Pathogenic copy number variants (pCNVs) in individuals diagnosed on the autism spectrum disorder (ASD): A closer look at candidate genes

Allan Dale Baniaga*[1]

[1]Department of Integrated Sciences, The University of British Columbia, 2329 West Mall, Vancouver, British Columbia, Canada V6T 1Z4
Guest Article

ABSTRACT

Introduction: The genetic basis for autism spectrum disorders (ASDs) is well established and its heterogenetic nature provides us with substantial evidence for the many chromosomal aberrations associated with this complex disorder *(5)*. However, little is known about the genes that occupy the different chromosomal regions and the gene networks they participate in as they relate to phenotypes associated with ASDs. **Methods**: Here, the author reports pathogenic copy number variants (pCNVs) validated through array-comparative genomic hybridization (CGH) and the candidate genes found on these affected regions that may be implicated in the observed clinical phenotypes in 9 patients diagnosed with an ASD. Formal clinical assessments, which include a full physical examination, a medical history report, and a family history, were administered by a clinical geneticist unaware of the array-CGH results. **Results**: The author's findings suggest a number of genes involved in neurodevelopment as well as craniofacial and systemic features that may account for the observed phenotypes in the nine affected patients. **Discussion**: Among the candidate genes found, the *CYFIP1* gene, which is involved in maturation and maintenance of dendrites, the Gamma acid receptor family (*GABA*) which exhibit linkage disequilibrium with autistic disorders, and the *PHF8* and *WNK3* genes, which have been shown to be associated with X-linked mental retardation (XLMR), present the most interesting findings as they may account for most of the neurodevelopmental pathogenesis observed in the affected patients. Future studies need to be conducted in order to precisely determine the networks these genes participate in and how they are regulated to gain a deeper understanding in the roles they play in the clinical presentations of affected individuals with ASDs.

KEYWORDS

Autism, chromosome, disorder, genetics, interactions, networks, phenotype

*Corresponding author:
baniaga@interchange.ubc.ca
Received: 25 November 2010
Revised: 28 February 2011

INTRODUCTION

Autism Spectrum Disorders (ASDs) are neurodevelopmental disorders characterized by impairments ranging from communication, social interaction, and repetitive behavior impairments, and have a prevalence as high as 60 per 10,000 individuals by recent estimates *(1-4)*. The heritability of ASDs has also been well established over the years, particularly through the comparison monozygotic (MZ) and dizygotic (DZ) concordance rates that, when examined collectively with family studies, clearly point to an important genetic basis for autism *(5)*. Although current estimates strongly show that clinical genetic evaluation can identify

a specific etiology in up to 40% of individuals with an ASD *(6)*, little is understood in terms of the genotype-phenotype correlations for the vast majority of genetic abnormalities found in affected individuals. The current paper identifies possible candidate genes associated with particular pathogenic copy number variants (pCNVs) in subjects with ASDs, particularly in 3 pairs of individuals, 2 pairs of which had *de novo* changes, and a pair of brothers who inherited the same X-linked deletion from their unaffected mother. Three other individuals with ASDs encompassing impairments in communication, social interaction, and repetitive behavior were also identified to have unique copy number changes not shared with other affected individuals in the cohort. Array-CGH, FISH, and RT-qPCR data were gathered from experiments, as well as detailed phenotypic descriptions from formal clinical assessments by a professional Clinical Geneticist (M.E.S.L.) for these affected individuals. From these, candidate genes in each unbalanced region were compiled through the use of various databases so as to provide a genetic basis for their clinical phenotypes.

METHODS AND MATERIALS

IDENTIFICATION OF PATHOGENIC COPY NUMBER VARIANTS (pCNVs)

A total of nine subjects were identified to have pCNVs through various genetic testing methods. Low resolution array-CGH findings were confirmed by FISH and RT-qPCR methods, which were also used to refine breakpoints and determine the origins of the changes where parents were available for testing.

Individuals identified to harbour pCNVs found at genomic loci associated with other disorders fulfilled CNV criteria strongly suggesting the pathogenicity of their pCNVs *(7)*. Several criteria were used to distinguish benign CNVs (bCNVs) from potentially pathogenic ones (pCNVs), including: a *de novo* origin (or maternally inherited X-linked in male probands), the involvement of multiple genes not known to vary in databases with regards to their expression and function, the overlap with a gene or region that leads to a clinical phenotype when there is a dosage imbalance of the gene product, or if the affected region is >1Mb and overlaps well characterized genes *(7)*, *(8)*. This is in contrast to benign CNVs (bCNVs) which are found in at least two healthy individuals in independent studies found on the Database of Genomic Variants website.

ARRAY-CGH

PUREGENE DNA Isolation Kits (Gentra, Minneapolis, MN) were utilized to extract DNA from the peripheral blood of the nine subjects examined and were matched to normal male and female control DNAs (Promega, Madison, WI) as a reference. Both sample and reference DNAs were then subsequently hybridized using the 1-Mb BAC array *(9)* [Spectral Genomics, Houston, TX] through dye swap methods. Data analysis was conducted with Spectralware 2 software (Spectral Genomics) and clones bearing a significant gain or loss were identified through the use of the experimentally established values of 1.2 and 0.8, respectively, as cut-offs.

FISH

Deletions and duplications of BAC DNA clones identified by array-CGH were confirmed through FISH analyses *(10)*. A Zeiss Axioplan 2 fluorescence microscope and the MacProbe software (Applied Imaging, Santa Clara, CA) were then utilized to view the slides and capture the images, respectively.

REAL-TIME QUANTITATIVE PCR (RT-qPCR)

RT-qPCR *(11)* was employed to confirm the pCNVs. The ABI Prism 7900HT system (Applied Biosystems) using SYBR Green I detection was utilized to assess the RT-qPCR products. The primers used can be made available upon request.

PHENOTYPIC DATA

A spectrum of clinical characteristics modified from the De Vries scoring method were taken into account in identifying the phenotypes, and includes prenatal and postnatal growth abnormalities that are indicative of subtelomeric rearrangements not in exclusion of other characteristics. A spectrum of clinical and physical characteristics was noted for each participant by a professional Clinical Geneticist (M.E.S.L.) blinded to the array results. These include micro- and macrocephaly, prenatal and postnatal growth abnormalities, craniofacial dysmorphisms, systemic anomalies, and the presentation of medical co-morbidities such as seizures, intellectual disabilities (ID), and gastro-intestinal (GI) problems. Pregnancy and postnatal histories were also collected at the time of the appointment.

IDENTIFICATION OF CANDIDATE GENES

A general list of candidate genes were compiled for each proband through the Database of Genomic Variants website (http://projects.tcag.ca/variation/), and were subsequently narrowed down further using the SUSPECTS database (http://www.genetics.med.ed.ac.uk/suspects/). The major premise of SUSPECTS is that genes associated with complex traits will participate in the same gene networks and exhibit similar expression patterns, and the program achieves this by ranking candidate genes according to their possible involvement with the trait of interest. Only candidate genes that SUSPECTS was able to positively confirm as accounting for the observed clinical phenotypes were included in the final list of genes compiled in Figure 1.

Genotype-phenotype relationship information on each candidate gene identified by SUSPECTS was collected through the NCBI

SUBJECT(S)	GENOMIC REGION IN (BP)[2] (START/END)	PATHOGENIC CNV(S) AND CYTOBAND(S)[2]	ORIGIN	CANDIDATE GENE(S)[1]	GENE ONTOLOGY[3]
A, B	**A**=56,800,000/63,200,000 **B**=55,500,000/63,400,000	del 2p15 – 16.1	de novo	**1**.PEX13 **2**.OTX1	**1**.Peroxisome biogenesis disorders **2**.Brain and sensory organ development, inner ear morphogenesis
C, D	**1**. 19,570,792/20,341,734 **2**. 20,536,416/30,830,821	**1.**del 14q11.2 **2**.dup 15q11 – 12	both translocations	**1a**.NP **2a**.NIPA1/NIPA2 **b**.CYFIP1 **c**.NDN **d**.SNRPN **e**.UBE3A **f**.ATP10A **g**.GABRB3 **h**.GABRA5 **i**.GABRG3 **j**.APBA2 **k**.CHRFAM7A **l**.GREM1 **m**.SCG5	**1a**.Purine nucleoside phosphorylase Activity **2a**.Prader-willi/angelman syndrome **b**.Nervous system development **c**.Neuron development **d**.Rna splicing, prader-willi Syndrome **e**.Brain development, angelman Syndrome **f**.Angelman syndrome **g**,**h**,**i**.Gaba-a receptor activity **j**.Nervous system development **k**.Role in failure to thrive in infants **l**.Bone and nervous system Development **m**.Neuropeptide signaling, hormone Secretion regulation
E, F	53,970,960/54,326,640	del xp11.2 (highly skewed x-inactivation)	familial inherited	**1**.PHF8 **2**.WNK3	**1**.X-linked mental retardation **2**.Protein amino acid phosphorylation
G	**1**. 15,780,358/15,940,642 **2.** 9,334,790/11,738,791	**1**.del 3p24.3 – 25 **2**.del 5p15.2	**1.**de novo **2.**de novo	**2a**.SEMA5A **b**.CCT5 **c**.CTNND2	**2a**.Axonal guidance, nervous system Development **b**.Chaperon protein binding **c**.Neuron adhesion, synaptic Plasticity
H	72,200,000/73,767,523	dup 7q11.23	unknown origin	**1**.GTF2I	**1**.General transcription factor, Williams-beuren syndrome
I	5,910,725/6,063,460	dup 18p11.3	de novo	**1**.L3MBTL4	**1**.Cell adhesion, platelet activation, Integrin complex component

Fig. 1. Observed Clinical Phenotypes
[1]Confirmed through Database of Genomic Variants website and subsequently narrowed down with SUSPECTS database
[2]Affected region confirmed by RT-qPCR and/or FISH
[3]Information gathered collectively from the NCBI, iHOP, and metalife databases

website (http://www.ncbi.nlm.nih.gov/), specifically looking at the Entrez Gene record and the database of Genotype and Phenotype (dbGaP) entry (when available) for each gene. Further information regarding genotype-phenotype relationships were gathered from the iHOP website (http://www.ihop-net.org/UniPub/iHOP/) which compiles a list of all known functions, interactions, and diseases the gene of interest is associated with. The metalife database (http://www.phenomicdb.de/) was also utilized to complement the gene ontology information collected from the iHOP website and gives a summary of all the collective research done on a gene of interest as well as linking it to various phenotypes associated with the gene.

RESULTS

Figure 1 summarizes the pCNVs found and validated, the candidate genes for each subject pair sharing the same pathogenic CNV, as well as the candidate genes for individuals bearing unique pCNVs. All candidate genes identified and confirmed by the SUSPECTS database has been reported in full in Figure 1. Known gene ontologies for each candidate gene are also listed in Figure 1. A list of shared and unique phenotypes was also compiled in Figure 2 for each pair and for unique individuals bearing a particular pathogenic CNV. Figure 1 is a schematic diagram summarizing all of the candidate genes found on the various chromosomes.

SUBJECT(S)	GENOMIC REGION IN (BP)[2] (START/END) PATHOGENIC CNV(S) AND CYTOBAND(S)[2] ORIGIN	CANDIDATE GENE(S)[1] GENE ONTOLOGY[3]
A, B	Severe ID[1], microcephaly (<2%), craniofacial dysmorphisms (short forehead, high and broad nasal root), other systemic dysmorphisms (bilaterally tight heel cords, oral motor dysfunction), abnormal brain imaging	**A**=Prenatal growth retardation **B**=Seizure disorder, postnatal small stature (<5%)
C, D	Craniofacial dysmorphisms (strabismus, flat occiput, prominent alar cartilage), no other systemic dysmorphisms	**C**=Moderate ID, respiratory distress and poor suck and feeding difficulties (postnatal) **D**=Mild ID, seizure disorder, floppy infant (postnatal)
E, F	Moderate ID, craniofacial dysmorphisms (flat occiput, coarse asymmetric face [right side > left side fullness], micrognathia, unilateral cleft lip), other systemic dysmorphisms (pes planus, bone anomalies)	**E**=Prominent metopic suture, prominent finger pads **F**=Long slender fingers, macrocephaly at birth, mild hyperoptic refractive error
G	N/A	**G**=Moderate ID, seizure disorder, macrocephaly (>98%), postnatal large stature (>98%), craniofacial dysmorphisms (coarse facial features, frontal bossing, prominent supra-orbital ridge), other systemic dysmorphsims (prominent finger pads, bilateral tight heel cords, slight toe walking)
H	N/A	**H**=Moderate ID, craniofacial dysmorphisms (plagiocephaly, brachycephaly, prognathia), other systemic dysmorphisms (GI unusual dark stool colour, walks on heels, occasional enuresis), normal brain imaging
I	N/A	**I**=Mild ID, hightened blood pressure (during pregnancy), craniofacial dysmorphisms (mild bilateral epicanthal folds, ears bilaterally protuberant, slight malar flattening), other systemic dysmorphisms (slight metatarsus varus when walking)

Fig. 2. Observed Clinical Phenotypes
[1]ID = Intellectual Disability

SUBJECTS A AND B

The affected, unrelated pair share almost identical 2p15-16.1 deletions, both of *de novo* origin. The affected region has been further defined and validated to be from positions 56,800,000 to 63,200,000 for *subject A*, and from 55,500,000 to 63,400,000 for *subject B* on chromosome 2 confirmed through RT-qPCR and FISH methods.

Candidate genes for both subjects within the overlapping regions include *PEX13*, involved in Peroxisome Biogenesis Disorders, and *OTX1*, involved in brain and sensory organ development as well as inner ear morphogenesis.

SUBJECTS C AND D (FAMILY #1)

Subject D is the aunt of *subject C*, and both individuals share the same deletion at the 14q11.2 locus and duplication at the 15q11-12 locus. An unbalanced product of a reciprocal cryptic 14q/15q translocation represents this finding. A perfect overlap has been confirmed at the 14q11.2 and 15q11-12 regions between both subjects (19,570,792 to 20,341,734 and 20,536,416 to 30,830,821, respectively).

Of particular note, several candidate genes are involved in nervous system development, which include *CYFIP1*, *NDN*, *UBE3A*, *APBA2*, *GREM1*, and *SCG5*, all of which are found on the shared 15q11-12 site. Several candidate genes found on the same locus are also clearly involved in Angelman syndrome as well as Prader-Willi syndrome, and includes the *NIPA1*/*NIPA2*, *SNRPN*, *UBE3A*, and *ATP10A* genes.

SUBJECTS E AND F (FAMILY #2)

Subjects E and F are brothers who share the same Xp11.2 deletion that is familial inherited, and the precise genomic region was found and validated to be from 53,970,960 to 54,326,640 on the X chromosome. The mother was confirmed to have a relatively skewed X-inactivation, perhaps implying dosage imbalance problems in both siblings. Potential candidate genes in the affected region of both *subjects E and F* include *PHF8*, which is associated with X-linked mental retardation, as well as the *WNK3* gene, which is involved in protein amino acid phosphorylation.

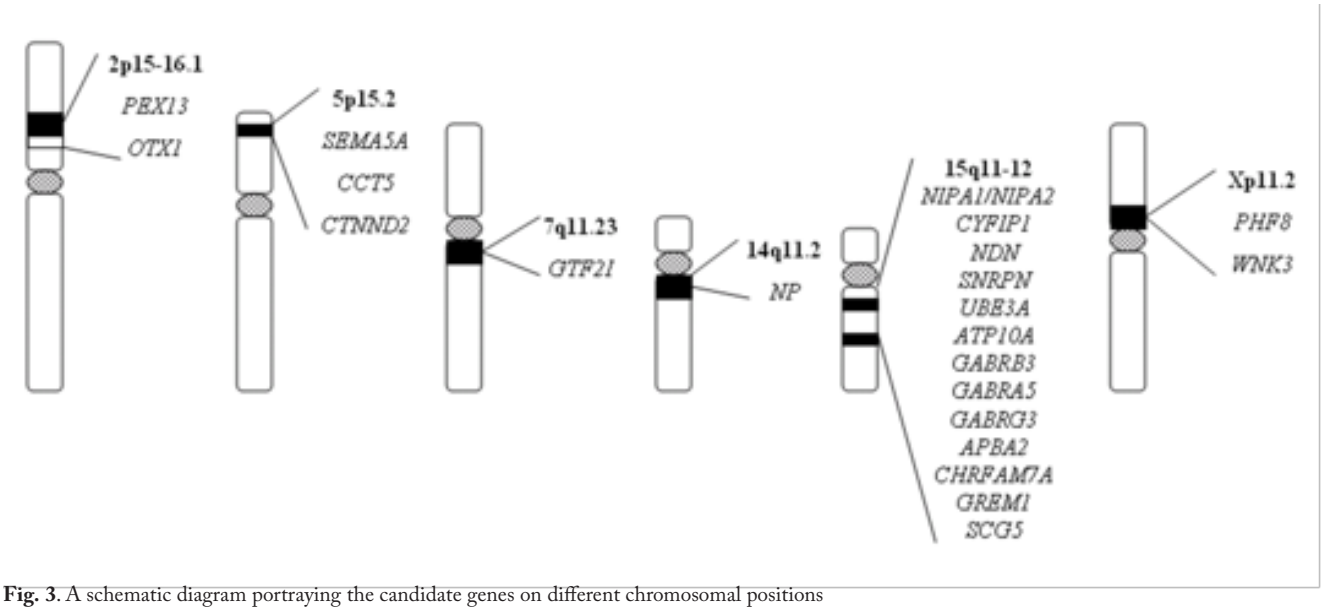

Fig. 3. A schematic diagram portraying the candidate genes on different chromosomal positions

SUBJECT G

The affected individual harbours two deletions at the 3p24.3-25 and 5p15.2 regions, with both regions of *de novo* origin. The precise affected region for the chromosome 3 deletion was confirmed to be from 15,780,358 to 15,940,642, and the chromosome 5 deletion is from 9,334,790 to 11,738,791, as validated by FISH and RT-qPCR.

In relation to the observed phenotypes, several candidate genes located on chromosome 5 that may play a role in the pathogenesis of classical autism include *SEMA5A*, which is involved in axonal guidance, and *CTNND2*, which is involved in neuron adhesion and synaptic plasticity. *CCT5* was also identified by SUSPECTS to be another candidate gene on the same chromosome, and is involved with chaperon protein binding in neurons.

SUBJECT H

Subject H is an individual with a unique 7q11.23 duplication of unknown origin that is not found in other subjects presenting with pathogenic CNVs. The affected genomic region was confirmed to be from 72,200,000 to 73,767,523 through FISH and RT-qPCR methods. Notably, the GTF2I gene was identified as a candidate gene by the SUSPECTS database, and is known to be as a general transcription factor, as well as being implicated in Williams-Beuren syndrome.

SUBJECT I

This affected individual was identified to have a duplication of *de novo* origin at the 18p11.3 region which was further refined to be from 5,910,725 to 6,063,460 on chromosome 18. *L3MBTL4* was the sole candidate gene identified that may account for the observed phenotypes, and is known to be involved in cell adhesion, platelet activation, as well as being a component of the integrin complex.

DISCUSSION

SUBJECTS A AND B: 2P15-16.1 DELETION

Several characteristics shared by both subjects include intellectual disability as well as poor oral motor skills (ie. speech) and poor muscle tone, which overlap with certain Peroxisome Biogenesis Disorders (PBD) phenotypes, thus making the *PEX13* gene on chromosome 2 a possible culprit for the observed phenotypes. Since both subjects are still alive, Refsum's disease is the most likely out of all the other PBDs as its prognosis is the most hopeful in terms of living past the early childhood years. However, both *subjects A* and *B* exhibited normal laboratory evidence of phytanic acid and long chain fatty acids, which are found in elevated levels in Refsum's affected individuals due to faulty enzymes during the alpha oxidation of phytanic acid and fatty acid oxidation *(10)*. Nonetheless, the possibility of *PEX13* contributing to part of the observed phenotypes, in particular those involved with neurodevelopment (in keeping with microcephaly in both subjects) and poor muscle tone, cannot be completely ruled out, as well as other candidate genes in the deleted region.

Another candidate gene that may be responsible for neurodevelopment as well as sensory organ formation is *OTX1*, a transcription factor that was recently found to be essential in cerebellum development *(12)*. The observed large ears in both subjects, relative to the microcephaly, may be in part due to the loss of

function of the *OTX1* gene. Furthermore, evidence suggests that the *OTX1* gene dictates the segregation of the saccule and the utricle during inner ear morphogenesis *(13)*, and thus its loss of function due to the deletion may perhaps be responsible for the hyperacusis observed in *subject A* as well as the bilateral sensorineural loss (mild to moderate in the left ear and slight to mild in the right ear) observed in *subject B*.

SUBJECTS C AND D (FAMILY 1): 14Q11.2 DELETION AND 15Q11-12 DUPLICATION

Both subjects present strikingly similar characteristics, most notably sharing several craniofacial dysmorphisms (strabismus, flat occiput, prominent alar cartilage) and intellectual disability (*subject C* has moderate ID while *subject D* has mild ID). These findings suggest an underlying genomic basis that may be responsible for the shared confirmed pathogenesis, and indeed, both a deletion and a duplication arising through a reciprocal cryptic 14q/15q translocation were found whose affected genomic areas were found to have perfect overlap between the two subjects (Fig. 1).

Candidate genes that may be involved with cephalic development were found in the duplicated 15q11-12 region and, as such, dosage effects may be at play here in terms of over-expression of a particular gene and/or altered regulation of a gene that may exert its effects on adjacent genes involved in the same gene network. The highly conserved *CYFIP1* gene may be one such gene. Through co-localization experiments, there is evidence that the products of the *CYFIP1* gene do indeed interact with *FMRPs* [Fragile X mental retardation proteins] *(14)*. Although the functions of the *CYFIP1* proteins are currently unknown, the extraction of *CYFIP1* proteins at the synaptosome of the distal portion of dendrites suggests that they also interact with the small GTPase *Rac1* where *CYFIP1* proteins also localize *(14)*. *Rac1* is known to be essential for dendritic spine maturation as well as maintenance *(15)*, and a duplication of the *CYFIP1* gene with which it interacts may have direct or indirect effects on the maturation and maintenance of these structures. One possibility is that an increased dosage of the *CYFIP1* gene product could directly alter expression levels of *Rac1* and other genes in the network involved in dendrite formation. This would have major implications for neurodevelopment and could thus be responsible for the shared ID and ASD observed in both subjects.

Several other candidate genes that may be responsible for the shared dysmorphisms include *UBE3A* and *APBA2*. Most notably, maternally derived duplications of the 15q11-13 region results in changes to *UBE3A* expression observed in autistic individuals, whereas the duplication is not present in normal individuals *(16)*. In addition, *GREM1* may also be the culprit gene for the observed bone fractures in *subject D* and the ligamentous laxity in *subject C*, as over-expression of the gene in transgenic mice analogous to that of a duplication event resulted in a 20-30% reduction in bone mineral density as well as formation of bone fractures *(17)*. It has been well established that deletions in the 15q11-13 region result in Prader-Willi syndrome (PWS) as well as Angelman syndrome (AS) *(18),* in which several of the phenotypes overlap with those present in both subjects. These include reduced fetal movements, respiration and feeding difficulties, strabismus, and intellectual disability *(19)*. *NIPA1/NIPA2*, *NDN*, *SNRPN*, *UBE3A*, *ATP10A*, and the Gamma acid receptor family (*GABA*) were identified and validated by SUSPECTS in the duplicated 15q11-12 region to be implicated in ASD in both subjects. There is evidence that a marker in the gene for the gamma aminobutyric acid receptor subunit of *GABRB3* was found to have linkage disequilibrium with autistic disorder, making this gene as well as other members of the gene family another prime candidate gene *(20)*. Furthermore, the role of benzodiazepine as a *GABA* receptor agonist in treating autistic phenotypes such as anxiety disorders and seizures suggest a potential role of the *GABA* gene family in the presentation of these phenotypes beyond the normal inhibitory neurotransmitter *GABA* function (20). Additional studies need to be conducted to precisely dissect the roles of the GABA gene family as well as others that exhibit linkage disequilibrium with autistic disorders.

SUBJECTS E AND F (FAMILY 2): XP11.2 DELETION

Both brothers share the same familial inherited Xp11.2 deletion, and a possible candidate gene that may account for the shared moderate ID is *PHF8*. The *PHF8* gene encodes a PHD finger protein that, when mutated through truncation mutation experiments, has been shown to cause X-linked mental retardation (XLMR) with or without cleft lip/cleft palate presentation *(21)*. The PHD finger protein has also been thought to regulate and modify chromatin structure *(22)*, which has major implications in terms of altered transcription levels in neurons and their maturation in individuals with mutations or deletions of the *PHF8* gene.

Another candidate gene that may account for the observed moderate ID in both subjects is *WNK3*, as it has been shown to occupy the critical linkage region on Xp11.2, and thus may also play a critical role in neurodevelopmental disorders such as XLMR *(23)*. However, future studies need to be conducted to determine whether the *WNK3* deletions could account for the difference in autism occurrences in comparison to the *PHF8* deletion cases, and whether deletion size differences between the *WNK3* and *PHF8* genes affect their interaction with neighboring genes.

SUBJECT G: 3P24.3-25 AND 5P15.2 DELETIONS

Subject G is an affected individual identified and validated to harbour unique deletions at the 3p24.3-25 and 5p15.2 regions. Several phenotypes unique to *subject G* include moderate ID, macrocephaly (>98%), postnatal large stature (>98%), and several craniofacial dysmorphisms (coarse facial features, frontal bossing). *SEMA5A*, a candidate gene in the 5p15.2 region identified by SUSPECTS, may account for the observed moderate ID and macrocephaly as it is known to be involved in axonal guidance and nervous system development *(24)*. Experimental evidence in the literature also shows that axonal development and formation of synapses may be affected by changes in *SEMA5A* expression *(25)*. Furthermore, deletions from the 5p band are also implicated in the Cri-du-chat phenotype, and haploinsufficiency of *SEMA5A* may be responsible for the intellectual disability (ID) in individuals exhibiting this phenotype *(26)*. A deletion at the 5p15.2 region may thus have a major impact on *SEMA5A* expression levels as not enough protein is made to maintain proper axonal development and synapse formation, possibly leading to the observed macrocephaly and moderate ID in *subject G*.

CCT5 is another candidate gene that was identified by SUSPECTS to be likely involved in neurodevelopment as the chaperon protein product of *CCT5* was found to have a role in polymerization of cytoskeletal proteins and their structural maintenance in neurons *(27)*. Deletions in the 5p15.2 region would thus have profound effects on proper neurodevelopment due to a lack of *CCT5* chaperone proteins essential for proper neuron functioning. The *CTNND2* gene, which is also found on 5p15.2, was also identified as a possible candidate for the shared macrocephaly and moderate ID phenotypes. Mutational experiments in the literature suggest a specialized role for the *CTNND2* protein as deletions in the gene result in problems in synaptic plasticity in neurons which may lead to learning deficits *(28)*. Furthermore, a strong correlation between a hemizygous deletion of the *CTNND2* gene and severe mental retardation in individuals with Cri-du-chat syndrome (*CDCS*) was found *(29)*, further underlying the critical role of *CTNND2* in intellectual disability in subjects with *CDCS* harbouring a deletion in the 5p15.2 region. Moreover, delta catenin (the protein product of *CTNND2*) was also found to co-bind with kaiso to the promoter sites of rapsyn, a synapse protein necessary for segregating acetylcholine receptors at the neuromuscular area *(30)*.

Deletion of the *CTNND2* gene on 5p15.2 would thus have adverse effects on proper rapsyn functioning, which may in turn contribute to the seizures, delays in motor milestones, as well as failure to thrive observed in the affected individual.

SUBJECT H: 7Q11.23 DUPLICATION

Subject H is an isolated case that does not share any pathogenic CNVs or cytogenetic bands with the other cases here but, nonetheless, the subject's candidate genes have been reported here, as they present distinctive findings. A unique 7q11.23 duplication of unknown origin was found, and *GTF2I* was the sole candidate gene identified by SUSPECTS that may account for the unique phenotypes observed such as plagiocephaly, brachycephaly, and prognathia linked to this affected genomic region. Structural features of 7q11.23 render this region susceptible to genomic rearrangement and deletions, yielding various CNVs that are also involved with Williams-Beuren syndrome (WBS) *(31)*. Several characteristics are shared by the subject with WBS including intellectual disability, hyperacusis, and genito-urinary problems (subject has a history of enuresis) *(32)*. Furthermore, *GTF2I* was also shown to be involved in tooth development at the bud and early bell stage *(33)*, and thus may also account for the carious as well as early loss of teeth in the subject. Moreover, hemizygosity of *GTF2I* was found to be sufficient to account for a number of features associated with WBS, including visuospatial deficits *(34)*, which may contribute to the astigmatism and myopia observed in *subject H*.

SUBJECT I: 18P11.3 DUPLICATION

The affected subject harbours a unique duplication at the 18p11.3 region of *de novo* origin and presents several phenotypes that may be associated with the duplication. Most notably, the candidate *L3MBTL4* gene is known to be involved in platelet activation, and a dosage imbalance arising through a duplication may have direct or indirect consequences with regards to the observed heightened blood pressure during pregnancy. The role of the *L3MBTL4* gene in cell adhesion, in particular as a component of the integrin complex (which is known to be involved in mediating various intracellular signals), may indeed account for the craniofacial and systemic dysmorphisms as well as the mild ID observed in *subject I*. The implications of dosage effects through duplication of the 18p11.3 region need to be further investigated as well as the specific roles of *L3MBTL4*, as research on this region and its genes is limited to make any conclusive statements regarding genotype-phenotype relationships at this time.

CONCLUSION

By identifying and compiling a list of candidate genes on various chromosomes, changes occurring at the gene level affecting phenotype at the organismal level are better understood. This provides us with an overall map of the candidate genes and their respective protein products that may account for the presentation of ASD phenotypes. Autism spectrum disorders present an especially daunting task, as their varied phenotypes between different indi-

viduals suggests a multitude of genes that may interact with other candidate genes involved in the same or different gene networks. Every effort was made to account for the clinical manifestations of the patients presenting with an ASD with their candidate genes through the extensive use of databases in order to hypothesize and explore their genotype-phenotype relationships. In addition to this, the SUSPECTS database utilized in narrowing the list of candidate genes in this paper has limitations in that the matches for the genes are weighed differently for different types of matches. The weights assigned are arbitrary, and so there is concern with regards to how consistent and accurate the program ranks the candidate genes. Additional experiments need to be conducted in the future so as to uncover the functions of these candidate genes, how their expression is regulated, and what gene networks they participate in in order to fully validate or reject their involvement in the observed clinical phenotypes compiled in this paper. The author urges for further research on candidate genes identified in this paper involved in neurodevelopment, which include the Gamma acid receptor family (*GABA*), the *CYFIP1* gene, and the the *PHF8* and *WNK3* genes. Possible experiments that can be conducted in order to further investigate their functions include gain and loss of function experiments to explore the effects of protein dosage on phenotype expression, mutational experiments (such as recessive mutations in hemizygosity), stoichiometry and amplification experiments, as well as looking at the three-dimensional nuclear structure of the chromosomal regions involved. Only through understanding the finer details at the gene level through further experimental research can we then unravel the genetic bases for some of the phenotypes associated with classical autism.

ACKNOWLEDGEMENTS

The author would like to thank Noemie Riendeau and Ying Qiao at the Child and Family Research Institute (CFRI) wet lab for all of their help in assisting with the CNV data, as well as for their insightful comments. A warm thank-you also goes out to Priscilla Carrion and Lindsay Swinton at the dry lab for all of their help with the clinical files, all of which the author is grateful for having access to. And last, but certainly not least, the author would like to express his sincerest gratitude towards his mentor, Dr. Lewis, for allowing the author to undertake this directed studies project, as well as for all of her guidance without which this project would have never been possible without.

REFERENCES

1. G. Baird, T. Charman, S. Baron-Cohen, A. Cox, J. Swettenham *et al. J. Am. Acad. Child Adolesc.* **39**, 694–702 (2000).
2. J. Bertrand, A. Mars, C. Boyle, C. Bove, M. Yeargin-Allsopp *et al. Pediatrics*. **108**, 1155–1161 (2001).
3. S. Chakrabarti and E. Fombonne, *JAMA*. **285**, 3093–3099 (2001).
4. S. Chakrabarti and E. Fombonne, *Am. J. Psychiatry*. **162**, 1133–1141 (2005).
5. C.J. Newschaffer et *al. Annu. Rev. Public Health*. **28**, 235–258 (2007).
6. G.B. Schaefer and N.J. Mendelsohn, *Genetics in Medicine*. **10**(1), 4-12 (2008).
7. Y.S. Fan, P. Jayakar, H. Zhu, D. Barbouth, S. Sacharow *et al. Human mutation*. **28**(11), 1124-1132 (2007).
8. C. Lee, A.J. Iafrate, A.R. Brothman, *Nat. Genet*. **39**, S48-54 (2007).
9. C. Tyson, C. Harvard, R. Locker, J.M. Friedman, S. Langlois *et al. Am. J. Med. Genet. A*. **139**(3), 173-185 (2005).
10. E. Rajcan-Separovic, C. Harvard, X. Liu, B. McGillivray, J.G. Hall *et al. J. Med. Genet*. **44**(4), 269-276 (2007).
11. Y. Qiao, X. Liu, C. Harvard, S.L. Nolin, W.T. Brown *et al. BMC genomics*. **8**, 167 (2007).
12. T. de Haas *et al. J. Neuropathol Exp. Neurol*. **65**(2), 176-186 (2006).
13. K.W. Beisel *et al. Journal of Vestibular Research*. **15**(5-6), 225-241 (2005).
14. A. Schenck *et al. PNAS*. **98**(15), 8844-8849 (2001).
15. A.Y. Nakayama, M.B. Harms, L. Luo, *The Journal of Neuroscience*. **20**(14) 5329-5338 (2000).
16. J. Veenstra-VanderWeele, D. Gonen, B.L. Leventhal, E.H. Cook Jr, *Mol. Psychiatry*. **4**, 64-67 (1999).
17. E. Gazzerro *et al. Endocrinology*. **146**(2), 655–665 (2005).
18. B. Dittrich *et al. Hum. Genet*. **90**, 313-315 (1992).
19. V.A. Holm et *al. Pediatrics*. **91**, 398-402 (1993).
20. E.H. Cook Jr. *et al. Am. J. Hum. Genet*. **62**, 1077–1083 (1998).
21. F. Laumonnier *et al. J. Med. Genet*. **42**, 780–786 (2005).
22. L.R. Jensen *et al. Am. J. Hum. Genet*. **76**, 227–236 (2005).
23. H.H. Ropers *et al. Trends Genet*. **19**, 316-320 (2003).
24. R.H. Adams, H. Betz, A.W. Puschel, *Mech. Dev*. **57**, 33–45 (1996).
25. L. Jones, G. Lopez-Bendito, P. Gruss, A. Stoykova, Z. Molnar, *Development*. **129**: 5041–5052 (2002).
26. A.D. Simmons, A.W. Puschel, J.D. McPherson, J. Overhauser, M. Lovett, *Biochem. Biophys. Res. Commun*. **242**, 685–691 (1998).
27. G.J. Bourke, W. El Alami, S.J. Wilson, A. Yuan, A. Roobol *et al. J. Neurosci. Res*. **68**: 29–35 (2002).
28. I. Israely *et al. Curr. Biol*. **14**, 1657–1663 (2004).
29. M. Medina, R.C. Marinescu, J. Overhauser, K.S. Kosik, *Genomics*. **63**(2), 157–164 (2000).
30. M. Rodova, K.F. Kelly, M. VanSaun, J.M. Daniel, M.J. Werle, *Mol. Cell Biol*. **24**(16), 7188–7196 (2004).
31. I. Cuscó, R. Corominas, M. Bayés *et al. Genome Res*. **18**(5), 683-694 (2008).
32. Z. Sammour, C. Gomes, R. Duarte, F. Trigo-Rocha, M. Srougi, *The Journal of Urology.* **175**(4), 1472-1476 (2006).
33. A. Ohazama and P.T. Sharpe, *Developmental Dynamics*. **236**(10), 2884–2888 (2007).
34. L. Edelmann *et al. Journal of Medical Genetics*. **44**, 136-143 (2007).