Head size, percentile	>97th %	11	20%	6.99	5.08-8.20	0.0078
	3rd-97th	37	67%	3.34	0.22-9.22	Ref[c]
	<3rd %	6	11%	3.32	1.34-8.55	0.9556
Reflexes	Atypical	16	37%	6.51	1.85-8.96	0.0298
	Typical	27	63%	4.19	0.34-8.62	
Dolichocephaly	+	16	30%	6.38	2.23-8.96	0.0467
	-	38	70%	4.51	0.22-8.66	
Stature, percentile	>95th	5	11%	6.18	5.78-9.22	0.0480
	5-95th	36	80%	4.80	0.22-8.55	Ref[c]
	<5th	5	11%	8.06	2.23-8.96	0.1698
Facial asymmetry	+	5	9%	6.87	6.46-8.07	0.0218
	-	49	91%	4.52	0.22-8.96	

[a]Phenotypes were obtained from physical examinations or medical records.
[b]Two-sided, Wilcoxon rank-sum test.
[c]Reference group for statistical comparison

Table 2.2. Medical History Features Showing Significant Differences in 22q13 Deletion Sizes. [a]

Feature	Reported	N	%	Median deletion size, Mb	Range of deletion sizes, Mb	*P*-value [b]
Neonatal hypotonia	+	49	80%	5.96	1.34-9.22	0.0006
	-	12	20%	2.01	0.22-8.39	
Walked later than 15 months[c]	+	37	80%	5.77	1.34-8.96	0.0108
	-	9	20%	1.85	0.22-8.39	
Expressive speech/language[c]	No words	25	50%	6.72	0.34-9.22	Ref[d]
	< 40 words	14	28%	5.52	1.62-7.32	0.1555
	Sentences	11	22%	3.27	0.22-5.25	0.0094
Neonatal feeding problems	+	47	77%	5.96	0.34-9.22	0.0399
	-	14	23%	3.22	0.22-8.39	
Genital anomalies, males	+	7	35%	8.20	6.18-8.96	0.0102
	-	13	65%	4.19	0.22-8.55	
Aggressive behavior	+	20	33%	4.20	0.57-8.20	0.0365
	-	40	67%	6.02	0.22-9.22	
Autism spectrum disorder[c]	+	14	26%	3.39	0.22-7.19	0.0144
	-	39	74%	6.03	0.57-9.22	
Neonatal hyporeflexia	+	12	24%	7.15	3.27-9.22	0.0032
	-	39	76%	4.52	0.22-8.96	

[a]Features were obtained from parent-provided medical history or, for two individuals, a medical record.

[b] Two-sided, Wilcoxon rank-sum test.

[c]Among those three or more years of age.

[d]Reference group for statistical comparison.

Supplemental Table 1. Physical Examination Phenotypes and 22q13 Deletion Size. [a]

Phenotypes	Observed	N	%	Median Deletion Size, Mb	Range of Deletion sizes, MB	*P*-value [b]
Long eyelashes	+	51	96%	5.07	0.22-8.96	0.4702
	-	2	4%	6.23	5.75-6.72	
Dysplastic toenails	+	40	75%	5.89	0.22-8.96	**0.0210**
	-	13	25%	2.72	0.34-8.66	
Hypotonia	+	38	73%	5.09	0.67-8.96	0.5335
	-	14	27%	5.12	0.22-8.66	
Lax ligaments	+	36	68%	5.16	0.22-8.96	0.5636
	-	17	32%	3.92	0.34-8.66	
Full or puffy cheeks	+	35	65%	5.09	1.04-8.96	0.4185
	-	19	35%	4.49	0.22-8.66	
Bulbous nose	+	33	61%	5.75	0.57-8.96	0.0701
	-	21	39%	3.92	0.22-8.66	
Hyperextensible joints	+	32	60%	5.09	0.22-8.38	0.9064
	-	21	40%	5.08	0.34-8.96	
Full (prominent) brow	+	31	58%	5.78	1.62-8.96	**0.0287**
	-	22	42%	3.18	0.22-8.66	
Large fleshy hands	+	29	55%	6.03	1.65-8.96	**0.0030**
	-	24	45%	2.91	0.22-8.66	
Pointed chin	+	29	55%	3.92	0.22-8.96	0.2097
	-	24	45%	5.50	0.24-8.66	
Small or recessed jaw	+	27	52%	3.92	1.04-8.62	0.5011
	-	25	48%	5.09	0.22-8.96	
Fully or puffy eyelids	+	26	49%	5.08	0.34-8.96	0.9788
	-	27	51%	5.08	0.22-8.66	
Epicanthal folds	+	24	45%	4.50	1.34-8.96	0.5765
	-	29	55%	5.25	0.22-8.66	
Ptosis	+	23	43%	5.23	1.65-8.96	0.108
	-	31	57%	4.52	0.22-8.66	
High arched palate	+	21	42%	5.78	1.62-8.62	0.4242
	-	29	58%	5.08	0.22-8.66	
Reflexes	Atypical	16	37%	6.51	1.85-8.96	**0.0298**
	typical	27	63%	4.19	0.34-8.62	
Sacral dimple	+	19	37%	6.02	1.34-8.66	0.0544
	-	33	63%	4.19	0.22-8.96	
2/3 toe syndactyly	+	19	37%	6.03	1.34-8.38	0.1495
	-	33	63%	4.74	0.22-8.96	
Bitemporal narrowing	+	19	35%	6.03	1.85-8.96	0.3237
	-	35	65%	5.08	0.22-8.66	

Long philtrum	+	19	35%	6.03	1.65-8.96	0.1337
	-	35	65%	4.52	0.22-8.66	
Flat midface	+	17	31%	5.75	1.65-8.96	0.1943
	-	37	69%	4.74	0.22-8.66	
Upslanting palpebral fissures	+	17	32%	4.19	1.34-8.96	0.6303
	-	36	68%	5.16	0.22-8.66	
Deepset eyes	+	17	32%	5.25	1.62-8.96	0.8283
	-	36	68%	5.08	0.22-8.66	
Widely spaced teeth	+	16	31%	2.70	0.22-8.62	0.1000
	-	36	69%	5.24	0.24-8.66	
Dolichocephaly	+	16	30%	6.38	2.23-8.96	**0.0467**
	-	38	70%	4.51	0.22-8.66	
Malocclusion	+	14	27%	5.24	1.62-8.66	0.3959
	-	38	73%	4.91	0.22-8.38	
Dysplastic fingernails	+	13	25%	5.25	0.57-8.07	0.8783
	-	40	75%	5.08	0.22-8.96	
5^{th} finger clinodactyly	+	13	25%	3.92	1.34-8.20	0.5881
	-	40	75%	5.09	0.22-8.96	
Strabismus	+	13	25%	6.09	2.22-8.38	0.0750
	-	39	75%	4.52	0.22-8.96	
Lymphedema	+	11	21%	6.63	0.57-8.62	0.1814
	-	41	79%	4.49	0.22-8.96	
Head size, percentile	>97^{th} %	11	20%	6.99	5.08-8.20	**0.0078**
	3^{rd}-97^{th}	37	67%	3.34	0.22-9.22	Ref [c]
	<3^{rd} %	6	11%	3.32	1.34-8.55	0.9556
Deep nasolacrimal groove	+	10	20%	6.10	1.62-7.22	0.2856
	-	41	80%	4.49	0.22-8.66	
Downslanting palpebral fissures	+	7	15%	6.03	2.30-6.87	0.6424
	-	41	85%	4.74	0.22-8.96	
Stature, percentile	>95^{th}	5	11%	6.18	5.78-9.22	**0.0480**
	5-95	36	80%	4.8	0.22-8.55	Ref [c]
	<5^{th}	5	11%	8.06	2.23-8.96	0.1698
Single palmar crease	+	5	9%	6.46	2.23-8.62	0.3971
	-	48	91%	5.08	0.22-8.96	
Facial asymmetry	+	5	9%	6.85	6.46-8.07	**0.0218**
	-	49	91%	4.52	0.22-8.96	
Skin tags	+	2	4%	3.56	1.34-5.78	0.4887
	-	50	93%	5.08	0.22-8.96	

[a]Phenotypes were obtained from physical examinations or medical records.
[b]Two-sided, Wilcoxon rank-sum test. Bold font indicates significant (P<0.05) finding.
[c]Reference group for statistical comparison.

Supplemental Table 2. Medical History Features and Size of 22q13 Deletion. [a]

Feature	Reported	N	%	Median deletion size, Mb	Range of deletion sizes, Mb	*P*-value[b]
Neonatal hypotonia	+	49	80%	5.96	1.34-9.22	**0.0006**
	-	12	20%	2.01	0.22-8.39	
Walked later than 15 months[c]	+	37	80%	5.77	1.34-8.96	**0.0108**
	-	9	20%	1.85	0.22-8.39	
Expressive speech/language[c]	No words	25	50%	6.72	0.34-9.22	Ref[d]
	<40 words	14	28%	5.52	1.62-7.32	0.1555
	Sentences	11	22%	3.27	0.22-5.25	**0.0094**
Neonatal feeding problems	+	47	77%	5.96	0.34-9.22	**0.0399**
	-	14	23%	3.22	0.22-8.39	
Overheats or turns red easily	+	40	70%	5.86	0.34-9.22	0.4012
	-	17	30%	5.25	0.22-8.39	
Gastrointestinal reflux	+	31	57%	5.23	0.22-9.22	0.8075
	-	23	43%	5.25	1.34-8.55	
Decreased perspiration	+	33	57%	6.03	0.34-9.22	0.3663
	-	25	43%	5.08	0.22-8.96	
Large, dysplastic, or prominent ears	+	15	38%	5.23	1.65-8.20	0.0621
	-	24	62%	3.88	0.22-7.86	
Failure to gain weight, newborn period	+	22	36%	6.30	1.34-8.96	0.2093
	-	39	64%	5.09	0.22-9.22	
Genital anomalies, males	+	7	35%	8.20	6.18-8.96	**0.0102**
	-	13	65%	4.19	0.22-8.55	
Reported any brain abnormality and had an MRI or imaging	+	20	28%	6.51	1.85-8.96	0.0923
	-	51	72%	5.08	0.22-9.22	
Autism spectrum disorder[c]	+	14	26%	3.39	0.22-7.19	**0.0144**
	-	39	74%	6.03	0.57-9.22	
Preterm birth	+	15	25%	6.99	1.34-8.62	0.1298
	-	45	75%	5.09	0.22-9.22	
Neonatal hyporeflexia	+	12	24%	7.15	3.27-9.22	**0.0032**
	-	39	76%	4.52	0.22-8.96	
Thick lower lip	+	14	24%	5.26	1.04-8.38	0.7698
	-	45	76%	5.09	0.22-9.22	
Low birth weight	+	14	23%	6.84	1.34-8.62	0.2658
	-	47	77%	5.23	0.22-9.22	
Precocious puberty	+	9	20%	7.32	1.04-9.22	0.1280
	-	37	80%	5.09	0.22-8.62	
Sleep disturbance[e]	+	14	20%	3.67	1.04-8.96	0.2605
	-	57	80%	5.77	0.22-9.22	

Diarrhea	+ -	10 42	19% 81%	5.56 5.51	0.57-9.22 0.22-8.96	0.8806
Ever having a seizure and taking anti-seizure medication	+ -	10 45	18% 82%	6.34 5.78	0.22-8.96 0.34-9.22	0.5366
Any kidney disease or abnormality	+ -	8 41	16% 84%	5.77 5.09	1.34-8.62 0.22-8.96	0.6381
Arachnoid cyst	+ -	8 44	15% 85%	5.08 5.77	3.09-8.55 0.22-9.22	0.9296
Patent ductus arteriosus or ventricular septal defect	+ -	5 37	12% 88%	8.20 5.25	2.22-8.96 0.22-8.62	0.1583
Vesicouretal reflux	+ -	6 51	11% 89%	6.18 5.23	5.75-8.62 0.22-9.22	0.2521
Sleep apnea	+ -	5 50	11% 89%	6.72 5.51	0.57-8.20 0.22-9.22	0.4763
Cellulitis	+ -	5 47	10% 90%	6.72 5.08	4.52-8.20 0.22-9.22	0.0945
Neonatal dehydration	+ -	6 52	10% 90%	6.85 5.24	1.62-8.55 0.22-9.22	0.5014
Neonatal hyperreflexia	+ -	3 48	6% 94%	8.20 5.24	3.96-8.39 0.22-9.22	0.2432
Polycystic kidney disease	+ -	3 48	6% 94%	5.23 5.76	1.34-5.78 0.22-8.96	0.3854
Sensory/Behavioral Problems						
Increased tolerance to pain/Does not show pain	+ -	53 5	91% 9%	5.25 7.86	0.22-9.99 2.23-8.39	0.1402
Chews nonfood items	+ -	52 9	84% 15%	5.51 5.75	0.22-9.22 3.09-8.62	0.4134
Poor eye contact	+ -	40 21	64% 36%	5.77 5.23	0.22-9.22 0.34-7.32	0.2821
Biting	+ -	37 22	63% 27%	5.75 5.87	0.57-9.22 0.22-8.96	0.3795
Sensitive to touch	+ -	38 22	63% 37%	6.10 4.81	0.22-9.22 0.34-8.02	0.0610
Laughs at misbehavior	+ -	35 24	59% 41%	5.25 5.99	0.22-9.22 1.62-8.55	0.9694
Impulsive behavior	+ -	33 25	57% 43%	4.85 6.03	0.22-8.96 1.62-9.22	0.1631
Excessive screaming	+ -	26 34	43% 57%	4.80 5.99	0.57-8.62 0.22-9.22	0.1969

Hair pulling	+	14	41%	6.54	2.10-9.22	0.0991
	-	20	59%	4.79	0.22-8.55	
Pinching	+	23	38%	4.49	0.34-8.39	0.0550
	-	38	62%	6.03	0.22-9.22	
Non-stop crying	+	22	37%	5.44	0.34-9.22	0.6043
	-	38	63%	5.76	0.22-8.39	
Aggressive behavior	+	20	33%	4.20	0.57-8.20	**0.0365**
	-	40	67%	6.02	0.22-9.22	
Tongue thrusting	+	17	30%	5.96	0.57-8.96	0.9152
	-	39	70%	5.75	0.22-9.22	
Overly emotional	+	12	27%	5.17	1.34-8.96	0.3037
	-	32	73%	6.09	0.22-9.22	
Self destructive	+	11	19%	6.03	0.57-8.20	0.9151
behavior	-	48	81%	5.50	0.22-9.22	
Mistreats animals	+	11	19%	6.72	1.65-8.96	0.3591
	-	48	81%	5.76	0.22-9.22	

[a]Features were obtained from parent-provided medical history or medical record.

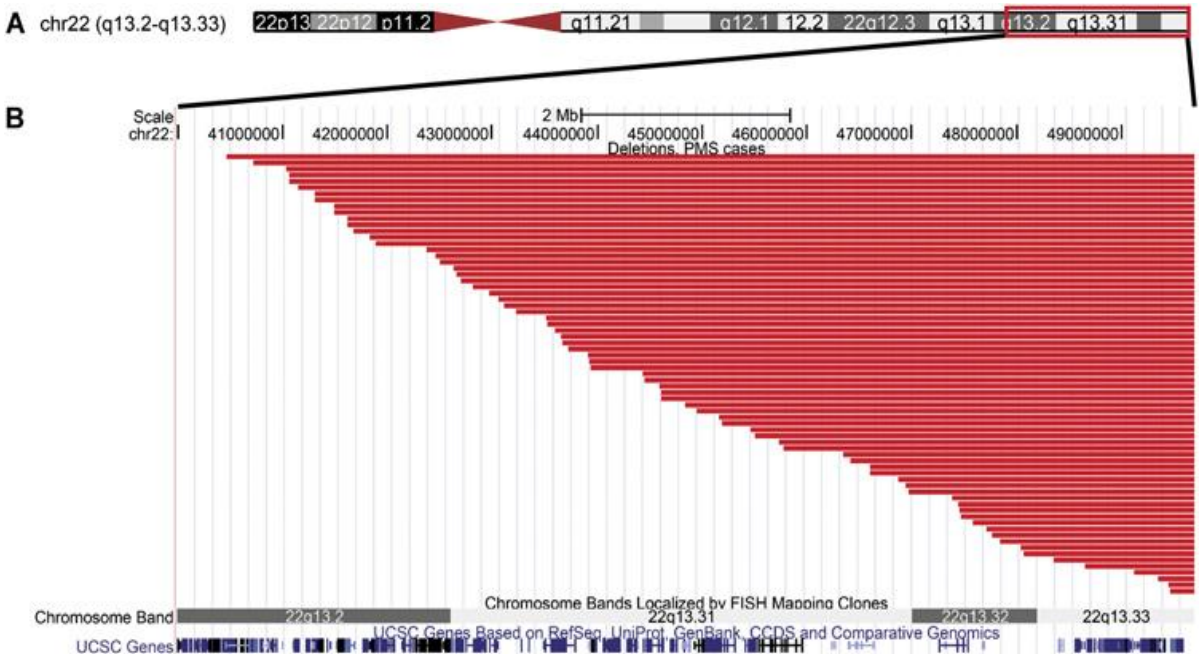
[b] Two-sided, Wilcoxon rank-sum test. Bold font indicates significant (P<0.05) finding.

[c]Among those three or more years of age.

[d]Reference group for statistical comparison.

[e]The medical history questionnaire did not ask specifically about sleep disturbances or problems. However, 14 individuals mentioned sleep difficulties in open-ended or "other problems" sections of the questionnaire or mentioned taking a medication to help with sleep. The remaining individuals did not mention sleep related problems or taking a medication for sleep.

Figure 2.1. Distribution of Deletion Sizes Among 71 Patients with Phelan-McDermid Syndrome. (A) Ideogram of chromosome 22q13.2-qter. (B) Horizontal bars represent deleted regions, sorted by deletion size. Terminal deletions ranged from 0.2 to 9.2 Mb in size and covered chromosome regions 22q13.2-22q13.3. Figures produced using the UCSC genome browser.[35]



References

[1] Phelan MC, Rogers RC, Saul RA, Stapleton GA, Sweet K, McDermid H, Shaw SR, Claytor J, Willis J, Kelly DP. 22q13 Deletion Syndrome. *Am J Med Genet* 2001;101:91-99.

[2] Luciani JJ, de Mas P, Depetris D, Mignon-Ravix C, Bottani A, Prieur M, Jonveaux P, Philippe A, Bourrouillou G, de Martinville B, Delobel B, Vallee L, Croquette M-F, Mattei G-G. Telomeric 22q13 deletions resulting from rings, simple deletions, and translocations: cytogenetic, molecular, and clinical analyses of 32 new observations. *J Med Genet* 2003;40:690-696.

[3] Havens JM, Visootsak J, Phelan MC, Graham JM,Jr. 22q13 Deletion Syndrome: an Update and Review for the Primary Pediatrician. *Clin Pediatr (Phila)* 2004;43:43-53.

[4] Cusmano-Ozog K, Manning MA, Hoyme HE. 22q13.3 Deletion Syndrome: a Recognizable Malformation Syndrome Associated with Marked Speech and Language Delay. *Am J Med Genet C Semin Med Genet* 2007;145C:393-398.

[5] Phelan MC. Deletion 22q13.3 syndrome. *Orphanet J Rare Dis* 2008;3:14-19.

[6] Wilson HL, Wong AC, Shaw SR, Tse WY, Stapleton GA, Phelan MC, Hu S, Marshall J, McDermid HE. Molecular characterisation of the 22q13 deletion syndrome supports the role of haploinsufficiency of SHANK3/PROSAP2 in the major neurological symptoms. *J Med Genet* 2003;40:575-584.

[7] Philippe A, Boddaert N, Vaivre-Douret L, Robel L, Danon-Boileau L, Malan V, de Blois M-C, Heron D, Colleaux L, Golse B, Zilbovicius M, Munnich A. Neurobehavioral profile and brain imaging study of the 22q13.3 deletion syndrome in childhood. *Pediatrics* 2008;122:e376-82.

[8] Phelan MC, Stapleton GA, Rogers RC. Deletion 22q13 Syndrome: Phelan-McDermid Syndrome. In: Cassidy SB, Allanson JE, editors. Management of Genetic Syndromes. Third Edition ed. Hoboken, New Jersey: Wiley-Blackwell; 2010. p. 285-297.

[9] Ravnan JB, Tepperberg JH, Papenhausen P, Lamb AN, Hedrick J, Eash D, Ledbetter DH, Martin CL. Subtelomere FISH analysis of 11 688 cases: an evaluation of the frequency and pattern of subtelomere rearrangements in individuals with developmental disabilities. *J Med Genet* 2006;43:478-489.

[10] Maitz S, Gentilin B, Colli AM, Rizzuti T, Brandolisio E, Vetro A, Zuffardi O, Guerneri S, Lalatta F. Expanding the phenotype of 22q13.3 deletion: report of a case detected prenatally. *Prenat Diagn* 2008;28:978-980.

[11] Bonaglia MC, Giorda R, Borgatti R, Felisari G, Gagliardi C, Selicorni A, Zuffardi O. Disruption of the ProSAP2 gene in a t(12;22)(q24.1;q13.3) is associated with the 22q13.3 deletion syndrome. *Am J Hum Genet* 2001;69:261-268.

[12] Jamsheer A, Smyk M, Wierzba J, Kolowska J, Wozniak A, Skolozdrzy J, Fischer M, Latos-Bielenska A. Subtle familial translocation t(11;22)(q24.2;q13.33) resulting in Jacobsen syndrome and distal trisomy 22q13.3: further details of genotype-phenotype maps. *J Appl Genet* 2008;49:397-405.

[13] Toruner GA, Kurvathi R, Sugalski R, Shulman L, Twersky S, Pearson PG, Tozzi R, Schwalb MN, Wallerstein R. Copy number variations in three children with sudden infant death. *Clin Genet* 2009;76:63-68.

[14] Battini R, Battaglia A, Bertini V, Cioni G, Parrini B, Rapalini E, Simi P, Tinelli F, Valetto A. Characterization of the phenotype and definition of the deletion in a new patient with ring chromosome 22. *Am J Med Genet A* 2004;130A:196-199.

[15] Jeffries AR, Curran S, Elmslie F, Sharma A, Wenger S, Hummel M, Powell J. Molecular and phenotypic characterization of ring chromosome 22. *Am J Med Genet A* 2005;137:139-147.

[16] De Mas P, Chassaing N, Chaix Y, Vincent MC, Julia S, Bourrouillou G, Calvas P, Bieth E. Molecular characterisation of a ring chromosome 22 in a patient with severe language delay: a contribution to the refinement of the subtelomeric 22q deletion syndrome. *J Med Genet* 2002;39:e17.

[17] Phelan MC, Brown EF, Rogers RC. Prenatal diagnosis of mosaicism for deletion 22q13.3. *Prenat Diagn* 2001;21:1100.

[18] Bonaglia MC, Giorda R, Beri S, Bigoni S, Sensi A, Baroncini A, Capucci A, De Agostini C, Gwilliam R, Deloukas P, Dunham I, Zuffardi O. Mosaic 22q13 deletions: evidence for concurrent mosaic segmental isodisomy and gene conversion. *Eur J Hum Genet* 2009;17:426-433.

[19] Delahaye A, Toutain A, Aboura A, Dupont C, Tabet AC, Benzacken B, Elion J, Verloes A, Pipiras E, Drunat S. Chromosome 22q13.3 deletion syndrome with a de novo interstitial 22q13.3 cryptic deletion disrupting SHANK3. *Eur J Med Genet* 2009;52:328-332.

[20] Ching TT, Maunakea AK, Jun P, Hong C, Zardo G, Pinkel D, Albertson DG, Fridlyand J, Mao J-H, Shchors K, Weiss WA, Costello JF. Epigenome analyses using BAC microarrays identify evolutionary conservation of tissue-specific methylation of SHANK3. *Nat Genet* 2005;37:645-651.

[21] Beri S, Tonna N, Menozzi G, Bonaglia MC, Sala C, Giorda R. DNA methylation regulates tissue-specific expression of Shank3. *J Neurochem* 2007;101:1380-1391.

[22] Boeckers TM, Bockmann J, Kreutz MR, Gundelfinger ED. ProSAP/Shank proteins - a family of higher order organizing molecules of the postsynaptic density with an emerging role in human neurological disease. *J Neurochem* 2002;81:903-910.

[23] Feng J, Schroer R, Yan J, Song W, Yang C, Bockholt A, Cook EH Jr, Skinner C, Schwartz CE, Sommer SS. High frequency of neurexin 1beta signal peptide structural variants in patients with autism. *Neurosci Lett* 2006;409:10-13.

[24] Alarcon M, Abrahams BS, Stone JL, Duvall JA, Perederiy JV, Bomar JM, Sebat J, Wigler M, Martin CL, Ledbetter DH, Nelson SF, Cantor RM, Geschwind DH. Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. *Am J Hum Genet* 2008;82:150-159.

[25] Arking DE, Cutler DJ, Brune CW, Teslovich TM, West K, Ikeda M, Rea A, Guy M, Lin S, Cook EH, Chakravarti A. A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism. *Am J Hum Genet* 2008;82:160-164.

[26] Bakkaloglu B, O'Roak BJ, Louvi A, Gupta AR, Abelson JF, Morgan TM, Chawarska K, Klin A, Ercan-Sencicek AG, Stillman AA, Tanriover G, Abrahams BS, Duvall JA, Robbins EM, Geschwind DH, Biederer T, Gunel M, Lifton RP, State MW. Molecular cytogenetic analysis and resequencing of contactin associated protein-like 2 in autism spectrum disorders. *Am J Hum Genet* 2008;82:165-173.

[27] Jamain S, Quach H, Betancur C, Rastam M, Colineaux C, Gillberg IC, Soderstrom H, Giros B, Leboyer M, Gillberg C, Bourgeron T, Paris Autism Research International Sibpair Study. Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat Genet* 2003;34:27-29.

[28] Chubykin AA, Liu X, Comoletti D, Tsigelny I, Taylor P, Sudhof TC. Dissection of synapse induction by neuroligins: effect of a neuroligin mutation associated with autism. *J Biol Chem* 2005;280:22365-22374.

[29] Betancur C, Sakurai T, Buxbaum JD. The emerging role of synaptic cell-adhesion pathways in the pathogenesis of autism spectrum disorders. *Trends Neurosci* 2009;32:402-412.

[30] Bozdagi O, Sakurai T, Papapetrou D, Wang X, Dickstein DL, Takahashi N, Kajiwara Y, Yang M, Katz AM, Scattoni ML, Harris MJ, Saxena R, Silverman JL, Crawley JN, Zhou Q, Hof PR, Buxbaum JD. Haploinsufficiency of the autism-associated Shank3 gene leads to deficits in synaptic function, social interaction, and social communication. *Mol Autism* 2010;1:1-15.

[31] Peca J, Feliciano C, Ting JT, Wang W, Wells MF, Venkatraman TN, Lascola CD, Fu Z, Feng G. Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature* 2011;472:437-442.

[32] Wilson HL, Crolla JA, Walker D, Artifoni L, Dallapiccola B, Takano T, Vasudevan P, Huang S, Malone V, Yobb T, Quarrell O, McDermid HE. Interstitial 22q13 deletions: genes other than SHANK3 have major effects on cognitive and language development. *Eur J Hum Genet* 2008;16:1301-1310.

[33] Giza J, Urbanski MJ, Prestori F, Bandyopadhyay B, Yam A, Friedrich V, Kelley K, D'Angelo E, Goldfarb M. Behavioral and cerebellar transmission deficits in mice lacking the autism-linked gene islet brain-2. *J Neurosci* 2010;30:14805-14816.

[34] International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* 2004;431:931-945.

[35] Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. *Genome Res* 2002;12:996-1006.

[36] Dhar SU, del Gaudio D, German JR, Peters SU, Ou Z, Bader PI, Berg JS, Blazo M, Brown CW, Graham BH, Grebe TA, Lalani S, Irons M, Sparagana S, Williams M, Phillips JA III, Beaudet AL, Stankiewicz P, Patel A, Cheung SW, Sahoo T. 22q13.3 deletion syndrome: clinical and molecular analysis using array CGH. *Am J Med Genet A* 2010;152A:573-581.

[37] Fujita Y, Mochizuki D, Mori Y, Nakamoto N, Kobayashi M, Omi K, Kodama H, Yanagaw Y, Abe T, Tsuzuku T, Yamanouchi Y, Takano T. Girl with accelerated growth, hearing loss, inner ear anomalies, delayed myelination of the brain, and del(22)(q13.1q13.2). *Am J Med Genet* 2000;92:195-199.

[38] Spiteri E, Konopka G, Coppola G, Bomar J, Oldham M, Ou J, Vernes SC, Fisher SE, Ren B, Geschwind DH. Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. *Am J Hum Genet* 2007;81:1144-1157.

[39] Lindquist SG, Kirchhoff M, Lundsteen C, Pedersen W, Erichsen G, Kristensen K, Lilquist K, Smedegaard HH, Skov L, Tommerup N, Brondum-Nielsen K. Further delineation of the 22q13 deletion syndrome. *Clin Dysmorphol* 2005;14:55-60.

[40] Manning MA, Cassidy SB, Clericuzio C, Cherry AM, Schwartz S, Hudgins L, Enns GM, Hoyme HE. Terminal 22q deletion syndrome: a newly recognized cause of speech and language disability in the autism spectrum. *Pediatrics* 2004;114:451-457.

[41] Rutter, M., LeCouteur, A., Lord, C. Autism Diagnostic Interview, Revised. Los Angeles: Western Psychological Services; 2003.

[42] Rollins JD, Sarasua SM, Phelan MC, DuPont BR, Rogers RC, Collins JS. Growth in Phelan-McDermid Syndrome. *Am J Med Genet* 2011;155:2324-6.

[43] Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JM, Eichler EE, Epstein CJ, Faucett WA, Feuk L, Friedman JM, Hamosh A, Jackson L, Kaminsky EB, Kok K, Krantz ID, Kuhn RM, Lee C, Ostell JM, Rosenberg C, Scherer SW, Spinner NB, Stavropoulos DJ, Tepperberg JH, Thorland EC, Vermeesch JR, Waggoner DJ, Watson MS, Martin CL, Ledbetter DH. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 2010;86:749-764.

CHAPTER THREE

22Q13.2Q13.32 GENOMIC REGIONS ASSOCIATED WITH SEVERITY OF SPEECH DELAY, DEVELOPMENTAL DELAY, AND PHYSICAL FEATURES IN PHELAN-MCDERMID SYNDROME

[This paper is in submission as: Sarasua SM, Dwivedi A, Boccuto L, Chen CF, Sharp JL, Rollins JD, Collins JS, Rogers RC, Phelan K, DuPont BR. 2012. 22q13.2q13.32 Genomic Regions Associated with Severity of Speech Delay, Developmental Delay, and Physical Features in Phelan-McDermid Syndrome.]

Abstract

Phelan-McDermid Syndrome (PMS), also known as 22q13.3 deletion syndrome, commonly presents with varying degrees of hypotonia, speech and other developmental delays, autistic traits, and mild dysmorphic features. Deletion breakpoints are variable with terminal deletion sizes ranging from 0.1 to 9 Mb. Prior research found specific phenotypes were associated with larger deletion sizes. This study tested the hypothesis that specific genomic regions within 22q13 are associated with specific phenotypes of interest. In a patient cohort of 71 individuals, we paired clinical information with high density oligo array CGH to identify specific 22q13 regions associated with 22 phenotypes. In this cohort everyone has a terminal deletion encompassing *SHANK3* (SH3 and multiple ankyrin repeat domains 3, located at 22q13.33), the major candidate gene for PMS neurological findings. We find that additional gene loss proximal to *SHANK3* is positively associated with severity of speech/language delay, neonatal hypotonia, neonatal feeding problems, delayed age at walking, dysplastic toenails, large/fleshy hands, macrocephaly, tall stature, short stature, facial asymmetry, dolichocephaly, full brow, bulbous nose, sacral dimple, strabismus, abnormal reflexes, sensitivity to touch, hair-pulling behaviors, and male genital anomalies; and negatively associated with autism spectrum disorders, aggressive behavior, and pinching behavior. Our use of statistical methods along with

protein interaction networks highlight the potential role of the genomic region 22q13.2q13.32 in PMS, explain some of the variability in observed phenotypes, and suggest additional genes and genomic regions for investigation as causal for speech and developmental delay. The statistical methods may be useful in genotype-phenotype analyses for other microdeletion or microduplication syndromes.

Introduction

Phelan-McDermid syndrome (PMS [MIM 606232]) is a rare condition typically caused by deletions of chromosome 22q13. Common features of the syndrome include developmental delay, absent or impaired speech, neonatal hypotonia, autistic-like behaviors, and mild dysmorphic features (Bonaglia and others 2011; Cusmano-Ozog, Manning, Hoyme 2007; Havens and others 2004; Luciani and others 2003; Phelan and others 2001; Phelan 2008; Philippe and others 2008; Wilson and others 2003). Loss of one copy of *SHANK3* (SH3 and multiple ankyrin repeat domains 3), a gene in the telomeric portion of 22q13.33, is likely responsible for many of the neurological features of the PMS phenotype (Bonaglia and others 2001; Delahaye and others 2009; Durand and others 2007; Grabrucker and others 2011; Phelan 2008; Wilson and others 2003). An additional candidate gene is *IB2* (islet-brain 2*)*, also known as *MAPK8IP2* (mitogen-activated protein kinase 8-interacting protein 2). *IB2* maps 70 kb proximal to *SHANK3,* is deleted in most PMS patients, and may play an important role in synaptic stability and neuronal transmission (Giza and others 2010). Initial work with the present cohort identified phenotypes associated with deletion size (Sarasua and others 2011).

This paper addresses the hypothesis that specific genomic regions of chromosome 22q13 are associated with PMS-related phenotypes.

Materials and Methods

Study subjects and methods have been previously described (Sarasua and others 2011) and are briefly summarized below.

Subjects

The cohort included 71 individuals, 42 females and 29 males, with deletion sizes ranging from 0.2 to 9.2 Mb and a median deletion size of 5.25 Mb. Ages spanned from 5 months to 40 years, with a mean age of 7 years 6 months. Most subjects attended one or more PMS family support conferences held between 2001 and 2008 and came from more than 30 states and 12 countries. Blood samples were collected at these meetings or were collected by personal physicians and sent to the investigators. Clinical information was obtained either from standardized physical examinations performed at the conferences by trained clinical geneticists (n=54), medical record review (3 cases), and/or parent-completed medical history questionnaires (n=61). The study was approved by the Institutional Review Board of Self Regional Healthcare (Greenwood, South Carolina), and all participants' parents or guardians provided informed consent. We excluded from analysis individuals with chromosomal anomalies other than simple terminal deletions ascertained by our arrays or previous cytogenetic exams. All participants have a terminal deletion encompassing the *SHANK3* gene and all but two are also missing one copy of *IB2*.

Genetic analysis

Genetic deletions were measured from specimens of whole blood using a custom 4x44K 60-mer oligo array designed to cover 22q12.3-terminus by Oxford Gene Technology (Oxford, UK). Array CGH genomic coordinates of breakpoints were established according to the 2006 human genome build (NCBI 36/HG 18) (International Human Genome Sequencing Consortium 2004) with terminal deletion breakpoints ranging from chromosomal position 40.1 Mb to 49.5 Mb. The terminus of chromosome 22 is located at position 49.69 Mb.

Statistical analysis

Statistical analyses were used to examine 21 phenotypes identified from prior work (Sarasua and others 2011) as having different deletion sizes (P<0.10) between those with and without the given phenotype (Tables 3.1 and 3.2). As a comparison, two phenotypes that were previously found to be unassociated with deletion size were also examined: microcephaly (P=0.9556) and seizures (considered as parent reporting the child had at least one episode of seizure and also used an anti-seizure medication, P=0.5366).

Three exploratory and complementary methods were used to identify the genomic regions most associated with phenotypes. The first two methods seek to identify the genomic breakpoint most associated with any given phenotype, and this breakpoint will be defined as the optimal cutpoint in the genotype data. The first method is called the "minimum P-value" method (Altman and others 1994; Williams and others 2006) and was implemented using SAS v. 9.2 (SAS Institute 2009). Sequentially at each breakpoint, the proportion of individuals with a given phenotype was compared between those with a given deletion size or larger to those with

a smaller deletion size. Fisher's Exact Test 2-sided *P*-value was calculated for each breakpoint comparison. The genomic region bounded by the most distal and most proximal breakpoints which had a nominal *P*-value < 0.05 was identified as a genomic region of potential association. Within this region, the breakpoint resulting in the smallest *P*-value was identified as the optimal data cutpoint. Relative risk (RR) and 95% confidence intervals were calculated at the optimal cutpoint. Bonferroni-adjusted *P*-values were also calculated to adjust for the fact that *n*-1 statistical tests were calculated for each phenotype. The distribution of age and gender were examined for each phenotype to look for significant differences.

Secondly, receiver operator characteristic (ROC) methods (Bewick, Cheek, Ball 2004) were used to examine sensitivity and specificity for all possible breakpoints using a logistic regression model in SAS (SAS Institute 2009). The Youden Index was used to identify the optimal cutpoint for a particular phenotype in the genomic data. The Youden Index is calculated as J= (Sensitivity + Specificity - 1) and has a range of 0 to 1 (Bewick, Cheek, Ball 2004). For each clinical feature, sensitivity was calculated as the proportion of all cases having a given deletion size or larger out of the total number of cases. Specificity was calculated as the proportion of noncases with a deletion less than the given breakpoint out of all noncases. The breakpoint where the maximum Youden Index was achieved was identified as the optimal cutpoint. The area under the ROC curve (AUC) was calculated to determine whether genomic breakpoint position explained the data more than chance (AUC > 0.5) (Hanley and McNeil 1982; Hanley and McNeil 1983). While an AUC value of 1.0 would indicate a perfect predictor, AUC values above 0.9 are considered to be highly accurate and AUC values from 0.7 to 0.9 can be considered moderately accurate (Greiner, Pfeiffer, Smith 2000; Swets 1988). Additionally as part of this

method, age and gender effects were evaluated using logistic models for all phenotypes, but they did not confound the association between genomic breakpoint position and phenotype.

Thirdly, the genomic region of common deletion for each phenotype was also identified (the traditional approach). These three approaches were used simultaneously to identify genomic regions and to narrow the search for potential genes most likely associated with a given phenotype. Lastly, we compared the prevalence of various conditions among the four cytogenetic bands in our region of interest (22q13.2, 22q13.31, 22q13.32, 22q13.33).

Protein interaction networks to annotate 22q13 genes

To further identify 22q13 genes of interest in the regions highlighted by the association analysis, known genes related to autism spectrum disorders, intellectual disability, hypotonia, and head size were used as seeds in a protein interaction network in order to identify interacting partners located in the 22q13 deletion region under study. As listed in Table 3.3, curated gene lists were used as seeds for autism spectrum disorders (Basu, Kollu, Banerjee-Basu 2009; Kou and others 2012; Pinto and others 2010; Sakai and others 2011) and intellectual disability (Chiurazzi and others 2008; Kou and others 2012; Lubs, Stevenson, Schwartz 2012; Pinto and others 2010). The OMIM database (omim.org) was searched for the term “hypotonia” for genes with known locus, genes with known sequence and phenotype, phenotype description with molecular basis known, and having a gene map locus. Similarly, “macrocephaly” and “microcephaly” were searched. These lists contained genes with 22q13 genes (noted in Table 3.3). The gene lists were submitted for each phenotype separately to the online gene interaction tool GeneMANIA (Warde-Farley and others 2010) to search against known protein-

protein interaction databases using the iRefIndex which includes BIND (Alfarano and others 2005; Bader, Betel, Hogue 2003), BioGRID (Stark and others 2006; Stark and others 2011), CORUM (Ruepp and others 2008), DIP (Salwinski and others 2004), HPRD (Mishra and others 2006; Peri and others 2003), IntAct (Hermjakob and others 2004; Kerrien and others 2007; Kerrien and others 2012), MINT (Chatr-aryamontri and others 2007), MPact (Guldener and others 2006), MPPI (Pagel and others 2005) and OPHID (Brown and Jurisica 2005). We used GeneMANIA to search only physical interactions, up to 100 related genes per query gene, biological process based weighting, and submitting 100 genes in a batch. Genes identified as interacting partners with the seed genes were then compared to the list of protein coding genes on 22q13 using a Venn diagram maker.

Results

Specific genomic regions were associated with each phenotype assessed (Tables 3.1-3.2, Figures 3.1-3.4, Supplemental Figure 3.1). The regions are described by genomic position based upon the 2006 human genome build 18 (International Human Genome Sequencing Consortium 2004) where base position 49.69 Mb identifies the distal end of chromosome 22. The location of minimum *P*-value and maximum Youden Index were almost always identical and the location of the maximum Youden Index was always within the range of significant *P*-values obtained from the association analysis. Using the smallest common deletion as an indication of optimal cutpoint identified the same genomic regions as the association analysis for 7 of 21 features (facial asymmetry, male genital anomalies, tall stature, macrocephaly, delayed age at walking,

speech delay, and strabismus). In these cases, the relative risk was either undefined (no cases observed below the optimal breakpoint) or 0 (no cases observed above the optimal breakpoint). For the remaining 14 features, a genomic location was found to be associated with either increased risk of the feature (RR ranging from 1.5 to 8.5) or decreased risk (RRs ranging from 0.3 to 0.4). For most phenotypes, those with the smallest deletions (22q13.33) were less severely affected than those with the largest deletions (22q13.2) (Table 3.2). For selected clinical features of interest, the genomic regions significantly associated with the feature is presented graphically in Figure 3.5.

Speech/language delay and developmental delay

The 3.4 Mb genomic region surrounding genomic position 43.9 Mb was associated with speech ability (Figures 3.1 and 3.2). While all individuals presented with speech delay, there were differences in verbal communication abilities. Of the 50 individuals over age 3 years with information about speech development, 25 had absent speech (0 words), 14 had minimal speech (spoke 1-39 words, but no known sentences or phrases), and 11 had "sentences" (spoke 40 or more words or spoke in phrases or sentences). The subjects in the minimal speech group were not included in the analysis to reduce misclassification and to better differentiate speech abilities between absent speech and verbal communication ability. Subjects with absent speech had deletion breakpoints ranging from position 40.4 to 49.3 Mb with a median deletion size of 6.7 Mb. Subjects with "sentences" had deletion breakpoints ranging from position 44.4 to 49.4 Mb, with a median deletion size of 3.3 Mb. The distribution of breakpoints is illustrated in Figures 3.1 and 3.2A. As shown in Table 3.1 and Figure 3.2B, deletion breakpoints at position 41.8 to 45.2 Mb are significantly associated with speech ability, with the smallest *P*-value

(*P*=0.0007) occuring at base position 43.9 Mb. None of the 15 individuals with deletion breakpoints at 43.9 Mb or more proximal had "sentences," whereas 11 of the 21 (52%) subjects with deletions at 44.4 Mb or more distal had "sentences" (RR=0; Bonferroni-adjusted *P*<0.05). The cutpoint of 44.4 Mb was identified using the Youden Index (Figure 3.2C). The area under the curve (AUC) was 0.79 ("moderately accurate"; Figure 3.2D). As shown in Table 3.2, none of the 12 individuals with deletions in 22q13.2 were able to speak in sentences compared to 47% with deletions of 22q13.31 or 60% with deletions occurring at 22q13.33. Age and gender were not significant predictors or confounders of speech.

Speech ability was also examined in relation to autism spectrum disorders (ASDs) and degree of developmental delay. The proportion of subjects forming sentences was similar for those who were reported to have an ASD (3 out of 11) compared to those who were not (8 out of 24, Fisher's Exact Test *P*=1.0). Speech ability was associated with parent report of degree of developmental delay (rank score of 1=mild to 7=profound). Those with sentences had a median developmental delay score of 3 ("moderate") while those without sentences had a median developmental delay score of 6 ("severe to profound", Wilcoxon Rank-Sum Exact Test *P*=0.026). Those with deletion breakpoints at 43.9 Mb or more proximal had a median developmental delay score of 6 ("severe to profound") while those with a deletion breakpoint more distal had a median developmental delay score of 5 ("severe", Wilcoxon Rank-Sum Exact Test *P*=0.0495). Deletion size and developmental delay score were significantly correlated (Spearman rank correlation coefficient ρ=0.52, *P*=0.0045).

Neonatal features

Neonatal hypotonia and neonatal feeding problems, as reported by parents, were significantly associated with 22q13.31 to 22q13.32 deletion regions (Tables 3.1 and 3.2, Figure 3.3). In the case of neonatal hypotonia, the Youden Index is maximum at chromosome 22 position 45.7 Mb whereas position 47.4 Mb is the location of the minimum *P*-value ($P<0.0001$). Neonatal feeding problems were identified to have similar associated genomic regions with the minimum *P*-value ($P=0.0009$) and maximum Youden Index occuring at base position 45.5 Mb.

Abnormal growth

The presence of short stature (<5th percentile) and tall stature (>95th percentile) were moderately associated with distinct deletion regions (Tables 3.1 and 3.2, Figure 3.3). Short stature was associated with deletions of the genomic region 41.0 to 42.5 Mb, with optimal cutpoints at position 41.6 Mb. Tall stature was associated with the genomic region bounded by position 43.9 to 44.6 Mb, with position 43.9 Mb having the optimal cutpoint. Macrocephaly was associated with the genomic position 42.5 to 47.4 Mb, with optimal cutpoint at position 44.6 Mb and overlapping the genomic region associated with tall stature (Tables 3.1 and 3.2, Figure 3.3). Having large or fleshy hands identified the same peak genomic region as macrocephaly (Tables 3.1 and 3.2, Figure 3.3). No genomic region was identified as being associated with microcephaly and the AUC for the ROC curve was 0.51 (similar to random chance, Tables 3.1 and 3.2, Figure 3.4). This lack of association was expected given that microcephaly was unassociated in the preliminary analysis ($P=0.9556$) and was included in this analysis only for comparison purposes.

Autism spectrum disorders (ASDs) and aggressive behavior

The genomic region from position 41.6 to 46.6 Mb was found to be associated with reduced prevalence of parent-reported diagnosis of Autism Spectrum Disorders (ASDs, Tables 3.1 and 3.2, Figure 3.4). The graphs depicting association statistics and Youden Index are broad and inconsistent. Aggressive behavior and pinching behavior (themselves or others) were also associated with smaller deletions (Table 3.1 and Figure 3.4), although the statistical support for these associations is less compared to ASDs as neither had an AUC > 0.7 and neither achieved Bonferroni corrected levels of statistical significance (Table 3.1).

Other features

Other features, including male genital anomalies, atypical reflexes, dolichocephaly, sacral dimple, bulbous nose, strabismus, and full brow, were associated with specific genomic regions from 41.5 to 48.7 Mb (Tables 3.1 and 3.2). No genomic region was associated with seizures and the AUC for the ROC curve was 0.56 (close to random chance, Figure 3.4). This lack of association between genomic region and seizures was expected given that seizures were unassociated in the preliminary analysis (P=0.5366) and assessment of the seizure phenotype was included in this analysis only for comparison purposes.

22q13 genes identified as interacting partners with known developmental disability genes

The use of GeneMANIA to search existing physical protein interactions identified several genes across the 22q13 deletion region (Table 3.3) not otherwise immediately known as being candidate genes. In particular, WNT7B and PARVB, both located in 22q13.31, were identified as interacting partners for ASD, ID, and hypotonia and WNT7B for macrocephaly. The location of

22q13 genes used either as seeds or found to be interacting partners of ASD, ID, hypotonia, and macrocephaly associated genes are provided in Figure 3.5. Also in Figure 3.5 are the genomic regions found to be most strongly statistically associated with features related to the central nervous system. The location of known genes and micro RNAs as well as predictions of haploinsufficiency (Huang and others 2010) are also provided.

Discussion

This study is the first to identify specific chromosome 22q13.2q13.32 genomic regions, in addition to the terminal 22q13.33 genomic region encompassing *SHANK3*, associated with key phenotypes in Phelan-McDermid syndrome. Strengths of this study design include a large sample size for a rare condition, high resolution genotyping and widely dispersed breakpoints allowing for resolution between individuals exhibiting different phenotypes. The location of the maximum Youden Index was always within the genomic region identified by the association analysis and was usually identical to the location with the minimum *P*-value (Table 3.1). These methods identified genomic regions of interest for phenotypes which had shown crude deletion size differences in the first stage of analysis and did not spuriously identify genomic regions for phenotypes which were not associated with deletion size. The use of both ROC characteristics and statistical association allowed us to examine genomic regions associated with phenotypes that may have many causes. Individual genetic background will influence the appearance of many physical phenotypes such as growth and hypotonia. Typically, genotype-phenotype studies of chromosomal deletions/duplications involve examining rare phenotypes and identifying the genomic region of common deletion/duplication (Feenstra and others 2007;

Korbel and others 2009). However, those methods will not work for common or multifactorial phenotypes and genes with incomplete penetrance. The combined use of association and ROC analysis allows for the identification of contributory genomic regions in multifactorial phenotypes and does not fail to identify critical genomic regions for phenotypes with strong genomic clustering.

In previous works on PMS, the *SHANK3* gene was successfully identified as constantly deleted in reported cases. Haploinsufficiency of this gene was frequently associated with a core set of phenotypes including speech delay, developmental delay, hypotonia, and minor dysmorphic features (Bonaglia and others 2001; Bonaglia and others 2006; Bonaglia and others 2011; Delahaye and others 2009; Jeffries and others 2005; Luciani and others 2003; Phelan and others 2001; Wilson and others 2003). While there was a general impression that patients with larger deletions were more seriously affected, genotype-phenotype studies were hampered by small sample size, low resolution genotyping, or reliance on statistical measures of linear association (correlation coefficients, linear regression) (Dhar and others 2010; Jeffries and others 2005; Jeffries and others 2005; Koolen and others 2005; Phelan 2008; Philippe and others 2008; Wilson and others 2003). We hypothesized that for each phenotype, there was at least one gene that increased the risk of being affected, when disrupted along with *SHANK3*. As part of this hypothesis, we predicted that the effect of an additional causal gene would be discrete rather than continuous.

Because our patient cohort includes only those with terminal deletions and all patients are missing one copy of *SHANK3*, we cannot distinguish whether the more proximal genomic regions we identified have independent or additive effects along with *SHANK3*. The literature

reports three cases of interstitial deletions which have intact *SHANK3* and phenotypes similar to those in PMS (Fujita and others 2000; Wilson and others 2008), suggesting an independent role for genes in these genomic regions. In particular, two individuals had speech delay (two words each and no sentences), macrocephaly, tall stature, hypotonia, delayed walking, and developmental delay yet had two copies of *SHANK3* (Wilson and others 2008). Their deletion breakpoints are reported to be between 40.42 and 44.00 Mb for one patient and between 41.22 and 45.37 Mb for the second. Our results are consistent with these reports of phenotypes being caused by deletions of 22q13.31-13.32, particularly for speech, tall stature, and macrocephaly. As shown in Figure 3.5, *CYB5R3*, *PARVB*, and has-mir-1249 all map in the region deleted in these interstitial cases.

Speech and language delay

The 3.4 Mb genomic region identified as being associated with lack of speech contains an estimated 45 protein coding genes. Within this region, the most strongly associated segment of chromosome 22 position 43.0 to 44.5 Mb contains the 13 protein coding genes *KIAA1644, LDOC1L, LOC388910, LOC553158, PHF21B, DJ031123, NUP50, C22orf9, UPK3A, FAM118A, SMC1B, RIBC2,* and *FBLN1* along with a micro RNA *miR-1249* and other miRNA and several non-coding *RNAs*. The genes *PARVB* and *WNT7B*, found to be interacting partners of genes known to be associated with ASD, ID, and hypotonia, are also in the region associated with speech. The findings of severe speech impairment among two published cases with interstitial deletions overlapping our genomic region of interest and intact *SHANK3* (Wilson and others 2008) support the presence of genes affecting speech in this region. None of the genes in this region were found to be transcriptional targets of FOXP2, a transcription factor known to be associated with

speech (Spiteri and others 2007; Vernes and others 2007). Beyond *FOXP2*, little is known about genes related to speech (Fisher and Scharff 2009; Kang and Drayna 2011), although recent studies have added *CNTNAP2, CMIP, ATP2C2, RIT2*, and *SYT4* as potential genes of interest (Bouquillon and others 2011; Newbury, Fisher, Monaco 2010). The analysis of speech ability is made difficult given the variable language abilities of those with known *SHANK3* deletions or mutations. For instance, a study of autistic patients with *SHANK3* mutations identified individuals with varying language abilities from no speech problems to those with absent speech (Waga and others 2011). The additional loss of or mutation in causal genes may be the "second hit" in addition to *SHANK3* alterations, as proposed for a different microdeletion syndrome (16p12.1) which also has speech and developmental delay (Girirajan and others 2010). We found correlation between speech ability and degree of developmental delay. The association between 22q13.31 and speech ability may reflect an association with intellectual disability. Future research into speech and language abilities in PMS would benefit from having detailed evaluations to better characterize the types of language delay specific to this syndrome.

Abnormal growth

We recently reported that both tall (>95th percentile) and short stature (< 5th percentile) as well as macrocephaly (>97th percentile) are more common in PMS than expected (Rollins and others 2011). The present analysis provides evidence of distinct deletion regions associated with these growth parameters which support earlier findings by others (Dhar and others 2010; Wilson and others 2003). The genomic region we identified as being associated with short stature coincides with the genomic region identified as a human stature quantitative trail locus

(QTL) (Sammalisto and others 2005) as well as a body mass QTL in rats (Rat Genome Database, (Rapp 2000) accessed on the UCSC genome browser).

Autism spectrum disorders

In our analysis we noted that ASDs were associated with smaller deletions. *SHANK3* mutations have been found to be associated with ASDs (Durand and others 2007; Gauthier and others 2009; Moessner and others 2007). In the situation of PMS, the effect of *SHANK3* may be attenuated as the deletion size increases and additional genes are codeleted. It may also be more difficult to evaluate ASDs in patients with severe developmental and speech delay, both of which are associated with larger deletions. ASDs are a heterogeneous group of neurodevelopmental conditions and recent investigations have suggested different gene expression patterns associated with different domains of impairment (Hu and Steinberg 2009). Future research is needed to better delineate the autism phenotype in PMS patients and those with *SHANK3* deletions or mutations to better identify particular domains affected.

Review of potential genes of interest

A large number of potentially interesting candidate genes were identified in this analysis, most with expression observed in relevant tissue types as described in the EST profile database (http://www.ncbi.nlm.nih.gov/nucest) and GNF Expression Atlas (Su and others 2002; Su and others 2004). Few of these genes have been previously implicated as causative for human phenotypes. 22q13.31 genes predicted to be haploinsufficient according to prediction models include *NUP50* (3% predicted probability of being haplosufficient), *FBLN1* (9%), *SMC1B* (13%), *CELSR1* (19%), *PPARA* (32%), *SCUBE1* (42%), *PHF21B* (48%), and *CERK* (49%) (Huang and

others 2010). Predictions were not available for all protein coding genes in our area, including *c22orf9, KIAA1644,* and *WNT7B.*

The genomic region containing *NUP50* was associated with neonatal hypotonia, atypical reflexes, tall stature, sensitivity to touch, macrocephaly, and lack of speech. NUP50, nucleoporin 50kD, expressed in brain and muscle, is part of the nuclear pore complex regulating traffic between the nucleus and cytoplasm and may be involved in regulating transcription (Akhtar and Gasser 2007; Kalverda and others 2010; Ogawa and others 2010). CERK, ceramide kinase, phosphorylates ceramide, an important membrane associated lipid. *CERK* is highly expressed in brain tissue and may be responsible for neuronal function and emotional behavior. Both of these genes have family members (*NUP35,* nucleoporin 35, and *CERKL*, ceramide kinase-like) which were among the genes deleted in the critical genomic region of 2q31.2q32.32 deletion syndrome, a condition presenting with intellectual disability and speech impairment as prominent features (Cocchella and others 2010).

Two poorly annotated genes, *C22orf9* (also known as *LOC23313* or *KIAA0930*) and *KIAA1644* are in the genomic region associated with speech delay, tall stature, macrocephaly, dolichocephaly, and facial asymmetry and are highly expressed in the brain. C22orf9 has been found to interact with a class of genes that are expressed in brain, the 14-3-3 family of proteins involved in signal transduction (Ewing and others 2007; Jin and others 2004). The 14-3-3ε protein is involved in Miller-Dieker syndrome and plays a direct role in brain development and neuronal migration (Toyo-oka and others 2003).

The genomic region containing *PHF21B*, PHD finger protein, was associated with neonatal hypotonia, sensitivity to touch, hair-pulling behavior, dolichocephaly, macrocephaly,

and speech delay among others. Zhang et al. determined that *PHF21B* was a nuclear calcium regulated gene and expressed in hippocampal neurons (Zhang and others 2009). Yang et al. found *PHF21B*, along with *ATXN10*, to be diurnally regulated in mouse prefronatal cortex and may be related to sleep cycles and mood (Yang and others 2007). The genomic region containing *FBLN1* was associated with neonatal hypotonia, speech delay, macrocephaly, delayed age at walking, hair-pulling behavior, atypical reflexes, dysplastic toenails, and male genital anomalies. FBLN1 (Fibulin-1), is an extracellular matrix protein that has been found to be important in morphogenesis of neural crest cells in mice (Cooley and others 2008). *Fbln1* null mice had cardiac wall thinning, ventricular septal defects, under-development of the thymus, thyroid, and skull bones, among other anomalies (Cooley and others 2008). FBLN1 has been shown to interact with the autism associated proteins FXR1, HOXA1, and NUF1P2 (Sakai and others 2011). Knockout mice for *CELSR1*, cadherin EGF LAG seven-pass G-type receptor 1, demonstrated abnormal neuronal migration and neural development (Formstone and others 2010; Qu and others 2010). *CELSR1* was also identified as an interacting partner of a hypotonia associated protein.

Two genes that were identified using protein interaction networks were *PARVB* (Beta parvin) and *WNT7B* (wingless-type MMTV integration site family). Both proteins were found to be physical interacting partners with proteins associated with ASD, ID, and hypotonia and WNT7B was also found to interact with macrocephaly associated proteins (Table 3.3). *PARVB* is deleted in the regions associated with lack of ASD, lack of speech, touch sensitivity, hair pulling behavior, abnormal reflexes, and macrocephaly. It is also deleted in both cases of interstitial deletions (Wilson and others 2008). *WNT7B* is deleted in one case of interstitial deletion and is

located in the regions associated with lack of ASD, lack of speech, walking late, hair pulling behavior, neonatal hypotonia, and macrocephaly. The Database of Genomic Variants reports CNVs affecting coding regions of *WNT7B*, but not for *PARVB.* Interestingly, PARVB and SHANK3 are both scaffolding proteins important in post synaptic structures (Govek, Newey, Van Aelst 2005). PARVB binds ARHGEF6 (a Rho guanine nucleotide exchange factor 6; a guanine nucleotide exchange factor for Rho GTPases) (Rosenberger and others 2003). Mutations in the PARVB binding site on ARHGEF6 are associated with X-linked intellectual disability (Kutsche and others 2000). ARHGEF6 heterodimerizes with ARHGEF7 (Rosenberger and others 2003), a binding partner of SHANK3, a candidate gene for PMS as well as autism spectrum disorders (Bonaglia and others 2001; Delahaye and others 2009; Durand and others 2007; Grabrucker and others 2011; Phelan 2008; Wilson and others 2003). Figure 3.6 illustrates the physical interactions between PARVB and SHANK3 as obtained from GeneMANIA. Thus, PARVB may have overlapping and interacting roles with SHANK3 (Govek, Newey, Van Aelst 2005).

Finally, we note the presence of several micro RNAs in 22q13.31 including hsa-mir-1249 which is located within an intronic region of *C22orf9.* This micro RNA is predicted to target the 22q13 genes *SHANK3*, *PHF21B,* and *SERHL2* as well as a number of other brain and development associated genes (TargetScan(Friedman and others 2009)). Loss of hsa-mir-1249 may lead to misregulation of the preserved copy of *SHANK3*, affecting the severity of neurological phenotypes associated with *SHANK3* haploinsufficiency, such as hypotonia, speech, and developmental delay. The fact that hsa-mir-1249 maps within the regions of interstitial deletions associated with PMS (Fujita and others 2000; Wilson and others 2008) probably represents

further evidence of the potential role played by these micro RNAs in neurodevelopmental disorders.

The remaining genes, micro RNAs, and non-coding RNA in genomic regions associated with various phenotypes, particularly in the chromosome 22q13.31 region, deserve further examination in future studies.

Limitations

There are several limitations in the present analysis. Reporting of medical history phenotypes relied upon parent recall and may be subject to recall or information bias (Rothman and Greenland 1998). Association statistics are heavily influenced by sample size and this study had small statistical power to examine associations with genomic regions near the telomere (the smallest deletions) and most proximal (those with deletions > 8 Mb in size). Bonferroni-adjustment of *P*-values was excessively conservative because each statistical comparison along the chromosome is not independent, given that the data are terminal deletions and a breakpoint at one position implies that all genes distal to it are also deleted. However, the Bonferroni-adjusted method helped to illuminate the likely critical genomic regions to examine. No statistical adjustment was made for the fact that multiple phenotypes were examined in the same population. While age and gender were not found to confound the genotype-phenotype associations in our analysis, we cannot rule out the possibility that age or gender distribution differences between affected and unaffected individuals could affect the genomic regions identified with various phenotypes.

Conclusion

This study demonstrates how different genomic regions may play a critical role in the phenotypic heterogeneity that characterizes PMS. All 71 subjects used in this study share haploinsufficiency of *SHANK3* and present with a wide spectrum of phenotypes and terminal deletion sizes. We identified 22q13.2q13.32 genomic regions associated with speech/language delay, developmental delay, neonatal hypotonia, abnormal growth, behavioral features, and physical features. These results indicate that, although the terminal 22q13.33 region encompassing *SHANK3* is critical for the core of the PMS phenotype, additional, more proximal genomic regions are important to determine the severity of some symptoms and the variable occurrence of many secondary features of the syndrome. Therefore, we believe that future studies on the role of *SHANK3*, *IB2*, and other telomeric genes in PMS should also include work to determine the independent and additive effects of loss of 22q13.2q13.32 genes, micro-RNAs, or regulatory elements.

ACKNOWLEDGMENTS

We thank the patients and family members who participated in this study and made this work possible. We thank the Phelan-McDermid Syndrome Foundation which organized the biannual family conferences where much of the study data collection took place. We thank Dr. Amy Lawton-Rauh and Dr. Charles E. Schwartz for helpful comments on the manuscript. We thank Gail Stapleton and Cindy Skinner who coordinated study operations. We dedicate this paper to the late Dr. Julianne Collins.

FUNDING

This work was supported by the Phelan-McDermid Syndrome Foundation [to SMS]; the Genetics Endowment of South Carolina; and the South Carolina Department of Disabilities and Special Needs.

Abbreviations

Mb million base pairs

PMS Phelan-McDermid Syndrome

ASD Autism spectrum disorder

ID Intellectual disability

Table 3.1. PMS Phenotypes and Chromosome 22q13 Genomic Positions Identified by Minimun *P*-Value Method, MaximumYouden Index and ROC Analysis, and Smallest Common Deletion. Phenotypes are Listed in Order of Bonferroni Adjusted Minimum *P*-value.

Phenotypes	Associated Genomic Region (Location of Min. P-value) [a]	Minimum P-value (Fisher's Exact)	Bonferroni Adjusted Minimum P-value	Relative Risk (95% CI)[b]	Max Youden Index[c]	Genomic Location of maximum Youden Index (Mb)[a]	Sensi-tivity	Speci-ficity	Area Under the ROC[d]	Smallest Common Deletion Position (Mb)	n
Clinical results from physical exams											
Large/fleshy hands	42.7-48.7 (44.6)	0.0009	0.0454	2.5 (1.4-4.7)	0.47	44.6	0.72	0.75	0.75	48	53
Macrocephaly (>97th)	42.7-47.5 (44.6)	0.0010	0.0481	Undefined	0.55	44.6	1	0.56	0.76	44.6	47
Facial asymmetry	43.1-44.6 (43.2)	0.0020	0.1037	Undefined	0.76	43.2	1	0.76	0.81	43.2	54
Full brow	45.2-49.1 (48.4)	0.0033	0.1690	Undefined	0.33	45.8	0.74	0.59	0.68	48.1	53
Tall stature (>95th)	43.9-44.6 (43.9)	0.0094	0.3667	Undefined	0.66	43.9	1	0.66	0.77	43.9	40
Abnormal reflexes	42.7-44.6 (43.7)	0.0101	0.4259	3.1 (1.3-7.3)	0.43	43.7	0.69	0.74	0.71	47.8	43
Dysplastic toenails	43.2-47.0 (43.7)	0.0084	0.4340	1.5 (1.1-2.0)	0.42	43.7	0.5	0.92	0.71	49.5	53
Short stature (<5th)	41.1-42.5 (41.6)	0.0178	0.6956	8.5 (1.8-40.6)	0.51	41.6	0.6	0.91	0.7	47.5	40
Dolichocephaly	43.5-47.8 (43.9)	0.0137	0.7283	3.2 (1.3-7.9)	0.4	43.9	0.69	0.71	0.68	47.5	54
Bulbous nose	44.6-47.6 (47.5)	0.0154	0.8162	2.6 (0.96-6.8)	0.33	45.0	0.66	0.66	0.65	49.1	54
Sacral dimple	45.0-47.0 (46.6)	0.0273	1.0000	3.8 (0.99-14.5)	0.32	46.6	0.89	0.42	0.66	48.4	52
Strabismus	47.5	0.0466	1.0000	Undefined	0.28	47.5	1	0.28	0.67	47.5	52

Microcephaly (<3rd)	None	--		--	--	--	--	--	0.51	48.4	42
Features derived from parent-provided medical history questionnaires											
Neonatal hypotonia	43.1-49.5 (47.5)	6.50E-06	0.0004	4.6 (1.3-16.0)	0.67	45.8	0.84	0.83	0.84	48.4	61
Speech (sentences v. absent speech)[e]	41.8-45.2 (43.9)	0.0007	0.0230	0	0.60	43.9	1	0.6	0.79	44.4[f]	36
Walk > 15 months[e]	44.4-49.1 (48.4)	0.0008	0.0347	Undefined	0.57	45.2	0.68	0.89	0.79	48.4	46
Neonatal feeding problems	44.7-47.5 (45.5)	0.0009	0.0549	1.8 (1.2-2.8)	0.50	45.5	0.79	0.71	0.68	49.4	61
Hair pulling behavior (age < 12 years)	42.4-47.8 (43.7)	0.0022	0.0560	3.0 (1.6-5.8)	0.57	43.7	0.57	1	0.82	47.6	26
Genital anomalies (male)	41.1-45.5 (43.5)	0.0031	0.0589	Undefined	0.77	43.5	1	0.77	0.9	43.5	20
Sensitivity to touch	41.9-43.9 (42.5)	0.0054	0.3173	1.8 (1.3-2.4)	0.33	43.9	0.61	0.73	0.65	49.5	60
Autism spectrum disorders[e]	41.7-46.6 (45.2)	0.0093	0.4834	0.3 (0.1-0.7)	0.43	45.2	0.72	0.72	0.73	49.5	53
Aggressive behavior	43.7-45.2 (44.8)	0.0108	0.6378	0.4 (0.2-0.8)	0.38	43.9 44.8	0.75 0.65	0.63 0.73	0.67 0.67	49.1	60
Pinching	43.9-45.8 (44.8)	0.0140	0.8418	0.4 (0.2-0.8)	0.35	44.8	0.61	0.74	0.65	49.4	60
Seizures requiring medication	None	--		--	--	--	--	--	0.56	49.5	55

[a] Genomic position according to the 2006 Human Genome Build 18. The terminus is position 49.69 Mb.

[b]Confidence Interval
[c] Youden Index = (Sensitivity + Specificity - 1)
[d] Receiver Operator Characteristic Curve. The curve plots sensitivity v. 1-specificity to obtain the ROC curve.
[e]Among those over 3 years of age
[f]The largest deletion breakpoint among those with sentences was at base position 44.4 Mb. The next observed deletion breakpoint occurs at position 43.9 Mb (an individual with absent speech). In the situation of evaluating speech ability, we calculate sensitivity as identifying everyone with speech at a given breakpoint or more telomeric rather than more centromeric because speech is observed to be associated with smaller deletions rather than larger deletions and the maximum sensitivity (1.0) and maximum Youden index occur at position 44.4 Mb. The minimum P-value was obtained when comparing those with deletions at 43.9 Mb or more centromeric to those with deletions at 44.4 Mb or more telomeric. Thus, because there are no cases with breakpoints between 43.9 and 44.4 Mb, the minimum P-value at 43.9 Mb and the maximum Youden index at 44.4 Mb are equivalent cutpoints.

Table 3.2. Prevalence of PMS Phenotypes by 22q13 Deletion Band. Phenotypes presented in descending order of prevalence in largest deletion group.

Phenotype	Cytogenetic Location of Deletion				
	22q13.2	22q13.31	22q13.32	22q13.33	*P*-value[a]
Physical exam features					
Gender (% male)	7/9 (78%)	15/30 (50%)	2/9 (22%)	2/5 (40%)	0.0321
Age in years, median (range)	8.4 (3.1-17.2)	5.2 (0.9-40)	4.2 (1.5-8.8)	6.9 (4.4-9.2)	0.1430[b]
Dysplastic toenails	8/9 (89%)	24/30 (80%)	5/9 (56%)	3/5 (60%)	0.0770
Large/fleshy hands	7/9 (78%)	19/30 (63%)	3/9 (33%)	0/5 (0%)	0.0028
Atypical reflexes	5/7 (71%)	8/25 (32%)	3/8 (38%)	0/3 (0%)	0.0843
Full brow	6/9 (67%)	19/29 (66%)	6/9 (66%)	0/6 (0%)	0.0663
Bulbous nose	6/9 (67%)	21/30 (70%)	4/9 (44%)	2/4 (50%)	0.2594
Sacral dimple	5/9 (56%)	12/29 (41%)	1/9 (11%)	1/5 (20%)	0.0402
Macrocephaly (>97th %)	5/10 (50%)	6/24 (25%)	0/8 (0%)	0/5 (0%)	0.0060
Strabismus	4/8 (50%)	7/30 (23%)	2/8 (25%)	0/6 (0%)	0.0688
Dolichocephaly	4/9 (44%)	9/30 (30%)	3/9 (33%)	0/6 (0%)	0.1669
Short stature (<5th %)	3/8 (38%)	1/20 (5%)	1/5 (17%)	0/5 (0%)	0.0904
Tall stature (>95th %)	2/7 (29%)	3/23 (13%)	0/5 (0%)	0/5 (0%)	0.1301
Microcephaly (<3rd %)	1/6 (17%)	3/21 (14%)	1/9 (11%)	1/6 (17%)	0.9870
Facial asymmetry	1/9 (11%)	4/30 (13%)	0/9 (0%)	0/6 (0%)	0.3079
Parent-Provided Medical History Questionnaire Phenotypes					
Gender (%) male	7/15 (47%)	14/35 (40%)	1/6 (17%)	2/5 (40%)	0.4053
Age in years, median (range)	5.0 (2.0-17.2)	6.5 (0.9-40)	3.9 (1.5-4.8)	6.9 (2.5-9.3)	0.1679[b]
Hair-pulling behavior (age < 12 years)	6/6 (100%)	6/13 (46%)	2/4 (50%)	0/3 (0%)	0.0091
Neonatal hypotonia	14/15 (93%)	32/35 (91%)	2/6 (33%)	1/5 (20%)	<0.0001
Walk later than 15 months[c]	11/12 (92%)	20/23 (87%)	5/6 (83%)	1/5 (20%)	0.0161
Sensitivity to touch	14/16 (88%)	18/35 (51%)	4/5 (80%)	2/4 (50%)	0.0993
Neonatal feeding problems	13/16 (81%)	29/35/ (83%)	2/6 (33%)	3/4 (75%)	0.1592
Genital anomalies (male)	5/8 (63%)	2/10 (20%)	0/1 (0%)	0/1 (0%)	0.0349
Pinching behavior	4/16 (25%)	13/35 (37%)	4/6 (67%)	2/4 (50%)	0.0945
Aggressive behavior	3/15 (20%)	13/35 (37%)	2/6 (33%)	2/4 (50%)	0.2358
Seizures requiring medication	3/16 (19%)	5/28 (18%)	0/6 (0%)	2/5 (40%)	0.8739
Autism spectrum disorders[c]	1/15 (7%)	9/25 (33%)	1/6 (17%)	3/5 (60%)	0.0468
Speech (sentences)[c]	0/12 (0%)	8/17 (47%)	0/2 (0%)	3/5 (60%)	0.0101

[a]Mantel Haenzel Chi-Square using rank scores

[b]Kruskal-Wallis test

[c]Among those 3 years of age or older

Table 3.3. Source of Seed Genes and Additional Genes Identified Through Protein Interaction Networks to Annotate 22q13 Genes.

Phenotype	Gene seed list source and number of genes	PMS genes in seed list and location	Additional PMS genes found in interactions and location
Autism spectrum disorders	(Basu, Kollu, Banerjee-Basu 2009) 328 (Pinto and others 2010) 139 (Kou and others 2012) 114 (Sakai and others 2011) 35 Total of 434 genes	SHANK3 (22q13.33)	FAM109B, POLR3H, BIK, TSPO 22q13.2 PARVB, WNT7B 22q13.31
Intellectual Disability	(Chiurazzi and others 2008) 89 (Lubs, Stevenson, Schwartz 2012) 29 (Kou and others 2012) 223 (Pinto and others 2010) 110 Total of 302 genes	ALG12, CHKB (22q13.33) CYB5R3 (22q13.2)	FAM109B, POLR3H (22q13.2) PARVB, WNT7B (22q13.31) MAPK12 (22q13.33)
Hypotonia	OMIM searched "hypotonia" in * gene with known locus, + gene with known sequence and phenotype, # phenotype description with molecular basis known, gene map locus Total 557 genes	ACO2 (22q13.2) TRMU (22q13.31) ALG12, SCO2, SHANK3 (22q13.33)	FAM109B (22q13.2) CELSR1, PARVB, WNT7B (22q13.31)
Macrocephaly	OMIM searched "macrocephaly" in * gene with known locus, + gene with known sequence and phenotype, # phenotype description with molecular basis known, gene map locus Total 181 genes	MLC1, SHANK3 (22q13.33)	POLR3H (22q13.2) WNT7B (22q13.31) CHKB, ARSA (22q13.33)
Microcephaly	OMIM +, *, # 423 genes	CYB5R3 (22q13.2) ALG12 (22q13.33)	POLR3H (22q13.2)

Figure 3.1. Comparison of 22q13 deletion regions for those with absent speech compared to those with sentences. Snapshot from the UCSC genome browser showing chromosome band positions and locations of 22q13 genes and micro RNA using the March 2006 (NCBI36 /hg18) assembly.

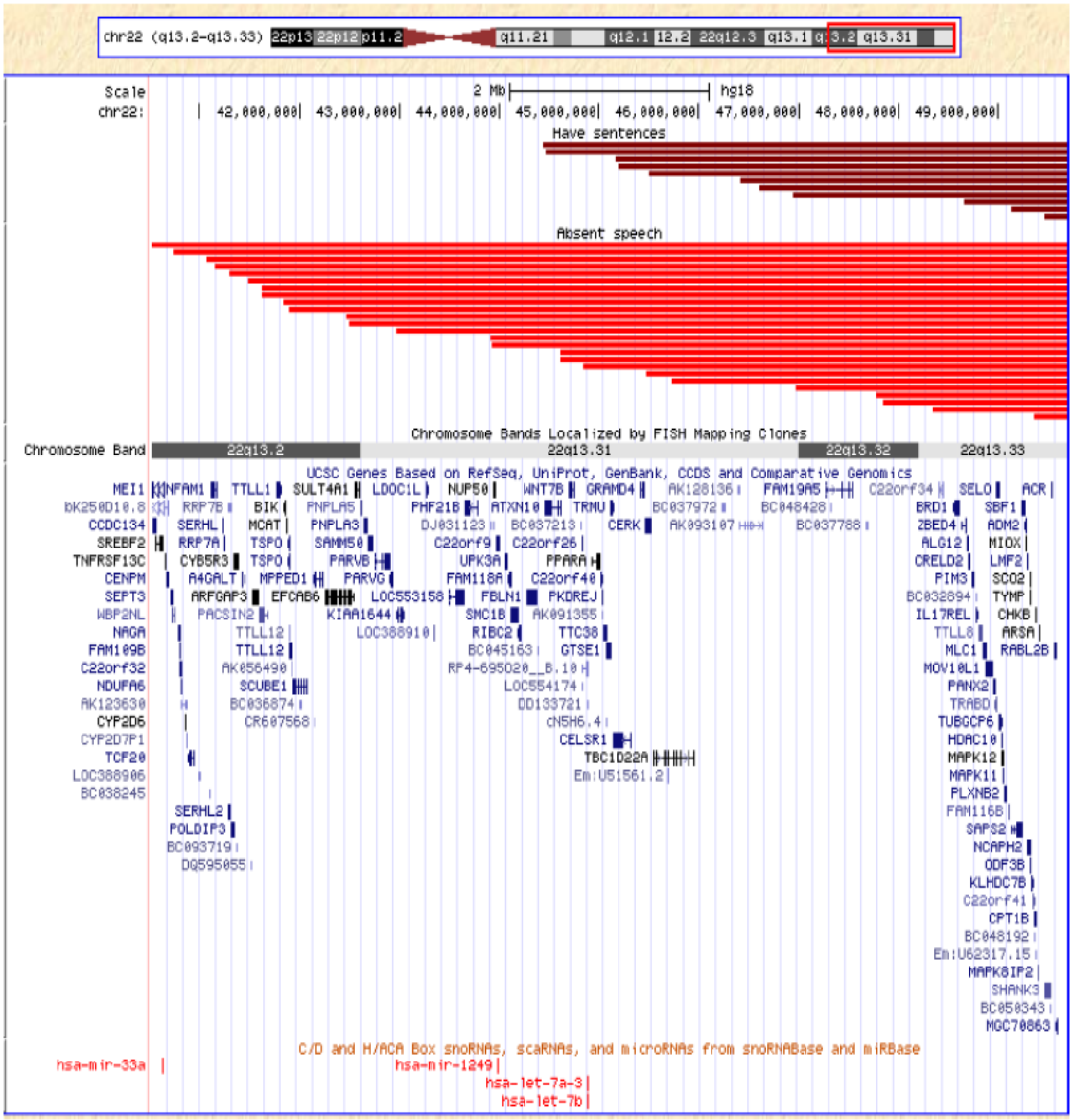


Figure 3.2. Association and ROC analysis for speech delay**.** Panel A: Distribution of chromosome 22q13 breakpoints by genomic position for those with sentences and those with absent speech. Panel B: minus log10 (*P*-value) of association between genomic breakpoint position and speech. The blue line represents a *P*-value < 0.05 ; the red line is significant after Bonferroni correction (*P* < 0.0014). Panel C: Youden Index by breakpoint position. Panel D: ROC curve.

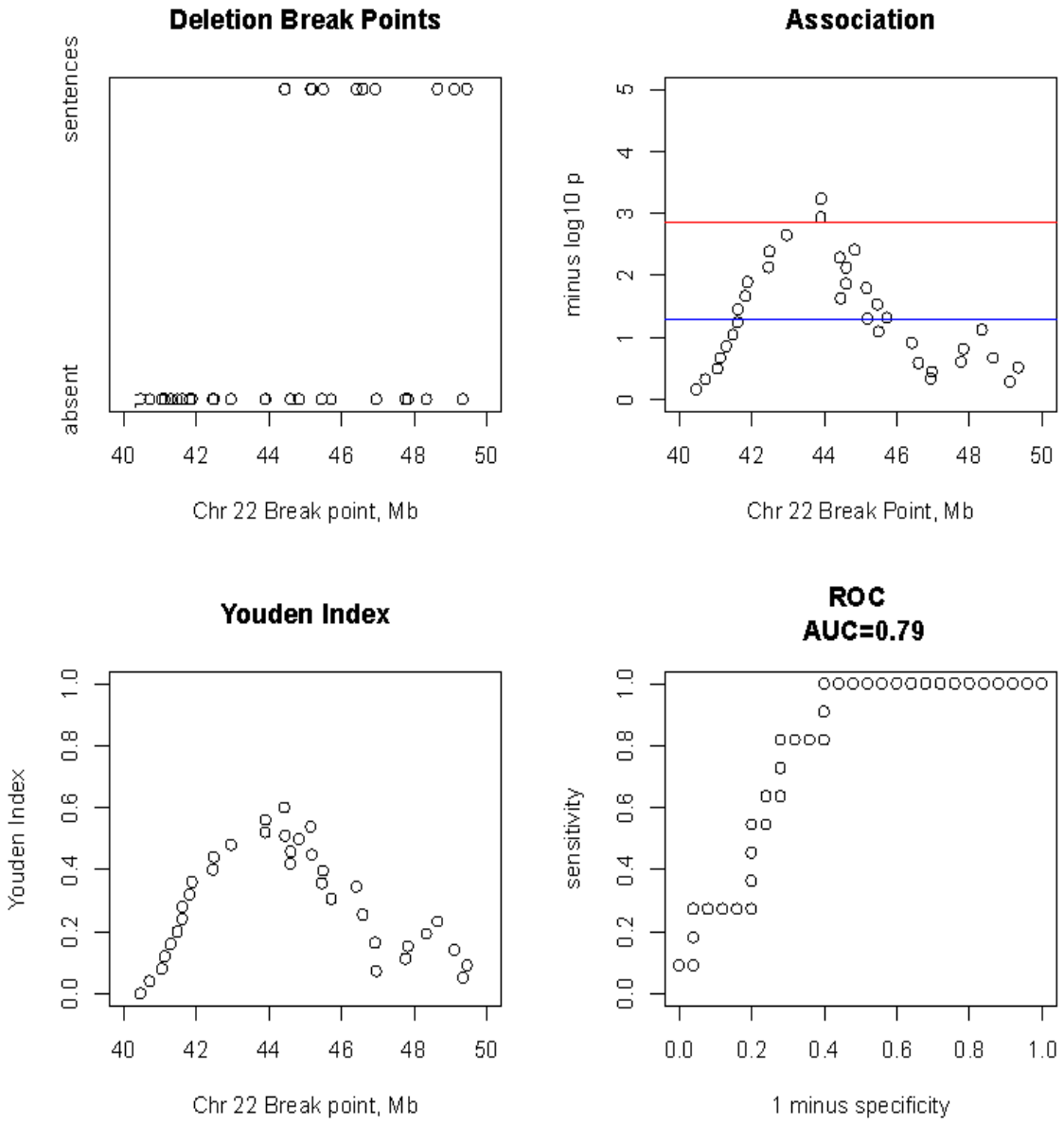


Figure 3.3. Distribution of chromosome 22q13 breakpoints by genomic position for phenotypes significant after bonferroni correction and having an AUC ≥ 0.75. Each column describes a different phenotype. Row A displays the breakpoint positions for those with and without the given phenotype. Row B displays the minus log10 (*P*-value) of association between genomic breakpoint position and phenotype. The blue line represents a *P*-value < 0.05 ; the red line indicates significance after Bonferroni correction. Row C displays the ROC curve.

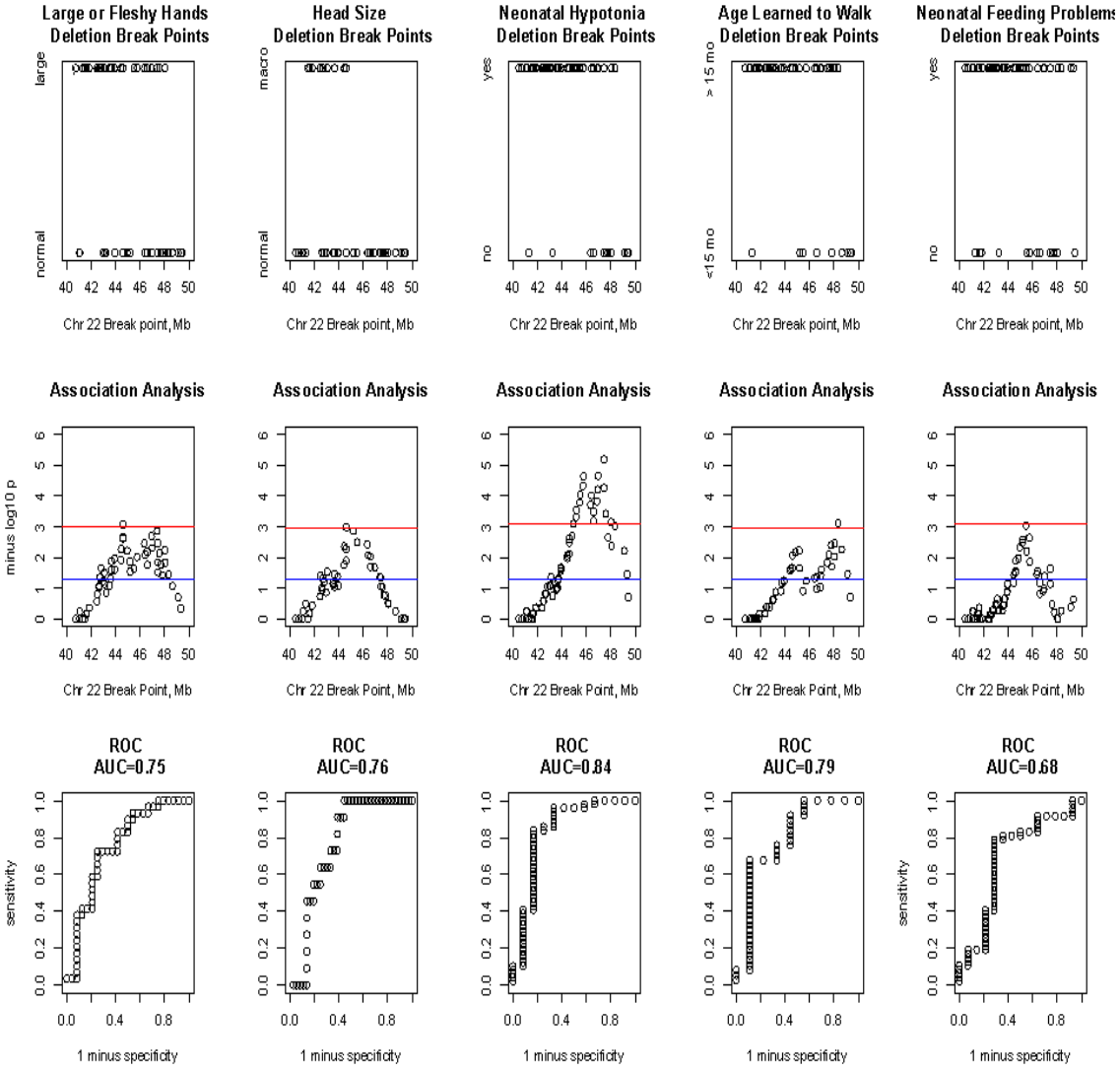


Figure 3.4. Distribution of chromosome 22q13 breakpoints by genomic position for ASDs, microcephaly, and seizures. ASDs were less likely in those with larger deletions. Microcephaly and Seizures do not show association with deletion size. Each column describes a different phenotype. Row A displays the breakpoint positions for those with and without the given phenotype. Row B displays the minus log10 (*P*-value) of association between genomic breakpoint position and phenotype. The blue line represents a *P*-value < 0.05 ; the red line indicates significance after Bonferroni correction. Row C displays the ROC curve.

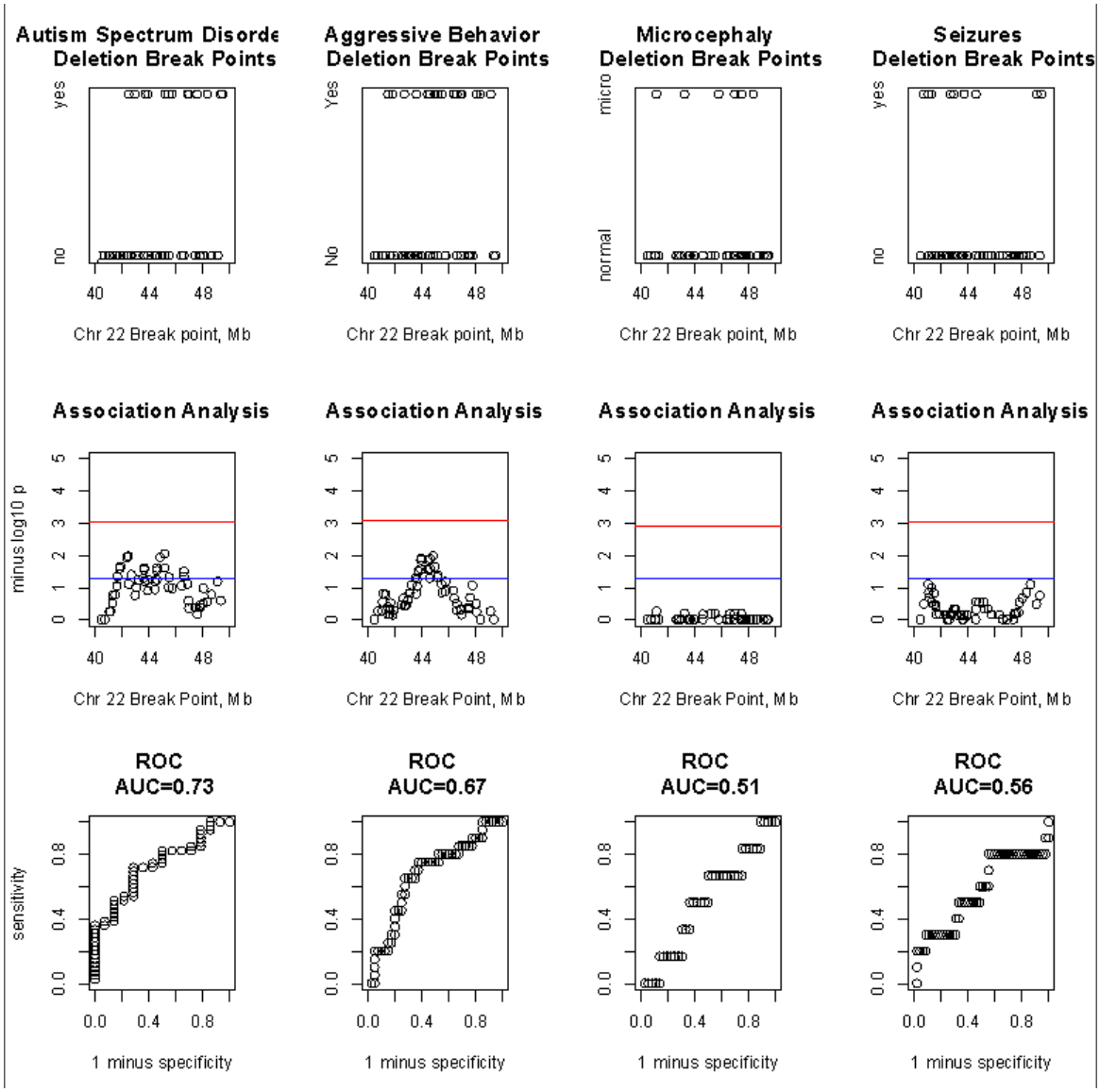


Figure 3.5. The PMS deletion region illustrating the genomic regions statistically associated with medical and physical phenotypes of interest. The dark green bands represent the features obtained from medical history and the light green bands represent features obtained from physical exam. The bands represent regions associated at P<0.05. The thick band represents the regions with the smallest P-value. Also shown are locations of known genes and miRNAs, locations of genes found to be associated with ASD, ID, hypotonia, or macrocephaly in protein interaction databases using GeneMANIA (Warde-Farley and others 2010). In addition, predictions of haplosufficiency (Huang and others 2010) as shown with red indicating predicted haploinsufficient and green indicating predicted haplosufficient. Finally, the interstitial deletions as given by (Wilson et al. 2008) are provided.

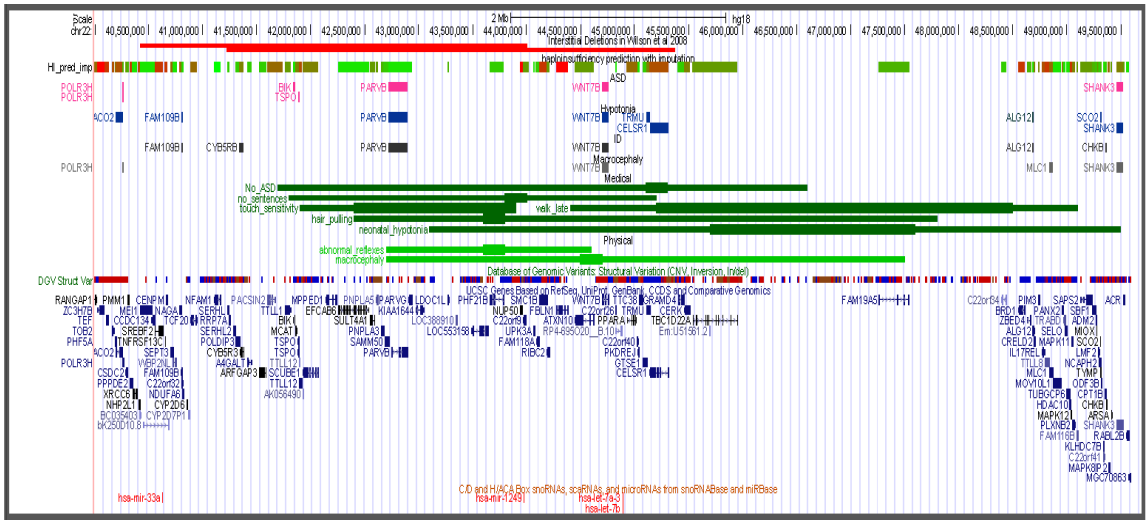
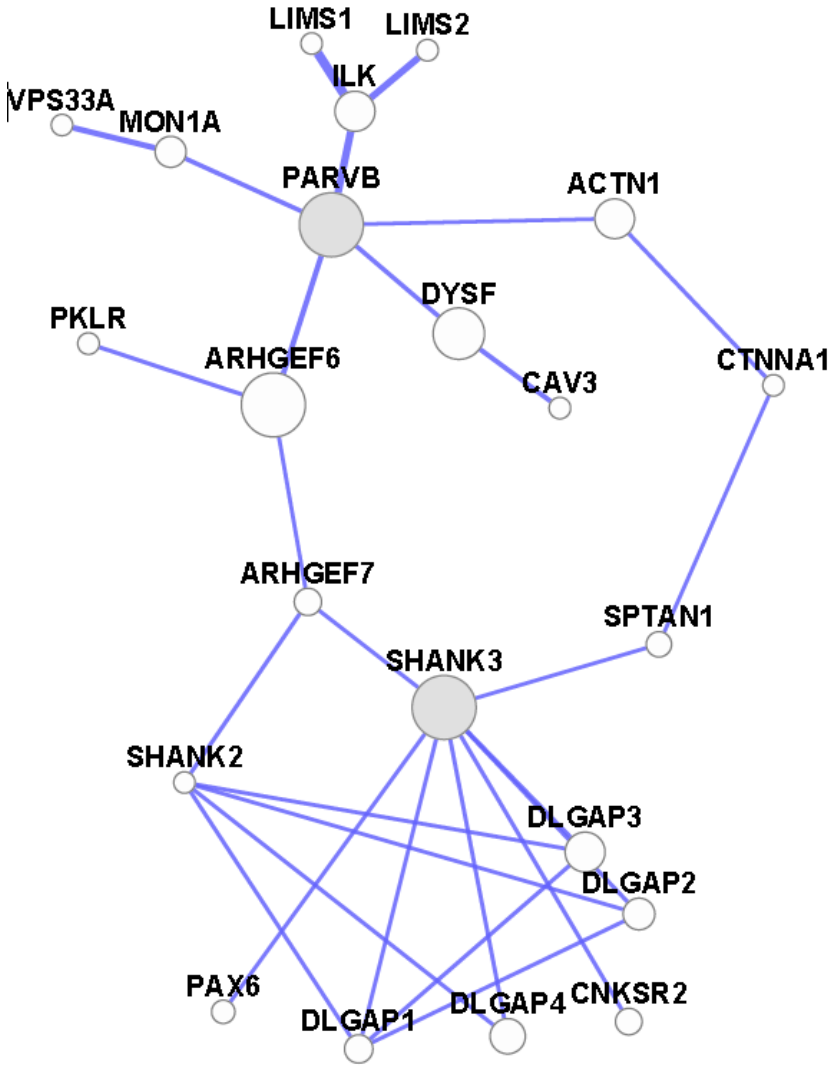
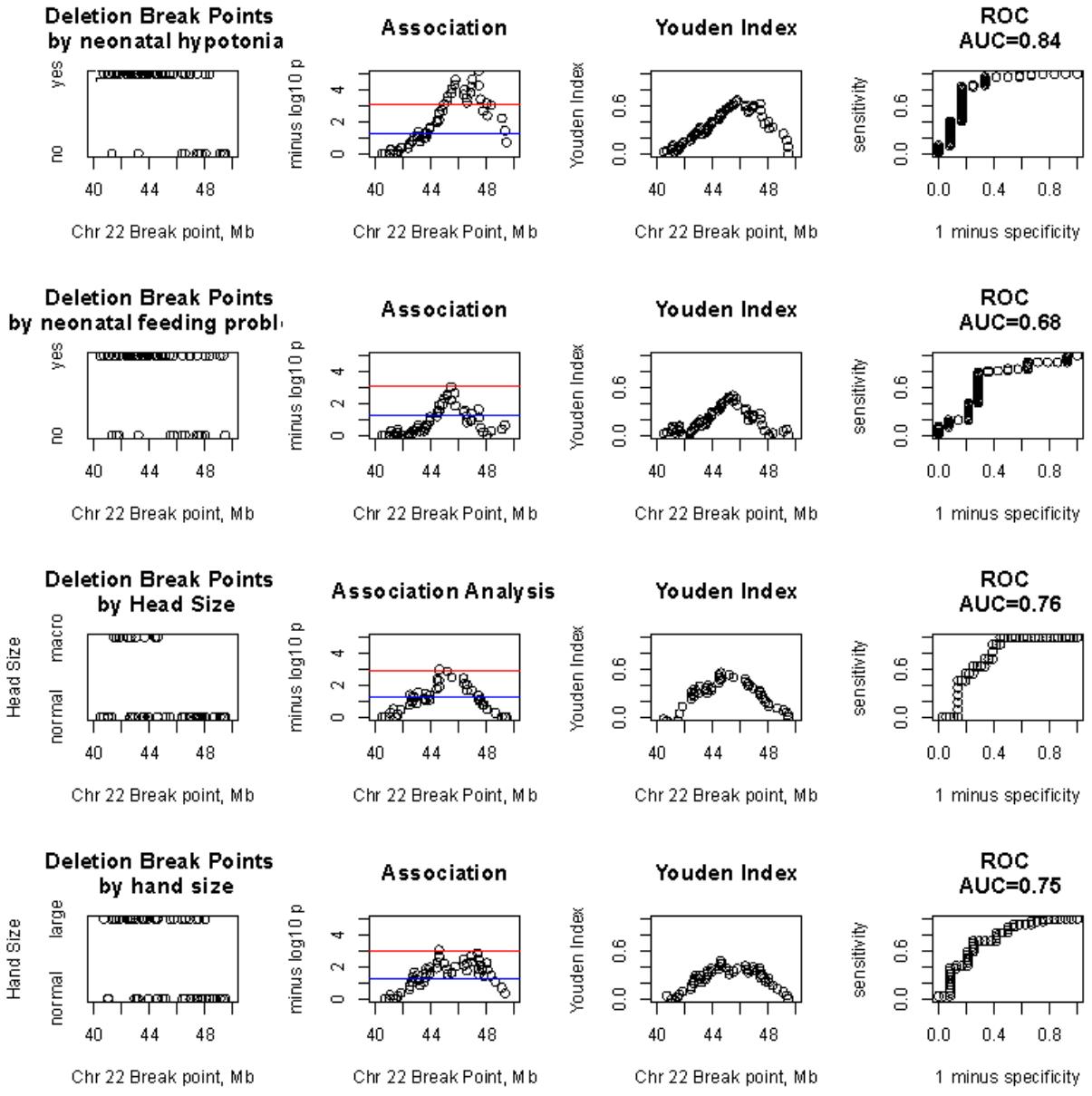
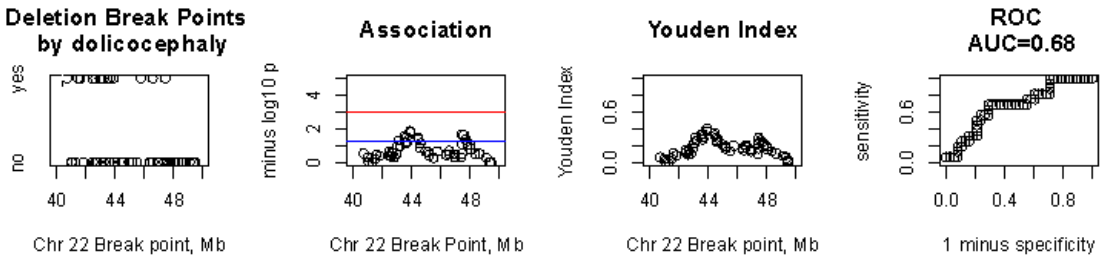


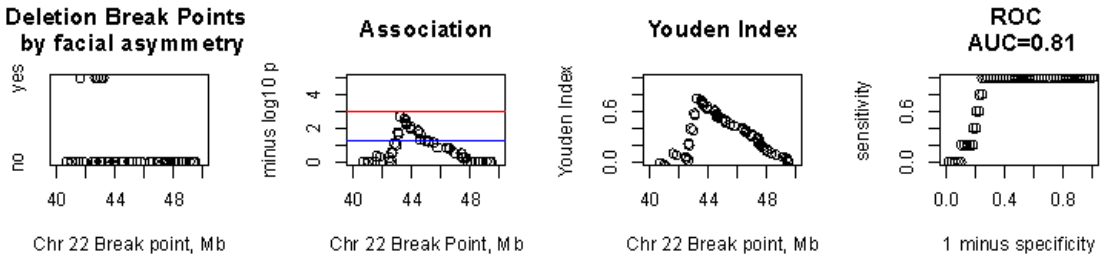
Figure 3.6. Physical protein-protein interaction network between PARVB and SHANK3 as created by GeneMANIA (Warde-Farley and others 2010). Both proteins share interacting partners. Intersitital deletions are missing a copy of PARVB and not SHANK3 while the remaining cases of PMS are all missing SHANK3.

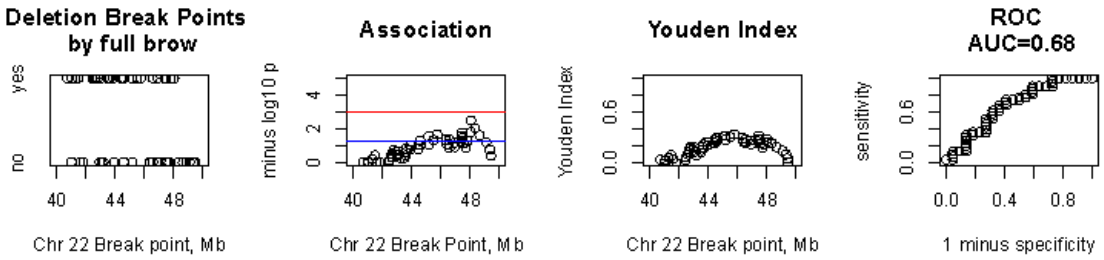


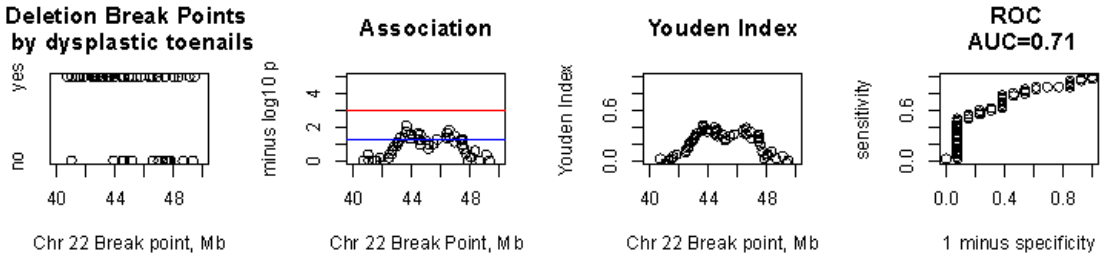
Supplemental Figure 3.1. Association Analysis Results for all phenotypes. The first column depicts the distribution of chromosome 22q13 breakpoints by genomic position. The second column displays the minus log10 (*P*-value) of association between genomic breakpoint position and phenotype. The third column displays the Youden Index by genomic breakpoint position. The fourth column displays the ROC curve. The figure extends 6 pages.

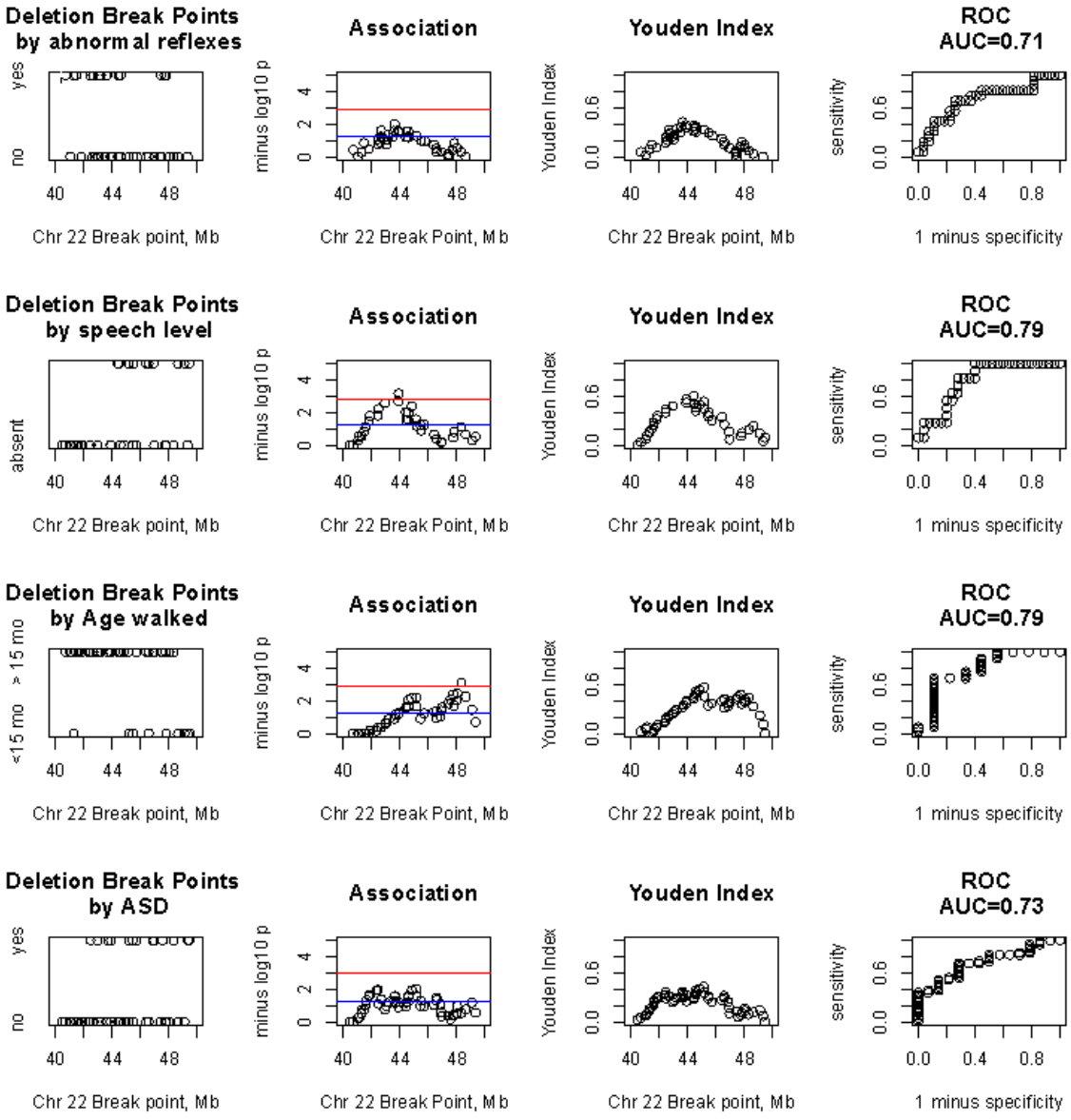


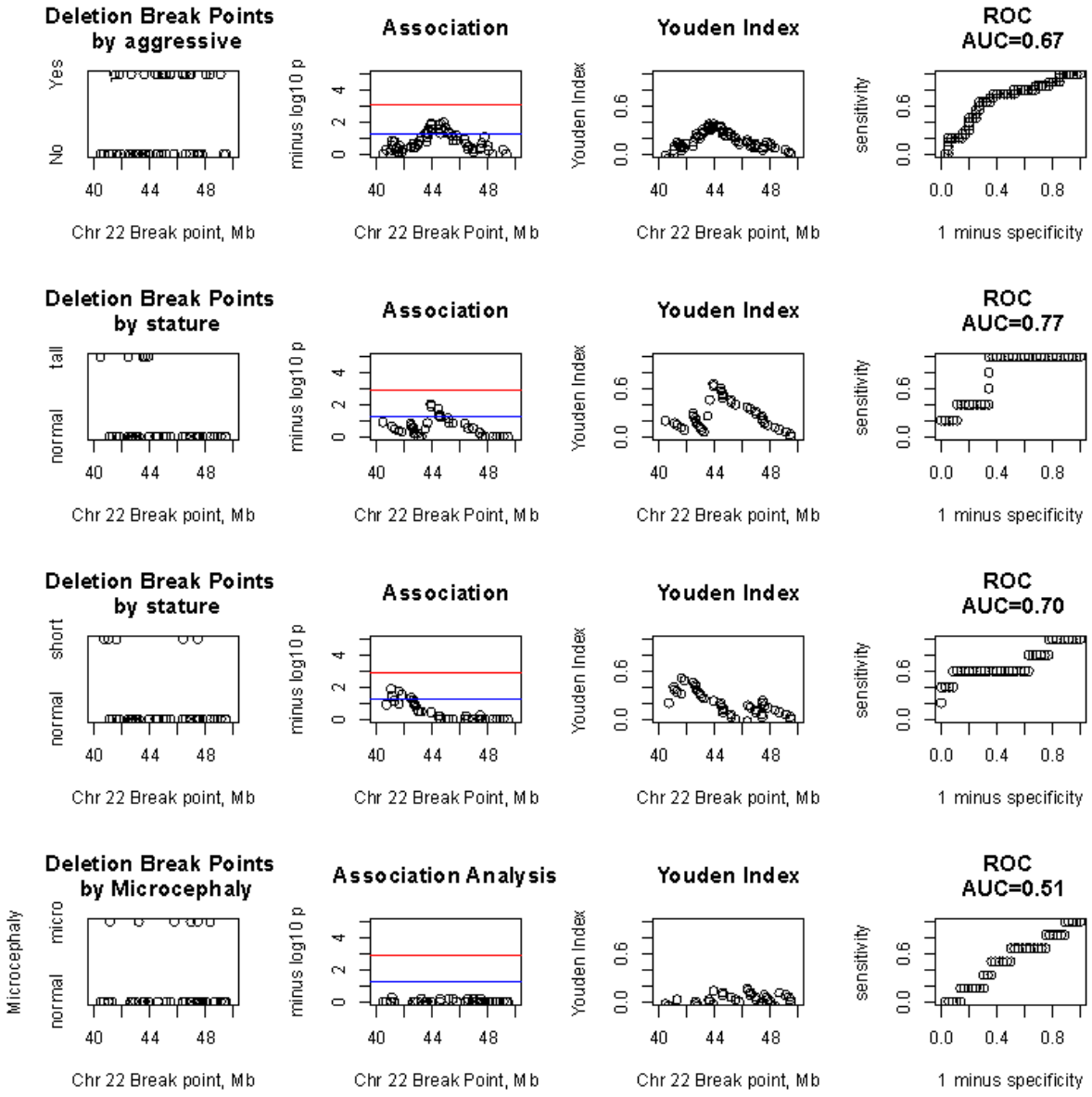

Deletion Break Points
by dolicocephaly
yes
no
40 44 48
Chr 22 Break point, Mb
Association
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC
AUC=0.68
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity

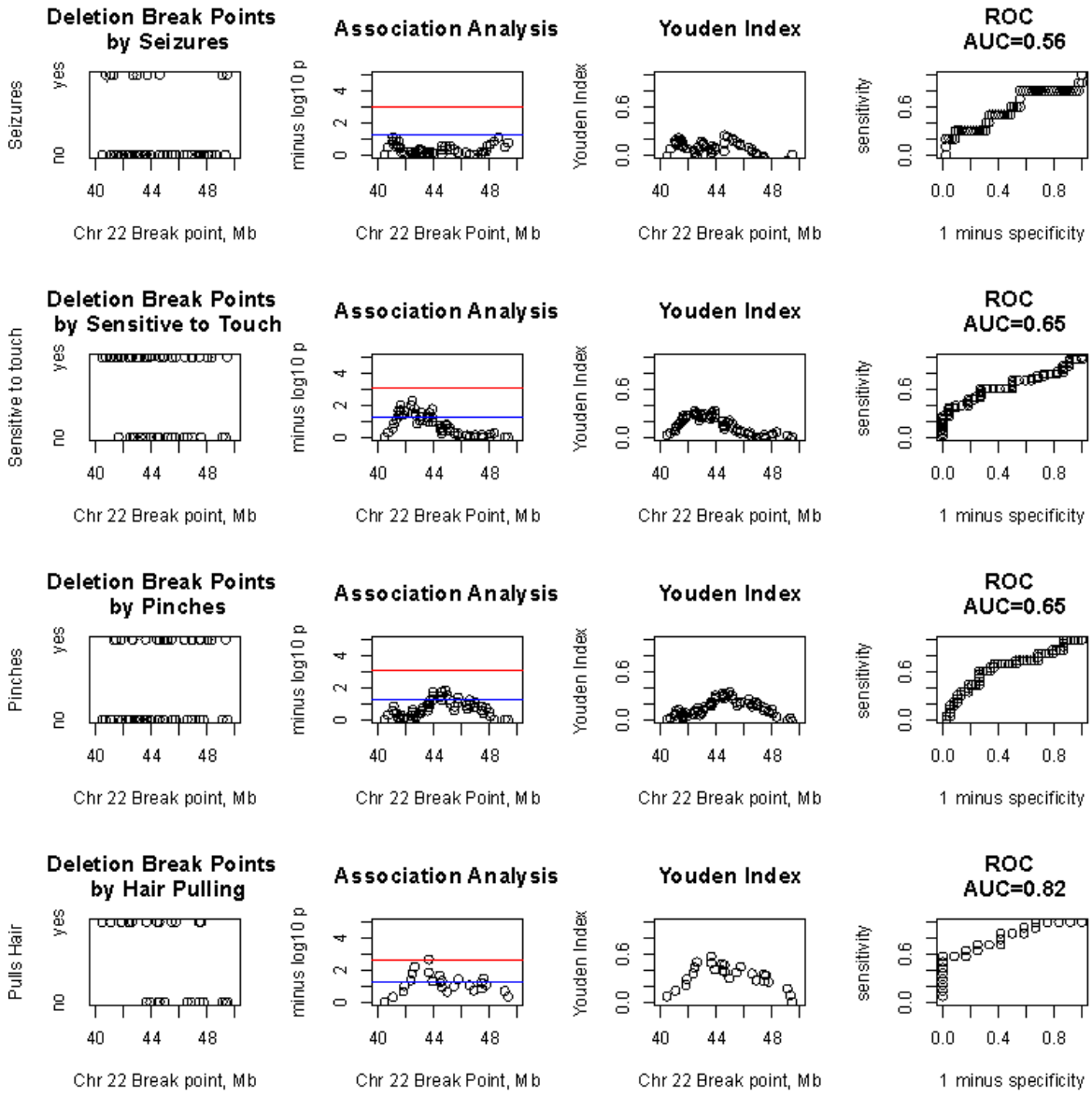

Deletion Break Points
by facial asymmetry
yes
no
40 44 48
Chr 22 Break point, Mb
Association
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC
AUC=0.81
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity

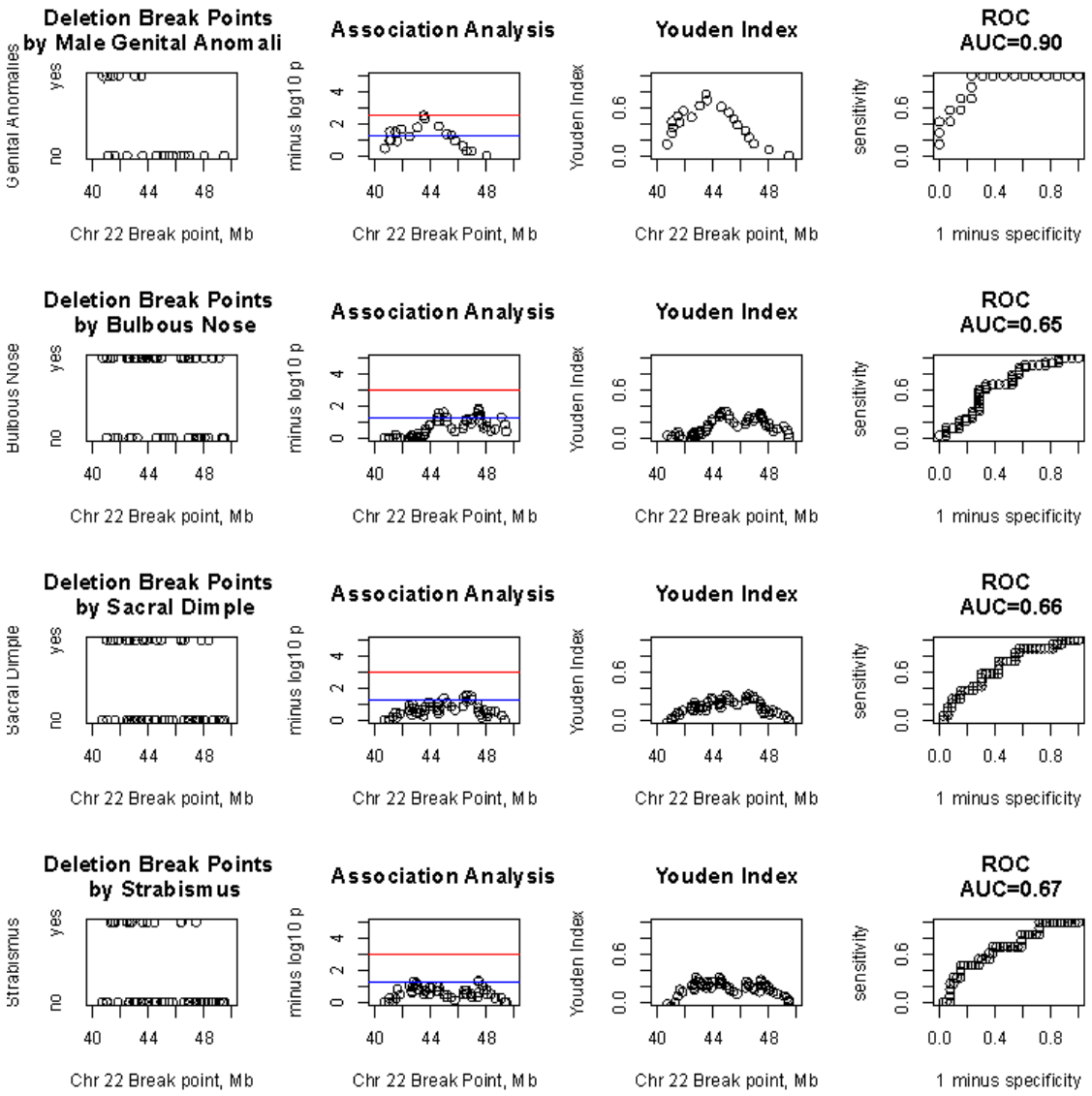

Deletion Break Points
by full brow
yes
no
40 44 48
Chr 22 Break point, Mb
Association
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC
AUC=0.68
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity


Deletion Break Points
by dysplastic toenails
yes
no
40 44 48
Chr 22 Break point, Mb
Association
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC
AUC=0.71
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity


Deletion Break Points
by abnormal reflexes
yes
no
40 44 48
Chr 22 Break point, Mb
Association
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC
AUC=0.71
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity
Deletion Break Points
by speech level
absent
40 44 48
Chr 22 Break point, Mb
Association
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC
AUC=0.79
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity
Deletion Break Points
by Age walked
<15 mo
> 15 mo
40 44 48
Chr 22 Break point, Mb
Association
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC
AUC=0.79
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity
Deletion Break Points
by ASD
yes
no
40 44 48
Chr 22 Break point, Mb
Association
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC
AUC=0.73
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity


Deletion Break Points by aggressive
Yes
No
40 44 48
Chr 22 Break point, Mb
Association
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC AUC=0.67
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity
Deletion Break Points by stature
tall
normal
40 44 48
Chr 22 Break point, Mb
Association
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC AUC=0.77
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity
Deletion Break Points by stature
short
normal
40 44 48
Chr 22 Break point, Mb
Association
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC AUC=0.70
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity
Deletion Break Points by Microcephaly
Microcephaly
micro
normal
40 44 48
Chr 22 Break point, Mb
Association Analysis
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC AUC=0.51
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity


Deletion Break Points by Seizures
Seizures
yes
no
40 44 48
Chr 22 Break point, Mb
Association Analysis
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC AUC=0.56
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity
Deletion Break Points by Sensitive to Touch
Sensitive to touch
yes
no
40 44 48
Chr 22 Break point, Mb
Association Analysis
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC AUC=0.65
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity
Deletion Break Points by Pinches
Pinches
yes
no
40 44 48
Chr 22 Break point, Mb
Association Analysis
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC AUC=0.65
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity
Deletion Break Points by Hair Pulling
Pulls Hair
yes
no
40 44 48
Chr 22 Break point, Mb
Association Analysis
minus log10 p
0 2 4
40 44 48
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0 0.6
40 44 48
Chr 22 Break point, Mb
ROC AUC=0.82
sensitivity
0.0 0.6
0.0 0.4 0.8
1 minus specificity


Deletion Break Points
by Male Genital Anomali
Genital Anomalies
yes
no
40
44
48
Chr 22 Break point, Mb
Association Analysis
minus log10 p
0
2
4
Chr 22 Break Point, Mb
Youden Index
Youden Index
0.0
0.6
Chr 22 Break point, Mb
ROC
AUC=0.90
sensitivity
0.0
0.4
0.8
1 minus specificity
Deletion Break Points
by Bulbous Nose
Bulbous Nose
Association Analysis
Youden Index
ROC
AUC=0.65
Deletion Break Points
by Sacral Dimple
Sacral Dimple
Association Analysis
Youden Index
ROC
AUC=0.66
Deletion Break Points
by Strabismus
Strabismus
Association Analysis
Youden Index
ROC
AUC=0.67

References

Akhtar A and Gasser SM. 2007. The nuclear envelope and transcriptional control. Nat Rev Genet 8(7):507-1.

Alfarano C, Andrade CE, Anthony K, Bahroos N, Bajec M, Bantoft K, Betel D, Bobechko B, Boutilier K, Burgess E, and others. 2005. The biomolecular interaction network database and related tools 2005 update. Nucleic Acids Res 33(Database issue):D418-24.

Altman DG, Lausen B, Sauerbrei W, Schumacher M. 1994. Dangers of using "optimal" cutpoints in the evaluation of prognostic factors. J Natl Cancer Inst 86(11):829-35.

Bader GD, Betel D, Hogue CW. 2003. BIND: The biomolecular interaction network database. Nucleic Acids Res 31(1):248-50.

Basu SN, Kollu R, Banerjee-Basu S. 2009. AutDB: A gene reference resource for autism research. Nucleic Acids Res 37(Database issue):D832-6.

Bewick V, Cheek L, Ball J. 2004. Statistics review 13: Receiver operating characteristic curves. Crit Care 8(6):508-12.

Bonaglia MC, Giorda R, Borgatti R, Felisari G, Gagliardi C, Selicorni A, Zuffardi O. 2001. Disruption of the ProSAP2 gene in a t(12;22)(q24.1;q13.3) is associated with the 22q13.3 deletion syndrome. Am J Hum Genet 69(2):261-8.

Bonaglia MC, Giorda R, Mani E, Aceti G, Anderlid BM, Baroncini A, Pramparo T, Zuffardi O. 2006. Identification of a recurrent breakpoint within the SHANK3 gene in the 22q13.3 deletion syndrome. J Med Genet 43(10):822-8.

Bonaglia MC, Giorda R, Beri S, De Agostini C, Novara F, Fichera M, Grillo L, Galesi O, Vetro A, Ciccone R, and others. 2011. Molecular mechanisms generating and stabilizing terminal 22q13 deletions in 44 subjects with Phelan/McDermid syndrome. PLoS Genet 7(7):e1002173.

Bouquillon S, Andrieux J, Landais E, Duban-Bedu B, Boidein F, Lenne B, Vallee L, Leal T, Doco-Fenzy M, Delobel B. 2011. A 5.3Mb deletion in chromosome 18q12.3 as the smallest region of overlap in two patients with expressive speech delay. Eur J Med Genet 54(2):194-7.

Brown KR and Jurisica I. 2005. Online predicted human interaction database. Bioinformatics 21(9):2076-82.

Chatr-aryamontri A, Ceol A, Palazzi LM, Nardelli G, Schneider MV, Castagnoli L, Cesareni G. 2007. MINT: The molecular INTeraction database. Nucleic Acids Res 35(Database issue):D572-4.

Chiurazzi P, Schwartz CE, Gecz J, Neri G. 2008. XLMR genes: Update 2007. Eur J Hum Genet 16(4):422-34.

Cocchella A, Malacarne M, Forzano F, Marciano C, Pierluigi M, Perroni L, Faravelli F, Di Maria E. 2010. The refinement of the critical region for the 2q31.2q32.3 deletion syndrome indicates candidate genes for mental retardation and speech impairment. Am J Med Genet B Neuropsychiatr Genet 153B(7):1342-6.

Cooley MA, Kern CB, Fresco VM, Wessels A, Thompson RP, McQuinn TC, Twal WO, Mjaatvedt CH, Drake CJ, Argraves WS. 2008. Fibulin-1 is required for morphogenesis of neural crest-derived structures. Dev Biol 319(2):336-45.

Cusmano-Ozog K, Manning MA, Hoyme HE. 2007. 22q13.3 deletion syndrome: A recognizable malformation syndrome associated with marked speech and language delay. Am J Med Genet C Semin Med Genet 145C(4):393-8.

Delahaye A, Toutain A, Aboura A, Dupont C, Tabet AC, Benzacken B, Elion J, Verloes A, Pipiras E, Drunat S. 2009. Chromosome 22q13.3 deletion syndrome with a de novo interstitial 22q13.3 cryptic deletion disrupting SHANK3. Eur J Med Genet 52(5):328-32.

Dhar SU, del Gaudio D, German JR, Peters SU, Ou Z, Bader PI, Berg JS, Blazo M, Brown CW, Graham BH, and others. 2010. 22q13.3 deletion syndrome: Clinical and molecular analysis using array CGH. Am J Med Genet A 152A(3):573-81.

Durand CM, Betancur C, Boeckers TM, Bockmann J, Chaste P, Fauchereau F, Nygren G, Rastam M, Gillberg IC, Anckarsater H, and others. 2007. Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. Nat Genet 39(1):25-7.

Ewing RM, Chu P, Elisma F, Li H, Taylor P, Climie S, McBroom-Cerajewski L, Robinson MD, O'Connor L, Li M, and others. 2007. Large-scale mapping of human protein-protein interactions by mass spectrometry. Mol Syst Biol 3:89.

Feenstra I, Vissers LE, Orsel M, van Kessel AG, Brunner HG, Veltman JA, van Ravenswaaij-Arts CM. 2007. Genotype-phenotype mapping of chromosome 18q deletions by high-resolution array CGH: An update of the phenotypic map. Am J Med Genet A 143A(16):1858-67.

Fisher SE and Scharff C. 2009. FOXP2 as a molecular window into speech and language. Trends Genet 25(4):166-77.

Formstone CJ, Moxon C, Murdoch J, Little P, Mason I. 2010. Basal enrichment within neuroepithelia suggests novel function(s) for Celsr1 protein. Mol Cell Neurosci 44(3):210-22.

Friedman RC, Farh KK, Burge CB, Bartel DP. 2009. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19(1):92-105.

Fujita Y, Mochizuki D, Mori Y, Nakamoto N, Kobayashi M, Omi K, Kodama H, Yanagawa Y, Abe T, Tsuzuku T, and others. 2000. Girl with accelerated growth, hearing loss, inner ear anomalies, delayed myelination of the brain, and del(22)(q13.1q13.2). Am J Med Genet 92(3):195-9.

Gauthier J, Spiegelman D, Piton A, Lafreniere RG, Laurent S, St-Onge J, Lapointe L, Hamdan FF, Cossette P, Mottron L, and others. 2009. Novel de novo SHANK3 mutation in autistic patients. Am J Med Genet B Neuropsychiatr Genet 150B(3):421-4.

Girirajan S, Rosenfeld JA, Cooper GM, Antonacci F, Siswara P, Itsara A, Vives L, Walsh T, McCarthy SE, Baker C, and others. 2010. A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. Nat Genet 42(3):203-9.

Giza J, Urbanski MJ, Prestori F, Bandyopadhyay B, Yam A, Friedrich V, Kelley K, D'Angelo E, Goldfarb M. 2010. Behavioral and cerebellar transmission deficits in mice lacking the autism-linked gene islet brain-2. J Neurosci 30(44):14805-16.

Govek EE, Newey SE, Van Aelst L. 2005. The role of the rho GTPases in neuronal development. Genes Dev 19(1):1-49.

Grabrucker AM, Schmeisser MJ, Schoen M, Boeckers TM. 2011. Postsynaptic ProSAP/shank scaffolds in the cross-hair of synaptopathies. Trends Cell Biol 21(10):594-603.

Greiner M, Pfeiffer D, Smith RD. 2000. Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Prev Vet Med 45(1-2):23-41.

Guldener U, Munsterkotter M, Oesterheld M, Pagel P, Ruepp A, Mewes HW, Stumpflen V. 2006. MPact: The MIPS protein interaction resource on yeast. Nucleic Acids Res 34(Database issue):D436-41.

Hanley JA and McNeil BJ. 1983. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 148(3):839-43.

Hanley JA and McNeil BJ. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143(1):29-36.

Havens JM, Visootsak J, Phelan MC, Graham JM,Jr. 2004. 22q13 deletion syndrome: An update and review for the primary pediatrician. Clin Pediatr (Phila) 43(1):43-5.

Hermjakob H, Montecchi-Palazzi L, Lewington C, Mudali S, Kerrien S, Orchard S, Vingron M, Roechert B, Roepstorff P, Valencia A, and others. 2004. IntAct: An open source molecular interaction database. Nucleic Acids Res 32(Database issue):D452-5.

Hu VW and Steinberg ME. 2009. Novel clustering of items from the autism diagnostic interview-revised to define phenotypes within autism spectrum disorders. Autism Res 2(2):67-7.

Huang N, Lee I, Marcotte EM, Hurles ME. 2010. Characterising and predicting haploinsufficiency in the human genome. PLoS Genet 6(10):e1001154.

International Human Genome Sequencing Consortium. 2004. Finishing the euchromatic sequence of the human genome. Nature 431(7011):931-45.

Jeffries AR, Curran S, Elmslie F, Sharma A, Wenger S, Hummel M, Powell J. 2005. Molecular and phenotypic characterization of ring chromosome 22. Am J Med Genet A 137(2):139-47.

Jin J, Smith FD, Stark C, Wells CD, Fawcett JP, Kulkarni S, Metalnikov P, O'Donnell P, Taylor P, Taylor L, and others. 2004. Proteomic, functional, and domain-based analysis of in vivo 14-3-3 binding proteins involved in cytoskeletal regulation and cellular organization. Curr Biol 14(16):1436-50.

Kalverda B, Pickersgill H, Shloma VV, Fornerod M. 2010. Nucleoporins directly stimulate expression of developmental and cell-cycle genes inside the nucleoplasm. Cell 140(3):360-71.

Kang C and Drayna D. 2011. Genetics of speech and language disorders. Annu Rev Genomics Hum Genet 12:145-64.

Kerrien S, Aranda B, Breuza L, Bridge A, Broackes-Carter F, Chen C, Duesbury M, Dumousseau M, Feuermann M, Hinz U, and others. 2012. The IntAct molecular interaction database in 2012. Nucleic Acids Res 40(Database issue):D841-6.

Kerrien S, Alam-Faruque Y, Aranda B, Bancarz I, Bridge A, Derow C, Dimmer E, Feuermann M, Friedrichsen A, Huntley R, and others. 2007. IntAct--open source resource for molecular interaction data. Nucleic Acids Res 35(Database issue):D561-5.

Koolen DA, Reardon W, Rosser EM, Lacombe D, Hurst JA, Law CJ, Bongers EM, van Ravenswaaij-Arts CM, Leisink MA, van Kessel AG, and others. 2005. Molecular characterisation of patients with subtelomeric 22q abnormalities using chromosome specific array-based comparative genomic hybridisation. Eur J Hum Genet 13(9):1019-24.

Korbel JO, Tirosh-Wagner T, Urban AE, Chen XN, Kasowski M, Dai L, Grubert F, Erdman C, Gao MC, Lange K, and others. 2009. The genetic architecture of down syndrome phenotypes revealed by high-resolution analysis of human segmental trisomies. Proc Natl Acad Sci U S A 106(29):12031-6.

Kou Y, Betancur C, Xu H, Buxbaum JD, Ma'ayan A. 2012. Network- and attribute-based classifiers can prioritize genes and pathways for autism spectrum disorders and intellectual disability. Am J Med Genet C Semin Med Genet 160C(2):130-42.

Kutsche K, Yntema H, Brandt A, Jantke I, Nothwang HG, Orth U, Boavida MG, David D, Chelly J, Fryns JP, and others. 2000. Mutations in ARHGEF6, encoding a guanine nucleotide exchange factor for rho GTPases, in patients with X-linked mental retardation. Nat Genet 26(2):247-50.

Lubs HA, Stevenson RE, Schwartz CE. 2012. Fragile X and X-linked intellectual disability: Four decades of discovery. Am J Hum Genet 90(4):579-90.

Luciani JJ, de Mas P, Depetris D, Mignon-Ravix C, Bottani A, Prieur M, Jonveaux P, Philippe A, Bourrouillou G, de Martinville B, and others. 2003. Telomeric 22q13 deletions resulting from rings, simple deletions, and translocations: Cytogenetic, molecular, and clinical analyses of 32 new observations. J Med Genet 40(9):690-6.

Mishra GR, Suresh M, Kumaran K, Kannabiran N, Suresh S, Bala P, Shivakumar K, Anuradha N, Reddy R, Raghavan TM, and others. 2006. Human protein reference database--2006 update. Nucleic Acids Res 34(Database issue):D411-4.

Moessner R, Marshall CR, Sutcliffe JS, Skaug J, Pinto D, Vincent J, Zwaigenbaum L, Fernandez B, Roberts W, Szatmari P, and others. 2007. Contribution of SHANK3 mutations to autism spectrum disorder. Am J Hum Genet 81(6):1289-97.

Newbury DF, Fisher SE, Monaco AP. 2010. Recent advances in the genetics of language impairment. Genome Med 2(1):6.

Ogawa Y, Miyamoto Y, Asally M, Oka M, Yasuda Y, Yoneda Y. 2010. Two isoforms of Npap60 (Nup50) differentially regulate nuclear protein import. Mol Biol Cell 21(4):630-8.

Pagel P, Kovac S, Oesterheld M, Brauner B, Dunger-Kaltenbach I, Frishman G, Montrone C, Mark P, Stumpflen V, Mewes HW, and others. 2005. The MIPS mammalian protein-protein interaction database. Bioinformatics 21(6):832-4.

Peri S, Navarro JD, Amanchy R, Kristiansen TZ, Jonnalagadda CK, Surendranath V, Niranjan V, Muthusamy B, Gandhi TK, Gronborg M, and others. 2003. Development of human protein reference database as an initial platform for approaching systems biology in humans. Genome Res 13(10):2363-71.

Phelan MC. 2008. Deletion 22q13.3 syndrome. Orphanet J Rare Dis 3:14-9.

Phelan MC, Rogers RC, Saul RA, Stapleton GA, Sweet K, McDermid H, Shaw SR, Claytor J, Willis J, Kelly DP. 2001. 22q13 deletion syndrome. Am J Med Genet 101(2):91-9.

Philippe A, Boddaert N, Vaivre-Douret L, Robel L, Danon-Boileau L, Malan V, de Blois MC, Heron D, Colleaux L, Golse B, and others. 2008. Neurobehavioral profile and brain imaging study of the 22q13.3 deletion syndrome in childhood. Pediatrics 122(2):e376-82.

Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, and others. 2010. Functional impact of global rare copy number variation in autism spectrum disorders. Nature 466(7304):368-72.

Qu Y, Glasco DM, Zhou L, Sawant A, Ravni A, Fritzsch B, Damrau C, Murdoch JN, Evans S, Pfaff SL, and others. 2010. Atypical cadherins Celsr1-3 differentially regulate migration of facial branchiomotor neurons in mice. J Neurosci 30(28):9392-401.

Rapp JP. 2000. Genetic analysis of inherited hypertension in the rat. Physiol Rev 80(1):135-72.

Rollins JD, Sarasua SM, Phelan MC, DuPont BR, Rogers RC, Collins JS. 2011. Growth in Phelan-McDermid syndrome. Am J Med Genet (155):2324-6.

Rosenberger G, Jantke I, Gal A, Kutsche K. 2003. Interaction of alphaPIX (ARHGEF6) with beta-parvin (PARVB) suggests an involvement of alphaPIX in integrin-mediated signaling. Hum Mol Genet 12(2):155-67.

Rothman KJ and Greenland S. 1998. Modern epidemiology. 2nd ed. Philadelphia, PA: Lippincott-Raven.

Ruepp A, Brauner B, Dunger-Kaltenbach I, Frishman G, Montrone C, Stransky M, Waegele B, Schmidt T, Doudieu ON, Stumpflen V, and others. 2008. CORUM: The comprehensive resource of mammalian protein complexes. Nucleic Acids Res 36(Database issue):D646-50.

Sakai Y, Shaw CA, Dawson BC, Dugas DV, Al-Mohtaseb Z, Hill DE, Zoghbi HY. 2011. Protein interactome reveals converging molecular pathways among autism disorders. Sci Transl Med 3(86):86ra49.

Salwinski L, Miller CS, Smith AJ, Pettit FK, Bowie JU, Eisenberg D. 2004. The database of interacting proteins: 2004 update. Nucleic Acids Res 32(Database issue):D449-51.

Sammalisto S, Hiekkalinna T, Suviolahti E, Sood K, Metzidis A, Pajukanta P, Lilja HE, Soro-Paavonen A, Taskinen MR, Tuomi T, and others. 2005. A male-specific quantitative trait locus on 1p21 controlling human stature. J Med Genet 42(12):932-9.

Sarasua SM, Dwivedi A, Boccuto L, Rollins JD, Chen CF, Rogers RC, Phelan K, DuPont BR, Collins JS. 2011. Association between deletion size and important phenotypes expands the genomic region of interest in Phelan-McDermid syndrome (22q13 deletion syndrome). J Med Genet 48(11):761-6.

SAS Institute. 2009. SAS [computer program]. Version 9.2. Cary, North Carolina: .

Spiteri E, Konopka G, Coppola G, Bomar J, Oldham M, Ou J, Vernes SC, Fisher SE, Ren B, Geschwind DH. 2007. Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. Am J Hum Genet 81(6):1144-57.

Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. 2006. BioGRID: A general repository for interaction datasets. Nucleic Acids Res 34(Database issue):D535-9.

Stark C, Breitkreutz BJ, Chatr-Aryamontri A, Boucher L, Oughtred R, Livstone MS, Nixon J, Van Auken K, Wang X, Shi X, and others. 2011. The BioGRID interaction database: 2011 update. Nucleic Acids Res 39(Database issue):D698-704.

Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, Zhang J, Soden R, Hayakawa M, Kreiman G, and others. 2004. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A 101(16):6062-7.

Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, Wiltshire T, Orth AP, Vega RG, Sapinoso LM, Moqrich A, and others. 2002. Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci U S A 99(7):4465-70.

Swets JA. 1988. Measuring the accuracy of diagnostic systems. Science 240(4857):1285-93.

Toyo-oka K, Shionoya A, Gambello MJ, Cardoso C, Leventer R, Ward HL, Ayala R, Tsai LH, Dobyns W, Ledbetter D, and others. 2003. 14-3-3epsilon is important for neuronal migration by binding to NUDEL: A molecular explanation for miller-dieker syndrome. Nat Genet 34(3):274-85.

Vernes SC, Spiteri E, Nicod J, Groszer M, Taylor JM, Davies KE, Geschwind DH, Fisher SE. 2007. High-throughput analysis of promoter occupancy reveals direct neural targets of FOXP2, a gene mutated in speech and language disorders. Am J Hum Genet 81(6):1232-50.

Waga C, Okamoto N, Ondo Y, Fukumura-Kato R, Goto Y, Kohsaka S, Uchino S. 2011. Novel variants of the SHANK3 gene in japanese autistic patients with severe delayed speech development. Psychiatr Genet 21(4):208-11.

Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes CT, and others. 2010. The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res 38(Web Server issue):W214-20.

Williams BA, Mandrekar JN, Mandrekar SJ, Cha SS, Furth AF. 2006. Finding optimal cutpoints for continuous covariates with binary and time-to-event outcomes. Technical Report Series #79:27.

Wilson HL, Wong AC, Shaw SR, Tse WY, Stapleton GA, Phelan MC, Hu S, Marshall J, McDermid HE. 2003. Molecular characterisation of the 22q13 deletion syndrome supports the role of haploinsufficiency of SHANK3/PROSAP2 in the major neurological symptoms. J Med Genet 40(8):575-84.

Wilson HL, Crolla JA, Walker D, Artifoni L, Dallapiccola B, Takano T, Vasudevan P, Huang S, Maloney V, Yobb T, and others. 2008. Interstitial 22q13 deletions: Genes other than SHANK3 have major effects on cognitive and language development. Eur J Hum Genet 16(11):1301-10.

Yang S, Wang K, Valladares O, Hannenhalli S, Bucan M. 2007. Genome-wide expression profiling and bioinformatics analysis of diurnally regulated genes in the mouse prefrontal cortex. Genome Biol 8(11):R247.

Zhang SJ, Zou M, Lu L, Lau D, Ditzel DA, Delucinge-Vivier C, Aso Y, Descombes P, Bading H. 2009. Nuclear calcium signaling controls expression of a large gene pool: Identification of a gene program for acquired neuroprotection induced by synaptic activity. PLoS Genet 5(8):e1000604.

CHAPTER FOUR

CLINICAL AND GENOMIC FEATURES OF PHELAN-MCDERMID SYNDROME IN A LONGITUDINAL ASSESSMENT OF 201 PATIENTS.

Sarasua SM, Boccuto L, Sharp JL, Dwivedi A, Chen CF, Rollins JD, Rogers RC, Phelan K, DuPont BR.

Abstract

This study is the first to describe age-related changes in a large cohort of Phelan-McDermid syndrome patients. The clinical and cytogenetic characteristics of Phelan-McDermid syndrome (PMS) were studied over a period of up to12 years follow-up. 201 patients attended one or more Phelan-McDermid Syndrome Family Support conferences and 130 also provided a whole blood specimen for high resolution oligo array CGH to determine 22q13 breakpoints. Genomic anomalies included terminal deletions (85%), terminal deletions and interstitial duplications (9%), interstitial deletions (2%), a 22q13.3 duplication and no deletion (1%), and 4% who were mosaic preventing detection on array CGH. Individuals diagnosed prior to the year 2005, and especially those diagnosed prior to the year 2000, have larger deletions on average than those diagnosed more recently. Behavioral problems (chewing behaviors, biting, hair pulling, crying) subsided with increasing age and developmental abilities (independent walking and toileting) improved with increasing age. The proportion reporting having an autism spectrum disorder, seizures, or cellulitis increased with age. Among physical features, the presence of large or fleshy hands, hypotonia, lax ligaments, and hyperextensible joints improved with increasing age. Lymphedema and abnormal reflexes increased with age. Younger children tended to be of typical or tall stature while older children tended to be of typical or short stature. Among those over age 3 years, speech abilities did not vary significantly with age, but did vary significantly by deletion size (P=0.0064): absent speech (median deletion size 6.53Mb, range 0.34-9.22), less than 50 words (median deletion size 5.1 Mb, range 1.62-7.45), 50 or more words or phrases (median deletion size 4.36 Mb, range 3.09-5.23), and sentences or being fully

verbal (median deletion size 2.42 Mb, range 0.22-4.52). Larger deletion sizes were associated with greater degree of developmental delay, later age to walk, more hair pulling behaviors, strabismus, facial asymmetry, abnormal reflexes, macrocephaly, dolicocephaly, large or fleshy hands, skin rashes, and 2/3 toe syndactyly. Those with larger deletions were less likely to report an autism diagnosis or aggressive or impulsive behaviors. Hypotonia and seizures were not associated with larger deletions. Deletion size did not vary by parent of origin of the affected chromosome but 73% of affected chromosomes were of paternal inheritance. Seizures were the only feature found associated with maternal parent of origin. This analysis demonstrates the importance in assessing age and deletion size in understanding the clinical presentation of PMS.

Introduction

A substantial body of work is accumulating to better understand the clinical features of Phelan-McDermid syndrome (PMS), which generally consist of developmental delay including significant speech delay, hypotonia, and minor dysmorphic features (Bonaglia and others 2011; Phelan and McDermid 2012). The syndrome is associated with deletions, and occasionally duplications, of up to 9 Mb in size in the 22q13 region. Chromosomal abnormalities include simple terminal deletions, ring chromosomes, translocations, interstitial deletions, as well as duplications. In almost all cases, the *SHANK3* gene, mapping to the distal end of 22q13.33, is affected (Bonaglia and others 2010; Phelan and McDermid 2012) or in rarer cases, disrupted (Anderlid and others 2002; Bonaglia and others 2001; Bonaglia and others 2006; Delahaye and others 2009; Misceo and others 2011) and *SHANK3* is a candidate gene for many of the

neurologic features of the syndrome (Bonaglia and others 2001; Bonaglia and others 2006; Phelan and McDermid 2012). *SHANK3* deletions or point mutations have also been reported in isolated cases of autism spectrum disorders (ASDs) (Betancur, Sakurai, Buxbaum 2009; Boccuto and others 2012; Durand and others 2007; Grabrucker and others 2011). ASD is frequently a co-morbid condition of PMS, with different studies reporting highly variable rates ranging from 0 to 94% (Dhar and others 2010; Jeffries and others 2005; Lindquist and others 2005; Manning and others 2004; Phelan and others 2001; Philippe and others 2008; Sarasua and others 2011). Work is ongoing to determine the effects of *Shank3* haploinsufficiency in mouse models (Bangash and others 2011; Bozdagi and others 2010; Peca and others 2011; Verpelli and others 2011; Wang and others 2011). While the role of *SHANK3* is being actively investigated, little work has been done to examine the independent or contributory roles of the approximately 120 protein coding genes and additional miRNAs or noncoding RNAs co-deleted in the larger deletion patients. Several lines of evidence suggest that deletion size is predictive of phenotypic severity (Jeffries and others 2005; Sarasua and others 2011; Wilson and others 2003; Wilson and others 2008).

Clinical studies of PMS have generally been based upon case reports and small case series, with the largest study describing 61 patients (Phelan and others 2001). One area of research that has not been performed until now is a longitudinal assessment of age-related changes in PMS phenotypes. Because the syndrome is rare and published studies have analyzed small cohorts, clinical features tend to have been examined in aggregate without looking at age-specific differences. However, understanding age-related changes is important for clinical

management and genotype-phenotype assessments may need to account for age if clinical features change with age.

In this study, we examine the prevalence of 60 phenotypes across ages, deletion sizes, and parent-of-origin to better inform patients, researchers, and clinicians on the clinical profile of PMS.

Methods

Patient Recruitment

A total of 201 individuals participated in the study at least once from 2004 to 2010 at four biannual family conferences of the Phelan-McDermid syndrome Family Support Conference. Of these, 55 attended two or more times, including from the earliest conferences in 1998 and 2000. At the conferences, parents or guardians answered a health history survey and patients were offered a standardized physical examination by a trained clinical geneticist. In five cases, medical records were abstracted to obtain clinical information. Height and head circumference were obtained at the physical examinations and were converted to age- and gender- specific percentiles using standard growth charts (Greenwood Genetic Center 2011; Kuczmarski and others 2002; Rollins, Collins, Holden 2010; World Health Organization 2006). In addition, 130 individuals provided a blood sample for array comparative genomic hybridization (CGH) between 2006 and 2010, and of these, 118 also answered the health history questionnaire or had a physical exam. Informed consent was provided and the study was approved by the Institutional Review Board of the Self Regional Health System.

Genetic analysis

22q13 deletions and duplications were delineated from whole blood specimens using a custom 4x44K 60-mer oligo array designed to cover 22q12.3-terminus by Oxford Gene Technology (Oxford, UK) as described previously (Sarasua and others 2011). The 2006 human genome build (NCBI 36/HG 18) was used to establish array CGH genomic breakpoint coordinates (International Human Genome Sequencing Consortium 2004). Whole blood specimens were obtained from 92 patients and their parents to determine the parent of origin of the affected chromosome. Genomic DNA was isolated from peripheral blood by high salt precipitation. Purified DNA was diluted to a concentration of 105 ng/µl and stored at 4°C in TE (10 mmol/l Tris-HCl, pH 7.6, 1 mmol/l EDTA). Because *SHANK3* is deleted in all cases tested, this gene was selected for parent of origin testing. The Primer 3 Input website (http://frodo.wi.mit.edu/primer3/) was used to design the primers to amplify our target region on the *SHANK3* gene: the sequence of the forward primer is 5'- GCCTGGGCAAACTGGACAAGT-3', and the sequence of the reverse primer is 5'-TCCCCAACCAGGAAGCCCTAG-3'. The two primers flank a 491 base pair region including the last part of intron 9, exon 10 and the first part of intron 10 of the *SHANK3* gene. The region was selected for the presence of two highly variable SNPs: c.1304+42 G>A (NG_008607.1:g.25449 G>A, rs13055562) and c.1304+103 C>A (NG_008607.1:g.25511 C>A, rs2341009). The rs13055562 allele frequencies, according to the 1000 genome database (http://www.1000genomes.org/home), are 62.4% G and 37.6% A; while the rs2341009 allele frequencies are 73.7% A and 26.3% C. The region also includes a third SNP, c.1304+48 C>T (NG_008607.1:g.25455 C>T, rs76224556), which is more rare in the normal population (96.2% C and 3.8% T) and has been reported in association with autism spectrum

disorders (Boccuto and others 2012). Each sample was amplified by PCR, purified and then sequenced using the DYEnamiv ET Dye Terminator Cycle Sequencing Kit on the MegaBACE 1000 Analysis System (Amersham Biosciences, Sunnyvale, CA). Sequencing was performed in both forward and reverse directions. We also analyzed the highly polymorphic tandem TA repeats of the microsatellite *D22S1169*, mapping within *SHANK3*. The forward primers (5'-GCACACACATGCACATAATC-3') were synthesized and labeled with fluorescein amidite (Sigma-Aldrich, St. Louis, MO), while the reverse primers (5'-AACAACTTCCAGCAGACG-3') were not labeled. The region amplified by these primers ranges between 113 and 131 base pairs, depending on the number of repeats. The polymorphisms were detected by CEQ™ 8800 Genetic Analysis System (Beckman Coulter, Brea, CA).

Cytogenetic analysis of buccal specimen

A buccal swab collection kit was mailed to a participant with a non-mosaic >7 Mb terminal deletion. The collection kit requested the patient rinse his or her mouth with water, gently scrape the inside of the cheek with a swab, and then wipe the swab on two enclosed glass slides. After air drying for 30 minutes, the slides were placed in a cardboard mailer and returned by mail to the Greenwood Genetic Center for cytogenetic analysis. A control specimen was also collected using the same procedures. For *in situ* hybridization of interphase chromosomes, fluorescent labeled commercial probes for the terminus of 22q (N85A3) and 22q11.2 (D22S75) were obtained from Abbott. Probes were used per manufacturer's specifications. Slides were fixed with a 50% methanol:50% acetic acid dip, then incubated in 2x SSC at 37^{o}C for 15 minutes, serially dehydrated in 70%, 85%, and 100% ethanol at room temperature, denatured in 70% formamide/2 x SSC at 79^{o}C for two minutes, then serially

dehydrated in 70%, 85%, 100% ethanol. *In situ* hybridization was performed at 37^{o}C for a minimum of 12 hours in a humidity chamber. The slides were post washed at 42 ^{o}C using a Post Wash I (50% formimide/2xSSC) solution three times for 5 minutes each and a 2xSSC solution two times for 2 minutes each. Then the slides were place at room temperature in a 2xSSC/0.1%NP-40 solution for 3 minutes. The labeled probes were visualized with FITC-labeled anti-digoxigenin, and chromosomes were counterstained with DAPI. Images were examined at 100x magnification under a Zeiss Axiophot fluorescent microscope equipped with FITC, DAPI and dual band pass filter sets. Digital images were captured by computer using Applied Imaging Cytovision software (Pittsburgh), and photographs were printed on a Kodak XL 7700 color image printer.

Statistical Analysis

A total of 201 individuals participated in the study and 55 of these individuals participated more than once across the different conferences. Most of the analyses in this study are cross-sectional and include only the data from the most recent visit such that each observation is independent. In a subset of analyses designed to specifically look longitudinally, only those individuals who attended multiple times were included in a repeated measures analysis as explained below. SAS software was used for all statistical analysis (SAS Institute 2009). When available, currently having a condition was assessed rather than ever having a condition. Wilcoxon Rank-Sum or Kruskal Wallis *P*-values were calculated when comparing continuous variables such as age and deletion size. Spearman correlation coefficients were calculated for continuous measures. Results were judged to be statistically significant at a $P<0.05$ and of borderline significance at $0.05 < P < 0.10$.

Cross-sectional analysis. To make use of the full sample size of 201 individuals, a cross-sectional analysis by age group was performed. The prevalence of 62 conditions was compared across four age groups: 0.4 to 4.9 years, 5 to 9.9 years, 10 to 17.9 years, and 18 to 64 years. The age categories were selected to represent the pre-school, school age, adolescent, and adult phases of life. Statistical significance was assessed with the Cochran-Mantel-Haenszel chi-square test using rank scores or Fisher's Exact Test. The effect of deletion size (independent variable) on phenotypes (dependent variable) was assessed in regression models and included age and gender covariates. For continuous phenotypes (age at learning to walk, level of developmental delay, and head circumference percentile), linear regression models were used. For dichotomous phenotypes, logistic regression was used. For the models assessing the effect of deletion size, the sample was restricted to those with simple terminal deletions and excluded those with duplications, interstitial deletions, or those with known translocations as these genomic rearrangements may have independent effects from those of the terminal deletions. The association between parent of origin of the affected chromosome and dichotomous outcomes was assessed by Fisher's Exact Test.

Longitudinal analysis. For those individuals who attended two or more times, a longitudinal analysis was conducted to examine changes within the same group of individuals. The longitudinal analysis was conducted for those physical or medical features that showed statistically significant ($P<0.05$) age-related differences in prevalence in the cross sectional analysis. Only those individuals with two or more visits (n=55) were included in the longitudinal analysis. A repeated measures logistic regression procedure was used to model the log odds of

medical/behavioral features. A longitudinal analysis was not conducted of physical exam features because too few individuals had multiple visits for these conditions (n≤12).

Results

Cytogenetic findings

Of 201 individuals with at least one physical exam or health record, the mean age was 8.4 years, median age was 6.15 years with a range of 0.4 to 64.2 years (Table 4.1). The majority (85%) of 22q13 anomalies were terminal deletions; however, deletions accompanied by duplications (9%), and interstitial deletions (2%) were also observed (Table 4.1, Figure 4.1). There was also one case of a duplication without a deletion and in 4% of cases, no deletion was observed on array CGH. These final cases were previously known to carry a mosaic deletion. Among those with a measured deletion, there was no difference in median deletion size between those with terminal deletions, terminal deletions and duplications, or interstitial deletions (Table 4.1). Deletion breakpoints were highly varied across the 9 Mb terminal region of 22q13.

Detection of 22q13 Deletion in Buccal Swab Specimen. The loss of 22q13.3 was demonstrated in an individual with both an array CGH confirmed deletion and a FISH confirmed deletion in a buccal specimen. No loss of the 22q11.2 probe (red) was observed in any of 110 interphase spreads, but in 105 (95%) of these spreads, only one probe for the 22q13.33 terminus (green) was observed, consistent with a terminal deletion of 22q13.3. This individual

was found to have a >7 Mb deletion by array CGH. In a control specimen, 97 of 100 cells showed no deletion.

Age Correlation with Genetic Features

Age of patient and deletion size did not vary by gender of the subject. No difference was observed in age of participants who did and did not have a microarray test performed and no difference was observed in age by the type of chromosomal anomaly. Individuals diagnosed prior to 2000 tended to be older and tended to have larger deletion sizes, however age and deletion size were not correlated. The Spearman rank correlation coefficient between age at assessment and deletion size (among those with terminal deletions) was nonsignificant at r=0.013, *P*=0.887 (n=118). The correlation remained small and nonsignificant even when stratifying by year of diagnosis. Year of diagnosis was correlated with deletion size (r=-0.257, *P*=0.0156, n=88). Among the 30 patients missing year of diagnosis, there was no difference in deletion size (Krusak-Wallis *P*=0.621) or age of participant (Kruskal-Wallis *P*=0.7385) for those with missing or known year of diagnosis. For those with missing deletion size data, there was no difference in age of participant (Wilcoxon *P*=0.4335).

Cross-sectional Analysis of Development, Health, and Physical Features

Features based on medical history

Developmental, speech, and neurological features. By three years of age, the majority of patients (88%) could walk independently (Tables 4.2 and 4.3). The mean age when this skill was acquired was 28.2 months and ranged from 10 to 98 months of age. Fewer patients were reported to be toilet trained (24%) and this skill took longer to acquire, ranging from 36 to 240

months (Table 4.2). Speech abilities, among those over age three years, did not improve significantly with age (*P*=0.6179), except that the strongest verbal abilities were more prevalent in those over 5 years of age (Table 4.3). In total, 50% of the patient group (72/144) had no speech, 27% (39/144) reported having a vocabulary of 40 or fewer words, 10% (15/144) were reported to have 50 or more words in their vocabulary or the ability to use phrases, and 13% (18/144) were reported to have large vocabularies, use full sentences, and to use speech as a primary means of communication. Speech ability, categorized into four levels was inversely correlated with other measures of development including age at learning to walk (Spearman r=-0.3189, *P*=0.0015) and level of developmental delay (r=-0.5016, *P*<0.0001).

The prevalence of seizures increased with age from 11% among those under age 5 years to 60% among adults. Finally, the prevalence of having a high pain threshold increased with age from 69% to 89% between the youngest and older age groups.

Behavioral features and major systems. The number of PMS patients reported to have an autism spectrum disorder or to exhibit autistic-like features increased with age ranging from 19% in the 3-4.9 year old age group and 60% among those over 18 years. As seen in Table 4.3, many adverse behaviors improved with age including decreased prevalence of chewing behaviors, nonstop crying, biting, and hair pulling. Some conditions remained problematic such as gastroesophageal reflux which was reported for 30-50% of individuals regardless of age. Precocious puberty was reported for 14% of those from age 5 to 9 and for 41% among those ages 10 to 17. The prevalence of kidney problems did not change with age while the prevalence of skin rashes and cellulitis increased with increasing age. Diabetes and thyroid disease were rare.

Physical Exam Features

Commonly reported features include long eyelashes, dysplastic toenails, pointed chin, full or puffy eyelids, and large or fleshy hands (Table 4.4). Most physical features remained unchanged by age of the patient other than height and size of hands. While the majority of patients are of typical stature (5-95th percentile, mean height percentile=53%, range from 0.5 to 99.5%), there was a trend of decreasing height percentile with age. Among those under age 5, 19% are tall for age and 7% are short for age, but among those 10 to 17 years of age, none are tall for age and 28% were short for age. The proportion with large or fleshy hands also decreased with age. The proportion with atypical head circumference did not vary by age, but 18% had macrocephaly and 11% had microcephaly. The mean head circumference was 54th centile, range of 0.5 to 99.5th centile.

Neurological/muscular features. Several neurological features improved with age, including fewer individuals presenting with hypotonia, lax ligaments, hyperextensible joints, or weak reflexes as the age of the patient increased (Table 4.3). The prevalence of having strong reflexes increased with age.

Parent of Origin Effects

Of the 92 trios assessed, 64 (70%) were informative for parent of origin of the affected chromosome. While none of the assessed parents had any chromosomal rearrangements, the 22q13 deletion originated from the paternal chromosome in 73% of cases (47/64) and from the maternal chromosome in 27% of cases (17/64) (Table 4.1). No difference was observed in age of the child or deletion size by parent of origin. There was no difference in parental age at

conception between those transmitting an affected chromosome and those transmitting an unaffected chromosome (Table 4.6). The proportion of paternal origin was similar for sons (71%) as for daughters (74%) (Table 4.7). Of the 60 features examined, only the proportion reporting seizures differed by parent of origin (Table 4.8). Having seizures was more common among offspring inheriting a maternally derived chromosome deletion (7/13, 54%) than a paternally derived chromosome deletion (5/30, 17%), P=0.0241 Fisher's Exact Test. The association remained after adjustment for age and gender. None of the other 60 features assessed was statistically significant at P<0.05, although pointed chin and lax ligaments were of borderline significance (Table 4.8).

Longitudinal Analysis

Among 55 patients who had two or more visits and could be assessed longitudinally, the mean time between assessments was 3.7 years, median was 3.0 years (range 0.3 years to 12.1 years). Of these, 41 came for two visits, 13 came for 3 visits, and 1 came for 4 visits. Some clinical data were missing at different visits. The longitudinal change in prevalence of clinical features was assessed with a repeated measures logistic regression model (Table 4.9). This assessment was performed on those features that showed a statistically significant difference in the cross-sectional analysis. Too few individuals came for multiple physical exams ($n \leq 12$) to model physical features. The longitudinal assessment supported the observations derived from the cross-sectional analysis. In general, however, there was little observed change in prevalence of the assessed features between observation time periods. The direction of effect was in agreement with the cross-sectional analysis, although the level of statistical significance was typically worse. Prevalence of high pain tolerance increased with increasing age (P=0.0062).

22q13 Deletion Size Effects

Size of terminal deletion was assessed in a logistic regression model adjusting for potential age and gender effects. Deletion size was statistically significant for 18 features (Tables 4.5 and 4.10) out of 54 features that were assessed (listed in Tables 4.2 and 4.3). In particular, features related to developmental delay (level of developmental delay, speech ability, walking ability), growth (macrocephaly, large/fleshy hand size), dysmorphic features (facial asymmetry, 2/3 toe syndactyly, dolichocephaly), and neurological features (abnormal reflexes, strabismus) were associated with larger deletion sizes. All the macrocephalic patients with terminal deletions had a deletion size > 5.07 Mb and in the macrocephalic cases with a deletion and a duplication, deletions were all > 5.02 Mb. Stature was not associated with deletion size. Autism spectrum disorders and the behavioral features of impulsiveness and aggressive behavior were inversely associated with deletion size: those with larger deletions were less likely to report these features. Seizures and hypotonia were not statistically significantly associated with deletion size in this analysis. Figures 4.2 and 4.3 show deletion size distributions for those with varying speech abilities and for head circumference demonstrating the marked differences observed by phenotype. Speech ability was inversely correlated with size of terminal deletion ($r=-0.3832$, $P<0.0001$). When represented graphically, size of deletion was significantly associated with speech ability and deletion sizes varied greatly for each level of speech ability (Figure 4.2). While the largest deletion among those with absent speech was 9.22 Mb, the largest deletion for those with full sentences and functional language was 4.52 Mb (Table 4.5, Figure 4.2).

Discussion

Overall Clinical Findings

The prevalence of phenotypes observed in this study generally agrees with prior published reports (Bonaglia and others 2010; Jeffries and others 2005; Luciani and others 2003; Phelan and McDermid 2012; Phelan and others 2001; Sarasua and others 2011). Other than developmental and speech delay, the most common features observed in our cohort (Tables 4.3 and 4.4) were: hypotonia (75%), distinctive neurological abnormalities, such as overheating (68%) and high pain threshold (77%), and minor dysmorphic traits, like long eyelashes (93%), dysplastic toenails (73%), and large or fleshy hands (63%). All patients had speech delay with half having no speech, 27% having single words, and 23% having phrases or full sentences. Other common problems included seizures (27%), ASDs (31%), gastroesophageal reflux (42%), kidney problems (26%), frequent constipation (41%), and skin rashes (39%), among other features. Diabetes, thyroid abnormalities, and enzyme deficiencies were rare (≤6%). Among physical features, both tall and short stature were common, as was macrocephaly. Abnormal reflexes and lax ligaments were also common.

Clinical Features Change with Age

This is the first study of PMS to examine age-related changes in a large sample size. A large number of features were found to either increase or decrease in prevalence with age. Certain features such as hypotonia, lax ligaments, and weak reflexes decreased with increasing age. Reflexes were observed to become stronger with age. Tall stature was more prevalent among younger patients, but was absent in children and adults over age 10, and short stature

became more common with age. Macrocephaly was common (~18%) across all age groups. An earlier analysis, restricted to only those with pure non-mosaic terminal deletions, also observed the prevalence of both macrocephaly and microcephaly as well as tall and short stature (Rollins and others 2011). The larger sample size in the less restrictive sample allowed us to perform the age-related analysis. Acceleration or deceleration in growth has been reported in other syndromes. Examples of syndromes with accelerated growth include Sotos syndrome (Douglas and others 2003), Weaver syndrome (Douglas and others 2003), Simpson-Golabi-Behmel syndrome (Pilia and others 1996), and Beckwith-Wiedemann syndrome (Choufani, Shuman, Weksberg 2010) and Russell-Silver syndrome is an example of a syndrome showing growth deceleration (Eggermann 2010).

Longitudinal vs. Cross-Sectional Analysis

The longitudinal analysis agreed with the results derived from the cross-sectional analysis when sufficient sample size was available to perform the assessment. The short longitudinal follow-up period, 3.7 years on average, may not have been long enough to observe changes that may be due to development. Further, some individuals may not have reached the ages most at risk for some features or may have already passed the highest risk time periods for any given clinical feature. Longer follow-up time periods with larger sample sizes are needed to better assess the longitudinal compared to cross-sectional approach to assessing age-related changes in PMS phenotypes. The agreement between age effects noted in the cross-sectional analysis was supported by the repeated measures longitudinal follow-up in the smaller sample supporting the use of the cross-sectional data as a good proxy for longitudinal follow-up of individuals.

Comparison to Other Syndromes Where Longitudinal Assessment Has Been Performed

Few longitudinal assessments have been performed in rare genetic syndromes given the inherent difficulty in studying a large sample size of a rare syndrome. One study to assess cross-sectional and longitudinal changes in a neurodevelopmental disorder caused by a microdeletion was a study of William syndrome (Elison, Stinton, Howlin 2010). They assessed age-related changes among adults and found that increased age was not correlated with diminished cognitive, language, or adaptive functioning, and that some areas improved with age among adults. Our study used methods most similar to the Elison study by comparing prevalence of outcomes by age groups in a cross-sectional manner, as well as including a longitudinal assessment in a subset. Our study included children as well as adults, however. In a study of Fragile X syndrome, boys from ages 9 months to 68 months were assessed multiple times to examine longitudinal changes in development (Roberts and others 2009). The primary aim of the study was to determine age when the delays first became evident and to trace age-related changes in development during early childhood. Unlike in our study, they focused on young children and did not follow them through adolescence or adulthood. In a study of predictors of psychosis in 22q11.2 deletion syndrome, imaging studies and measures of psychiatric function were taken in late childhood and early adolescences in a group of 22q11.2 deletion syndrome patients and controls (Kates and others 2011). The authors were able to correlate brain-related changes to measures of prodromal symptoms and psychosis. The authors employed continuous outcome measures rather than the dichotomous outcome measures used in our study, and they focused on a key period of development, while our study also examined early and late childhood, adolescence, and adulthood to assess key ages of development. We did not include

brain imaging studies, which would be of great benefit to improve specificity and reliability of measurements of neurological and cognitive function. One study of PMS that did examine brain imaging was a study of eight individuals from ages 5 to 8 years (Philippe and others 2008), but longitudinal imaging was not included.

Identification of Deletion in Buccal Specimens

This study successfully identified the deletion in a buccal specimen (cells from internal cheek mucosa). Buccal and central nervous system tissues share the same ectodermal origin while blood cells derive from the mesoderm (Sadler 2004). While testing is typically performed on blood specimens, deletions have been also been detected in amniotic fluid (Chen and others 2005; Maitz and others 2008; Phelan, Brown, Rogers 2001) and umbilical cord blood or specimens (Phelan, Brown, Rogers 2001). The finding of deletions in multiple specimen types and of differing developmental germ layers establishes that the 22q13.3 deletions are not limited to a blood-specific anomaly. With improved laboratory testing methods, less invasive specimens such as saliva or buccal cells may be more frequently employed. Our method used FISH on interphase buccal smears to confirm the presence of the deletion, but adaptations of array CGH to saliva or buccal specimens may allow for the measurement of deletion size, break points, and genome-wide anomalies in the future. Further, future assessments of the potential effects of mosaicism on severity of the phenotype may contribute to the understanding of the clinical variability of PMS.

Parent of origin effects

In our study, no parents were observed to have a chromosomal rearrangement involving 22q13. However we determined that 73% of deleted chromosomes in the PMS patients were derived from a paternally inherited chromosome. This figure agrees with previous reports indicating a greater frequency of paternal inheritance (74% (Luciani and others 2003); 69% (Wilson and others 2003); 59% for ring chromosomes (Jeffries and others 2005); and 74% for terminal deletions and 60% for ring chromosomes (Bonaglia and others 2011). We observed no difference in age between the parents transmitting the affected chromosome and the parents transmitting the normal one (Table 4.6), in agreement with others (Bonaglia and others 2011). No difference in sizes of deletions by parent of origin was observed, also in agreement with others (Bonaglia and others 2011; Jeffries and others 2005). Finally, we observed only one significant feature (seizures) associated with maternal inheritance and no features significantly associated with paternal origin (Table 4.8). Other studies report lack of association with parent of origin and phenotypes (Jeffries and others 2005; Luciani and others 2003). Wilson and others found two features (measures of community living and a high palate) to be associated with paternal origin, and found seizures to be unassociated with parent of origin (Wilson and others 2003). The lack of an observed parent of origin difference in phenotypes implies a lack of imprinting of causal genes in the deletion region, although the studies were small ($n<50$) and did not attempt to include interaction between size of deletion and parent of origin. This lack of a parent of origin effect is in contrast with the well-known parent of origin effects observed in Angelman and Prader-Willi syndromes (Buiting 2010; Nicholls and Knepper 2001). While no 22q13.3 genes are known to be imprinted according to the Catalog of Parent of Origin Effects

(Morison, Paton, Cleverley 2001), prediction algorithms suggest the potential for several 22q13.31q13.33 genes to be imprinted (Luedi and others 2007).

Effect of Deletion Size

As noted by Sarasua and others, deletion size is associated with speech abilities and other measures of development, neurologic function, and physical characteristics (Sarasua and others 2011). In this more detailed analysis, the stepwise decrease in speech ability with increased deletion size could indicate the presence of causal genes in the deletion regions. Individuals with duplications and deletions had speech abilities consistent with their deletion sizes and the size of the duplication did not appear to impact speech abilities. Prior investigations identified genes of particular interest (Chapter 3). In particular, the genes *TRMU*, *WNT7B*, *ATXN10* and micro RNAs *hsa-let-7a-3* and *hsa-let7b* are not deleted in the group with functional speech/full sentences, the genes *FBLN1*, *NUP50*, *C22orf9*, *KIAA1644*, and *PARVB* and *hsa-mir-1249* are not deleted in the group with more than 50 words or phrases, and the genes *MPPED1* and *CYB5R3* are not deleted in the group with single words (Figure 4.2).

Regarding macrocephaly, all patients with macrocephaly have a deletion size > 5 Mb and all simple terminal deletion cases are missing one copy of *WNT7B*, a candidate gene for macrocephaly identified in Chapter 3, while two cases of macrocephaly among those with duplication and deletions are either missing *WNT7B* or have a duplication covering *WNT7B*. WNT7B has been found to interact with GPC3 (Capurro and others 2005), the protein involved in the macrocephaly syndrome Simpson-Golabi-Behmel syndrome (Pilia and others 1996). The large number of genes and microRNAs in the deletion region merit further investigation for contributing to the phenotype of PMS as there are clear deletion size effects. Too few

individuals (n=2) were missing only *SHANK3* to be able to distinguish *SHANK3*-specific phenotypes, although a previous analysis (Chapter 3) identified phenotypic differences by deletion size and deletion band.

This study found decreased prevalence of autism spectrum disorders with increased deletion size. It may be that the more severely impaired individuals with more physical or cognitive impairments may not be appropriate for evaluation of autism.

Changes in Diagnostic Sensitivity: Changing Picture of PMS

Diagnostic sensitivity has increased dramatically in the past 10 or more years. Now that chromosome array CGH is a first tier test for children with developmental disabilities as well as congenital anomalies (Miller and others 2010), not only will more individuals be tested than would have previously, but more small deletions will be detected with this technique than previously. These two factors may affect the severity and types of disabilities present in individuals diagnosed with PMS, likely identifying individuals with smaller deletions and more mild features. Because it has been found in other studies (Dhar and others 2010; Jeffries and others 2005; Sarasua and others 2011; Wilson and others 2003) and in ours that size of deletion affects the constellation and severity of phenotypes found in a patient and that age of evaluation affects the degree to which certain phenotypes are manifested, future studies of PMS will need to assess both age and deletion size in future genotype-phenotype studies and clinical therapeutic trials. As more research is conducted into the effects of *SHANK3* specific mutations and deletions, the *SHANK3*-specific phenotypes will be able to be compared with the remainder of the 22q13.3 deletion patients to separate out effects due to *SHANK3* or other genes.

Conclusion

In our cohort, we found that patients with PMS have widely varying deletion sizes from 0.2 Mb to 9.2 Mb, with no apparent common breakpoints. PMS can be due to simple terminal deletion of 22q13.3ter, interstitial deletions of 22q13.2q13.31, 22q13.3 deletions accompanied by 22q13.3 duplications, and 22q13 duplication without a deletion. Patients present with a diversity of speech, developmental, behavioral, neurologic, and health conditions. The manifestation of these features varies by age at evaluation and deletion size, but not parent of origin of the affected chromosome. Patients diagnosed more recently tend to have smaller deletions and thus may present with less severe phenotypes than reported in the earlier literature on PMS. Research is needed to distinguish between phenotypes caused by *SHANK3* deletion alone and the additional impact of other 22q13.3 genes or genomic regions.

Acknowledgments

I thank the patients and families who participated in this study and made this work possible and the Phelan-McDermid Syndrome Foundation who organized the biannual family conferences where much of the data collection took place. I thank Dr. Amy Lawton-Rauh for helpful comments on the manuscript and the late Dr. Julianne Collins who provided guidance for this study.

Funding

This work was supported by the Phelan-McDermid Syndrome Foundation [to SMS]; the Genetics Endowment of South Carolina; and the South Carolina Department of Disabilities and Special Needs.

Table 4.1. Age and Size of Deletions by Gender, Age Group, Type of Chromosomal Anomaly, Year of Diagnosis and Parent Of Origin.

Patient Information	Sample size	Median age in years (min-max)	*P-Value	22q13.3 Deletion Size (Mb) n=118	*P-value
At least one health history interview or physical exam From 2004 to 2010	201	6.2 (0.4-64.2)		5.23 (0.22-9.22)	
Gender					
Male	81 (40%)	6.5 (0.4-44.6)	0.1481	6.03 (0.22-9.22)	0.2087
Female	120 (61%)	6.2 (1.1-64.2)		5.06 (0.41-9.22)	
Age at most recent visit					
0.3-4.9 years	77 (38%)	3.2 (0.4-4.9)	<0.0001	6.03 (0.34-9.22)	0.1206
5-9.9 years	67 (33%)	6.9 (5.0-9.8)		4.41 (0.22-9.22)	
10-17.9 years	36 (18%)	12.3 (10.0-17.7)		5.23 (1.04-8.94)	
18-64 years	21 (10%)	21.3 (18.3-64.2)		6.63 (2.75-8.96)	
Microarray result					
None available	77 (38%)	5.6 (0.4-64.2)	0.4386	NA	NA
Available	124 (62%)	6.5 (0.9-44.6)			
Type of result					
Terminal deletion	105 (85%)	6.4 (0.9-44.6)		5.3 (0.22-9.22)	0.3300**
Interstitial deletion	2 (2%)	18.5(15.7-21.2)		4.4 (2.72-6.04)	
Duplication	1 (1%)	4.1		4.7 (duplication size)	
Del and duplication*	11 (9%)	6.7 (2.3-12.2)		4.8 (0.41-7.18)	
				1.3 (0.02-6.84) dup size	
No deletion visualized (mosaics)	5 (4%)	18.7 (2.8-33.8)		NA	
Year of Diagnosis					
Unavailable	43 (21%)	5.6 (0.9-36.6)	0.9552	5.0 (0.34-9.22)	0.6232
Available	158 (79%)	6.3 (0.4-64.2)		5.4 (0.22-9.22)	
Before 2000	22 (14%)	13.7 (7.9-23.5)	<0.0001	8.5 (1.58-9.22)	0.0058
2000-2004	62 (39%)	7.0 (2.6-44.6)		5.8 (0.41-8.78)	
2005-2010	74 (47%)	4.2 (0.4-64.2)		4.1 (0.22-8.94)	
Parent of origin of deleted chromosome	17 (27%)	6.9 (0.9-21.3)	0.6759	5.2 (1.78-9.22)	0.2978
Mother	47 (73%)	5.3 (1.1-36.6)		5.1 (0.34-8.94)	
Father					

*Wilcoxon Rank Sum 2-Sided *P*-value for two-level analysis, Kruskal-Wallis *P*-value for multilevel analysis

** the test compared deletion sizes only among the three types of deletions (terminal deletions, interstitial, deletions with co-occurring duplications in 22q13.3)

Table 4.2. The Ability To Walk and Use the Toilet Independently, Among PMS Patients Over Three Years of Age.

Skill	Proportion with ability	Age acquired the ability		
		Mean (months)	Median (months)	Minimum-Maximum (months)
Walk unassisted	88% (n=145)	28.2[1]	22	10-98
Use toilet	24% (n=139)	92.5[2]	78	36-240

[1]Of the 122 patients with an age indicated when the skill was acquired.
[2]Of the 32 patients with an age indicated when the skill was acquired.

Table 4.3. Current Health, Developmental, and Behavioral Features of Individuals with Phelan-McDermid Syndrome, by Age.

Feature	Total (%)	Age 0.4-4.9 years		Age 5-9.9 years		Age 10-17.9 years		Age 18-64 years		P-value[1]
		Yes/total	%	Yes/total	%	Yes/total	%	Yes/total	%	
Developmental/Neurological										
Speech (age > 3)										
None	72/144 (50%)	20/37	54	29/59	49	16/29	55	7/19	37	0.6179
1-40 words	39/144 (27%)	13/37	35	13/59	22	5/29	17	8/19	42	
50+ words or phrases	15/144 (10%)	4/37	11	5/59	8	3/29	10	3/19	16	
Verbal communication	18/144 (13%)	0/37	0	12/59	20	5/29	17	1/19`	5	
Walking unassisted	136/174 (78%)	40/68	59	51/58	88	27/30	90	18/18	100	<0.0001
Toilet trained	33/168 (20%)	3/67	4	10/55	18	9/28	32	11/18	60	<0.0001
Any Seizures	41/151 (27%)	6/55	11	14/53	26	10/23	43	12/20	60	<0.0001
Sleep problems	12/26 (46%)	5/11	45	3/6	50	1/1	100	3/7	43	1.0000[2]
Overheats or turns red easily	105/105 (68%)	33/58	57	41/55	75	20/24	83	11/18	61	0.0816
Decreased perspiration	89/149 (60%)	33/61	54	35/58	60	17/25	68	11/18	61	0.3655
Touch Sensitivity	80/175 (46%)	33/67	49	24/58	41	14/32	44	9/18	50	0.6990
High pain threshold	131/170 (77%)	44/64	69	44/56	79	26/31	84	17/19	89	0.0256
Arachnoid cyst	24/129 (19%)	6/46	13	11/44	25	5/21	24	2/18	11	0.6091
Gastroesophageal reflux	62/149 (42%)	24/56	43	22/52	42	7/23	30	9/18	50	0.8502
Behavioral Features										
ASD (age ≥ 3 years)	39/127 (31%)	7/36	19	15/50	30	9/25	36	8/16	50	0.0270
ASD + autistic like features (age ≥ 3 years)	44/127 (35%)	7/36	19	17/50	34	9/25	36	11/16	69	0.0027
Chewing non-food items	153/181(85%)	67/76	88	58/64	91	27/34	79	12/20	60	0.0121
Impulsiveness	78/166 (47%)	26/61	43	29/55	53	15/31	48	8/19	42	0.7085
Biting (self or others)	82/179 (46%)	41/70	58	26/58	45	13/32	41	2/19	11	0.0006
Hair pulling	48/118 (41%)	23/43	53	15/38	39	6/21	29	4/16	25	0.0163
Excessive screaming	54/174 (31%)	24/68	35	17/56	30	10/32	31	3/18	17	0.2231
Aggressive behavior	49/127 (28%)	18/74	24	15/61	25	11/33	33	5/20	25	0.5534
Nonstop crying	38/178 (21%)	20/69	29	14/58	24	4/32	13	0/19	0	0.0051
Genitourinary										
Genital anomalies	8/146 (5%)	3/56	5	1/51	2	2/23	9	2/16	13	0.2302

Precocious puberty	15/121 (12%)	0/48	0	5/35	14	9/22	41	1/16	6	0.0017
Frequent urinary tract infections	12/158 (8%)	4/59	7	3/54	6	2/26	8	3/19	16	0.5182[2]
Vesicouretal reflux	18/133(14%)	9/46	20	7/50	14	2/21	10	0/16	0	0.2593[2]
Polycystic kidneys	6/132 (5%)	1/51	2	5/46	11	0/19	0	0/16	0	0.1650[2]
Duplicate kidney	1/135 (1%)	0/51	0	1/47	2	0/21	0	0/16	0	0.6222[2]
Dilated renal pelvis	7/129 (5%)	5/51	10	1/43	2	1/19	5	0/16	0	0.3760[2]
Increased kidney size	11/126 (9%)	3/50	6	6/42	14	2/18	11	0/16	0	0.3176[2]
Other kidney trouble	25/133 (19%)	12/49	24	9/50	18	1/18	6	3/16	19	0.3752[2]
Any kidney problem	39/148 (26%)	18/56	32	14/53	26	5/22	23	2/17	12	0.1094
Other Clinical Features										
Frequent constipation	11/27 (41%)	4/10	40	1/8	13	2/2	100	4/7	57	0.1023[2]
Skin rashes	60/152 (39%)	15/58	26	22/50	44	15/26	58	8/18	44	0.0088
Cellulitis	9/137 (7%)	2/54	4	1/48	2	1/19	5	5/16	31	0.0027[2]
Diabetes	2/129 (2%)	1/48	2	0/42	0	0/22	0	1/17	6	0.2751[2]
Hypothyroid	7/121 (6%)	0/44	0	3/42	7	2/19	11	2/16	13	0.0707[2]
Hyperthyroid	1/122 (1%)	0/44	0	0/42	0	0/19	0	1/17	6	0.1391[2]
Enzyme deficiency	4/107 (4%)	0/37	0	1/37	3	1/17	6	2/16	13	0.0806[2]
Immune deficiency	14/113 (12%)	7/44	16	4/39	10	0/15	0	3/15	20	0.2910[2]

[1]*P*-value from Cochran-Mantel-Haenszel Chi-Square (Row Mean Scores Differ) Statistic, rank scores

[2]Fisher's Exact 2-sided *P*-value

Table 4.4. Physical Exam Features by Age of Individuals with Phelan-McDermid Syndrome.

Feature	Total (%)	Age 0.4-4.9 years		Age 5-9.9 years		Age 10-17.9 years		Age 18-44 years		*P*-value[1]
		Yes/total	%	Yes/total	%	Yes/total	%	Yes/total	%	
Physical features										
Height										
<5th percentile	11/96 (11%)	3/42	7	3/29	10	5/18	28	0/7	0	0.0296
5-95th percentile	76/96 (79%)	31/42	74	25/29	86	13/18	72	7/7	100	
95th percentile	9/96 (9%)	8/42	19	1/29	3	0/18	0	0/7	0	
Head circumference										
<3rd percentile	12/110 (11%)	4/52	8	5/33	15	3/20	15	0/5	0	0.4352
3-97th percentile	78/110 (71%)	38/52	73	26/33	79	11/20	55	3/5	60	
>97th percentile	20/110 (18%)	10/52	19	2/33	6	6/20	30	2/5	40	
Long eyelashes	105/113 (93%)	45/49	92	35/35	100	19/22	86	6/7	86	0.0790[2]
Dolicocephaly	36/113 (32%)	19/50	38	7/35	20	7/21	33	3/7	42	0.2923[2]
Pointed chin	58/111 (52%)	31/51	61	15/32	47	9/21	43	3/7	43	0.4003[2]
Facial asymmetry	9/110 (8%)	2/48	4	2/34	6	3/21	14	2/7	29	0.0948[2]
High or arched palate	49/104 (47%)	20/44	45	13/34	38	11/21	52	5/5	100	0.2495
Full or puffy eyelids	60/111 (54%)	32/49	65	16/34	47	11/21	52	1/7	14	0.0548[2]
Epicanthal folds	52/111 (47%)	28/49	57	16/34	47	6/21	29	2/7	29	0.1204[2]
Deep set eyes	34/111 (31%)	18/49	37	9/34	26	6/21	29	1/7	14	0.6113[2]
Large or fleshy hands	71/112 (63%)	39/49	80	14/35	40	14/21	67	4/7	57	0.0023[2]
*2/3 toe syndactyly	53/110 (48%)	23/48	48	17/35	49	10/21	48	3/6	50	1.0000[2]
Dysplastic toenails	81/111 (73%)	39/48	81	24/36	67	15/21	71	3/3	50	0.2268[2]
Dysplastic fingernails	26/111 (23%)	15/48	31	7/35	20	3/21	14	1/7	14	0.4368[2]
Single Palmar crease	12/109 (11%)	4/47	9	2/34	6	4/21	19	2/7	29	0.1374[2]
Neurological/Muscular										
Strabismus	28/109 (26%)	11/47	23	9/34	26	7/21	33	1/7	14	0.6874
Hypotonia	82/110 (75%)	41/48	85	25/34	74	12/21	57	4/7	57	0.0428[2]

Lax ligaments	72/110 (65%)	34/48	71	25/34	74	12/21	57	1/7	14	0.0427
Hyperextensible joints	68/111 (61%)	32/49	65	24/34	71	11/21	52	1/7	14	0.0338[2]
Lymphedema	26/108 (24%)	8/47	17	6/34	18	7/20	35	5/7	71	0.0122
Abnormal reflexes	44/91 (48%)	15/41	37	12/27	44	14/18	78	3/5	60	0.0230[2]
Reflexes										
Weak	28/91 (31%)	12/41	29	8/27	30	7/18	39	1/5	20	0.0119[2]
Typical	47/91 (52%)	26/41	63	15/27	30	4/18	22	2/5	40	
Mixed	2/91 (2%)	0/41	0	1/27	4	0/18	0	1/5	20	
Strong	14/91 (15%)	3/41	7	3/27	11	7/18	39	1/5	20	
Ptosis	53/112 (47%)	25/50	50	10/33	30	13/22	59	5/7	71	0.0722[2]

[1]*P*-value from Cochran-Mantel-Haenzel Chi-Square (Row Mean Scores Differ) Statistic, rank scores

[2]Fisher's Exact 2-sided *P*-value

Table 4.5. Terminal 22q13 Deletion Size by Level of Speech Ability Among Patients Age 3 Years and Older.

Speech level	Sample Size	Median Deletion Size (Mb)	Mininum (Mb)	Maximum (Mb)	Kruskal-Wallis *P*-value
Absent	39	6.53	0.34	9.22	0.0050
1-49 words	18	5.81	1.62	7.45	
50+ words or phrases	6	4.36	3.09	5.23	
Sentences, verbal communication	8	2.42	0.22	4.52	

Table 4.6. Ages at Conception for Parents Transmitting an Affected or Unaffected Chromosome.

Parent of origin of affected chromosome 22q13	Age of Parent transmitting affected chromosome Mean (range in years)	Age of Parent transmitting unaffected chromosome Mean (range in years)	T-Test *P*-value
Mother	31.4 (23-42)	31.4 (19-42)	0.99
Father	33.4 (20-45)	33.5 (20-46)	0.97

Table 4.7. Parent of Origin Transmission by Gender of Child.

Transmission	Number	Proportion Paternal inheritance	Risk ratio (95% confidence interval)
Father to son	20	20/28 (71%)	0.97 (0.72-1.31)
Mother to son	8		
Father to daughter	28	28/38 (74%)	
Mother to daughter	10		

Table 4.8. Phenotypes with Significant Differences by Parent of Origin of Deletion.

Phenotype	Maternally inherited	Paternally inherited	*P*-value Fisher's Exact
Pointed chin	4/12 (25%)	22/34 (65%)	0.0914
Lax ligaments	5/11 (45%)	27/35 (77%)	0.0654
Seizures	7/13 (58%)	5/30 (17%)	**0.0241***

*This association remained significant (*P*=0.0218) after adjusting for age and gender in a logistic model.

Table 4.9. Effect of Increasing Age on Selected Features Assessed in a Repeated Measures Logistic Regression Model.

Feature	Sample Size	Beta Coefficient[1]	*P*-value	Odds ratio (95% CI) [2]
Walk alone	47	0.6623	0.0459	1.9 (1.01-3.71)
Use toilet alone	45	0.4363	0.0037	1.5 (1.16-2.06)
Autism spectrum disorder (age > 3 years)	31	-0.0968	0.7085	0.9 (0.54-1.53)
High pain tolerance	46	0.3348	0.0062	1.4 (1.10-1.77)
Seizures	30	0.0693	0.3666	1.1 (0.92-1.25)
Chewing behavior[3]	51	-0.0701	0.2850	0.9 (0.82-1.06)
Excessive crying	49	-0.0775	0.1638	0.9 (0.83-1.08)
Biting	48	-0.0880	0.0801	0.9 (0.83-1.01)
Hair pulling	12	-0.7665	0.0920	0.5 (0.19-1.16)
Skin rashes	25	-0.1014	0.2572	0.9 (0.755-1.08)
Cellulitis	20	0.1983	0.0789	1.2 (0.93-1.52)

[1]The Beta coefficient represents the change in log odds between a one year increase in follow-up age.

[2]The odds ratio (with 95% confidence interval) represents the increased odds of having the feature between a one year increase in follow-up age.

[3]For chewing behavior, the model improved with adding a quadratic term. The coefficient for age is β=0.4100 (*P*=0.0951) and the coefficient for age^2 is β=-0.0141 (*P*=0.0602).

Table 4.10. Association Between Size of Terminal Deletion (In Mb) and Phenotypes, Adjusting for Age and Gender in a Regression Model (Results With *P*<0.10) [1]

Phenotype	Sample Size	Coefficient for size of terminal deletion (Mb)	*P*-value	Odds ratio (95% CI) comparing a 1 Mb difference in deletion size
Linear regression models[1]				
Age learned to walk (months) among those age > 3 years	58	4.439	<0.0001	
Level of delay (among those age > 3 years)	61	0.220	0.0108	
Head circumference, percentile	72	5.266	0.0030	
Logistic regression models[2]				
Large or fleshy hands	73	0.3860	0.0017	1.47 (1.16-1.87)
Ability to walk alone (among those age > 3 years)	69	-0.9193	0.0028	0.40 (0.22-0.73)
Abnormal reflexes	60	0.3925	0.0059	1.48 (1.12-1.96)
Dolichocephaly	74	0.3112	0.0125	1.37 (1.07-1.74)
Macrocephaly (>97th percentile)	72	0.3826	0.0136	1.47 (1.06-1.97)
Aggressive behavior	85	-0.2479	0.0253	0.78 (0.63-0.97)
Hair pulling	47	0.3463	0.0303	1.41 (1.03-1.93)
2/3 toe syndactyly	72	0.2326	0.0373	1.26 (1.01-1.57)
Impulsiveness	79	-0.2120	0.0462	0.81 (0.66-1.00)
Strabismus	71	0.2552	0.0462	1.29 (1.00-1.66)
ASD (age > 3)	62	-0.2513	0.0540	0.78 (0.60-1.00)
ASD + ASD features (age > 3)	62	-0.2313	0.0710	0.79 (0.62-1.02)
Cellulitis	65	1.0553	0.0617	2.87 (0.95-8.69)
Facial asymmetry	72	0.5579	0.0691	1.75 (0.96-3.19)
Skin rashes	75	0.1896	0.0837	1.21 (0.98-1.50)

[1]For linear regression models the hypothesis tested was no linear relationship between deletion size and phenotype. The Beta coefficient represents the predicted mean change in phenotype for a 1 Mb change in deletion size.

[2]For the logistic regression models, the hypothesis tested was no change in log odds of phenotype for a 1 Mb change in deletion size.

Figure 4.1. Deletions (red bars) and Duplications (blue bars) Observed in Unrelated Individuals with Phelan-McDermid Syndrome. The location of 22q13 cytogenetic bands, known genes, and the location of previously reported interstitial deletions (Kent and others 2002; Wilson and others 2008) are also presented in this figure produced using the UCSC genome browser (Kent and others 2002) using the 2006 (hg18) genome build (International Human Genome Sequencing Consortium 2004).

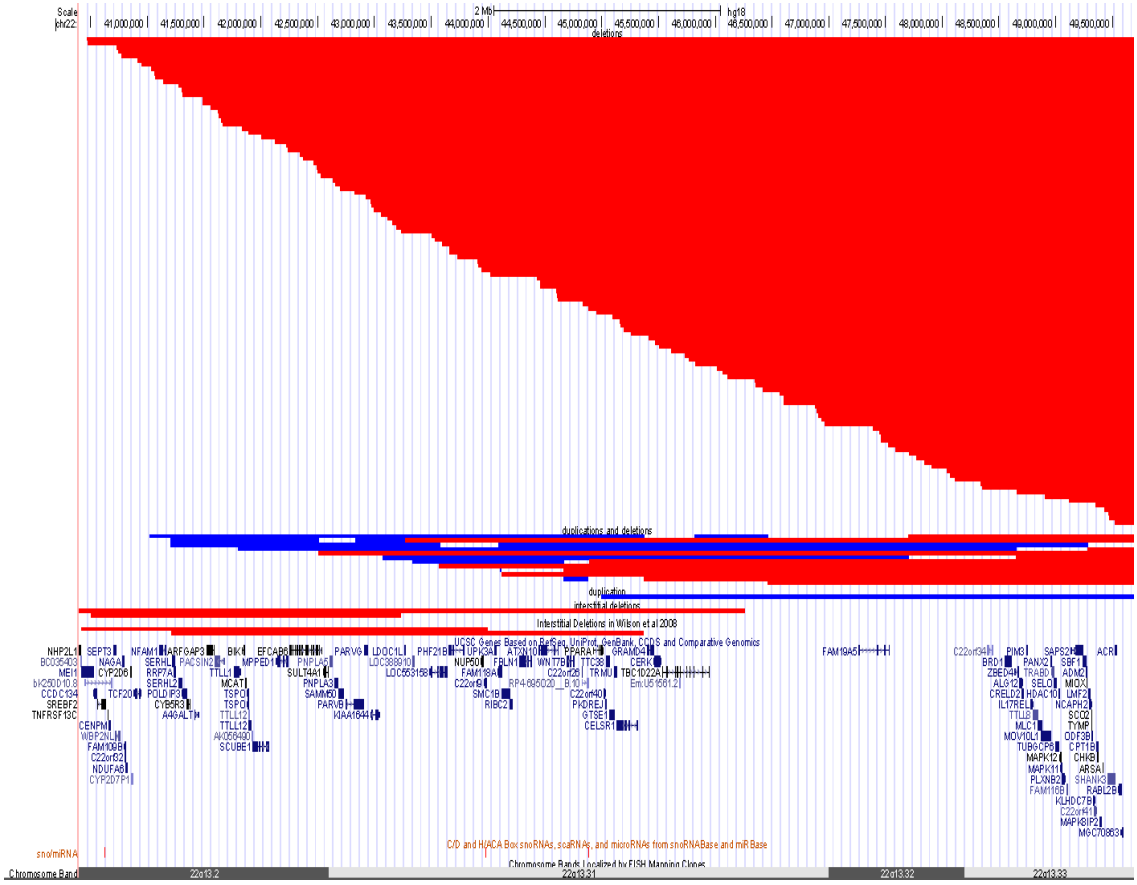


Figure 4.2. Deletions (red bars) and Duplications (blue bars) Grouped by Speech Ability. The numbers to the left of the bars indicate the approximate number of words in the patient's vocabulary. The location of 22q13 cytogenetic bands, known genes, and the location of previously reported interstitial deletions (Kent and others 2002; Wilson and others 2008) are also presented in this figure produced using the UCSC genome browser (Kent and others 2002) using the 2006 (hg18) genome build (International Human Genome Sequencing Consortium 2004)

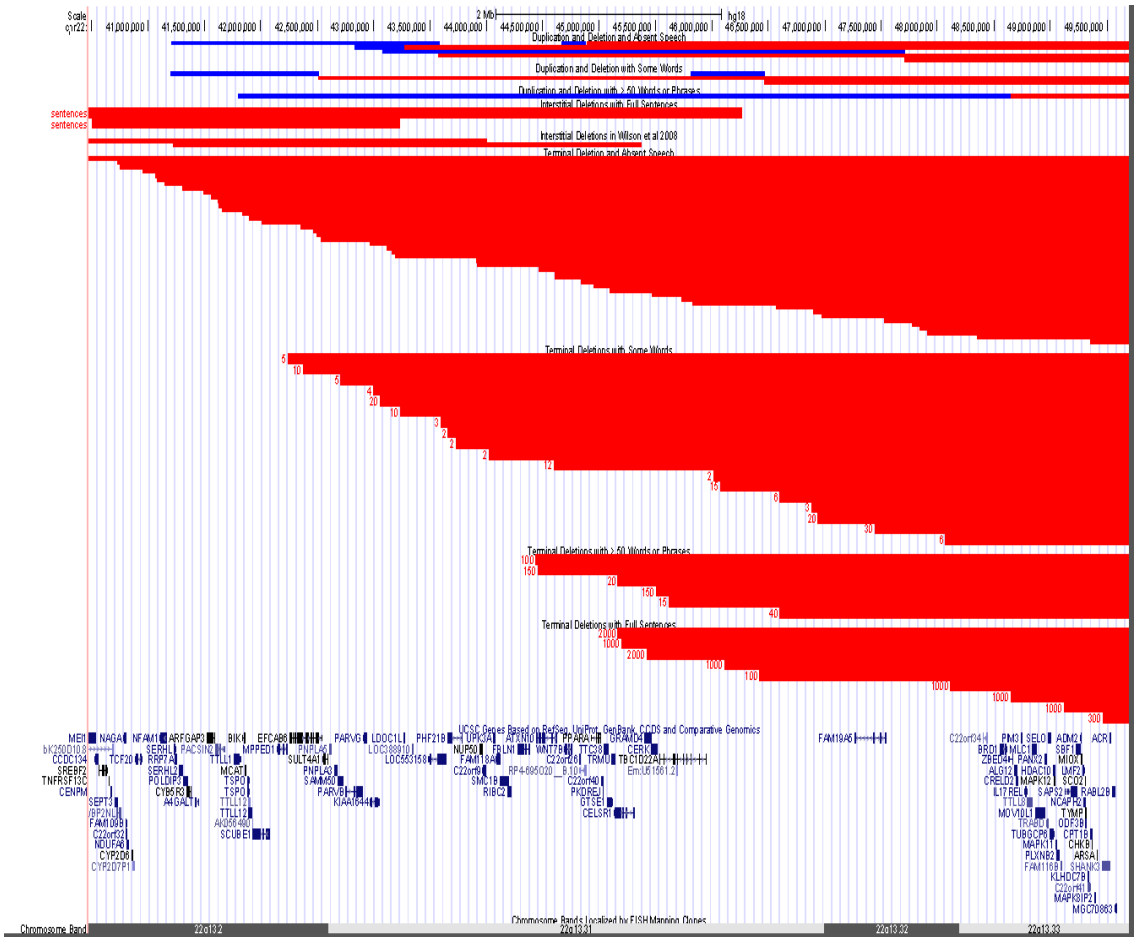


Figure 4.3. Deletions (red bars) and duplications (blue bars) for those with macrocephaly (>97th percentile) and those with normocephaly (3rd-97th percentile of head size for age and gender).

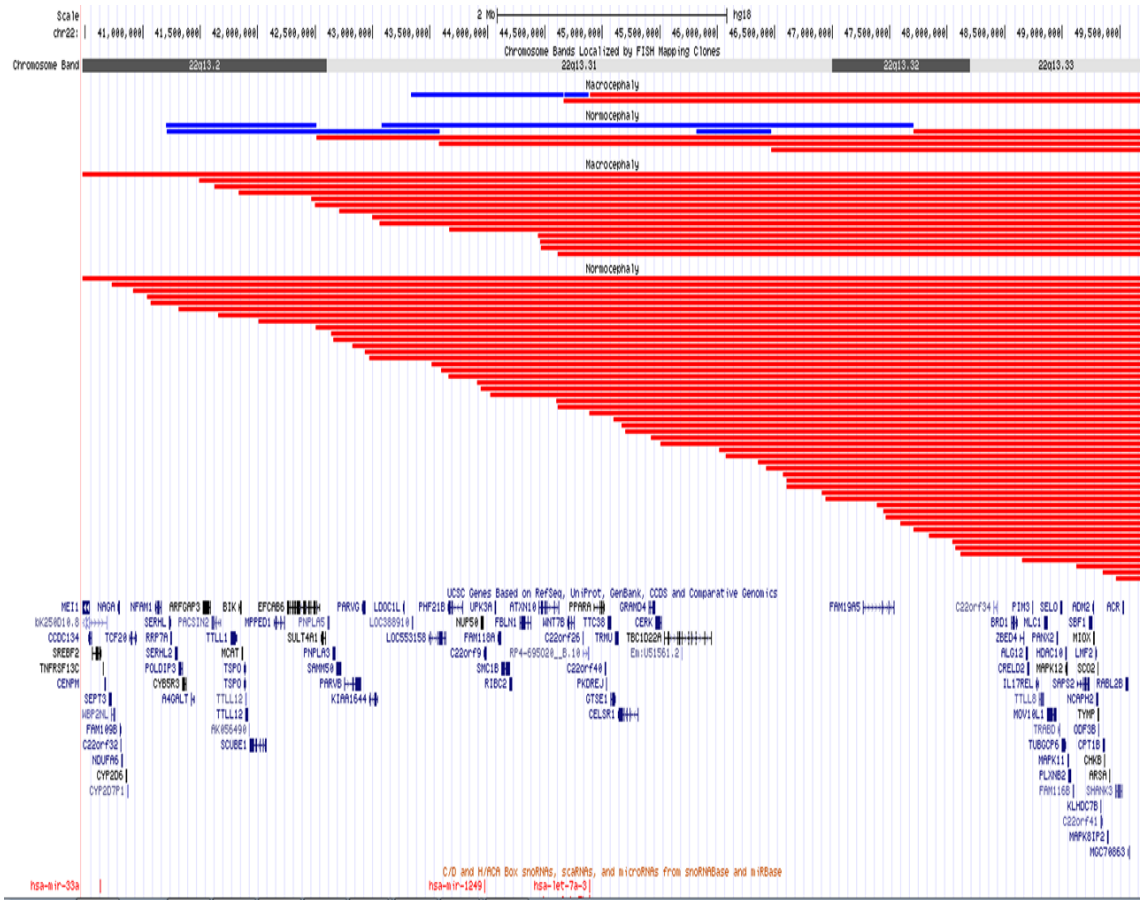
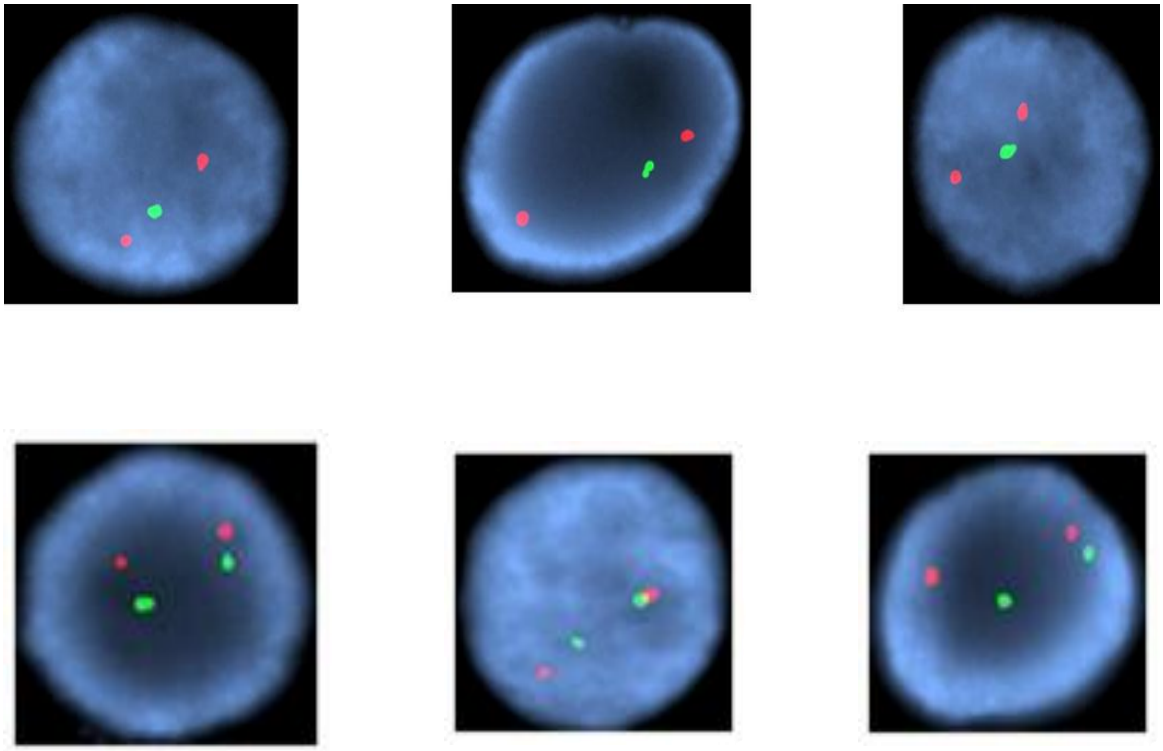


Figure 4.4. Fluorescence I*n Situ* Hybridization of Interphase Chromosomes from Buccal Cells. Red probes label 22q11.2 and green probes label 22q13.33.The top panel shows results from a PMS patient with a >7Mb deletion showing deletion of one terminus of chromosome 22. The bottom panel shows results from an unaffected control.



References

Anderlid BM, Schoumans J, Anneren G, Tapia-Paez I, Dumanski J, Blennow E, Nordenskjold M. 2002. FISH-mapping of a 100-kb terminal 22q13 deletion. Hum Genet 110(5):439-43.

Bangash MA, Park JM, Melnikova T, Wang D, Jeon SK, Lee D, Syeda S, Kim J, Kouser M, Schwartz J, and others. 2011. Enhanced polyubiquitination of Shank3 and NMDA receptor in a mouse model of autism. Cell 145(5):758-72.

Betancur C, Sakurai T, Buxbaum JD. 2009. The emerging role of synaptic cell-adhesion pathways in the pathogenesis of autism spectrum disorders. Trends Neurosci 32(7):402-1.

Boccuto L, Lauri M, Sarasua SM, Skinner CD, Buccella D, Dwivedi A, Orteschi D, Collins JS, Zollino M, Visconti P, and others. 2012. Prevalence of SHANK3 variants in patients with different subtypes of autism spectrum disorders. Eur J Hum Genet [in press].

Bonaglia MC, Giorda R, Ciccone R, Zuffardi O. 2010. Chromosome 22q13 rearrangements causing global developmental delay and autistic spectrum disorder. Knight SJL, editor. In: Genetics of mental retardation. Karger. 137 p.

Bonaglia MC, Giorda R, Borgatti R, Felisari G, Gagliardi C, Selicorni A, Zuffardi O. 2001. Disruption of the ProSAP2 gene in a t(12;22)(q24.1;q13.3) is associated with the 22q13.3 deletion syndrome. Am J Hum Genet 69(2):261-8.

Bonaglia MC, Giorda R, Mani E, Aceti G, Anderlid BM, Baroncini A, Pramparo T, Zuffardi O. 2006. Identification of a recurrent breakpoint within the SHANK3 gene in the 22q13.3 deletion syndrome. J Med Genet 43(10):822-8.

Bonaglia MC, Giorda R, Beri S, De Agostini C, Novara F, Fichera M, Grillo L, Galesi O, Vetro A, Ciccone R, and others. 2011. Molecular mechanisms generating and stabilizing terminal 22q13 deletions in 44 subjects with Phelan/McDermid syndrome. PLoS Genet 7(7):e1002173.

Bozdagi O, Sakurai T, Papapetrou D, Wang X, Dickstein DL, Takahashi N, Kajiwara Y, Yang M, Katz AM, Scattoni ML, and others. 2010. Haploinsufficiency of the autism-associated Shank3 gene leads to deficits in synaptic function, social interaction, and social communication. Mol Autism 1(1):1-15.

Buiting K. 2010. Prader-willi syndrome and angelman syndrome. Am J Med Genet C Semin Med Genet 154C(3):365-76.

Capurro MI, Xiang YY, Lobe C, Filmus J. 2005. Glypican-3 promotes the growth of hepatocellular carcinoma by stimulating canonical wnt signaling. Cancer Res 65(14):6245-54.

Chen CP, Hsu CY, Huang JK, Lee CC, Chen WL, Wang W. 2005. Prenatal diagnosis of partial trisomy 16q and distal 22q13 deletion associated with dolichocephaly and frontal bossing on second-trimester ultrasound. Prenat Diagn 25(10):964-6.

Choufani S., Shuman C. and Weksberg R. 2010. Beckwith–Wiedemann syndrome. American journal of medical genetics part C: Seminars in medical geneticsWiley Online Library. 343 p.

Delahaye A, Toutain A, Aboura A, Dupont C, Tabet AC, Benzacken B, Elion J, Verloes A, Pipiras E, Drunat S. 2009. Chromosome 22q13.3 deletion syndrome with a de novo interstitial 22q13.3 cryptic deletion disrupting SHANK3. Eur J Med Genet 52(5):328-32.

Dhar SU, del Gaudio D, German JR, Peters SU, Ou Z, Bader PI, Berg JS, Blazo M, Brown CW, Graham BH, and others. 2010. 22q13.3 deletion syndrome: Clinical and molecular analysis using array CGH. Am J Med Genet A 152A(3):573-81.

Douglas J, Hanks S, Temple IK, Davies S, Murray A, Upadhyaya M, Tomkins S, Hughes HE, Cole TRP, Rahman N. 2003. NSD1 mutations are the major cause of sotos syndrome and occur in some cases of weaver syndrome but are rare in other overgrowth phenotypes. Am J Hum Genet 72(1):132.

Durand CM, Betancur C, Boeckers TM, Bockmann J, Chaste P, Fauchereau F, Nygren G, Rastam M, Gillberg IC, Anckarsater H, and others. 2007. Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. Nat Genet 39(1):25-7.

Eggermann T. 2010. Russell–Silver syndrome. American journal of medical genetics part C: Seminars in medical geneticsWiley Online Library. 355 p.

Elison S, Stinton C, Howlin P. 2010. Health and social outcomes in adults with williams syndrome: Findings from cross-sectional and longitudinal cohorts. Res Dev Disabil 31(2):587-99.

Grabrucker AM, Schmeisser MJ, Schoen M, Boeckers TM. 2011. Postsynaptic ProSAP/shank scaffolds in the cross-hair of synaptopathies. Trends Cell Biol 21(10):594-603.

Greenwood Genetic Center. 2011. Growth references. Rollins JD, Tribble LM, Collins JS, and others, editors. 3rd ed. Greenville, South Carolina: Greenwood Genetic Center.

International Human Genome Sequencing Consortium. 2004. Finishing the euchromatic sequence of the human genome. Nature 431(7011):931-45.

Jeffries AR, Curran S, Elmslie F, Sharma A, Wenger S, Hummel M, Powell J. 2005. Molecular and phenotypic characterization of ring chromosome 22. Am J Med Genet A 137(2):139-47.

Kates WR, Antshel KM, Faraone SV, Fremont WP, Higgins AM, Shprintzen RJ, Botti JA, Kelchner L, McCarthy C. 2011. Neuroanatomic predictors to prodromal psychosis in velocardiofacial syndrome (22q11.2 deletion syndrome): A longitudinal study. Biol Psychiatry 69(10):945-52.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. Genome Res 12(6):996-1006.

Kuczmarski RJ, Ogden CL, Guo SS, Grummer-Strawn LM, Flegal KM, Mei Z, Wei R, Curtin LR, Roche AF, Johnson CL. 2002. 2000 CDC growth charts for the united states: Methods and development. Vital Health Stat 11 (246)(246):1-190.

Lindquist SG, Kirchhoff M, Lundsteen C, Pedersen W, Erichsen G, Kristensen K, Lillquist K, Smedegaard HH, Skov L, Tommerup N, and others. 2005. Further delineation of the 22q13 deletion syndrome. Clin Dysmorphol 14(2):55-60.

Luciani JJ, de Mas P, Depetris D, Mignon-Ravix C, Bottani A, Prieur M, Jonveaux P, Philippe A, Bourrouillou G, de Martinville B, and others. 2003. Telomeric 22q13 deletions resulting from rings, simple deletions, and translocations: Cytogenetic, molecular, and clinical analyses of 32 new observations. J Med Genet 40(9):690-6.

Luedi PP, Dietrich FS, Weidman JR, Bosko JM, Jirtle RL, Hartemink AJ. 2007. Computational and experimental identification of novel human imprinted genes. Genome Res 17(12):1723-30.

Maitz S, Gentilin B, Colli AM, Rizzuti T, Brandolisio E, Vetro A, Zuffardi O, Guerneri S, Lalatta F. 2008. Expanding the phenotype of 22q13.3 deletion: Report of a case detected prenatally. Prenat Diagn 28(10):978-80.

Manning MA, Cassidy SB, Clericuzio C, Cherry AM, Schwartz S, Hudgins L, Enns GM, Hoyme HE. 2004. Terminal 22q deletion syndrome: A newly recognized cause of speech and language disability in the autism spectrum. Pediatrics 114(2):451-7.

Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ, and others. 2010. Consensus statement: Chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. Am J Hum Genet 86(5):749-64.

Misceo D, Rodningen OK, Baroy T, Sorte H, Mellembakken JR, Stromme P, Fannemel M, Frengen E. 2011. A translocation between Xq21.33 and 22q13.33 causes an intragenic SHANK3 deletion in a woman with Phelan-McDermid syndrome and hypergonadotropic hypogonadism. Am J Med Genet A 155A(2):403-8.

Morison IM, Paton CJ, Cleverley SD. 2001. The imprinted gene and parent-of-origin effect database. Nucleic Acids Res 29(1):275-6.

Nicholls RD and Knepper JL. 2001. Genome organization, function, and imprinting in prader-willi and angelman syndromes. Annual Review of Genomics and Human Genetics 2(1):153-75.

Peca J, Feliciano C, Ting JT, Wang W, Wells MF, Venkatraman TN, Lascola CD, Fu Z, Feng G. 2011. Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. Nature 472:437-42.

Phelan K and McDermid HE. 2012. The 22q13.3 deletion syndrome (Phelan-McDermid syndrome). Mol Syndromol 2(3-5):186-201.

Phelan MC, Brown EF, Rogers RC. 2001. Prenatal diagnosis of mosaicism for deletion 22q13.3. Prenat Diagn 21(12):1100.

Phelan MC, Rogers RC, Saul RA, Stapleton GA, Sweet K, McDermid H, Shaw SR, Claytor J, Willis J, Kelly DP. 2001. 22q13 deletion syndrome. Am J Med Genet 101(2):91-9.

Philippe A, Boddaert N, Vaivre-Douret L, Robel L, Danon-Boileau L, Malan V, de Blois MC, Heron D, Colleaux L, Golse B, and others. 2008. Neurobehavioral profile and brain imaging study of the 22q13.3 deletion syndrome in childhood. Pediatrics 122(2):e376-82.

Pilia G, Hughes-Benzie RM, MacKenzie A, Baybayan P, Chen EY, Huber R, Neri G, Cao A, Forabosco A, Schlessinger D. 1996. Mutations in GPC3, a glypican gene, cause the simpson-golabi-behmel overgrowth syndrome. Nat Genet 12(3):241-7.

Roberts JE, Mankowski JB, Sideris J, Goldman BD, Hatton DD, Mirrett PL, Baranek GT, Reznick JS, Long AC, Bailey DB,Jr. 2009. Trajectories and predictors of the development of very young boys with fragile X syndrome. J Pediatr Psychol 34(8):827-36.

Rollins JD, Sarasua SM, Phelan MC, DuPont BR, Rogers RC, Collins JS. 2011. Growth in Phelan-McDermid syndrome. Am J Med Genet (155):2324-6.

Rollins JD, Collins JS, Holden KR. 2010. United states head circumference growth reference charts: Birth to 21 years. J Pediatr 156(6):907,13, 913.e1-2.

Sadler TW. 2004. Langman's medical embryology. Ninth Edition ed. Baltimore, Maryland: Lippincott Williams & Wilkins. 534 p.

Sarasua SM, Dwivedi A, Boccuto L, Rollins JD, Chen CF, Rogers RC, Phelan K, DuPont BR, Collins JS. 2011. Association between deletion size and important phenotypes expands the genomic region of interest in Phelan-McDermid syndrome (22q13 deletion syndrome). J Med Genet 48(11):761-6.

SAS Institute. 2009. SAS [computer program]. Version 9.2. Cary, North Carolina.

Verpelli C, Dvoretskova E, Vicidomini C, Rossi F, Chiappalone M, Schoen M, Di Stefano B, Mantegazza R, Broccoli V, Bockers TM, and others. 2011. Importance of Shank3 protein in regulating metabotropic glutamate receptor 5 (mGluR5) expression and signaling at synapses. J Biol Chem 286(40):34839-50.

Wang X, McCoy PA, Rodriguiz RM, Pan Y, Je HS, Roberts AC, Kim CJ, Berrios J, Colvin JS, Bousquet-Moore D, and others. 2011. Synaptic dysfunction and abnormal behaviors in mice lacking major isoforms of Shank3. Hum Mol Genet 20(15):3093-108.

Wilson HL, Wong AC, Shaw SR, Tse WY, Stapleton GA, Phelan MC, Hu S, Marshall J, McDermid HE. 2003. Molecular characterisation of the 22q13 deletion syndrome supports the role of haploinsufficiency of SHANK3/PROSAP2 in the major neurological symptoms. J Med Genet 40(8):575-84.

Wilson HL, Crolla JA, Walker D, Artifoni L, Dallapiccola B, Takano T, Vasudevan P, Huang S, Maloney V, Yobb T, and others. 2008. Interstitial 22q13 deletions: Genes other than SHANK3 have major effects on cognitive and language development. Eur J Hum Genet 16(11):1301-10.

World Health Organization. 2006. WHO child growth standards: Methods and development: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age.. Geneva: World Health Organization.

CONCLUSION

In this body of work, I conducted an epidemiologic and cytogenetic investigation of Phelan-McDermid syndrome (PMS) to better understand the molecular basis of the syndrome which ultimately one hopes can be used to direct therapeutic strategies.

I began my work with a review of the relevant literature (Chapter One) to describe the history and context of research in Phelan-McDermid syndrome. I reviewed the most common and troubling phenotypes, types of genomic rearrangements, and the current understanding of the role of *SHANK3*, a candidate gene for many of the neurologic features of PMS, and an actively investigated gene for autism spectrum disorders. I also reviewed known associations with the 22q13.3 region with other phenotypes.

For my research, I first established the rationale for expanding the genomic region of investigations beyond *SHANK3* by demonstrating significant differences in deletion sizes between those with and without a long list of phenotypes (Chapter Two). I pursued this finding by applying statistical methods in a novel application to identify deletion regions most associated with 22 phenotypes (Chapter Three). I combined reviews of the literature and the use of protein interaction networks to highlight candidate genes within these associated genomic regions. Finally, I conducted a cross-sectional analysis on the largest cohort to date (n=201), as well as a longitudinal assessment of 55 individuals to assess age related differences in phenotype (Chapter Four). I established that the prevalence of many phenotypes varies by age and that genotype-phenotype studies need to account for this in their analyses. In addition, I confirmed the strong association between deletion size and phenotype with a strong correlation between deleted

regions and varying degrees of speech impairment from absent speech to functional language. While the deletions are *de novo*, I established that parent of origin of the affected chromosome was more often paternal (73%). No differences were observed in parental age or transmission to a son or daughter. Maternal parent of origin was associated with seizures, a finding that needs confirmation through larger studies. Lastly, I demonstrated that this chromosomal anomaly can be detected in buccal specimens, a tissue derived from the same embryonic ectodermal layer as the central nervous system. The comprehensive analysis of cytogenetic abnormalities demonstrated the preponderance of simple terminal deletions in this syndrome, but also demonstrated the presence of interstitial deletions, deletions accompanied by duplications, and duplications without deletions.

In reviewing what I have learned about PMS, I see several areas of future research that follow directly from my findings. In Chapter Three, I identified genomic regions associated with neurologic and growth related phenotypes that warrant further investigation. Several genes, including *CYB3R5*, *PARVB*, and *WNT7B* are potential candidates for speech, growth, hypotonia, and intellectual ability. PARVB, in particular, appears to have many overlapping functions and even interactions with SHANK3 in the post synaptic density. As has been done with *SHANK3*, mouse models investigating the impact of haploinsuficiency of these genes would be of interest. Further, other investigators are currently using induced pluripotent stem cell techniques to transform fibroblast cells into neuronal cell lines to investigate gene expression. Better understanding of the interactions between the genes located along 22q13 would help to understand the impact of the deletions. Does deletion of one copy of *SHANK3* affect gene expression elsewhere on 22q13 or elsewhere in the genome? While it is known that telomere position effects may extend up to 100 kb, are these effects observed in the neuronal cells derived from patients with varying deletion breakpoints? Finally, it has not been established in humans

whether loss of one copy of *SHANK3* alone causes clinical features or rather, whether loss of one copy exposes recessive mutations in the remaining allele. Sequencing of the entire 22q13 haploid chromosome may provide information on this aspect of the syndrome. As more patients are identified, particularly those with deletions of only *SHANK3* or deletions within *SHANK3*, we may be able to parse the differences between haploinsufficiency and mutation in the gene and contributions from the remainder of 22q13 genes and noncoding elements. Finally, based upon my interactions with parents of affected children, the severe speech delay is of primary importance. A comprehensive speech evaluation of PMS patients would be useful in identifying the most important domains of speech affected.

APPENDICES

Appendix A

Effect of Population Stratification on the Identification of Significant Single-Nucleotide Polymorphisms in Genome-Wide Association Studies

[This paper was published in BMC Proceedings 2009: Sarasua S. M., Collins J. S., Williamson D. M., Satten G. A. and Allen A. S. 2009. Effect of population stratification on the identification of significant single-nucleotide polymorphisms in genome-wide association studies. BMC proceedings, 3(Suppl 7):S13]

Abstract

The North American Rheumatoid Arthritis Consortium case-control study collected case participants across the United States and control participants from New York. More than 500,000 single-nucleotide polymorphisms (SNPs) were genotyped in the sample of 2000 cases and controls. Careful adjustment for the confounding effect of population stratification must be conducted when analyzing these data; the variance inflation factor (VIF) without adjustment is 1.44. In the primary analyses of these data, a clustering algorithm in the program PLINK was used to reduce the VIF to 1.14, after which genomic control was used to control residual confounding. Here we use stratification scores to achieve a unified and coherent control for confounding. We used the first 10 principal components, calculated genome-wide using a set of 81,500 loci that had been selected to have low pair-wise linkage disequilibrium, as risk factors in a logistic model to calculate the stratification score. We then divided the data into five strata based on quantiles of the stratification score. The VIF of these stratified data is 1.04, indicating substantial control of stratification. However, after control for stratification, we find that there are no significant loci associated with rheumatoid arthritis outside of the HLA region. In particular, we find no evidence for association of *TRAF1-C5* with rheumatoid arthritis.

Background

Population stratification occurs when a population is composed of subpopulations that have varying allele frequencies. When these subpopulations also have differing baseline risks for a trait, then population stratification can lead to spurious allele-trait associations. To control for confounding by population stratification in case-control studies, statistical methods have been developed that use genetic markers to provide information on population structure. Among such methods are genomic control [1,2],structured association [3,4], and principal components [5,6].

A new statistical approach for controlling for population stratification in case-control studies was recently proposed by Epstein et al. [7]. This method involves modeling the odds of disease, given data on substructure-informative loci. For each participant the stratification score, which is that participant's estimated odds of disease calculated using his or her substructure-informative-loci data, is calculated using the diseaseodds model. Next, subjects are assigned to (typically five) strata defined by quantiles of the stratification score. Finally, the association between genotypes and the trait is ascertained using a stratified test. This approach is similar in spirit to the use of the propensity score to control for confounding in an observational study [8,9]. Epstein et al. showed that testing using the stratification score could control for confounding by population stratification in some situations where other methods fail [7]. The goal of this study was to assess the effect of controlling for population stratification in a genome-wide association study using the stratification score described above.

Methods

We analyzed the genome-wide association study data from the North American Rheumatoid Arthritis Consortium (NARAC) provided as Problem 1 for Genetic Analysis Workshop 16 [10,11]. This dataset is composed of cases from several sources: families, sib-pairs, sporadic cases, persons with long time disease, and new onset cases. Control participants were selected from a population-based cancer study in New York, frequency-matched to case participants for self-reported ethnic origin. Genotyping was performed with the Illumina Infinium HumanHap550 (version 1.0) platform (San Diego, CA) with 545,080 single-nucleotide polymorphisms (SNPs) for all case participants and 48% of control participants; 33% of controls were genotyped using HumanHap550 version 3.0 and 20% with the HumanHap300 and HumanHap240S arrays. The multiple sources of case and control participants in these data argues for careful examination of the role of population stratification in any associations found.

We followed the basic quality control procedures outlined by Fellay et al. [12], excluding data from SNPs that had extensive missingness (missingness > 5%), deviations from Hardy-Weinberg equilibrium (p-value < 0.001 in controls), and low minor allele frequency (<1%). After removing duplicated and contaminated samples, information was available for 2058 individuals (868 cases; 1190 controls). Of these, 568 individuals were male and 1490 were female. A total of 501,228 SNPs were used in subsequent analyses. The average genotyping rate for subjects was 0.994. PLINK [13] was used for data cleaning and to calculate both the unstratified and stratified Mantel-Haenszel allelic association test. p-Values of the max(T) were computed using both the Bonferroni method and 10,000 permutation datasets.

We used the stratification score of Epstein et al. to adjust our analyses for confounding due to population stratification [7]. The authors focus on adjusting association tests using a

limited number of ancestry-informative markers and, therefore, partial least squares (PLS) was used to estimate the stratification score. Here, no such marker panel was readily available; hence, we utilized markers from across the genome. Applying PLS to these data would likely result in substantial overfitting of the stratification score, leading to a loss of power [14,15]. In order to appropriately use this genome scale information, a different approach was needed. Thus we used a modified principal-component (PC) approach based on Fellay et al. [12] in place of PLS. Starting with the 501,228 SNPs that passed our quality control procedure, this modified PC approach captures the large-scale genetic variation in the data while minimizing the influence of a few regions high in linkage disequilibrium (LD) from dominating the PCs. This is accomplished by excluding SNPs from the PC analysis that reside in regions of known high LD and then further pruning the PC SNP set to minimize the LD between the remaining SNPs. After this pruning procedure 81,500 SNPs remained. Using the first few PCs, four individuals (D0009459, D0011466, D0012257, and D0012446) were found to be significant outliers, suggesting appreciable non-European ancestry. These individuals were excluded from subsequent analyses and, when the PC analysis was repeated, no further outliers were identified. The first 10 PCs were then used in a logistic model of disease to estimate each individual's stratification score--their predicted probability of being a case given the genomic information contained in their PCs. Five strata were then formed based on the quantiles of the stratification scores, for use in a stratified association analysis. We note that the computation demands presented by this procedure are quite minimal; it took approximately 30 minutes to generate the principal components and calculate the stratification score using a Linux workstation with two dual core 2.39-GHz opteron processors and 6 GB of RAM. We measured confounding by population stratification using the variance inflation factor (VIF), defined as the

median of the observed χ2 test statistics divided by the expected value of this median under the null hypothesis of no association of any SNP with rheumatoid arthritis (RA) [1].

Results

The unstratified analysis has a VIF of 1.44, while the VIF of the stratified analysis using the method of Epstein et al. was 1.034. In this context, it is worth noting that the identity-by-state (IBS) clustering approach to controlling for confounding by population stratification that is implemented in PLINK, and that was used by Plenge et al. [11], only attained a VIF of 1.14. For this reason, Plenge et al. also used genomic control [1,2] to control the residual confounding.

Aside from SNPs in the HLA region on chromosome 6, genome-wide we found no SNPs that were significantly associated with RA at the α=0.05 level (Figure 1). Interestingly, rs2900180 and rs3761847 on chromosome 9 in the *TRAF1-C5* gene (reported by Plenge et al. [11]) and rs2476601 on chromosome 1 in the *PTPN22* gene (reported by Begovich et al. [16]), were far from significant genome-wide (empirical adjusted $p = 1$, $p = 1$ and $p = 0.21$, respectively). To further investigate, we examined the five 2×3 tables for rs3761847 (Figure 2) and noted that there are only 12 cases in stratum 5. We then pooled strata 4 and 5 and recalculated the VIF to be 1.035. Pooling these strata did not increase the significance of these three SNPs (empirical adjusted $p = 1$, $p = 1$, and $p = 0.084$) and lack of statistical significance was not due to small strata size. The top three SNPs ranked by *p*-values, outside chromosome 6, were rs2476601 (chromosome 1, empirical *p*-value = 0.08), rs6596147 (chromosome 5, empirical *p*-value = 0.09), and rs1038848 (chromosome 8, empirical *p*-value = 0.21).

Conclusions

Differences in recruitment of cases and controls suggest that control of population stratification is crucial for a proper analysis of these data. This is confirmed by the large VIF for the unadjusted analysis. Stratification score analysis dramatically reduces the VIF, increasing confidence in any associations that are found. Interestingly, once we controlled for population stratification, we found no SNPs outside the HLA region on chromosome 6 that were associated with rheumatoid arthritis at the genome-wide significance level of $\alpha = 0.05$.

Like all stratified analyses, the stratification score approach will tend to lose power relative to a pooled (unadjusted) analysis when there is no confounding. Thus, our failure to replicate the associations found previously in these data may result from a loss of power from using the stratification score approach. However, the large VIF for these data makes confounding highly likely and, therefore, a competing explanation is that residual stratification in the primary analyses led to false associations. Further, Epstein et al. found that the stratification score approach had comparable power compared with other methods for control of population stratification [7]. Finally, we note that a spurious association may replicate if population stratification is not fully controlled in each analysis.

Figure 1 - Comparison of GWA Results for Unstratified, Stratified Analyses (5 Strata). Horizontal line is the Bonferroni threshold for genome-wide significance at α=0.05.

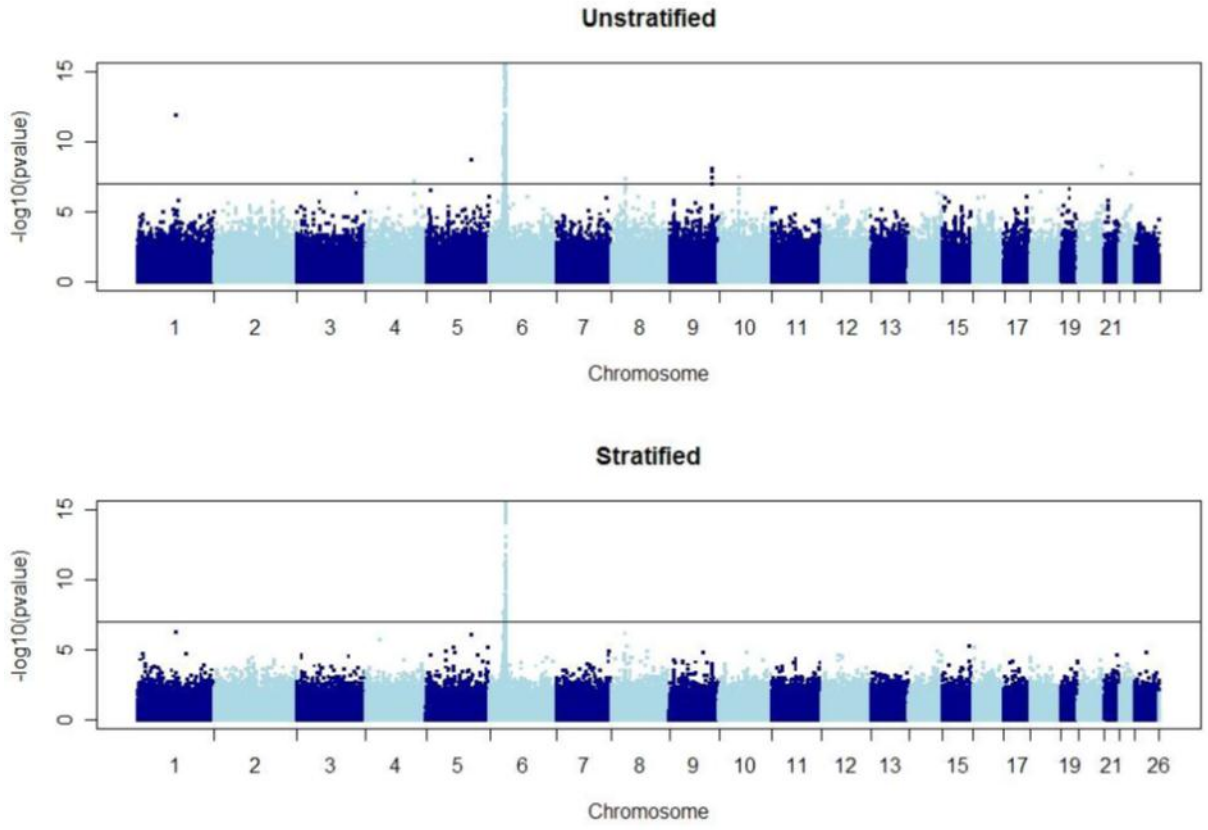


Figure 2 - Stratification score tables for association analysis of SNP rs3761847

Stratum 1	G=0	G=1	G=2	Total
Case	103	174	73	350
Control	20	32	8	60
Stratum 2				
Case	84	128	69	281
Control	50	57	22	129
Stratum 3				
Case	41	79	45	165
Control	88	121	38	247
Stratum 4				
Case	21	28	9	58
Control	138	171	44	353
Stratum 5				
Case	3	8	1	12
Control	161	178	60	399

List of Abbreviations

IBS: Identity-by-state

LD: Linkage disequilibrium

NARAC: North American Rheumatoid Arthritis Consortium

PC: Principal-component

PLS: Partial least squares

RA: Rheumatoid arthritis

SNP: Single-nucleotide polymorphism

VIF: Variance inflation factor

Authors' Contributions

SMS and ASA cleaned and analyzed the data. All authors participated in the design of the study and the writing of the manuscript.

Acknowledgements

The Genetic Analysis Workshops are supported by NIH grant R01 GM031575 from the National Institute of General Medical Sciences. SMS received a travel award from the Genetic Analysis Workshop. SMS and JSC were supported in part by a grant from the South Carolina Department of Disabilities and Special Needs. ASA acknowledges support from grants R01MH084680 and K25HL077663 from the National Institutes

of Health. The authors thank Min He for useful discussions and assistance performing computations.

References

1. Devlin B, Roeder K: Genomic control for association studies. *Biometrics* 1999, 55: 997-1004.

2. Devlin B, Roeder K, Wasserman L: Genomic control, a new approach to genetic-based association studies. *Theor Popul Biol* 2001, 60: 155-1663.

3. Pritchard JK, Rosenberg NA: Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet* 1999, 65:220-228.

4. Pritchard JK, Stephens M, Rosenberg NA, Donnelly P: Association mapping in structured populations. *Am J Hum Genet* 2000, 67: 170-181.

5. Chen H-S, Zhu X, Zhao H, Zhang S: Qualitative semiparametric test to detect genetic association in case-control design under structured population. *Ann Hum Genet* 2003, 67: 250-264.

6. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D: Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006, 38: 904-909.

7. Epstein MP, Allen AS, Satten GA: A Simple and improved correction for population stratification in case-control studies. *Am J Hum Genet* 2007, 80:921-930.

8. Rosenbaum PR, Rubin DB: The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983, 70: 41-55.

9. Rosenbaum PR, Rubin DB: Reducing bias in observational studies using subclassification on the propensity score. *J Am Stat Assoc* 1984, 79: 516-524.

10. Amos CI, Chen WV, Seldin MF, Remmers E, Criswell LA , Lee AT, Plenge RM, Kastner DL, Gregersen PS: Association analysis of rheumatoid arthritis data. *BMC Proc* 2009, 3(Suppl 7): S2.

11. Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B, Liew A, Khalili H, Chandrasekaran A, Davies LR, Li W, Tan AK, Bonnard C, Ong RT, Thalamuthu A, Pettersson S, Liu C, Tian C, Chen WV, Carulli JP, Beckman EM, Altshuler D, Alfredsson L, Criswell LA, Amos CI, Seldin MF, Kastner DL, Klareskog L, Gregersen PK: TRAF1-C5 as a risk locus for rheumatoid arthritis--a genomewide study. *N Engl J Med* 2007, 357: 1199-1209.

12. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, Zhang K, Gumbs C, Castagna A, Cossarizza A, Cozzi-Lepri A, De Luca A, Easterbrook P, Francioli P, Mallal S, Martinez-Picado J, Miro JM, Obel N, Smith JP, Wyniger J, Descombes P, Antonarakis SE, Letvin NL, McMichael AJ, Haynes BF, Telenti A, Goldstein DB: A whole-genome association study of major determinants for host control of HIV-1. *Science* 2007, 317: 944-947.

13. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC: PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet* 2007, 81: 559-575.

14. Lee S, Sullivan P, Zou, F, Wright F: Comment on a simple and improved correction for population stratification. *Am J Hum Genet* 2008, 82: 524-526.

15. Epstein MP, Allen AS, Satten GA: Response to Lee et al. *Am J Hum Genet* 2007, 82: 526-528.

16. Begovich AB, Carlton VEH, Honigberg LA: A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (*PTPN22*) is associated with rheumatoid arthritis. *Am J Hum Genet* 2004,75: 330-337.

Appendix B

List of Abbreviations

ASD Autism spectrum disorder

CNV Copy number variant

GWAS Genome wide association study

IBS Identity-by-state

ID Intellectual disability

Kb kilobase pairs

CGH Comparative genomic hybridization

LD Linkage disequilibrium

Mb million base pairs

NARAC North American Rheumatoid Arthritis Consortium

PC Principal-component

PLS Partial least squares

PMS Phelan-McDermid Syndrome

PSD Post synaptic density

RA Rheumatoid arthritis

SNP Single-nucleotide polymorphism

VIF Variance inflation factor