Sub-microscopic Chromosomal Imbalances in Idiopathic Autism Spectrum Disorder (ASD)

by

Maryam Koochek

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

The Faculty of Graduate Studies

(Medical Genetics)

The University Of British Columbia
(Vancouver)

November 2008

© Maryam Koochek, 2008
Abstract

Autism spectrum disorders (ASDs) are a group of neurodevelopmental conditions characterized by clinical variability, genetic heterogeneity and a male to female ratio of 4:1. The frequency of the disorders is approximately 1/150 (US Centre for Disease Control) and heritability in idiopathic autism is estimated at over 90%. Several genome-wide linkage analyses support the hypothesis of complex inheritance with involvement of many genes of small effect. Further, numerical and structural chromosomal anomalies involving almost all of the chromosomes have been previously reported. One class of chromosomal abnormalities that may have a significant involvement in the aetiology of ASDs include sub-microscopic genomic deletions and duplications also known as DNA Copy Number Variations (CNV). In this study, nineteen probands with a confirmed diagnosis of nonsyndromic ASD and additional complex phenotypic features (dysmorphisms +/- intellectual disability (ID)) with normal karyotype/Fragile X screening were evaluated for submicroscopic genomic imbalances using a commercially available 1Mb BAC microarray platform (Spectral Genomics). Additionally, two probands with a confirmed diagnosis of ASD and a previously identified chromosomal abnormality were used as positive controls. Probands from three families were determined to have a pathogenic or potentially pathogenic CNV following confirmation of their array findings by a secondary independent method such as FISH and/or qPCR. In the first multiplex (MPX), multigenerational family two probands had array profiles which suggested a gain of proximal 15q11-13, loss of 14q11 and gain of 6q22. Extensive FISH analysis for these probands and their relatives revealed a balanced cryptic translocation t(14;15)(q11.2;q13.3) segregating within the maternal lineage that was
expressed in an unbalanced form in the affected individuals. In two other families, the probands had single clone gains at 18p11.3 (de novo) and 21q22.12 (maternal) respectively. Overall 3/19 (16%) of families with ASD probands with a normal karyotype were found to harbour an autism-specific CNV or cryptic chromosomal rearrangement that plays a pathogenic or potentially pathogenic role in the aetiology of ASD.
# Table of Contents

Abstract ................................................................................................................. ii

Table of Contents ................................................................................................. iv

List of Tables ........................................................................................................... vi

List of Figures .......................................................................................................... vii

Abbreviations .......................................................................................................... ix

Acknowledgements ................................................................................................. xi

Dedications ................................................................................................................. xii

## Chapter 1. Introduction

1.1 Autism Spectrum disorders (ASD) ........................................................................ 1
1.2 Endophenotypes: On the Path to Reducing Phenotypic Heterogeneity ............... 3
1.3 Genetics of ASD: An Overview .......................................................................... 4
  1.3.1 Strategies for localizing ASD culprit genes ................................................. 5
1.4 Cytogenetic Regions of Interest in ASD (CROIs) ............................................... 10
  1.4.1 Chromosome 7 .......................................................................................... 10
  1.4.2 Duplications of Proximal 15q ..................................................................... 11
  1.4.3 Chromosome 22 ......................................................................................... 14
  1.4.4 Subtelomeric Regions ................................................................................ 15
1.5 Specific Aims ....................................................................................................... 24

## Chapter 2. Materials and Methods

2.1 Patient Selection ............................................................................................... 25
2.2 DNA isolation .................................................................................................. 29
2.3 DNA Quantification and Quality Control ......................................................... 29
2.4 Array Comparative Genomic Hybridization (array-CGH) .................................. 31
  2.4.1 The Principle of array-CGH ....................................................................... 31
  2.4.2 Array-CGH Protocol ................................................................................ 32
2.5 Fluorescence *in situ* Hybridization (FISH) ..................................................... 38
  2.5.1 FISH using Nick Translation ..................................................................... 38
  2.5.2 Metaphase Slide Preparation ..................................................................... 40
  2.5.3 Hybridization ............................................................................................ 40
  2.5.4 Post Hybridization Washes ....................................................................... 41
  2.5.5 Slide Evaluation and Image Capturing ....................................................... 41
2.6 Dual (Multi-coloured) FISH ............................................................................ 41
2.7 Quantitative PCR (qPCR) ................................................................................ 42
List of Tables

Table 1(a). Summary of findings from five studies searching for ASD associated CNVs using array genomic hybridization ..........................................................20

Table 1(b). Summary of recurrent and/or overlapping ASD associated CNV regions reported by two or more studies.................................................................22

Table 2. Clinical criteria for selecting patients for Array-CGH studies (adapted from de Vries, 2001). Only probands with a score of 3 or greater were selected for the microarray study..........................................................28

Table 3. Summary of clinical features for probands that were found to have a confirmed micro-deletion/duplication based on their array-CGH profile.............................51

Table 4. Summary of ASD cases with confirmed CNVs of potentially pathogenic and uncertain clinical relevance.................................................................52

Table 5. Summary of the main clinical findings observed in II:3 (proband 1) and III:1 (proband 2)..........................................................................................61
List of Figures

Figure 1. Positive control profiles for chromosome 22 in a proband with a known 22q deletion (a) and for chromosome 15 in a proband with a known duplication of proximal 15q (b) ..................................................46

Figure 2. Three clones were found to be duplicated in proband 3; (a) a single clone duplication at 18p11.3 (array values: 1.40 and 0.58), (b) a single clone duplicated at 20p12 (array values: 1.28 and 0.69) and (c) a single clone duplicated at Xp22.32 (array values 1.30 and 0.63) .............................................49

Figure 3. In proband 4 a single clone corresponding to 21q22 was observed to be duplicated (array values: 1.49 and 0.69) ..........................................................50

Figure 4. A multi-generation pedigree illustrates the incidence of autism in this family. Array-CGH studies were performed for probands 1 and 2 (II:3 and, III:1) as well as for II:2 (unaffected). I:2 and II:2 were found to be carriers of a cryptic balanced translocation t(14;15)(q11.2;q13.3). Both I:2 and II:2 transmitted the translocation in an unbalanced manner, leading to the gain of proximal 15q in probands 1 and 2 ........................................60

Figure 5. The ratio plots from array CGH for chromosomes 15, 14 and 6 for proband 1 (II:3) and proband 2 (III:1) showed the exact same chromosomal changes after array analysis (only profile from proband 1 illustrated here). (a) A seven clone duplication (RP11-80H14, RP11-385H1, RP11-339C21, AC090696.6, RP11-368G21, CTB-21M6 and RP11-420B6) on 15q11-13 resulting in a gain of ~10Mb; (b) A two clone deletion (RP11-98N22 and RP11-89F2) on 14q11 was ~0.7Mb in size and (c) A two clone duplication (RP11-80B14 and RP3-329N18) on 6q22 resulting in a gain of ~0.4Mb .................................................................61

Figure 6. FISH confirmation of 15q gain and 14q loss in the probands. The gain of 15q was confirmed using the proximal and distal clones on either side of the duplicated region: (a) signals on three acrocentric chromosomes were observed for the proximal clone RP11-80H14 in both probands. (b) For the most distal probe RP11-420B6, one of the three signals was less intense, suggesting that the breakpoint of the duplication resided within this clone. (c) RP11-89F2 was used to confirm the deletion on chromosome 14. Only one signal was present in all 10 metaphase spreads analyzed for both probands ..................................................66

Figure 7. FISH analysis for subjects I:2 and II:2 (the mothers of the two probands) is illustrated. Probe RP11-80H14 (a) showed two signals on two acrocentric chromosomes. To determine the identity of the chromosomes, reverse G banding (b) was performed on the same metaphase
spread. Upon examination of the banding pattern, it was determined that the 15q specific probe mapped to one normal 15 and a derivative 15 (der(15)t(14;15)) which had a chromosome 14 banding pattern since it consists of centromere and proximal end of chromosome 15 and most of the long arm of chromosome 14 (all other acrocentric chromosomes were also identified).

Figure 8. Presence of a cryptic balanced translocation was confirmed through dual FISH in I:2 and II:2. The centromere 15 probe (green) and a probe mapping to the distal end of chromosome 14 (red) were used. The latter probe hybridized to two acrocentric chromosomes, one of which was marked by the centromere 15 probe, indicating the translocation of the 14q arm on to the centromere and proximal end of 15 (der15).

Figure 9. Illustration of the mechanism by which the presence of two derivative chromosomes (der(14) and der(15)) in the translocation carriers has lead to the gain of proximal 15q and loss of proximal 14q in the probands.
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABI</td>
<td>Applied Biosystems</td>
</tr>
<tr>
<td>Abs</td>
<td>Absorbance</td>
</tr>
<tr>
<td>AD</td>
<td>Autistic Disorder</td>
</tr>
<tr>
<td>AS</td>
<td>Angelman Syndrome</td>
</tr>
<tr>
<td>ADHD</td>
<td>Attention-Deficit Hyperactivity Disorder</td>
</tr>
<tr>
<td>ADI-R</td>
<td>Autism Diagnostic Interview-Revised</td>
</tr>
<tr>
<td>ADOS-G</td>
<td>Autism Diagnostic Observation Generic</td>
</tr>
<tr>
<td>AGPC</td>
<td>Autism Genome Project Consortium</td>
</tr>
<tr>
<td>ASD</td>
<td>Autism Spectrum Disorder</td>
</tr>
<tr>
<td>ASD-CARC</td>
<td>Autism Spectrum Disorders-Canadian American Research Consortium</td>
</tr>
<tr>
<td>BAC</td>
<td>Bacteria Artificial Chromosomes</td>
</tr>
<tr>
<td>CNS</td>
<td>Central Nervous System</td>
</tr>
<tr>
<td>CDD</td>
<td>Childhood Disintegration Disorder</td>
</tr>
<tr>
<td>CGH</td>
<td>Comparative Genomic Hybridization</td>
</tr>
<tr>
<td>CNV</td>
<td>Copy Number Variant</td>
</tr>
<tr>
<td>Cy3</td>
<td>Cyanine 3-dCTP</td>
</tr>
<tr>
<td>Cy5</td>
<td>Cyanine 5-dCTP</td>
</tr>
<tr>
<td>DAPI</td>
<td>4',6-diamidino-2-phenylindole</td>
</tr>
<tr>
<td>ddH2O</td>
<td>Double Distilled water</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxy thymidine triphosphate</td>
</tr>
<tr>
<td>dTTP</td>
<td>Deoxyribonucleotide triphosphate</td>
</tr>
<tr>
<td>dUTP</td>
<td>Dioxigenin-labeled deoxyuridine triphosphate</td>
</tr>
<tr>
<td>DZ</td>
<td>Dizygotic</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic Acid</td>
</tr>
<tr>
<td>FISH</td>
<td>Fluorescence in situ Hybridization</td>
</tr>
<tr>
<td>FOXP2</td>
<td>Forkhead Box P2 gene</td>
</tr>
<tr>
<td>GABA</td>
<td>Gamma amino butyric acid</td>
</tr>
<tr>
<td>HDAC</td>
<td>Histone deacetylase inhibitor</td>
</tr>
<tr>
<td>ID</td>
<td>Intellectual Disability</td>
</tr>
<tr>
<td>Kb</td>
<td>Kilo base</td>
</tr>
<tr>
<td>LCR</td>
<td>Low Copy Repeats</td>
</tr>
<tr>
<td>LD</td>
<td>Linkage Disequilibrium</td>
</tr>
<tr>
<td>LOH</td>
<td>Loss of Heterozygosity</td>
</tr>
<tr>
<td>Mb</td>
<td>Mega base</td>
</tr>
<tr>
<td>ml</td>
<td>Milliliter</td>
</tr>
<tr>
<td>MZ</td>
<td>Monozygotic</td>
</tr>
<tr>
<td>MPX</td>
<td>Multiplex</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium Chloride</td>
</tr>
<tr>
<td>NaHep</td>
<td>Sodium Heparin</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Centre for Biotechnology information</td>
</tr>
<tr>
<td>nm</td>
<td>Nanometer</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>OMIM</td>
<td>Online Mendelian Inheritance in Man</td>
</tr>
<tr>
<td>PDD</td>
<td>Pervasive Developmental Disorder</td>
</tr>
<tr>
<td>PDD-NOS</td>
<td>Pervasive Developmental Disorder-Not Otherwise Specified</td>
</tr>
<tr>
<td>PSD</td>
<td>Postsynaptic Density</td>
</tr>
<tr>
<td>PWS</td>
<td>Prader-Willi Syndrome</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>Real Time quantitative Polymerase Chain Reaction</td>
</tr>
<tr>
<td>SPX</td>
<td>Simplex</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris-Acetate</td>
</tr>
<tr>
<td>UCSC</td>
<td>University of California Santa Cruz (Genome Browser)</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>VCFS</td>
<td>Velocardiofacial Syndrome</td>
</tr>
<tr>
<td>VPA</td>
<td>Valproate</td>
</tr>
<tr>
<td>WBS</td>
<td>Williams-Beuren Syndrome</td>
</tr>
<tr>
<td>WGA</td>
<td>Whole Genome Array</td>
</tr>
<tr>
<td>µl</td>
<td>Micro liter</td>
</tr>
</tbody>
</table>
Acknowledgements

First and foremost I would like to thank Dr. Suzanne Lewis and Dr. Evica Rajcan-Seperovic for their constant support and encouragement throughout the last five years. Even though it has taken me a long time to come back and write up this thesis, you have welcomed me back with open arms and for that I am ever so grateful. Suzanne, your drive and passion for the field of Autism and your complete dedication to the patients and their families has been truly inspiring. Thank you for being my mentor and friend. Evica, this project would not have been possible without your help and guidance. Thank you for welcoming me into your lab and doing everything in your power to help me navigate this project to completion.

I would also like to acknowledge the hard working genetic counsellors, especially Jeanette Hilderbrand, Jane Hurlburt and Susan Creighton for their dedication to the ASD patients and this study. I would also like to thank the members of the lab, mainly Ms. Chansonette Harvard and Dr. Ying Qiao for training me in Array CGH and FISH and for their continued support technically and otherwise, and above all for their friendship. I'd also like to thank the members of my committee, Dr. Wendy Robinson and Dr. Barb McGillivray for always making themselves available despite their busy schedules and for their advice and guidance.
This work is dedicated to my parents.

Love

Encouragement

Sacrifice

For all of it, I am forever grateful.
1. INTRODUCTION

1.1 Autism Spectrum Disorders (ASD)

Autism spectrum disorders (ASDs) are a group of neurodevelopmental conditions with significant clinical variability, genetic heterogeneity and a male to female ratio of 4:1. Deficits in reciprocal social interaction, verbal and non-verbal communication and the presence of restricted, repetitive and stereotyped patterns in behaviour, interests and activities are the hallmarks that have classically defined autistic disorder. ASDs include PDD-Not Otherwise Specified (PDD-NOS), Asperger syndrome and Autistic Disorder and form part of a larger group of Pervasive Developmental Disorders (PDD) that also include: Rett syndrome and Childhood Disintegration Disorder (CDD). Autistic disorder does not have a distinct, categorical phenotype; instead it represents part of the greater spectrum of social and communication deficits and behavioural restrictions associated with different genetic, epigenetic and environmental/teratogenic influences. For this reason, autism may be more appropriately referred to as a spectrum disorder.

Currently, the diagnosis of ASD is based primarily on tests that evaluate behaviour such as the ADI-R that are dependent on components such as verbal performance skills or cognitive developmental level. It is the interviewer’s responsibility to determine whether for each category or item on the test the behaviour truly meets the specific diagnostic criteria. Interviews usually focus on caregivers’ description of the subject’s actual behaviour as it occurs in his/her daily life and scoring is made on the basis of the interviewer’s judgement based on the caregivers’ account of behaviours. This subjective methodology makes the diagnosis and understanding of causative mechanisms very
difficult, and new tools that are dependent on more objective, definitive biological markers are desperately needed.

With the exception of Rett syndrome, which is attributed to mutations of the methyl-CpG-binding protein 2 gene (MeCP2) in most cases, ASDs are not linked to any one particular genetic or non-genetic cause. ASDs may, however, be divided into two phenotypic subgroups: essential and complex. Complex ASDs contribute to no less than 20% of cases and are defined by the presence of a significant number of physical anomalies with or without intellectual disability. Known causes of autism such as environmental agents and insults, chromosomal abnormality or single gene disorder such as Fragile X syndrome and tuberous sclerosis are included in this subgroup. In essential ASDs comprising the majority of cases (80%), the clinical morphology is normal. Making the distinction between these subgroups is an essential first step towards a better understanding of the genetic etiology of the ASDs. In complex ASDs, the underlying biological insult(s) to morphogenesis are different from essential cases in which somatic development proceeds normally and therefore, a different set of susceptibility genes and biological pathways are most likely at play. For instance, dysmorphic individuals (complex subgroup) are more than twice as likely to have an abnormality in their brain MRI compared to individuals in the essential subgroup. This suggests that at least in part, the biological events and susceptibility loci leading to complex ASDs are different from those leading to essential ASDs. In this study, phenotypic heterogeneity between patients undergoing genetic screening for microdeletions/microduplications has been minimized through a selection strategy that identified phenotypic subgroups based on standardized and comprehensive morphological characterization. These determinations were performed by one clinical
geneticist (Dr. MES. Lewis) based on phenotypic selection criteria adapted from those initially correlated with increased liability to subchromosomal telomeric rearrangements as described by de Vries et al (2001)\(^6\). Details of the specific phenotypic selection criteria are described in the materials and methods section.

1.2 Endophenotypes: On the Path to Reducing Phenotypic Heterogeneity

Despite continuous and countless efforts from the scientific and medical communities, the genetic aetiology of the ASDs remain poorly understood. One fundamental reason for this is the great phenotypic variation that exists among affected individuals. Phenotypic heterogeneity is in part a reflection of genetic heterogeneity, for which presence in a given test population will “dilute” the significance of potential candidate genes and regions. Recently, some groups have begun addressing the issue of phenotypic heterogeneity by identifying a more homogeneous subset within the broader autistic phenotype based on endophenotypes. Endophenotypes are measurable components of a syndrome including anatomic, biochemical, neuropsychological and other possible measures\(^7\). For instance, by using language delay as an endophenotype, the support from linkage to chromosome 2q and 7q has been greatly enhanced,\(^8,9\) and using a subset of families with obsessive compulsive behaviour, linkage to chromosome 17 has been strengthened. These studies demonstrate how phenotypic heterogeneity can be responsible for the lack of power and discrepant results found in some whole genome scans.
1.3 Genetics of ASD: An Overview

The first line of evidence pointing to a genetic influence in ASDs was provided by twin and family studies. Early twin studies in Scandinavia and the UK found that in monozygotic twins (MZ), the concordance for the more strictly defined diagnosis of classical autism was >60% versus 0% in dizygotic twins (DZ). When the broader phenotype for social and/or cognitive impairments were examined, the concordance in MZ pairs was as high as 92% versus 10% in DZ pairs. These incidences are presumably due to the fact that monozygotic twins share all their genes, compared to dizygotic twins who share only half of their genes and do not take into account other factors such as environmental influences. Some authors believe this large decrease in risk for dizygotic twins relative to monozygotic twins to be consistent with a genetic model in which involvement of several genes interacting with one another leads to disease susceptibility. However, based on recent array-CGH findings it is becoming evident that this theory needs to be re-examined. These studies (including this work) are finding that in some affected individuals it is rare variants often arising from spontaneous defects (de novo) with large phenotypic consequence that are contributing to ASD susceptibility.

The exact mode of inheritance for ASDs is not clear for all cases. Based on the broad clinical spectrum observed in these disorders, it has traditionally been suggested that a multigenic mechanism involving simultaneous defects at multiple loci confers susceptibility. The estimate of the number of loci involved has ranged from 2-10 to greater than 15. More recent studies based on current copy number variant (CNV) findings suggest a different mechanism however. Finding from these studies suggest 1) the existence of rare variants with significant phenotypic effect as seen by de novo cases of
deletions and duplications that have only been reported once or in a few cases and 2) several recurrent regions found among patients across studies\textsuperscript{15-19}. These observations suggest that the estimated number of loci may be higher than previously anticipated and the involvement of different aetiologies between cases.

1.3.1 Strategies for Locating ASD Culprit Genes

A variety of different approaches to locate and identify genes for ASDs have been undertaken. These are described below, together with some relevant findings and comments regarding the difficulties in applying these methods to uncovering the locations of genes of modest to mild effect for behaviourally-defined disorders. This section is not intended as a full review but rather serves to illustrate current knowledge, approaches being undertaken, and efforts reflecting recent research contributions in each area.

Methodologies to identify genes or regions of interest in ASD include: 1) Linkage studies such as genome screens in affected sib pairs and multiplex families; 2) Gene association studies, including candidate gene studies and 3) Chromosomal methods such as karyotyping and Fluorescence in situ Hybridization (FISH) that may guide molecular studies by pointing out de novo or inherited chromosomal anomalies in affected individuals and their relatives\textsuperscript{1} and most recently 4) Copy Number Variant (CNV) discovery to target underlying ASD susceptibility genes.

1) Linkage Studies: Numerous linkage and genome scan approaches to identify important loci in ASD have been undertaken using both simplex (SPX) and multiplex (MPX) families. Linkage between autism and markers on several chromosomes has been demonstrated and over the years collective data point to chromosome loci of interest across
almost all chromosomes including: 1p, 1q, 2q, 3p, 3q, 4p, 4q, 5p, 5q, 6q, 7q, 8q, 11p, 13q, 15q, 16p, 17q, 19p, 19q, Xp and Xq. Unfortunately, no region has been unambiguously confirmed across all studies, and no gene has been conclusively identified for susceptibility to autism. This lack of reproducibility is due to locus heterogeneity (the observation that mutations at different loci lead to the same disease or phenotype). Regions that hold a greater promise for harbouring an ASD susceptibility gene based on reproducibility of results between independent studies include 2q, 7q, 15q, 17q and Xq.

Overall, linkage and genome scans used to map disease associated genes are valuable for narrowing down regions of interest. However, these regions often consist of entire chromosome bands and are simply too large, containing many possible candidate genes. Further, linkage studies are often underpowered because of the genetic and phenotypic heterogeneity found between probands. ASD is a complex disorder and whether it is caused by many genes of small effect or rare variants of large effect, it remains difficult to detect them by linkage. As seen from the recent Autism Genome Project Consortium (AGPC) genome scan of 1168 multiple incident ASD families using the 10K SNP array, simply increasing the number of families does not necessarily lead to more refined target regions for analysis; in fact, increasing the sample size to 1168 families led to a single peak of “suggestive” linkage. These findings argue for additional studies on more well-defined subgroups of families based on specific clinical endophenotypes to increase the likelihood of discovering the associated susceptibility genes.

2) **Autism Candidate Gene Screening via Genetic Association Studies:** Genetic association can be used for evaluating the role of positional and functional candidate genes in susceptibility to autism. Association methods are usually used to study specific candidate
genes for a given disease. Association is not a specifically genetic phenomenon; it is rather a statistical statement about the co-occurrence of alleles or phenotypes\textsuperscript{21}. The most common method is the transmission disequilibrium test (TDT) that starts with couples with one or more affected offspring. To test whether marker allele "A" is associated with a disease, parents that are heterozygous for A are selected. The test simply compares the number of cases where such a parent transmits A to an affected offspring with the number of cases when they transmit the other allele\textsuperscript{21}. In ASDs, numerous association and candidate gene studies suggest the involvement of the serotonin transporter gene (SLC6A4; 17q11-12)\textsuperscript{22,23}, the Gamma amino butyric acid (GABA) system (15q11-13)\textsuperscript{24,25} and some of the genes present within the autism susceptibility loci on chromosome 7q including: RELN (Reelin) (7q22)\textsuperscript{26} and FOXP2 (Forkhead Box P2) (7q31)\textsuperscript{27}. However, association studies have often yielded conflicting results. A classic example is candidate gene studies for the serotonin Transporter Gene (SLC6A4 or 5-HTT) that resides on 17q11.

Serotonin transporter in the brain and many peripheral tissues is responsible for the active transport of serotonin into neurons, platelets and other cells and mediates the removal and recycling of released serotonin following neuronal stimulation\textsuperscript{28}. In ASD, SLC6A4 is a prime candidate gene based on the observations that 1) hyperserotonemia is found in about one third of autistic probands (the most consistently replicated neurochemical finding)\textsuperscript{12}; 2) the reduction of anxiety, ritualistic behaviour and aggression following administration of selective serotonin transporter inhibitors \textsuperscript{29}, 3) involvement of the serotonergic pathway in neurodevelopment of the brain including regulation of cell proliferation, migration and differentiation in neuronal tissue \textsuperscript{30} and 4) linkage findings in the 17q11 region where SLC6A4 resides \textsuperscript{31,32}. Transcriptional activity of SLC6A4 is
modulated by a polymorphic repetitive element located upstream of the transcription start site (5-HTT gene linked polymorphic region-5HTTPR)\textsuperscript{33}. A deletion/insertion in the 5HTTPR creates 2 major alleles: long (L) and short (S) (i.e. 14 and 16 repeat alleles) which alter promoter activity\textsuperscript{34}. The short allelic variant is associated with reduced expression of \textit{SLC6A4} due to a decrease in transcriptional efficiency, resulting in lower serotonin uptake activity \textsuperscript{35}. Several studies have examined the transmission of the L and S alleles in autism trios. Significant over-transmission of the S allele has been reported in four studies \textsuperscript{12,36} \textsuperscript{22,32}. In contrast, over-transmission of the L allele has been reported in three studies \textsuperscript{37-39} with no evidence for transmission disequilibrium in several additional studies \textsuperscript{31,40-42}. This lack of consistency between studies may in part be due to differences between studies such as the ethnicity of the groups examined. These inter-study differences have severely hampered the identification of candidate genes for ASDs.

Association studies are without a doubt a valuable method for evaluating the correlation of specific known alleles to disease. However the approach is often complicated by heterogeneity of the autistic phenotype, the involvement of multiple interacting loci, presence of rare variants and the degree to which risk alleles are expressed.

3) Cytogenetic Studies: A third method for identifying susceptibility genes associated with the ASDs involves the study of the patients' chromosomes. Both numerical and structural chromosome aberrations have been documented, including translocations, inversions, duplications, deletions, markers and rings. Mosaicism has also been reported. A review of 15 worldwide studies\textsuperscript{20} has placed the mean rate of known, visible cytogenetic abnormalities at 7.4% (129/1749), with a range from 0 to 54% depending on the studies'
ascertainment methods. Approximately, 11% of cases with cytogenetic abnormalities were associated with classical chromosome syndromes including: 45,X (Turner syndrome), 47,XXY (Klinefelter syndrome), XYY, trisomy 13 and trisomy 21. Further, 22 of the 129 cases had an apparently balanced rearrangement (translocation or inversion). Several reviews of chromosome abnormalities co-occurring with autism report the involvement of most chromosomes\textsuperscript{43-45}. Chromosomal anomalies can lead to functional genetic changes through 1) gene dosage alterations due to a deletion or duplication, 2) positional effects that alter gene regulation and expression levels, 3) breakpoints that lead to gene disruption and 4) deletions that unmask recessive mutations. Cytogenetic studies would especially be valuable in identification of chromosomal defects, discovering new candidate genes of likely large effect that may only be present in a subset of cases or are \textit{de novo} in origin. They can assist in localizing culprit genes and associated pathways or contribute to the discovery of mechanisms of altered gene expression due to imprinting (as in the case of 15q11-13 maternal duplications). They can also point to mechanisms leading to recurrent deletions and duplications such as recombination between non-homologous low copy repeat sequences (\textit{LCRs}). A brief review of some of the cytogenetic regions of interest reported in the ASDs, is found below.
1.4 Cytogenetic Regions of Interest in ASD (CROIs):

1.4.1 Chromosome 7

Significant overlap has been found between candidate regions for autism identified from linkage analysis and reported cytogenetic abnormalities involving various regions along the long arm of chromosome 7\(^45\). A large collection of published and unpublished clinical cases (www.chr7.org) with autism and cytogenetic abnormalities of chromosome 7 suggest the involvement of various regions along the entire length of the chromosome including: 7p22, p21, p15.3, p12.2, p11, pter, q11.2, q21, q22, q31, q32 and q34. Of special interest are 7q21-22 and 7q31-32. Several candidate genes for autism reside in these regions, including PIK3CG (Phosphatidylinositol 3-Kinase, Catalytic Gamma) (7q22), RELN (7q22), NRCAM (Neuronal Cell Adhesion Molecule (7q31.1-31.2), LAMBL (Laminin Beta-1) (7q31.1), WNT2 (Wingless-Type MMTV Integration Site Family Member 1) (7q31.2), FOXP2 (7q31) and UBE2H (Ubiquitin-Conjugating Enzyme E2H) (7q32)\(^45\). More recently, the 7q11.2 deletion region associated with Williams-Beuren Syndrome (WBS) is being implicated for its role in the etiology of ASD. In 2002 Sultana and colleagues\(^46\) reported a pair of monozygotic twins concordant for autism and a translocation t(7;20)(q11.2;p11.2). Somerville et al\(^47\) described the first case of a duplication of the WBS critical region in a patient with severe language delay, developmental delay and ADHD. Although the most striking and common phenotype of cases with the microduplication of 7q11.2 is severe language delay\(^48\), the spectrum of behavioural phenotypes include many of those associated with ASD and several patients with a diagnosis of ASD have now been described to harbour this microduplication\(^48,49\). The striking difference between WBS and the duplication 7q11.2 syndrome phenotype is in language development, as the WBS deletion is associated with highly developed expressive
language skills and the duplication with severe language development delays. This lends support to a dosage sensitive nature of some of the genes in this region.

1.4.2 Duplications of Proximal 15q

Alterations in the proximal 15q region are the single most commonly observed autosomal change found in ASD (1-4% of all cases\textsuperscript{20}). Proximal 15q is a hotspot for chromosomal rearrangements, in part due to the high number of low copy repeat sequences (also known as duplicons) found in this region. Increased sequence identity between duplicons mediates rearrangements that often lead to chromosomal imbalances, including translocations, inversions, deletions, duplications and triplications\textsuperscript{50}.

Deletions of 15q11-13 are associated with Prader Willi and Angelman syndromes (PWS/AS). This is an imprinted region, thus whether the deletion is maternally or paternally derived accounts for the differences in the clinical phenotype in patients with AS or PWS respectively\textsuperscript{51}. The majority of 15q abnormalities found in autism are maternally derived duplications of 15q11-13\textsuperscript{52,53}. The origin of duplications includes supernumerary dicentric chromosome 15 (dic (15;15)) and interstitial duplications and triplications. Individuals with small dic(15;15) that is mostly composed of heterochromatic material usually have a normal phenotype. The larger dic(15;15) contains 15q euchromatin and is associated with an abnormal clinical phenotype similar to that found in cases of interstitial duplications. This phenotype often includes intellectual disability (ID), developmental delay and behavioural disturbances\textsuperscript{54,55}.

Some of the genes found in this region have biological relevance for ASD. Nurmi et al\textsuperscript{56} found significant evidence for linkage disequilibrium at marker D15S122 located at
the 5' end of *UBE3A* in a study involving 94 multiplex autism families. Mutations or the deletion of the maternal copy of *UBE3A* can cause AS\(^5^1\). Since *UBE3A* is under uniallelic maternal expression in the brain during development, it is feasible that *UBE3A* gene product imbalances occurring as a result of the duplication of this region on the maternal chromosome 15 may be a contributing factor in the aetiology of ASD in individuals with maternally inherited duplication of 15q11-13. Incidentally, there is significant overlap between AS and ASD phenotypes, and children with AS often meet the behavioural criteria for diagnosis of ASD\(^5^7\). Gene dosage alterations from the normal monosomic state as a result of a deletion (AS) and duplication (ASD) appears to be an important contributor in both disorders with some overlapping phenotypic outcome.

*ATP10C* is found 200Kb distal to *UBE3A* and has been reported to be under preferential maternal expression in the human brain\(^5^8\)\(^5^9\). In addition to its location on 15q and its imprinting status, *ATP10C* makes an intriguing candidate gene for autism based on its pattern of expression and biological function. In mouse the *Atp10c* homologue (Atp10c/pfatp) transcript is localized to cerebellar granule cells, the hippocampus and cells surrounding the corpus callosum\(^6^0\) and overlaps the region of *Ube3a* expression\(^6^1\). The protein product of this gene is an aminophospholipid transporting ATPase involved in maintaining cell membrane integrity and may therefore be an essential component for cell signalling in the central nervous system (CNS)\(^5^8\)\(^6^0\). Transmission disequilibrium studies have, however, yielded mixed results regarding the likelihood of *ATP10C* as a candidate gene. Nurmi et al.\(^6^2\) reported positive evidence for linkage disequilibrium (LD) using a dense SNP map of proximal 15q. Previous to this study, Kim et al.\(^6^3\) reported a lack of
transmission disequilibrium between \textit{ATP10C} and autism using 14 markers within \textit{ATP10C} in 115 autism trios.

Similarly, the GABA-A receptor subunits (\textit{GABRB3}, \textit{GABRA5} and \textit{GABRG3}) which lie within 15q11-13 have been examined for their role in the etiology of autism. Biologically, these receptor subunits make excellent candidates for neurological disorders. In the mature brain, GABA acts as an inhibitory neurotransmitter. In the developing brain, however, GABA may function as a neurotrophic factor affecting neural differentiation, growth and organization, as seen by the pattern and temporal expression of GABA-A subunits in murine development \textsuperscript{64}. Several studies have screened the GABA-A receptor subunits for linkage disequilibrium in both simplex and multiplex autism families. Cook et al.\textsuperscript{64} found linkage disequilibrium between autistic disorder and marker 155-CA2 in the \textit{GABRB3} receptor subunit. Buxbaum et al.\textsuperscript{65} further reported a positive association between autism disorder and the polymorphic marker 155-CA2 in \textit{GABRB3} in the 80 families (59 multiplex, 21 trios) studied. In contrast, Maestrini et al.\textsuperscript{40} failed to reveal any significant association between autism and GABR3-155CA2 and other markers within and outside of the \textit{GABRB3} gene in the 94 families studied (86 multiplex, 8 simplex). Curran et al.\textsuperscript{66} used 5 markers spread across the 4Mb PW/AS critical region to examine linkage disequilibrium (LD) in 148 ASD and 82 AD families. They found no significant LD in either set of families when looking at paternal and maternal meioses combined. However, when considered separately, they found a significant reduction of the paternal allele at marker D15S11 in the broader ASD families and an over transmission of the paternal alleles at 2 markers (D15S1506 and D15S1002) in the AD families.
The precise contribution of the genes on 15q11-13 in the neuropathology of ASD remains elusive. Given that ASD is a multigenic disorder, with several susceptibility genes of modest effect, alterations in the level of expression of UBE3, ATP10C and GABA-A receptor subunits may act in unison, resulting in the autistic phenotype in individuals with the 15q11-13 maternal duplication.

1.4.3 Chromosome 22

Microdeletions of 22q11.2 are associated with velocardiofacial (VCFS) and DiGeorge Syndromes. The phenotype includes various physical and cognitive characteristics, including cardiac defect, cleft palate/velopharyngeal insufficiency, immunodeficiency, mild intellectual disability, learning disabilities and language delay. The phenotype can be quite variable among patients, even though there is much less heterogeneity at the genetic level. The majority of cases have an essentially identical 3Mb deletion, with a smaller subset having a 1.5 to 2Mb deletion. A few case with translocations leading to the VCFS/DGS have also been reported. Such variability also includes co-morbidity for ASD, as approximately 14% of individuals with a deletion of 22q11.2 are diagnosed for an ASD, with 11% meeting the diagnostic criteria for autistic disorder (AD).

The 22q13 deletion syndrome is also a clinically recognizable syndrome whose features include craniofacial dysmorphism, normal to advanced growth, moderate to severe developmental delay, hypotonia, speech delay and behaviours that overlap with those in the autism spectrum. Deletions result from the loss of genetic material from the terminus of the long arm of one copy of chromosome 22. One of the genes believed to be involved
with the phenotypic outcome of 22q13.3 deletion is *SHANK3* (SH 3 and multiple ankyrin repeat domains 3). *SHANK3* is a scaffolding protein found in the excitatory synapses directly opposite to the presynaptic active zone. Shank proteins are believed to act as master organizers of the postsynaptic density (PSD). At the PSD complex *SHANK3* has been shown to bind with neurologins (*NLGN3* and *NLGN4*), which together with neurexins form a complex at the glutamatergic synapses \(^{73,74}\). Pathogenic mutations of *SHANK3* are estimated to occur in 1% of cases of ASD based on findings from two large scale studies in which 3/400 and 3/222 ASD patients were found to have a functional mutation which was absent in non-affected parents, sibs and control subjects \(^{73,74}\). It has been hypothesized that the protein complex including neuroligins and Shank participate in an assembly of specialized postsynaptic structures required for development of language and social communications both of which are absent or hampered in both ASD and 22q13 deletion syndrome \(^{73}\).

### 1.4.4 Subtelomeric Regions

The most distal regions of chromosomes just proximal to the tandem telomeric repeats are known as the subtelomeric regions. These regions are often gene rich, and several candidate genes for recognizable syndromes are known to reside in the subtelomeric regions. Subtelomeric abnormalities have been detected in 6-11.1% of probands with moderate to severe intellectual disability \(^{75-77}\). In autism, 10-14% of probands have been found to have a subtelomeric aberration \(^{78,79}\) including del(22)(q13.3)(as previously discussed), del(1)(p36) leading to monosomy 1p36 deletion syndrome and del(2)(q37). Del(1)(p36) is the most common terminal deletion observed
with a rate of 1/5000 live births. The hallmark features associated with it include intellectual disability, developmental delay, hearing and growth impairments, hypotonia, seizures and heart defects. Del(2)(q37) is commonly reported in ASD with breakpoints at 2q37.1, 2q37.2 or 2q37.3. Autistic features or a diagnosis of autism is found in 17% (2q37.1), 50% (2q37.2) and 32% (2q37.3) of cases, depending on the breakpoint while developmental delay and intellectual disabilities seem to be universal regardless of the breakpoint.

Due to the prevalence of subtelomeric abnormalities found in ASD, all patients recruited for this study had their chromosomes analysed by subtelomeric FISH. Any patient found to have a rearrangement was disqualified as a candidate for array CGH studies.

4) CNV Discovery: The most recent trend in ASD research has been in Copy Number Variant (CNV) discovery to target underlying ASD susceptibility regions/genes using high resolution array technology. A CNV is defined as a segment of DNA that is 1 Kb or larger and is present at a variable copy number in comparison with a reference genome. Classes of CNVs include deletions and duplications. Whole genome arrays (WGA) allow detection of DNA copy number changes at a much higher resolution than conventional cytogenetic techniques such as chromosomal comparative genomic hybridization (Chromosomal CGH) and karyotyping and facilitate the analysis of the whole genome in one experiment. WGA contain small segments of chromosomal DNA (oligonucleotides or BAC clones) sampled on a glass slide. The detection of copy number changes can be done using the principle of chromosomal comparative genomic hybridization (array CGH). One area in which whole genome arrays have proven to be a
powerful tool has been the detection of otherwise unidentifiable abnormalities in individuals with unexplained causes of intellectual disability +/- physical abnormalities. In a recent review by Stankiewitz and Beaudet\textsuperscript{84} summarizing the array findings in a large number of ID individuals, the frequency of submicroscopic deletions and duplication was from 9-16\%. Using a tiling array containing 32,447 BACs, DeVries et al.\textsuperscript{85} screened 100 patients with unexplained ID and, although DNA copy number changes were present in 97\% of cases, only 10\% of cases had alterations that were \textit{de novo} in origin and considered clinically relevant. Interestingly, both studies noted that while many aberrations were detected by increasing the coverage and resolution of the array, only 25\% of \textit{de novo} aberrations found were less than 1 Mb. This suggests that the majority of submicroscopic anomalies relating to ID tend to be larger than 1Mb and can be detected by arrays of 1Mb resolution or smaller.

The first reported application of whole genome arrays in ASD was reported in 2006 using the 1Mb array. Jacquemont et al.\textsuperscript{16} identified eight rearrangements (27\%) in 29 syndromic ASD patients of which six were deletions and two were duplications ranging in size from 1.4 to 16 Mb. No recurrent abnormalities were identified. Sebat et al.\textsuperscript{18} tested the hypothesis that \textit{de novo} copy number variation (CNV) is associated with ASDs. They performed high resolution (35 kb) genomic microarray analysis on all parents, patients and unaffected children from 264 families. They found that the frequency of \textit{de novo} CNVs was greater in ASD (14 out of 195) than unaffected individuals (2 out of 196). Further, \textit{de novo} CNVs were more likely to occur in sporadic cases (10\%) compared to patients with an affected first degree relative (3\%). This observation suggests sporadic (simplex) and familial (multiplex) cases of ASD to be genetically distinct. This is consistent with the idea
that two different genetic mechanisms contribute to risk: spontaneous mutations and their inheritance, with the latter being more common in multiplex families with two or more affected members. Further, the ASD susceptibility regions identified in this study were highly heterogeneous, as none were observed more than twice and most were seen only once. CNV rates of 7.1% and 2.0% in idiopathic simplex and multiplex families, respectively, were reported in another study. Among the CNVs unmasked, some occurred at previously reported ASD loci such as: 2q37.3, 15q11-13, 22q11.2, and 22q13.31.

The fourth and largest study of this kind was conducted by The Autism Genome Project Consortium using Affimatrix SNP array analysis of 1,168 families with at least two affected individuals. The study identified 254 CNVs in 196 cases from 173 families. Some of the CNV found overlapped previously published ASD chromosomal rearrangements and others were not previously reported, such as 11p12-13 (Table 1). One interesting finding from this study was the mutation of the neurexin gene (NRXNI) as a result of a hemizygous deletion of coding exons found in an affected sib pair. Neurexin is a major presynaptic protein partner (ligand) of the postsynaptic neuroligin family proteins. Alterations in the Neuroligin genes (NLGN 3 and NLGN 4) have been previously reported for their association with ASD. The neurexin-neuroligin interaction is thought to trigger postsynaptic differentiation and control the balance of inhibitory GABAergic and stimulatory glutamatergic inputs. Together with SHANK3 and contactin-associated-protein-like 2 (CNTNAP2), which have also been associated with ASD, neureligins and neurexin are part of a synaptogenic pathway whose disruption is involved in the development of ASD. Lastly, Christian et al. reported 51 confirmed ASD specific
CNVs in 46 of 397 patients (11.6%). Seven CNV were \textit{de novo} and 44 were inherited. Collectively, these studies demonstrate the power of high resolution genomic analysis for unmasking novel loci and candidate regions that may prove to have a role in the aetiology of ASDs.
Table 1(a). Summary of findings from five studies searching for ASD associated CNVs using array genomic hybridization

<table>
<thead>
<tr>
<th>Study</th>
<th>No. of cases</th>
<th>Technology</th>
<th>No. De Novo CNVs</th>
<th>No. Inherited CNVs</th>
<th>% ASD-Specific CNVs</th>
<th>Recurrent/Overlapping Regions Within Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jacquemont</td>
<td>29</td>
<td>1 Mb array-CGH</td>
<td>7</td>
<td>2</td>
<td>8(27.5%)</td>
<td>No recurring CNVs detected</td>
</tr>
<tr>
<td>Sebat (2007)</td>
<td>165 ASD</td>
<td>35 Kb Oligo</td>
<td>14 in ASD and 2 in controls</td>
<td>Not reported</td>
<td>10% SPX 3% MPX</td>
<td>2q37.2-2q37.3 (2), 3p14.2(2), 20p13(2)</td>
</tr>
<tr>
<td></td>
<td>Families: 118</td>
<td>arrays</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>195 affected</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2007)</td>
<td>families in</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>total; 350 met</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>stringent cut-off values (data reported here)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1(a) continued. Summary of Findings from five studies searching for ASD associated CNVs using array genomic hybridization

<table>
<thead>
<tr>
<th>Study</th>
<th>No. of cases</th>
<th>Technology</th>
<th>No. De Novo CNVs</th>
<th>No. Inherited CNVs</th>
<th>% ASD-Specific CNVs</th>
<th>Recurrent/Overlapping Regions Within Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Christian (2008)</td>
<td>397 ASD probands</td>
<td>19 k Tiling array</td>
<td>9</td>
<td>42</td>
<td>51(11.6%)</td>
<td>42/51 CNV only observed once; recurrent/overlapping CNV identified for 3 regions: 22q11(2), 15q11-13(3), 16p11.2(4)</td>
</tr>
<tr>
<td>Marshall (2208)</td>
<td>427 ASD families: 413 idiopathic, 14 known abnormal karyotypes</td>
<td>500k SNP array</td>
<td>27</td>
<td>196</td>
<td>7% SPX 2% MPX</td>
<td>13 regions in unrelated cases: 2q14.1(2), 2q32.1(2), 6q22.31(2), 7q36.2(4), 8q11.23(2), 9q24.1(2), 11p12(2), 13q21.32(2), 15q11.2-q13.3(2), 16p12.2(2), 16p11.2(3), 22q11.2(4), 22q13.31(2)</td>
</tr>
</tbody>
</table>

*34 refers to the number of CNV present in 2 or more sibs that are de novo or inherited. Bold numbers in ( ) refer to the number of cases reported.
Table 1(b). Summary of recurrent and/or overlapping ASD associated CNV regions reported by 2 or more studies

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Recurrence/Overlapping CNV Regions Reported Between Studies</th>
<th>Study*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1p36 (2)</td>
<td>Jaquemont and Szatmari</td>
</tr>
<tr>
<td></td>
<td>1q31(2)</td>
<td>Jaquemont and Szatmari</td>
</tr>
<tr>
<td>2</td>
<td>2p16.1 (2)</td>
<td>Sebat and Szatmari</td>
</tr>
<tr>
<td></td>
<td>2p25.3 (2)</td>
<td>Szatmari and Christian</td>
</tr>
<tr>
<td></td>
<td>2q32.1-q32.3 (4)</td>
<td>Szatmari, Christian and Marshall</td>
</tr>
<tr>
<td>3</td>
<td>3p14.2 (3)</td>
<td>Sebat and Szatmari</td>
</tr>
<tr>
<td>6</td>
<td>6q22.31 (3)</td>
<td>Szatmari and Christian</td>
</tr>
<tr>
<td>7</td>
<td>7p21.1 (2)</td>
<td>Sebat and Christian</td>
</tr>
<tr>
<td>8</td>
<td>8q11.23 (4)</td>
<td>Szatmari, Christian and Marshall</td>
</tr>
<tr>
<td>10</td>
<td>10q11.22-q11.23 (3)</td>
<td>Sebat, Szatmari and Christian</td>
</tr>
<tr>
<td></td>
<td>10p13 (2)</td>
<td>Szatmari and Christian</td>
</tr>
<tr>
<td>11</td>
<td>11p11.12 (7)</td>
<td>Szatmari and Christian</td>
</tr>
<tr>
<td>13</td>
<td>13q14.2 (2)</td>
<td>Szatmari and Christian</td>
</tr>
<tr>
<td></td>
<td>13q21.32 (4)</td>
<td>Szatmari and Marshall</td>
</tr>
<tr>
<td>15</td>
<td>15q11-13 (19)</td>
<td>Jaquemont, Sebat, Szatmari, Christian and Marshall</td>
</tr>
<tr>
<td>16</td>
<td>16pl1.2 (7)</td>
<td>Szatmari, Christian and Marshall</td>
</tr>
<tr>
<td></td>
<td>16q23.3 (2)</td>
<td>Szatmari and Christian</td>
</tr>
<tr>
<td>17</td>
<td>17p12 (4)</td>
<td>Szatmari and Christian</td>
</tr>
<tr>
<td>19</td>
<td>19q12 (2)</td>
<td>Szatmari and Christian</td>
</tr>
<tr>
<td></td>
<td>20p13 (3)</td>
<td>Sebat and Szatmari</td>
</tr>
<tr>
<td>22</td>
<td>22q11.2-q11.21(7)</td>
<td>Szatmari, Christian and Marshall</td>
</tr>
<tr>
<td></td>
<td>22q13.31-q13.33 (8)</td>
<td>Sebat, Szatmari, Christian and Marshall</td>
</tr>
</tbody>
</table>


-Number indicated in ( ) refers to the total number of cases reported for a given region.
The presence of CNVs including micro-deletions and micro-duplications clearly contribute to ASD. However, how these small aberrations contribute to the aetiology of ASD is largely unknown, and their association with more comprehensively evaluated clinical phenotypes beyond neurobehavioural features alone remains unexplored. The identification of ASD related micro-deletions and micro-duplications and their associated phenotypes will facilitate identification of genes that may be important in biological functions such as embryonic neurodevelopment, maintenance of neuronal cell integrity and signalling pathways or non-coding regulatory regions. Thus, the unmasking and characterization of these micro-aberrations at the whole genome level continues to be an important tool for uncovering new susceptibility loci.
1.5 Specific Aims and Hypothesis

The purpose of this project was to apply whole genome analysis using the array-CGH technology in order to test the hypothesis that pathogenic CNVs contributing to the aetiology of idiopathic ASD will be identified. These CNVs may pinpoint novel genes associated with ASD.

To test this hypothesis, 19 probands with a confirmed diagnosis of ASD and no known cytogenetic or molecular abnormalities and who met our phenotype selection criteria (as outlined in the materials and methods section) were selected. DNA was obtained from whole blood and used for array CGH studies using commercially available 1Mb arrays (Spectral Genomics, Houston TX, USA). All detected CNVs were compared against the data base of genomic variation (http://projects.tcag.ca/variation/) and only those not previously reported as being a common variant in the normal population were further validated by FISH and or qPCR studies.
2. MATERIALS AND METHODS

Note: All families who participated in this study did so voluntarily. A consent process approved by the Clinical Research Board of the University of British Columbia and the BC Children's and Women's Hospital were signed by each participating family member. For probands under the age of seven years, a consent form was signed by the parents or guardians on behalf of the proband.

2.1 Patient Selection

Families for the study were recruited via the Research Registry of the Autism Spectrum Disorders-Canadian American Research Consortium (ASD-CARC; www.AutismResearch.com) according to the family’s voluntary expressed interest to participate in the genetic screening for autism susceptibility genes in both single-incidence (simplex; SPX) and multiple-incidence (multiplex; MPX) families with an ASD. The process was initiated with an interview with a family member to obtain family and medical histories. Affected family member(s) evaluated at the UBC regional center of ASD-CARC underwent a consistent standardized and comprehensive clinical assessment by one geneticist (Dr. MES Lewis). The assessment included: (I) an interview with parents regarding the birth and medical history of each affected proband; (II) comprehensive morphometric and general physical examination for each proband including anthropometric cranio-facial measurements with 2D and 3D digital imaging of the proband’s craniofacies. Upon the completion of each examination blood was drawn from the proband(s) and participating family members for the purpose of DNA isolation and cell culture to obtain metaphase chromosomes. Preliminary screening genetic tests including routine karyotype, Fragile X syndrome testing, FISH screening for del (22)(q11), del(22)(q13), and the
subtelomeres were obtained for subjects with an ASD as a part of their clinical assessment. Routine dup(15)(q11-q13) testing by FISH was subsequently added when it became available on a clinical service basis. Patients found to be positive for any of the abnormalities mentioned above were excluded from entering the array study since these abnormalities are known to be associated with ASD.

To reduce phenotypic heterogeneity, a five item checklist adapted from de Vries et al. 6 (Table 2) was used to score each proband. This checklist was originally devised to provide physicians with an effective clinical pre-selection tool when screening for subtelomere deletions in children with intellectual disability. According to the study, a score of ≥3 as a cut-off for subtelomere testing will exclude some patients from receiving the test without missing a subtelomeric abnormality. In their study, 21 of 104 children were excluded on the basis of this score. When the cut off was increased to a score of ≥4, 36 of 104 children could have been excluded, however, 11% of subtelomeric cases would be missed. Based on these data, the cut off score for our study is a score of ≥3 out of the possible maximum score of 10. Table 2 summarizes the five categories used for scoring. Each category has a slightly different scoring system. For example, family history for intellectual disability gets a score of 1 if the pattern of inheritance is Mendelian and a score of 2 if it’s non-Mendelian. Evidence of prenatal growth retardation gets an automatic score of 2 while for categories of postnatal growth abnormalities and non-facial dysmorphism including congenital anomalies each individual abnormality gets a score of 1 and only a maximum of 2 points is allowed for each of these categories. For the category of facial dysmorphology, at least 2 dysmorphic features have to be present in order to get the automatic score of 2. Overall however, it is not necessary for the proband to have a
positive feature in all of the categories listed in order to get a score rendering access to array-CGH evaluation. Subjects meeting the criteria for a score of >3 who did not have a previously reported cytogenetic or molecular abnormality were recommended for the microarray study. This approach has allowed us to distinguish between more complex (but still idiopathic) forms of ASD (score >3) and those that are more simple idiopathic forms (without dysmorphology or family history of ASD/ID) thus creating a more homogeneous sub-group to study.
Table 2. Clinical criteria for selecting patients for Array-CGH studies (Adapted from de Vries et al (2001))⁶. Only probands with a score of 3 or greater were selected for the microarray study.

<table>
<thead>
<tr>
<th>Items</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family history of mental retardation</td>
<td></td>
</tr>
<tr>
<td>➢ Compatible with Mendelian Inheritance</td>
<td>1</td>
</tr>
<tr>
<td>➢ Incompatible with Mendelian inheritance</td>
<td>2</td>
</tr>
<tr>
<td>Prenatal onset growth retardation</td>
<td>2</td>
</tr>
<tr>
<td>Postnatal growth abnormalities</td>
<td>2</td>
</tr>
<tr>
<td>➢ For each of the following 1 point (2 maximum):</td>
<td></td>
</tr>
<tr>
<td>➢ Microcephaly</td>
<td></td>
</tr>
<tr>
<td>➢ Short Stature</td>
<td></td>
</tr>
<tr>
<td>➢ Macrocephaly</td>
<td></td>
</tr>
<tr>
<td>➢ Tall stature</td>
<td></td>
</tr>
<tr>
<td>≥2 Facial Dysmorphic Features</td>
<td>2</td>
</tr>
<tr>
<td>➢ Hypertelorism</td>
<td></td>
</tr>
<tr>
<td>➢ Nasal anomalies</td>
<td></td>
</tr>
<tr>
<td>➢ Ear anomalies</td>
<td></td>
</tr>
<tr>
<td>Non facial dysmorphism and congenital abnormalities</td>
<td>2</td>
</tr>
<tr>
<td>➢ For each anomaly 1 point (maximum 2):</td>
<td></td>
</tr>
<tr>
<td>➢ Hand anomaly</td>
<td></td>
</tr>
<tr>
<td>➢ Heart anomaly</td>
<td></td>
</tr>
<tr>
<td>➢ Hypospadias +/- undescended testis</td>
<td></td>
</tr>
</tbody>
</table>
Note: The work outlined in the sections that follow was done by me unless otherwise stated in the beginning of the section.

2.2 DNA Isolation

Blood was collected in EDTA from affected proband(s) and non-affected family members for the purpose of DNA isolation. Blood in NaHep tubes was also collected for cases where there was no previous cell pellet available for cytogenetic FISH studies.

Each sample was first assigned a unique laboratory identification number. Blood was transferred to properly labelled 50 ml conical tubes and DNA extracted using the Puregene (Gaithersburg, MD, USA) DNA Isolation Kit using the manufacturer’s protocol.

2.3 DNA Quantification and Quality Control

To determine the concentration of DNA in solution, each sample was analyzed using a spectrophotometer (Pharmacia Biotech Ultraspec 3000). 2.5 ul of DNA was diluted in 97.5 ul of ddH₂O. One at a time, samples were loaded into the holding well of the spectrophotometer and three readings were obtained; absorbance at 260nm, absorbance at 280nm and ratio of the two readings.

To obtain the concentration of the sample, the following formula was used:

\[ \text{Formula: } \text{Abs}_{260}^* \times 40 \times 50 \]

(Where \( \text{Abs}_{260}^* \) is the absorbance by the sample at 260nm, 40 is the dilution factor and 50 is the DNA constant)
Example: For DNA sample with $\text{Abs}^{260}$ of 0.300:

$$[\text{DNA}]=0.300 \times 40 \times 50$$

$$[\text{DNA}]=600 \text{ ng/μl}$$

To ensure that it was not degraded, DNA was analyzed by electrophoresis on 1% agarose gel. To prepare the gel, 0.50g of agarose was mixed with 50ml 1 X TAE buffer and heated until dissolved. Standard genomic female DNA was also loaded and used as control. Gels were run at 100V for 30 minutes in 1X TAE buffer. All gels were observed under UV light and an image was captured and printed.
2.4 Array Comparative Genomic Hybridization (array-CGH)

2.4.1 The Principle of Array-CGH

In BAC array-CGH, cloned DNA fragments or BAC clones (±100-200kb in size) for which the exact chromosomal locations are known are spotted on a glass platform. Genomic DNA from subject (test) and sex matched reference (normal) DNA are differentially labelled with fluorescent dyes such as cyanine-3 (red) and cyanine-5 (green). The test and reference DNA are co-hybridized to the array platform where DNA sequences from both sources compete for their targets. Two experiments were performed for each patient: in the forward experiment (Experiment 1) the array slide was hybridized with a mixture of patient DNA labelled with Cy3 and reference (control) DNA labelled with Cy5, while in the reverse experiment (Experiment 2) patient DNA labelled with Cy5 was co-hybridized with reference (control) DNA labelled with Cy3, to the second array slide. DNA copy number changes are subsequently measured by evaluating Cy5 to Cy3 fluorescence signal ratios for each clone on the array.

For analysis, the software (Spectralware 2.0) reads the fluorescence signal intensity from every clone and transfers this information into a spreadsheet which links the data to clone names and genomic positions. The information for each clone is coupled to that of its duplicate spot (every clone is represented twice on the slide). The software normalizes the Cy5:Cy3 intensity ratios for each slide (experiment 1 and 2) and each data point. These normalized intensity ratios are computed for each clone and its repeat from both, and the pair of values from the forward and reverse experiment diverging most from the balance ratio of 1 for each clone is shown in a chromosome specific profile. This method allows for an easy visual identification of potential areas of gain or loss. Only clones which had
all 4 data points were further assessed by averaging the two sets of data per clone obtained in the forward and reverse experiments. The average value for the two data points for a clone in Experiment 1 is referred to as ratio 1 and the average value for the two data points for the same clone obtained in Experiment 2 is referred to as ratio 2. These averages (ratio 1 and 2) were then used to determine the validity of a clone imbalance. The clone was considered deleted if the ratio 1 value was <0.8 and ratio 2 value was >1.2; while a duplication was suspected if the ratio 1 value was >1.2 and ratio 2 value was <0.8. These cut-off values were previously determined empirically.

2.4.2 Array-CGH Protocol

Array CGH platforms and reagents were purchased from Spectral Genomics (Houston, TX, USA), and the manufacturer’s protocol was followed. Patient DNA was sex matched to reference DNA for all cases. The same male and female reference DNA were used for the duration of this study. The reference DNA consisted of pooled DNA from 7 normal individuals. DNA samples were quantified to obtain exact concentrations prior to each experiment using the method outlined above.

(1) DNA Sonication

To ensure efficient random prime labelling, each DNA sample was sonicated to produce smaller fragments:

1) X μl for 2 μg of test DNA was added to 200-Xμl of ddH2O; same was done for sex matched reference DNA. (where X= the volume of DNA)

2) The sonicator probe was placed directly into the samples and sonicated for 15 seconds.
(2) DNA Purification

Following sonication, samples were ‘cleaned’ in order to eliminate traces of protein, RNA and other impurities using DNA Clean & Concentrator™ from Zymo Research (Orange, CA, USA). DNA cleaning involved the following procedure using the manufacturer’s protocol:

1) 2 volumes of DNA binding buffer was added to each volume of DNA sample (400 μl for 200 μl of sample).

2) Samples were loaded into a spin column and columns were placed into a 2 ml collection tube.

3) Columns were centrifuged at 13,000 rpm for 10 seconds to draw the excess fluid into the collection tube.

4) 200 μl of DNA wash buffer was added to the columns. Columns were centrifuged at 13,000 rpm for 10 seconds. This process was repeated twice.

5) 27.5 μl of water was added to the columns. Columns were placed in 1-5 ml tubes and centrifuged for 30 seconds to elute the DNA. This process was also repeated twice, to increase the recovery of DNA from the column matrix. The final volume after elution was 55 μl.

6) Following purification, 5 μl of each DNA sample was analyzed by agrose gel electrophoresis to ensure adequate sonication. Optimal sonication produced homogeneous smears extending from 600bp to approximately 2 kb.

(3) Labelling

Each sample was divided into two tubes of 25 μl each. Tubes were labelled Cy3 and Cy5 respectively (at this stage, for each patient there were 4 tubes: Cy3 Reference, Cy3 Test, Cy5 Reference and Cy5 Test).

Master Mixes for each dye were prepared according to the following composition for each patient:
• 12.5 μl Labelling buffer (Spectral Genomics, Houston, TX, USA)
• 7.5 μl Cyanine 3-dCTP / Cyanine 5-dCTP fluorescent dye (Perkin-Elmer, Turku, Finland)
• 5.0 μl Klenow enzyme (Invitrogen BioPrime labeling kit) (Invitrogen Life Technologies, Carlsbad, CA, USA)

Master Mixes were prepared and kept on ice. Before adding the Master Mix to each designated tube the following steps were undertaken:

1) 20μl of 2.5x random prime mix (Invitrogen’s Bio Prime labelling kit) was added to each tube. Tubes were vortexed and centrifuged briefly.

2) DNA was denatured into single strands by heating each sample for 5 minutes at 100°C.

3) Following denaturation, samples were cooled in an ice water slurry for 5 minutes.

5μl of Cy5 Master Mix was added to each of the tubes labelled Cy5. Similarly, 5μl of Cy3 Master Mix was added to the tubes designated Cy3. Samples were incubated in a 37°C water bath for 1 hour. Following incubation, steps 2 and 3 outlined above were repeated. A second dose (5μl) of the Cy3/Cy5 Master Mix was added to each of their designated samples and incubated at 37°C for 1 hour.

Following the second hour of incubation, the reaction was stopped by adding 5μl of 0.5M EDTA (pH 8) to each tube. The enzyme was denatured using heat (72°C for 10 minutes). To ensure that the reaction had taken place, a 5μl aliquot from each sample was removed and analyzed by gel electrophoresis on 1% agarose gel.
(4) Hybridization

At this junction there were 4 tubes for each sample: Cy3 and Cy5 Test samples and Cy3 and Cy5 Reference samples. Differentially labelled DNA samples were combined together in the following manner:

The objective of this step was to facilitate the co-hybridization between the Cy3 labelled test and Cy5 labelled reference sample and conversely the Cy5 labelled test and Cy3 labelled reference samples.

To Precipitate DNA, the following procedure was carried out:

1) 45μl of hybridization buffer (Spectral Genomics, Houston, TX, USA) was added to each of the two tubes.

2) 12.9 μl of 5M NaCl and 130μl of isopropyl alcohol were added to each sample to initiate the precipitation reaction. Samples were incubated in the dark at room temperature for 20 minutes.

3) Samples were centrifuged at 13,000rpm for 20 minutes to obtain DNA pellets. Purple pellets were desired, indicating equal amounts of Cy3 and Cy5 labelled DNA (in experiments where the pellet(s) did not turn out to be purple, the pellets were discarded and a new experiment was initiated).
4) Supernatant was discarded and pellets were washed with 500μl of 70% ethanol. Samples were centrifuged for 3 minutes at 13,000rpm and supernatant was removed. Pellets were air dried (in the dark) for ten minutes.

5) DNA was resuspended in 10μl of sterile water.

6) Following resuspension, 30μl of hybridization buffer II (Spectral Genomics) was added to each sample.

7) Samples were incubated at 72°C for 10 minutes to denature DNA to single strands.

8) Samples were incubated in a 37°C water bath for 30 minutes following denaturation.

9) Each sample was applied to a separate array platform. The sample was pipetted as a line of liquid down the centre of the platform. A 22x60 glass cover slip was immediately placed over the slide to spread the solution over the entire face of the slide.

10) Each slide was placed in a hybridization chamber. 10μl of sterile water was added to the wells on both sides of each chamber to ensure that the inside of the chamber remains moist throughout the incubation period.

11) Chambers were placed in a Kapak pouch lined with wet paper towel (to prevent evaporation). The pouch was heat sealed and placed in an incubator at 37°C for 16 hours.

(5) Post Hybridization Washing

Each slide was washed in an individual Petri dish containing 25 ml of wash solution in a rocking platform incubator.

Washes:

   (1) At room temperature: 2X SSC, 0.5% SDS
       The slide was washed for 5 seconds by gently rocking the Petri dish, and the cover slip off the array.

   (2) At 50°C: 2X SSC/50% Formamide*
       The slide was incubated in rocking incubator for 20 minutes.

   (3) At 50°C: 2XSSC/0.1%Igepal*
       The slide was incubated in rocking incubator for 20 minutes.

   (4) At 50°C: 0.2X SSC*
       The slide was incubated in rocking incubator for 10 minutes.

   (5) At room temperature: double distilled water
Each slide was washed twice (for 5 seconds each time).

*wash solutions were heated and maintained at 50°C prior to use.

Following the second water wash, the slides were IMMEDIATELY blown dry with a stream of nitrogen gas. Slides were placed in a dark slide box to protect from light.

(6) Scanning

Scanning of the slides took place on the same day as the washing to ensure minimum loss of signal. Slides were scanned using a Scan-Array scanner (Perkin-Elmer, Turku, Finland) courtesy of the Prostate Centre located at Jack Bell Laboratories.

(7) Data Analysis

Data analysis was facilitated by Spectralware software Version 2 (Spectral Genomics, Houston, TX, USA). For each experiment, the average ratio and standard deviation (SD) were calculated for all spots on the array, spots with a ratio greater than 1.5 SD away from the mean were removed, and the average ratio and SD recalculated. Clones with an average ratio value of less than 0.80 in experiment 1 and average value of more than 1.20 in experiment 2\(^9\) were considered deleted, and those with a ratio value of 1.2 in experiment 1 and average values less than 0.8 in experiment 2, duplicated.

All of the clones selected as being deleted or duplicated based on their ratio averages were compared with the Copy Number Variant (CNV) list (http://projects.tcag.ca/variation/) to determine whether the clone has been previously reported as copy number variant in the general population. Clones that had not been reported as a variant were further investigated in the proband and their families.
2.5 Fluorescence in situ Hybridization (FISH)

Fluorescence in situ Hybridization (FISH) is a technique used for the identification of chromosomes and chromosomal loci using DNA probes that can be detected by fluorescence tagging. The technique can be used for chromosomes in both metaphase and interphase cells. In this study, FISH was used as a secondary, independent method to confirm array-CGH findings for two reasons. First, it is used to ensure that the deleted and/or duplicated clone(s) of interest found by the array is (are) not an experimental artefact(s) and that each clone maps to the chromosomal location outlined by the array analysis software (Spectralware) and bioinformatics websites such as NCBI (http://www.ncbi.nlm.nih.gov/) and UCSC (http://genome.ucsc.edu/) (it has been our lab’s experience that some clones are mis-mapped). Second, although array-CGH is an effective tool for finding quantitative changes in DNA copy number, it does not provide any information regarding possible structural rearrangements that the deleted/duplicated clone(s) may be involved in. FISH on the other hand is a tool that provides information regarding copy number changes as well as possible structural alterations such as cryptic translocations involving the clones of interest.

2.5.1 FISH Using Nick Translation

Direct labelling using the Nick Translation Kit (Abbott Molecular Inc. Des Plaines, IL, USA) was used for all FISH experiments. For each BAC clone of interest the following reaction mix was prepared to create the probe in an amount sufficient for 4 labelling experiments:
17.5-μl nuclease-free water (where X= the volume containing 1μg of BAC)  
2.5 μl 0.2 mM SR dUTP  
5.0 μl 0.1 mM dTTP  
10.0 μl 0.1 mM dNTP mix  
5.0 μl Nick Translation Buffer  
10.0 μl Nick Translation enzyme  
Total 50.0 μl

The reaction mix was incubated at 15°C for 16 hours. The reaction was stopped by heating samples at 70°C for 10 minutes. To ensure that the reaction had been completed successfully, 10 μl of the reaction mix was run on 1% agarose gel.

From the Nick Translation reaction (above), 8μl of probe was aliquoted. 2 μl of liquid Cot-i DNA and 4 μl of sdH2O were added to the reaction mix. Cot-i DNA is used in order to remove the repetitive sequence fraction of genomic DNA by binding to it, thus inhibiting the probe from binding to it. Probe was precipitated using 1.4 μl 5M NaCl and 30 μl of 100% iso-propanol. The reaction was left at room temperature for 20 minutes followed by centrifugation at 13000rpm for 20 minutes also at room temperature. The supernatant was removed and the pellet was rinsed with 70% ethanol. Pellet was centrifuged at top speed for 1 minute followed by removal of the supernatant. Pellet was air dried and re-suspended in 3 μl dsH2O and 7 μl Hybrisol (Q-BIOgene, Irvine CA, USA). To ensure that the pellet was not lost during precipitation, 1 μl of the sample was run on 1% agarose gel. Probe was then denatured at 75°C for 5 minutes and incubated at 37°C for 30 minutes to re-anneal.
2.5.2 Metaphase Slide Preparation

Note: Slides were prepared by Dr. Evica Rajcan-Separovic in cases where we had a very limited amount of cell pellet available for FISH studies. Otherwise, slides were prepared by me or Chansonette Harvard (Technician). Protocol for this method was provided by Dr. Rajcan-Separovic.

Metaphase slides using cell pellets were prepared one day prior to hybridization. To prepare the slide, 2 drops of fixative solution (3:1 ratio of methanol and acetic acid) containing cells were placed on a pre-cooled blank slide. The slide was then covered with fresh fixative and placed in a humid chamber. Once dry, the slides were evaluated for quality.

During the period in which the probe was being incubated (at 37°C), slides were placed in 70% formamide/2X SSC at 75°C for 5 minutes in order to denature the chromosomal DNA, followed by 3 consecutive ethanol washes using 70, 80 and 100% ethanol respectively. Slides were drained and air dried following ethanol washes.

2.5.3 Hybridization

Following 30 minutes of incubation at 37°C (see above), the probe was placed on the slide and covered with a coverslip. Rubber cement was applied around the edges of the slide cover to ensure a secure seal. The slide was placed in a closed container lined with moistened paper towels and incubated at 37°C for 16 hours.
2.5.4 Post Hybridization Washes

Slides were removed from the incubator and the rubber cement and cover slip were taken off. Slides were placed in a solution of 0.4 x SSC/0.3%NP-40 at 74°C for 2 minutes followed by placement in a solution of 2 x SSC/0.1%NP-40 at room temperature for 1 minute. Following the second wash, slides were drained and air dried. 14µl of DAPI was applied to each portion of the slide containing a FISH experiment and covered with a cover-slip. The cover-slip was sealed using nail polish. Slides were placed in a -20°C freezer for 30 minutes in order for the DAPI to be absorbed.

2.5.5 Slide Evaluation and Image Capturing

Note: To ensure accuracy, slides were evaluated and counted by me, Chansonette Harvard (lab technician) and Dr. Evica Rajan-Separovic for all cases.

Slides were evaluated using Zeiss Axioplan 2 fluorescence microscope and images were captured using MacProbe software (Applied imaging software, Santa Clara, CA). For each FISH probe, at least 10 metaphase cells were analyzed. In case of a duplication detected with the array, up to 100 nuclei were scored for the number of FISH signals/nucleus.

2.6 Dual (multi-colour) FISH

Dual FISH was performed using probe of interest and a centromere probe (Vysis, Abbott Molecular Inc. Des Plaines, IL, USA) in order to better identify the chromosome of interest. For dual FISH, all the steps outlined above were taken. In addition: 0.5 µl centromere probe +1.0 µl ddH₂O + 3.5µl CEP buffer were mixed and denatured at 75°C for 5 minutes. The probe was then combined with the pre-annealed BAC probe and mixed.
This probe cocktail containing both the centromere and BAC probes were applied to the metaphase slides.

2.7 Quantitative PCR (qPCR)

Note: All confirmations and parent of origin studies completed by qPCR experiments were performed by collaborators at Queen’s university. I was not involved in this process at any point, and the content of the following section was provided by Dr. Xudong Liu.

Array detected duplications unconfirmed by FISH were subsequently validated using 3-5 non-polymorphic markers within the duplicated clone using qPCR with SYBR Green I detection. Bioinformatics tools including the publicly available human genome databases such as NCBI (http://www.ncbi.nlm.nih.gov/) and UCSC (http://genome.ucsc.edu/) were used to design the primers. If the rearrangement was not detected by any of these markers, a further group of markers (2-4) were tested. If the second group of markers were negative, no additional testing was performed. Real time detection of PCR products was performed using an ABI Prism 7900HT system. The system allows one to see the threshold cycle (CT) during the experimental phase of amplification (ie., when none of the PCR reagents are limiting) and quantify each allele such that a single allele at a test locus in a person with a deletion would show 50% less amplification than in a person with 2 copies of that allele. Non-contiguous markers (markers from other chromosomal loci) were used as controls in each experiment. The ratio of amplification is approximately 1 (range 0.8-1.2:1) if both loci have the equivalent of 2 alleles. If the test locus is deleted on one chromosome, the ratio of test to control
markers is expected to be 1:2 (range of 1:1.7-2.5). If the test locus is duplicated, the ratio of the test to control markers is expected to be 3:2 (range of 2.6-3.5:2).
3. RESULTS

Note: The clinical assessment for all the patients in this study was conducted by Dr. Lewis at the BC Children’s and Women’s Hospital Medical Genetics clinic. Following her clinical assessment, we were able to categorize and select subjects to enter the study based on the phenotype scoring system described in detail in the material and methods section, prioritized for those subjects with phenotype scores $\geq 3$. Genetic tests including routine karyotype, Fragile X syndrome testing, FISH for del(22)(q11), del(22)(q13) and dup(15)(q11-13) (once it was available on a clinical service basis) were conducted by the clinical laboratories.

All array-CGH experiments were conducted in Dr. Rajcan-Separovic’s lab. I conducted the array experiments (including data analysis) and FISH validation studies outlined in this section. The karyotypes mentioned for the abnormal control cases were determined by the clinical lab.

Overall Findings

A total of 21 probands from 20 families were tested. Of these probands, two had a previously known cytogenetic abnormality and were used as positive controls. Of the 19 probands studied, four from three families (16%) were found to have a CNV of pathogenic or potentially pathogenic nature. Two probands belonged to a multiplex, multigenerational family and were counted as one case.
3.1 Array-CGH Findings

3.1.1 Array Validation Experiments Using Cytogenetically Abnormal Controls

Two probands diagnosed with an ASD and known cytogenetic anomalies detected previously by G-banding were used as positive controls. The chromosomal anomalies for these subjects included 1) del(22)(q13.3) based on the karyotype 46, XX, der(22)t(D or G;22)(p11;q13.1).ish del(22)(q13.3)(ARSA-) and 2) 15q11-13 duplication based on the karyotype 47, XY, +mar[11].ish dic(15;15)(q13;q13) (CEP15++, SNRPN++) [10]. Both abnormalities were clearly identifiable by the array ratio plots (Figure 1), thus validating our methodology.
Figure 1. Positive control profiles for chromosome 22 in a proband with a known 22q deletion (a) and for chromosome 15 in a proband with a known duplication of proximal15q (b).
3.1.2 Sub-chromosomal Imbalances Detected in ASD Probands

Nineteen patients from 18 families diagnosed with an autism spectrum disorder (ASD) using standardized Autism Diagnostic Observation Generic (ADOS-G) Scale and Autism Diagnostic Interview-Revised (ADI-R) criteria and who had a de Vries phenotype score of ≥3 based on clinical and morphometric scales detailed in Table 2 were selected. Array-CGH was performed to test for the presence of a CNV (micro-deletion and/or micro-duplication) that may harbour an autism susceptibility gene using 1 Mb array-CGH (Spectral Genomics (SG)). Two types of CNVs were observed: benign or pathogenic. Benign CNVs (bCNV) were defined as benign based on complete overlap with a CNV reported in controls from 2 or more studies in the Database of Genomic Variants (http://projects.tcag.ca/variation/ based on genome assembly build 36). CNVs were considered to be pathogenic (pCNV) if they overlapped with CNVs reported in controls from <2 studies, were de novo, had relevant genetic content, or were larger in size (criteria reviewed in Lee et al 2008)92. The presence of bCNVs were found for every patient, but were not further investigated. CNVs that have not been reported in unrelated normal cases but occurred in our ASD subjects and their parents were termed “potentially pathogenic” as it remains uncertain whether they are a predisposing factor for ASDs.

Following array-CGH analysis, four probands from three families (16% of cases) were found to have an array profile with a pathogenic or potentially pathogenic CNV. In Proband 1 and 2 (aunt and niece in a multiplex family, Family 1) multi-clone gains and losses on chromosomes 15q11-13 and 14q11.2 respectively were found. This was confirmed by FISH and determined to be the result of an unbalanced segregation of a familial cryptic balanced translocation t(14;15)(q11.2;q13.1). The results of a detailed
investigation on these two subjects and their family were published (Koochek et al, 2006)\(^{40}\) and are provided separately in chapter 5.

Proband 3 had a qPCR confirmed \textit{de novo} gain on 18p11.3 (family 2, Figure 2). qPCR studies also confirmed the size of the duplication to be 0.5Mb The same proband had gains at 20p12 and Xp22.2, which met array cut-off values but could not be confirmed by qPCR. Proband 4 had a single clone gain of 21q22.12 (Family 3, Figure 3). The gain was confirmed by qPCR and determined to be approximately 0.11Mb. The pathogenicity of this gain remains uncertain as it was determined to be inherited from a phenotypically normal mother. A summary of clinical findings and the genetic content of the confirmed CNVs for these probands are shown in tables 3 and 4.
Figure 2. Three clones were found to be duplicated in proband 3; (a) a single clone duplication at 18p11.3 (array values: 1.40 and 0.58), (b) a single clone duplicated at 20p12 (array values: 1.28 and 0.69) and (c) a single clone duplicated at Xp22.32 (array values 1.30 and 0.63).
Figure 3. In proband 4 a single clone corresponding to 21q22 was observed to be duplicated (array values: 1.49 and 0.69)
Table 3. Summary of the clinical findings for probands that were found to have a confirmed micro-deletion/duplication based on their array-CGH profile.

<table>
<thead>
<tr>
<th>Family</th>
<th>Clinical and Molecular Findings</th>
<th>qPCR/FISH confirmation of array-CGH Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family 2</td>
<td>• Autistic disorder confirmed by ADOS and ADI-R.</td>
<td>0.5 Mb dup(18)(p11.3), de novo in origin</td>
</tr>
<tr>
<td></td>
<td>• Paternal history of ASD (cousins with confirmed diagnosis of ASD and ADHD)</td>
<td>(qPCR used for confirmation and determination of size)</td>
</tr>
<tr>
<td></td>
<td>• Mild intellectual disability</td>
<td>dup (X)(p22.32), (unconfirmed, origin not determined)</td>
</tr>
<tr>
<td></td>
<td>• Normocephalic with large stature</td>
<td>dup(20)(p12) (unconfirmed, origin not determined)</td>
</tr>
<tr>
<td></td>
<td>• Slight facial asymmetry</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Broad and high nasal root</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Protuberant ears with thickened helices</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Slight malar flattening</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Downslanting palpebral fissures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Karyotype and all other cytogenetic and molecular tests normal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• De Vries Score of 4 (score based on non-Mendelian inheritance of ASD and ID (2) plus multiple dysmorphic craniofacial features (2))</td>
<td></td>
</tr>
<tr>
<td>Family 3</td>
<td>• Autistic Disorder confirmed by ADOS and ADI-R</td>
<td>0.11Mb dup(21)(q22.12) Maternal in origin</td>
</tr>
<tr>
<td></td>
<td>• Older brother suspected of having high functioning autism, possibly Asperger – no firm diagnosis</td>
<td>(qPCR used for confirmation and determination of size)</td>
</tr>
<tr>
<td></td>
<td>• Father is diagnosed with depression and has a sister with bipolar disorder.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Mother has idiopathic epilepsy which requires chronic Valproate and Carbamazepine therapy.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• De Vries Score of 5 (Score based on Mendelian history of ASD/ID (1), macrocephaly (1), post natal growth abnormality (1) and multiple dysmorphic craniofacial features (2))</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Summary of ASD cases with confirmed CNVs of potentially pathogenic and uncertain clinical relevance

<table>
<thead>
<tr>
<th>Proband</th>
<th>SG Clone</th>
<th>Array- CGH Finding</th>
<th>Cyto band</th>
<th>OMIM genes</th>
<th>Associated phenotype</th>
<th>Origin</th>
<th>Pathogenicity in ASD</th>
</tr>
</thead>
</table>
| 3       | RP11-105C15  | Gain               | 18p11.3   | L3MBTL4
EPB41L3 (DAL-1)            | None
Primary non-small cell lung carcinoma | De novo   | Potentially pathogenic   |
| 4       | RP11-79A12   | Gain               | 21q22.12  | DSCR1
KCNE1                     | Down Syndrome
Jervell and Lang-Nielsen Syndrome | Maternal  | Uncertain               |
4. DISCUSSION

The majority of children diagnosed with an ASD suffer from an idiopathic form of this disorder. Given the evidence for a strong genetic component in ASD, it was the aim of this study to investigate further the prevalence of submicroscopic genomic aberrations that may harbour autism susceptibility genes and/or loci using 1 Mb CGH-microarrays. To reduce heterogeneity in our cohort, a clinical and dysmorphology assessment scale leading to a measurable phenotype score was used to characterize and refine a phenotypically homogeneous subgroups. Of the nineteen patients with apparently normal karyotypes investigated, three families were confirmed to have affected individuals with at least one clone deleted and/or duplicated of potential pathogenic consequence.

The array profile for proband 3 revealed single clone gains in 3 areas: 18p11.3, 20p12 and Xp22.32. We were unable to confirm the two latter gains, thus the contribution of these genomic changes to the proband’s ASD phenotype cannot be commented upon with any degree of certainty. The gain at 18p11.3, confirmed by qPCR, was determined by this method to be approximately 0.5 Mb in size and de novo in origin. The duplicated clone overlaps the L3MBTL4 gene. However, there is virtually no information regarding the biological function or pattern of expression of this gene and little can be said with regards to its contribution to the ASD phenotype. This gain was determined (by qPCR) to extend beyond the boundaries of one clone. There is a small overlap between the duplicated area and EPB41L3 (erythrocyte membrane protein band 4.1-like 3). This gene, also known as DAL-1 is normally expressed at high levels in the brain with lower levels of expression in kidney, intestine and testes. DAL-1 is lost or inactivated in ~60% of primary non-small cell lung carcinoma tumours. In addition, loss of heterozygosity (LOH) for the
DAL-1 locus at 18p11.3 has been found in lung, breast and brain tumours. Gutmann et al. observed the loss of DAL-1 in 60% of sporadic meningiomas. Given that DAL-1 has a similar protein structure to that of the neurofibromatosis 2 protein product (NF2 on 22q) which has been well documented for its involvement in meningioma tumorigenesis; it is believed that DAL-1 possesses growth-suppressing properties. It is possible that the overexpression of DAL-1 as a result of a gain (as is the case in this proband) may have an entirely different phenotype than the loss of DAL-1, and, although there is no current evidence in the literature suggestive of the involvement of DAL-1 in neurodevelopmental disorders such as ASD, given the size, origin and lack of evidence that this clone is a benign CNV found in the general population, we cannot rule out a potential pathogenic contribution to this subject’s phenotype.

The gains of clones RP11-366M24 and RP5-873P14, correspond to Xp22.32 and 20p12, respectively, remain unconfirmed (by both FISH and qPCR) but neither one has been found as a benign CNV in the general population. There are a few possible reasons as to why these clones could not be confirmed. The observed duplications on the array profile may be an experimental artefact, rendering them false positive. Alternatively, since array-CGH compares relative DNA quantities of a reference sample with the subject’s DNA, if the number of copies for a specific segment of DNA is greater in the subject (eg 2) compared to the reference sample (which theoretically can be 1), this will show up as a gain on the profile, but is not a true gain, and instead reflects the variability in the number of copies between the control pool and the patient. If the gain is a result of a small (less than 1 Mb) tandem duplication, interphase or metaphase FISH cannot be used to confirm due to poor resolution. If the gain involves only some sections of a given DNA segment, it may
not be detected by qPCR as typically only a small number of short sequences within non-repeated DNA segments within each region are used for analysis. The array profile for proband 4 revealed a single clone gain (RP11-79A12) corresponding to 21q22.12 by array-CGH (Figure 3). This gain was subsequently confirmed by qPCR in both the proband and his unaffected mother. This clone overlaps with KCNE1 and DSCR1 (UCSC genome browser). KCNE1, also known as potassium voltage gated channel subfamily E member 1, is a protein that assembles as a beta subunit of voltage gated potassium channel complex. It modulates the gating kinetics and enhances stability of the channel complex. Mutations in this gene are associated with Jervell and Lang-Nielsen Syndrome (JLNS1) (OMIM # 220400), an autosomal recessive disorder characterized by congenital deafness, prolongation of QT intervals, syncopal attacks due to ventricular arrhythmias and high risk of sudden death. Based on this proband’s clinical and phenotypic characterization, the gain of 21q22.12 at the location of this gene has not resulted in any of the clinical features associated with this disorder. The second gene in this region is DSCR1 (Down syndrome critical region gene 1) (OMIM # 602917) whose protein product is found in the human brain and heart and whose overexpression in the brain as a result of an extra copy of the gene such as in Down syndrome, leads to the inhibition of the calcineurin signalling pathways that is involved in both short and long term effects on neuronal excitability and function.

DSCR1 could indeed be involved in the etiology of ASD, given its pattern of expression and biological function. Further, Christian et al. have also reported one ASD proband with an inherited dup(21)(q22.12) spanning 210kb. Although it is not clear from their data whether their reported CNV overlaps with the one seen in proband 4, it makes
this an interesting region for further analysis. In this specific case however, the contribution of the CNV and DSCR1 to the patient’s ASD is not clear because it was confirmed to have been inherited from an unaffected parent. Presumably benign genomic variation overlapping this region has been reported in one person (Database of Genomic Variant- http://projects.tcag.ca/variation/) leaving the contribution of this CNV to ASD aetiology undetermined until further evidence becomes available supporting or refuting its role.

This case is further complicated by the fact that the proband was exposed to Valproate (VPA) throughout nine months of gestation required for treatment and seizure prevention of his mother’s chronic epilepsy. Developmental delay characterized by low verbal IQ as well as autism have been reported in children exposed to this drug in utero99,100. Many factors may contribute to the overall developmental outcome of a child exposed to VPA, including number of drugs co-administered, drug dosage, differences in maternal and infant metabolism, gestational age of fetus at exposure and hereditary susceptibility99. VPA has also been shown to alter the cell’s epigenetic state by acting as a histone deacetylase inhibitor (HDAC) as well as triggering demethylation of promoter regions101,102. The dynamic interrelation of DNA methylation and histone acetylation are essential for regulation of gene expression at the promoter level, and the alteration of their state by VPA results in deregulation of gene expression and gene dosage in some genes.

It is difficult to tease out whether this proband’s ASD is a consequence of genetic or environmental insults or a combination of both. Since his mother, who is not affected by ASD or any other neuropsychiatric disorder, also carries the single clone duplication at 21q22, the contributor to the autistic phenotype is more likely to be in-utero exposure to VPA. One could speculate that the sub-microscopic genomic imbalance (gain of 21q22)
could possibly have a potentiative effect or may increase the susceptibility of this proband for developing an ASD phenotype through gene-environment interaction.

Our full understanding of the role of CNV with respect to phenotypic variation and the aetiology of ASD remains limited. What is intriguing is that thousands of putatively functional sequences, including known disease-related genes, flank or fall within many of the CNVs unmasked in normal populations thus far. Redon et al.\textsuperscript{103} reported that 58% of the 1447 CNV regions they mapped through the study of 270 individuals from the HapMap collection overlapped known RefSeq genes, and over 99% overlapped conserved non-coding sequences. Furthermore, they found that 285 out of 1961 (14.5%) genes in the OMIM morbid map overlapped with CNVs. CNVs in normal population studies have also been identified within regions commonly deleted in genetic disorders such as DiGeorge and Williams-Beuren syndrome. What seems to be emerging out of the studies of normal and ASD populations is a subset of CNVs that are more common in ASD probands and their families. Data comparison between the five studies published so far\textsuperscript{15-19} reveal two major groups of ASD CNVs: the first is composed of those occurring in regions or genes previously implicated in ASD including 2q37.2-q37.3, 15q11-q13, 22q11-13 N\textit{LGN4} and \textit{SHANK3}, further lending support to the involvement of these regions and genes in the aetiology of some ASD cases. The second group consists of CNVs that have been unmasked but only once or in only a few cases, such as 13q21.32 (\textit{PCDH9}-Protocadherin 9) and 16q24.3 (\textit{ANKRD1} -Ankyrin repeat domain-containing protein -1)\textsuperscript{17}. These CNVs fall across all autosomes and X (but not Y) and are unique to the individual or family. In fact, in this second group, there are very few CNVs that have been observed to occur in several non-related individuals.
In the course of our investigation we identified a submicroscopic gain of 15q11-13, in two probands (Proband 1 and 2) from the same family. However, contrary to the typical and recurrent chromosomal rearrangements leading to gain of proximal 15q in ASDs (such as interstitial duplication of this region or presence of a marker chromosome), the gain of proximal 15q was due to an unbalanced segregation of a familial submicroscopic translocation between two acrocentric chromosomes. I have described this unique finding in the following chapter and have published my findings.
5. Cryptic Familial Translocation t(14;15)(q11.2;q13.3) Detected Using array-CGH

Note: This case may also be reviewed in its published form.


Overall Findings

Probands 1 and 2 (Family 1), were identified to have two identical array profiles. Through FISH studies, we characterized a familial translocation giving rise to the duplication of 15q11-13, which is well established as the leading chromosomal abnormality in the aetiology of ASD.

5.1 Pedigree and Clinical Findings

Family history was positive for autism on the maternal side across two and possibly three generations (see pedigree; Fig.4). The clinical presentation of the two probands is provided in Table 5. The mother (II:2) of the youngest autistic proband, (proband 2, III:1), and sister of autistic proband 1 (II:3) reported having mild learning problems as a child but did not have symptoms of autism. III:1’s maternal grandmother (and II:2’s and II:3’s mother), I:2, had a sister (I:3) who died in her 40’s with learning disabilities, psychiatric illness and symptoms of autism, the latter of which were never evaluated definitively. III:1’s father, sibling and maternal uncle presented with no developmental or other concerns.
Figure 4. A multi-generation pedigree illustrates the incidence of autism in this family. Array-CGH studies were performed for probands 1 and 2 (II:3 and, III:1) as well as for II:2 (unaffected). I:2 and II:2 were found to be carriers of a cryptic balanced translocation t(14;15)(q11.2;q13.3). Both I:2 and II:2 transmitted the translocation in an unbalanced manner, leading to the gain of proximal 15q in probands 1 and 2.
Table 5. Summary of the main clinical findings observed in II:3 (proband 1) and III:1 (proband 2).

<table>
<thead>
<tr>
<th>Proband</th>
<th>Clinical and Molecular Findings</th>
<th>Confirmed Array-CGH Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Female</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Diagnosed with ASD at age 28 (38 at the time of examination in 2005).</td>
<td>dup(15)(q11-13) and del(14)(q11.2)</td>
</tr>
<tr>
<td></td>
<td>• Intellectual disability (IQ 35-50)</td>
<td>FISH confirmation of unbalanced</td>
</tr>
<tr>
<td></td>
<td>• Language disorder</td>
<td>t(14;15)(q11.2;q13.1) leading to the gain of 15q and loss of 14q.</td>
</tr>
<tr>
<td></td>
<td>• Aggression and verbal abuse against others</td>
<td>Inherited from mother who has a balanced</td>
</tr>
<tr>
<td></td>
<td>• Hoarse voice with articulation problems</td>
<td>t(14;15) (q11.2;q13.1)</td>
</tr>
<tr>
<td></td>
<td>• Consistent gaze avoidance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• General seizure disorder</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Coarsened facial features</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Down-slanting palpebral fissures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Broad and high nasal root</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Normal karyotypes for both probands</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Phenotype Score of 5 (Score based on non-Mendelian history of ASD/ID (2), multiple dysmorphic craniofacial features (2) and seizures(1))</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Female</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Diagnosed with autistic disorder</td>
<td>dup(15)(q11-13) and del(14)(q11.2)</td>
</tr>
<tr>
<td></td>
<td>• Intellectual disability</td>
<td>FISH confirmation of unbalanced</td>
</tr>
<tr>
<td></td>
<td>• Significant language disorder</td>
<td>t(14;15)(q11.2;q13.1) leading to the gain of 15q and loss of 14q.</td>
</tr>
<tr>
<td></td>
<td>• Frequent gaze avoidance</td>
<td>Inherited from mother who has a balanced</td>
</tr>
<tr>
<td></td>
<td>• Dysmorphism</td>
<td>t(14;15) (q11.2;q13.1)</td>
</tr>
<tr>
<td></td>
<td>• Down-slanting palpebral fissures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Deep set eyes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Facial asymmetry</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Nasal root and bridge broad and high respectively</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Phenotype Score of 4 (Score based on non-Mendelian history of ASD/ID (2) and multiple dysomorphic craniofacial features(2))</td>
<td></td>
</tr>
</tbody>
</table>
5.2 Array CGH Findings

Analysis of ratio profiles for II:3 (proband 1) and III:1 (proband 2) revealed three identical chromosomal alterations, including a seven clone duplication of proximal 15q (Fig. 5a), a two clone deletion in 14q (Fig. 5b), and a two clone duplication of 6q sequences (Fig. 5c). Based on the array findings, the duplicated region of chromosome 15q11-15q13 was found to extend approximately 10Mb, with RP11-80H14 at 20.5 cM and RP11-420B6 at 30.7 cM setting the minimal boundaries. The 14q11.2 deletion, encompassing clones RP11-98N22 and RP11-89F2, is at least 0.7Mb in size. The two clone gain found at 6q22.31 is at least 0.4Mb, and includes clones RP11-80B14 and RP3-329N18. The gain of 15q and loss of 14q clones were not found on the array profile of II:2, although she did have a gain of the same two clones from 6q22.31.
Figure 5. The ratio plots from array CGH for chromosomes 15, 14 and 6 for proband 1 (II:3) and proband 2 (III:1) showed the exact same chromosomal changes after array analysis (only profile from proband 1 illustrated here). (a) A seven clone duplication (RP11-80H14, RP11-385H1, RP11-339C21, AC090696.6, RP11-368G21, CTB-21M6 and RP11-420B6) on 15q11-13 resulting in a gain of ~10Mb; (b) A two clone deletion (RP11-98N22 and RP11-89F2) on 14q11 was ~0.7Mb in size and (c) A two clone duplication (RP11-80B14 and RP3-329N18) on 6q22 resulting in a gain of ~0.4Mb.

63
5.3 Fluorescent in Situ Hybridization (FISH) analysis

**Note:** Chromosome identification for reverse G-banding and subsequent karyotypes were done by Dr. Evica Rajcan-Separovic.

FISH analysis of metaphase chromosomes from subjects II:3 (proband 1) and III:1 (proband 2) was performed. The 15q gain was confirmed using the most proximal and distal clones on either side of the seven clone duplication (RP11-80H14 and RP11-420B6). Three distinct signals on three larger acrocentric chromosomes were observed for the proximal clone RP11-80H14 (Fig. 6a). For the distal probe RP11-420B6, one of the three signals was less intense, which suggested that the breakpoint region of the duplication was within this most distal clone (Fig. 6b). The three signals mapped to two normal chromosome 15s and a derivative chromosome 15 (this fact was later established through reverse G-banding and multi colour FISH analysis). The deletion on chromosome 14 was confirmed using clone RP11-89F2 in both probands. A single chromosome 14 showed hybridization of this clone in each of the 10 metaphase spreads analyzed (Fig. 6c). To determine the parental origin of the abnormalities, FISH was performed for the mothers (I:2 and II:2) of both affected probands since family history suggested maternal transmission of the rearrangement. Probes RP11-80H14 for chromosome 15 and RP11-89F2 for chromosome 14 were used to determine the presence of duplication and deletion in subjects I:2 and II:2 respectively. For each probe tested, two signals on two acrocentric chromosomes were observed in both I:2 and II:2 (Figs. 7a,c); thus no evidence of 15q gain or 14q loss was observed. The identity of the acrocentric chromosome was determined using reverse DAPI banding (Fig 7 b and d). Based on the family history and collective
array, and FISH findings, it seemed likely that a cryptic balanced translocation t(14;15)(q11.2;q13.3) was present in the grandmother (I:2) and was transmitted in an unbalanced form to her affected daughter (II:3, proband 1) and in a balanced form to her unaffected daughter (II:2) who subsequently transmitted the translocation in an unbalanced form to her daughter (III:1, proband 2). Two color FISH was used to confirm this hypothesis (Figure 8).
Figure 6. FISH confirmation of 15q gain and 14q loss in the probands. The gain of 15q was confirmed using the proximal and distal clones on either side of the duplicated region: (a) signals on three acrocentric chromosomes were observed for the proximal clone RP11-80H14 in both probands. (b) For the most distal probe RP11-420B6, one of the three signals was less intense, suggesting that the breakpoint of the duplication resided within this clone. (c) RP11-89F2 was used to confirm the deletion on chromosome 14. Only one signal was present in all 10 metaphase spreads analyzed for both probands.
5.4 Reverse G-banding and Dual Colour FISH Analysis

Reverse G-banding following FISH with RP11-80H14 (chromosome 15) and RP11-89F2 (chromosome 14) confirmed the suspicion that I:2 and II:2 are carriers of a balanced translocation t(14;15)(q11.2;q13). (Figs. 7b,d). This was further confirmed by two colour FISH using a centromeric 15q probe (Vysis, Abbott Molecular Inc.) and a probe mapping to the distal end of chromosome 14 (RP11-73M18) in I:2 and II:2. In both individuals, the latter probes hybridized to the distal ends of two acrocentric chromosomes, one of which was also marked by the 15 centromere probe, thus indicating the translocation of the 14q arm onto the centromere and proximal end of 15. The translocation was detected in all 10 metaphase cells examined in both I:2 and II:2 (Figs. 8a,b). To determine whether other members of the family were carriers of the same cryptic translocation, the two colour FISH assay was performed for II:4 and III:2. Results showed four signals on four separate acrocentric chromosomes indicating that neither individual is a translocation carrier (results not shown).
Figure 7. FISH analysis for subjects I:2 and II:2 (the mothers of the two probands) is illustrated. Probe RP11-80H14 (a) showed two signals on two acrocentric chromosomes. To determine the identity of the chromosomes, reverse G banding (b) was performed on the same metaphase spread. Upon examination of the banding pattern, it was determined that the 15q specific probe mapped to one normal 15 and a derivative 15 (der(15)t(14;15)) which had a chromosome 14 banding pattern since it consists of centromere and proximal end of chromosome 15 and most of the long arm of chromosome 14 (all other acrocentric chromosomes were also identified).
Figure 7 (Continued): Probe RP11-89F2 (c) was used to test for the presence of a deletion on 14q. Two signals on two acrocentric chromosomes were observed. Reverse G-banding on the same metaphase spread (d) revealed the presence of one normal copy of 14 and a derivative 14 (der(14)t(14;15)) containing the centromere and only the proximal end of chromosome 14 fused to the long arm of 15.
Figure 8: Presence of a cryptic balanced translocation was confirmed through dual FISH in I:2 and II:2. The centromere 15 probe (green) and a probe mapping to the distal end of chromosome 14 (red) were used. The latter probe hybridized to two acrocentric chromosomes, one of which was marked by the centromere 15 probe, indicating the translocation of the 14q arm on to the centromere and proximal end of 15 (der15).
Figure 9. Illustration of the mechanism by which the presence of two derivative chromosomes (der(14) and der(15)) in the translocation carriers has lead to the gain of proximal 15q and loss of proximal 14q in the probands.
A re-examination of the pericentromeric regions of the # 14 chromosomes in the G
banded karyotypes of II:3 (proband 1) and III:1 (proband 2) suggested that both individuals
had an expansion of the region below the centromere, which was interpreted originally as a
pericentromeric polymorphism. Based on the array CGH, FISH and reverse banding
analysis, the karyotypes of proband 1 and 2 who carried the derivative chromosomes 15
resulting in cryptic chromosomal imbalances of proximal 14 and 15 were revised to:
46,XX,-14,+der(15)t(14;15)(q11.2;q13.3) mat. ish der(15)(RP11-80H14+, RP11-420B6+, RP11-89F2-)mat.
5.5 Discussion

Chromosome 15q11.2-13 is a hotspot for rearrangements including deletions, duplications, triplications, inversions and translocations due to the large number of low copy repeat sequences in this region. However, these rearrangements are primarily intrachromosomal with deletions or duplications involving only the altered 15 homologue\textsuperscript{104}. Deletions of 15q11.2-13 occurring as a result of translocation events with breakpoints within proximal chromosome 15q have been previously reported\textsuperscript{105,106}. Translocations resulting in a duplication such as the one reported here have not been previously described.

So far several mechanisms for the duplication of 15q11-13 have been proposed, including duplicon mediated unequal crossover leading to reciprocal duplication, unequal sister chromatid exchange and recombination between inverted repeats leading to inversion duplications (also known as isodicentric duplications)\textsuperscript{107}. The gain of 15q11-13 and autism occurring as a result of a cryptic unbalanced translocation with breakpoints within the pericentromeric area of an acrocentric chromosome has not been previously reported and deserves consideration as a causative mechanism predisposing to gain of proximal 15q in association with autism. The gain of 15q11-13.3 described here occurred as a result of an unbalanced cryptic translocation t(14;15)(q11.2;q13.3) transmitted through the maternal line and is the underlying cause of autism in both probands.

This region of proximal 15q is under imprinting control. Autism occurs when there is a maternally inherited gain of 15q11-13 (including the PW/AS critical region). Some studies have also found a deletion (maternally-derived) of proximal 15q occurring in conjunction with autism\textsuperscript{55,57}. In a population-based study, Steffenburg et al.\textsuperscript{57} found that all
four children in their study with a maternally inherited deletion of proximal 15q causing Angelman syndrome (AS) also met full criteria for the diagnosis of autism. Interestingly, none of 11 PWS cases due to paternal deletion of proximal 15q could be identified as having clear-cut autistic disorder. It appears from this example that, like AS, autism is caused primarily by maternally derived aberrations of proximal 15q. Thus, a dosage change, which alters the normal monosomic functional state of the maternally imprinted proximal 15q11-q13 genes, can result in autistic like behaviours. In PWS the deletion is on the paternal chromosome and maternally expressed genes are not affected. This interpretation is consistent with reports of phenotypically normal mothers who carry a duplication on their paternally derived chromosome 15, subsequently transmitting the same duplication to their autistic children\textsuperscript{52, 55}.

The two clone loss from 14q11 is approximately 0.7Mb. This gene rich region includes several of the olfactory receptor gene families as well as genes involved in cell cycle progression (Cyclin B1 interacting protein), DNA repair enzymes (ADPRTL2-ADP-Ribosyl transferase-like 2 and APEX- Apurinic endonuclease) and several members of the ribonuclease (RNase) superfamily. Although there are no genes that appear to be involved in neurogenesis and/or neurodevelopment, position effects of this translocation or deletion on the expression of important down stream genes, or the unmasking of recessive or imprinted (thereby functionally nullisomic) alleles on the intact homologue cannot be ruled out. Recently, Zahir et al.\textsuperscript{108} reported 3 cases of de novo 14q11.2 deletions in children with developmental delay, cognitive impairment and mild dysmorphic features. They determined the deletions to be 101 kb, 1.6 Mb and 1.079 Mb and of de novo origin. There was a 35 kb common deletion region between these three cases. In our probands, the
14q11.2 deletion falls within the broader deletion region reported by this study (ie. the deletions they reported fall within the 19 584 863 to 20 998 178 bp region and the two consecutive clones found in our probands are from 19 570 808 to 20 341 734 bp) but falls outside of the common deletion region (20 896 740 to 20 931 827 bp). It is very likely that the deletion of 14q11.2 is also a contributor to the dysmorphic features and developmental delay seen in our probands. The 0.4 Mb duplication found on chromosome 6 appears to be an inherited and potentially benign CNV in this family because unaffected members of the family also have the duplication. One of the two clones from 6q duplication (RP3-329N18) has also been previously reported in the normal population (http://projects.tcag.ca/variation/). It contains the TCBA1 gene (T-cell Lymphoma Breakpoint –Associate Target 1), which has been associated with developmental delay and neurological phenotypes such as epilepsy in patients with translocations that have had breakpoints within this gene109,110. The other clone (RP11-80B14) contains the TRDN gene (Triadin) which is involved in calcium release in muscles111 and has not been previously reported as normal variant. One cannot eliminate the possibility of this duplication as being linked to an underlying autism susceptibility locus for this complex genetic disorder or another phenotype in the family. Moreover, little is known regarding the contribution of suspected normal copy number variants in subjects in whom the presence of other genetic alterations may predispose that individual to a more vulnerable disease state. Thus, although the recognizable phenotype of autism in these probands is attributed to the gain of 15q11-13, the accompanying chromosomal changes involving 14q11 and 6q22 may contribute to the overall phenotype. This consideration is important given that autism is a
heterogeneous and multigenic disorder resulting from the combination of several susceptibility genes.

Surprisingly, both probands were determined to have a normal karyotype at a resolution which should allow detection of 5-10 Mb changes. The discrepancy between the results of routine cytogenetic analysis and research findings using array-CGH that identified a ~10Mb gain of proximal 15q is attributed to the naturally occurring and well-recognized variation typically seen in the pericentromeric areas of acrocentric chromosomes, which represents the transition area between euchromatin and centromeric heterochromatin\textsuperscript{112}. In these two affected individuals this “pericentromeric variability” occurred on a chromosome which was identified by initial G banding as 14, as it contained only the centromere and a small segment of proximal 15q, while the rest of this chromosome derived from the long arm of 14. Pericentromeric variability of acrocentric chromosomes is not known to be associated with a clinical phenotype, except for chromosome 15, in which cytogenetic expansion or contraction of the pericentromeric end can be the cause of autism or PW/AS, respectively. However, expansion of proximal 15q can also be associated with a normal phenotype, as described by Ritchie et al.\textsuperscript{113}, who reported that a partial duplication of GABRA5 was present in up to four copies in the pericentromeric region of chromosome 15 in unaffected individuals. The degree of amplification reached up to 20 copies in individuals with a cytogenetically identifiable elongation of the 15q region. This variable region, which also includes immunoglobulin heavy chain (IgH)\textsubscript{D}, neurofibromatosis (NF\textsubscript{1}) and \textit{BCL8A} pseudogenes\textsuperscript{114}, is roughly 1 Mb in size and may be repeated several times on the same chromosome, leading to different banding morphologies among individuals. An interesting observation was made
by Cockwell et al.\textsuperscript{115}, who detected cryptic rearrangements involving acrocentric pericentromeric regions in 5/100 phenotypically normal individuals who had 3 or more spontaneous miscarriages ((50 recurrent miscarriage couples). They suggested that these abnormalities may be causing abnormal pairing configurations during meiosis leading to non-disjunction during metaphase I. The finding that a suspected pericentromeric variability of chromosome 14 was in fact a result of a gain of proximal 15 and a loss of proximal 14, emphasizes the need that suspicious size change of the pericentromeric region of any acrocentric chromosome detected by G banding in individuals affected by autism should be further investigated by analyzing parental karyotypes and FISH with probes from the 15q proximal region to rule out the possibility of a cryptic translocation leading to gain of proximal 15q.

The benefit of our findings to the family is one of the most important aspects that resulted from this research. Although the patients' management may not change very much in the light of these findings, the determination of cause for their underlying genetic susceptibility to autism has important implications for anticipatory medical and genetic recurrence risk counselling, especially since proband 2 has a younger sister who through this research was determined not to be a translocation carrier.

Furthermore, this case is a classic example of the power of whole genome array CGH in detecting clinically relevant copy number changes in individuals with ASD and normal karyotypes. The use of the technology is becoming widely accepted for routine clinical testing, due to its ability to detect potentially pathogenic submicroscopic copy number changes (in addition to larger cytogenetically visible changes). There are some important factors to consider however, including the resolution of the array, sensitivity,
specificity, false negative and positive rates and interpretation of CNV data. For example, higher resolution platforms such as tiling arrays may detect more pathogenic CNVs but they also detect many other benign variants that must be distinguished from disease causing variants. This alone makes the interpretation of the results difficult since we are not currently able to distinguish between benign and pathogenic CNVs for all cases, nor do we fully understand the relationship between many presumably "benign" CNVs and ASD. On the other hand, targeted or diagnostic arrays representing a specific area of the genome will detect far fewer benign CNVs but will also fail to detect any pathogenic CNVs that fall outside of the specific region of the genome represented on the array. This type of array is not likely suitable for genetic analysis of ASD patients since there are already many different CNVs that cause or predispose to ASD and most of them are not represented on current targeted arrays.
6. Conclusions and Further Directions

Over the last few years, array-CGH have greatly contributed to our understanding of the aetiology of ASD and has given us a glimpse of the genomic diversity present in both normal and patient populations. Our global understanding of the aetiology of ASD remains limited, with no one cogent or unifying theory to explain how growing numbers of uniquely identified CNVs may converge to explain autism susceptibility or uncover common gene/metabolic pathways or gene networks. What is becoming more evident, however, is the importance of deeper understanding of the cryptic imbalances and submicroscopic chromosomal anomalies such as CNVs and how they affect gene expression by affecting coding regions or non-coding regulatory regions. To grasp the full spectrum of these genomic changes in a disorder as complex as ASD will require unmasking and characterization of ASD-specific CNV in many more probands and their families.

It was hypothesized that new pathogenic CNVs will be identified in subjects with idiopathic ASD. This hypothesis was tested in 19 patients with a confirmed diagnosis of autism and normal karyotype using 1Mb array-CGH. To reduce heterogeneity in the sample population, a scoring system that facilitated the distinction between patients with a more complex (but still idiopathic) forms of ASD (score >3) and those that are more simple idiopathic forms (without dysmorphology or family history of ASD/ID) was used.

The results from the 19 families with ASD subjects examined revealed three probands of interest. In one family we were able to determine the genetic cause and novel genomic mechanism leading to ASD in a multigenerational family by determining the presence of a cryptic translocation leading to the gain of 15q11-13 that had gone undetected by conventional karyotyping in both affected family members. This case further
establishes the importance of the use of high resolution technology for diagnosis of ASD and validation studies via FISH to rule out cryptic translocations or other inter-chromosomal rearrangement. In the two remaining ASD probands, we were able to detect and confirm ASD specific CNVs that are unique to the probands and/or their families. Both CNVs contain genes that maybe interesting to examine further for their involvement in ASD. One such gene is DSCR1 from ....can not remember the abnormality which has a biologically relevant function and pattern of expression. Future studies should screen for mutations and copy number changes in this gene and others belonging to the calcineurin signalling pathway. Based on the findings here, CNVs do contribute to a portion of idiopathic ASD cases. These results are consistent with findings in the literature which puts the rate of autism specific CNVs at 7-28% in idiopathic cases15-19.

For those CNVs that appear interesting but could not be confirmed, future studies may have to resort to other methods of confirmation such as other array-CGH platforms of higher resolution, or sequencing. Once it has been established that the CNVs are true positives, they should also be tested for in the greater ASD population. It is essential to be able to distinguish between ASD-specific and population-specific variation. Thus it is imperative to have a well designed database and algorithm that can maintain the information and cross reference between the two populations.

Given the complex nature of ASD with many genes playing a role in the overall phenotype, there are most likely many, many CNVs that have yet to be uncovered and their interactions explored. Rare variants could in fact be very valuable as they may point to novel pathways not previously suspected in ASD. Only through continuous data gathering using high resolution methods such as array-CGH in combination with comprehensive
clinical analysis and subgrouping of subjects can we begin to unravel the mystery of the ASDs.
References


110. Yue, Y. et al. Disruption of TCBA1 associated with a de novo t(1;6)(q32.2;q22.3) presenting in a child with developmental delay and recurrent infections. *J Med Genet* **43**, 143-7 (2006).

