

DNA Copy Number Variation in Autism

A Senior Honors Thesis

Presented in Partial Fulfillment for the Requirement for graduation
with research distinction in Biology in the undergraduate colleges of
Biological Sciences at The Ohio State University

By

Ashwin Adur

The Ohio State University

June 2009

Project Advisor: Dr. Carlos Alvarez, Department of Pediatrics

Abstract

Autism is a childhood neurodevelopmental and psychiatric disorder that involves the impairment of social skills and communication. The genetics behind the inheritance and susceptibility of this disorder are not well known but have recently been studied. Lately, DNA Copy number variation (CNV) in autistic individuals has been a topic of high interest, and has been shown to have an association with this disorder. CNV generally refers to differences in the germ line DNA content between individuals. CNVs span from 1-kb to multi-megabases, and represent a significant portion of genetic variation in normal humans. Significant work has been done in the field of Autism Spectrum Disorder (ASD) genetics to find CNVs implicated in the development of autism and the variation of phenotype. In order to identify new disease alterations, we are conducting CNV discovery studies on 10 autistic children and their families. We expect our findings will lead to a better understanding of pathophysiology and to new avenues of therapeutic treatment. We conducted CNV discovery in 10 autistic children using array Comparative Genome Hybridization (aCGH) on a tiling-resolution whole genome BAC microarray. This platform can detect CNVs >50-kb. From the high confidence CNV calls (based on standard deviation and signal-to-noise Ratios), we filtered out all variants that are found in 4% or more of normal individuals. We have identified over 100 high confidence candidate CNVs that have rarely or never been reported previously. In addition, we have begun to determine if any variants were created *de novo* in the affected child (i.e., were not inherited from parents). Some candidate *de novo* variants span many genes and could have functional effects. We

conclude that it is worth conducting more genome-wide CNV discovery in autism, and testing suggestive CNVs for association in large numbers of autism patients.

Introduction

Overview of Genomic Variation

Genomic structural variation is an all-encompassing term used to group genomic alterations involving segments of DNA that range in size from single nucleotide alterations to entire losses of chromosomes (Fig. 1). This variation may be based on orientation, position, or quantity of DNA segments. The more closely-related humans are, the more similar their genomes and their variations. However, it has been estimated that the genomes of non-related people differ at about 1 in every 1,200 to 1,500 DNA bases, (Eenennaam 2008) creating more than three million differences between these two individual genomes. This discrepancy is mainly due to smaller genetic variations known as single nucleotide polymorphisms or SNPs (Scherer *et al.* 2007).

The majority of genomic variations deal with relatively small alterations involving only a few bases. However, larger variations have been found in normal individuals that have no current disease implications. Many studies are analyzing larger numbers of normal individuals from different backgrounds to identify the variations between individual genomes that exist. However, due to the large numbers of possible genomic variations and the current platforms available to study variation, it remains a difficult, expensive, and time-consuming endeavor. Although it has been found that 99% of human DNA sequences are the same, the variation in the sequence can have a vital

role in human response to disease and to environmental factors as well as pharmaceutical therapies.

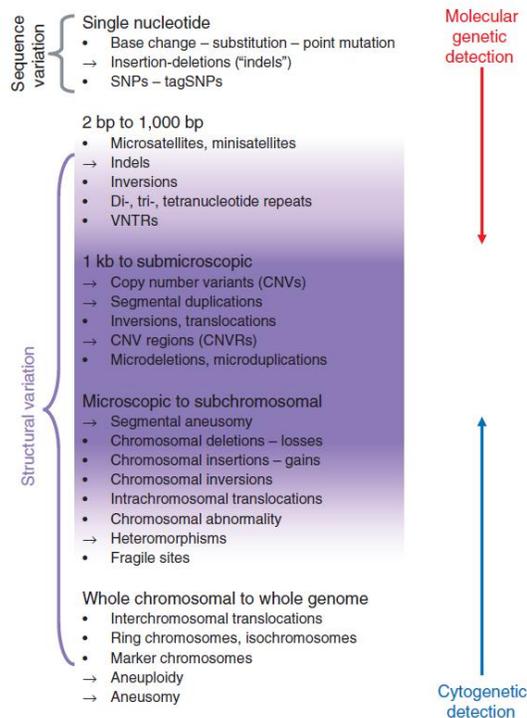


Figure 1 List of genetic variation found in humans. In a broad sense, structural variation has been used to refer to genomic segments both smaller and larger than the narrower operational definition, as illustrated by the large bracket. The focus of recent discoveries has been the subgroup in the midrange (indicated with strong highlighting), but the gradation of shading illustrates that the biological boundaries may really encompass some forms of variation previously recognized from either cytogenetic or molecular genetic approaches (Scherer *et al.* 2007).

Genomic variations can be found in all areas of the genome, including coding and non-coding gene regions. However, this variation is not distributed evenly throughout the genome. Various parts of the genome known as “hot spots” have a much higher chance of variability than others. Additionally, there are “stable” regions of the genome, which do not vary to a great extent between individuals and are said to be highly conserved throughout the species. Furthermore, the sizes of a structural variation have variable effects on the expression of genes. However, the effect is dependent on the

region in question. For example, an alteration in a very active and critical gene region which is necessary for life will have adverse effects despite size variability.

Forms of Structural Variation

There are many specific forms of genomic variation that have been classified and are known to be involved in altering genetic expression. One form of genomic structural variation that makes up approximately 90% of all human genetic variability is called single nucleotide polymorphism (SNP) (Scherer *et al.* 2007). SNPs are sequence variations that occur when a single nucleotide in the genome sequence is altered. However, in order for such a change to be called a polymorphism it must be found in at least 1% of the population. Over the entire human population, SNPs occur every 100 to 300 bases along the 3-billion-base human genome (i.e., more than 2 million SNPs) (Thorisson *et al.* 2003). The majority of SNPs are presumed to have no effect on cell function and are typically the replacement of cytosine with thymine, found in two of every three SNPs. SNPs are thought to have originally come about by errors in DNA replication or repair that occurred once in human history and are now shared among individuals by descent (Hinds *et al.* 2006). In addition, some SNPs have been found to have biological effects or to be associated with disease or drug responses.

Another form of genomic variation which is slightly larger in size are variable number of tandem repeats (VNTRs). VNTRs are a repeated sequence of DNA that tend to occur in non-coding DNA. Examples of VNTRs are microsatellites and minisatellites. In some VNTRs, the repeated unit may occur in differing numbers from 10 - 100 times. The repeated sequence is often simple, consisting of two, three, or four nucleotides. Cytosine and Adenine nucleotide repeats are very frequent in humans and are present

every few thousand base pairs (Blouin *et al.* 1996). There are usually many alleles found in VNTR regions, and the original allele from a parent can often be identified using VNTRs as markers. Due to this trait of VNTRs, they can be used in a multitude of applications such as determining paternity, population genetic studies and recombination mapping (Dakin *et al.* 2004). VNTRs have an increased rate of mutation compared to other neutral regions of DNA and are most frequently explained by slipped-strand mispairing, which is the mispairing of the complementary bases during DNA replication on a single DNA strand. VNTRs are structural variations which are found in normal individuals but are also found in individuals with disease. For example, Huntington's disease can be seen to be associated with a trinucleotide VNTR of CAG which repeats over 28 times (MacDonald *et al.* 1993). Studying these variations may help find additional genetic implications to various disorders and diseases.

DNA Copy Number Variation

Copy Number Variation (CNV) refers to larger genomic variation which represent a copy number change involving a DNA fragment that is ~1 kilobases (kb) or larger in germ line DNA (Redon *et al.* 2006). DNA copy number alterations have long been associated with specific chromosomal rearrangements and genomic disorders. The extent of this variation is still under investigation and it seems likely that in humans, CNVs account for a substantial amount of genetic variation. Approximately 4000 genetic loci are currently mapped on The Database of Genomic Variants (DVG) (<http://projects.tcag.ca/variation>) (Korbel *et al.* 2007). CNVs have been found to account for almost five times greater the variation than SNPs in the number of affected bases (de Smith *et al.* 2008). CNVs can occur in genic regions and can result in differential

levels of gene expression. Therefore, CNVs may account for a large proportion of normal phenotypic variation in the genome (Freeman *et al.* 2006). There are currently 800 CNVs which occur normally in humans (Korbel *et al.* 2007).

Until recently, the majority of structural variation was observed as segmental duplications or specific gene disorders (e.g., Lupski 1998; Ji *et al.* 2000; Inoue and Lupski 2002; Stankiewicz and Lupski 2002). In addition, they were thought to be primarily found in heterochromatin and repeat-rich areas which may not affect an individual's genetic expression. However, two landmark studies found that some large duplications may not be related to disease and are found in normal individuals (e.g., Barber *et al.* 1998; Engelen *et al.* 2000). Building on this work, Sebat *et al.* (2004) showed that there are extensive CNV present in healthy individuals (i.e., have no apparent disease association). However, the results were limited by the platforms and total number of subjects used. Following this study, another lab, Tuzun *et al.* (2005), found more than 80% of new CNVs were not identified previously. In addition, each CNV ranged from 8kb to 40kb which was a much higher resolution than previously identified CNVs. This study used a different platform that allowed for the greater resolution (Tuzun *et al.* 2005). A subsequent study by Sharp *et al.* (2005) targeted known duplicated regions of the human genome which may lead to rapid identification of CNVs. This study used different populations of individuals and identified 19 CNVs, of which only 39% had been described previously. The authors concluded that CNVs shared among several populations meant that specific genomic differences either predated the dispersal of modern humans out of Africa or recurred independently in different populations (Sharp *et al.* 2005).

More recently, three CNV studies that specifically interrogated human genomes for non-disease variants were published concurrently (Conrad *et al.* 2006; Hinds *et al.* 2006; McCarroll *et al.* 2006). Two of these studies used SNP data generated from the International HapMap Project, which characterized human genetic variation in a cohort of 269 individuals from different populations. The HapMap Project provided a SNP genotype at ~5-kb resolution in each of these 269 samples studied for a total of 1.2 million SNPs (The International HapMap Consortium 2005). Due to the abundance and availability of SNP data across the human genome, it was thought that SNP data could be used to find underlying CNVs by using the results of SNP genotyping assays (The International HapMap Consortium 2005; Conrad *et al.* 2006; McCarroll *et al.* 2006). In the third study 215 potential deletion variants ranging from 70 bp to 10 kb were identified (Hinds *et al.* 2006). A subset of 100 PCR-confirmed deletions was further characterized, with 41 of the deletions found to be present among the 24 individuals with an allelic frequency of greater than 10% (Hinds *et al.* 2006).

Due to the platforms and various procedures used in the past, there are discrepancies among the CNV discovery studies. This is based primarily on each study having its own bias toward specific types and sizes of CNVs. Using various platforms provides differing abilities to detect CNVs depending on their size, sequence class, or location. The high density oligo array used by Hinds *et al.* (2006) helped to detect deletions in a wide variety of sizes. However, their analysis avoided repetitive regions (e.g., segmental duplications) that may be more likely to be associated with larger size CNVs (Freeman *et al.* 2006). The DGV (<http://projects.tcag.ca/variation>) had an average sized CNV of 118kb; however the median size was 18 kb. (Freeman *et al.* 2006). This

discrepancy, as stated before, is due to the studies done by Conrad *et al.* (2006) and Hinds *et al.* (2006). These inconsistencies accounted for mainly smaller CNVs; the majority being <10 kb (Korbel *et al.* 2007). Using BAC aCGH methods, the copy number variation could entirely encompass a smaller CNV, overlap a CNV, or be totally within a CNV that is actually larger than the BAC clone itself. Because of this ambiguity, the size of the entire BAC clone is used in lieu of the actual size of the CNV (Freeman *et al.* 2006). Since then, there have been a large number of studies which have looked for CNVs in normal individuals and individuals with various genetic diseases. This variation between normal humans and diseased patients give additional information which can lead to disease associations with various gene regions. The progression of discovery has moved toward a more comprehensive cataloging and characterization of CNVs; this will provide the basis for determining how genomic diversity impacts biological function, evolution, and common human diseases. (Freeman *et al.* 2006).

CNVs have been found often in regions of larger homologous repeats or segmental duplications (Fredman *et al.* 2004; Sharp *et al.* 2005; Tuzun *et al.* 2005). Approximately 5% of the human genome is composed of duplicated sequence (Sharp *et al.* 2005). CNVs that are associated with segmental duplications may be susceptible to structural chromosomal rearrangements via non-allelic homologous recombination (NAHR) mechanisms (Lupski 1998). When segmental duplications are present on the same chromosome, they can facilitate changes in CNVs of the segmental duplicated regions along with intervening sequences (Inoue and Lupski 2002). NAHR can result in the formation of CNVs in normal individuals, in addition to large structural polymorphisms and chromosomal rearrangements (Lupski 1998; Ji *et al.* 2000;

Stankiewicz and Lupski 2002; Scherer *et al.* 2003; Eichler *et al.* 2004; Shaw and Lupski 2004; Lupski and Stankiewicz 2005). Approximately 50% of reported variant sequences have been found to overlap segmental duplications (de Smith *et al.* 2008). Some CNVs are also found in regions of little segmental duplication. The majority of smaller known CNVs are thought to be driven by non-homology-driven mutational mechanisms (Freeman *et al.* 2006). Using DGV, new CNVs which are found in diseased patients can be analyzed to see if the CNVs found are indeed disease associated or are CNVs found in the normal population (Wong *et al.* 2006).

Methods of CNV Discovery

The ability to detect genetic variation on a genome-wide scale is a relatively recent undertaking. CNVs were first observed by using microscopes; however the advances in genetics due to the human genome project led to the conception of CGH. Array CGH was initially used to identify segmental alterations in specific chromosomal regions associated with disease (Solinas-Toldo *et al.* 1997). The next generation of genomic microarrays examined a specific chromosome or chromosome arm. One of the first studies looked at a chromosome 20 array containing 22 cosmid, P1 phage artificial chromosome (PAC) and bacterial artificial chromosome (BAC) clones as interval markers covering chromosome 20 at 3 Mb resolutions (Pinkel *et al.* 1998). Although CGH studies using regional and chromosomal microarrays have yielded a great deal of information, these studies are naturally biased to specific areas of the genome or very large alterations (multi-Megabases) and require a prior knowledge of regions of interest. To overcome regional bias, genome-wide arrays were developed (Davies *et al.* 2005 array CGH).

The three main spotting elements used are complementary DNA (cDNA), oligonucleotides, and bacterial artificial chromosomes (Ylstra *et al.* 2006). In the past, laboratories had to manufacture their own microarrays which was a very complicated, time-consuming, and expensive process. However, recently these microarray slides have been commercially manufactured by various corporations, which make this form of DNA analysis much simpler and cost effective.

The use of cDNA as a spotting medium is not as prevalent due to a variety of disadvantages. One clear disadvantage is the low signal-to-noise ratio obtained from these spots. The variable signal intensities are also a major concern while using cDNA clones as targets for detecting copy-number alterations. In addition, in order to reliably detect single copy changes, a moving-average of clone intensities must be calculated, thereby reducing the resolving power of the arrays (Davies *et al.* 2005). Moreover, larger quantities (micrograms) of sample genomic DNA are required in order to generate a robust signal, thus limiting the utility of cDNA array CGH (Pollack *et al.* 1999).

More recently oligonucleotide aCGH platforms have become available. Oligo aCGH platforms generally were single-stranded 25 to 85mer oligonucleotide elements on the array. An inherent problem of cross hybridization of oligonucleotide targets to multiple genomic loci is the need for complexity reduction of the sample genomic DNA (Davies *et al.* 2005). In order to increase signal to noise ratios, the whole-genome sampling assay (WGSA) was developed to greatly reduce the genomic complexity of the sample probes by about 98% to improve hybridization kinetics (Davies *et al.* 2005). Hybridization of complexity-reduced probes to SNP arrays is able to detect high-level

copy number changes; however, since the resolution of the method is dependent on the inconsistent SNP-density on the array, over a fifth of the chromosomes measured were inconsistent with the segmental genomic alterations detected by BAC-array CGH due to poor SNP representation (Bignell *et al.* 2004). In addition, the data generated by the SNP-arrays showed high variability, probably due to the requirement of PCR amplification of genomic sample DNA. This required the use of a moving average to detect copy-number changes, further lowering the effective resolution of the method (Bignell *et al.* 2004).

There are two main commercial oligo aCGH platforms available from NimbleGen, and Agilent. However currently there are higher resolution arrays being manufactured. The first commercial oligo aCGH platform is manufactured by Agilent Technologies. They created their arrays specifically for array CGH, and include oligonucleotides covering intergenic regions. At the time of the current study, the Agilent array platform used 244k 60mer oligonucleotides, (very recently the single slide probe density was increased to include 1M probes per array). The direct labeling protocol is similar to the one used for the cDNA arrays and requires 1 µg of input DNA, which hampers the use of small clinical samples (Barrett *et al.* 2004). To overcome this problem, an additional PCR amplification procedure was developed allowing as little as 10 ng of input DNA. Apart from the necessity to do amplifications of test and reference sample(s) in parallel, any PCR-based DNA amplification introduces some level of additional variation and adds to the overall cost of arrays (Barrett *et al.* 2004).

The second oligonucleotide platform is commercially offered by NimbleGen. They provide arrays containing 385 K and a new 2.1 M oligonucleotides which are photo-

lithographically synthesized on the array. The array production is extremely flexible, such that each array produced can have a different set of oligonucleotides on it. The oligo aCGH oligonucleotides are designed to be isothermal and vary between 45 and 85 bp in length (Ylstra *et al.* 2006). For their labeling and hybridization procedures, NimbleGen adopted essentially the same direct labeling conditions as Agilent and the cDNA CGH platforms. Lucito *et al.* applied an alternative oligo aCGH array method. The method called ROMA (representational oligonucleotide micro array analysis) uses a 98% complexity reduction of test and reference DNA (Lucito *et al.* 2003). This is carried out by a digestion-amplification step, which allows starting with as little as 50 ng of input DNA. For reproducibility, this labeling and amplification procedure requires test and reference samples to be amplified in parallel. ROMA combined with the NimbleGen array provides a high-resolution alternative that allows low amounts of sample DNA input and has already proven its value in the field of human genetics (Ylstra *et al.* 2006).

Different oligo arrays are combined with various labeling and hybridization techniques and all yield high-resolution copy number measurements. Yet, none of the current oligo aCGH platforms can make a definite call for loss or gain using a single oligonucleotide; but rather at least 3–5 adjacent oligonucleotides and preferably greater or equal to 10 are necessary for a reliable call. As seen in the literature, there is no “gold standard” that has been used to find structural variation. However, each lab uses each platform to subsequently provide the most relevant information for their particular study.

The most widely used aCGH arrays in the past have been based on bacterial artificial chromosomes or BACs. DNA yielded for BACs vary in length from 150 to 200 kb and are generally isolated from *Escherichia coli*. The DNA is then amplified by PCR

prior to spotting the arrays. This is done since high DNA concentration is mandatory for high quality results. BAC platforms can be outstandingly sensitive and, in the range of 50kb-150kb, precise (Ylstra *et al.* 2006). One can detect a single-copy gain or loss on a single arrayed BAC element. Although BAC arrays have relatively few printed elements on the arrays compared to oligo arrays, their spatial resolution is high. The higher signal compared to oligo arrays, allows for single copy transition boundaries (Ylstra *et al.* 2006). The use of BACs also confers high and consistent binding specificity (Pinkel *et al.* 1998; Hodgson *et al.* 2001), and thus more accurate copy number determination. The breakpoint resolution that can be obtained with BAC arrays, however, is finite because of their large size (Ylstra *et al.* 2006). Another variable in array CGH is the Cot-1 DNA which is used in the procedure to block repetitive DNA sequences in the arrayed elements (Pinkel *et al.* 2005). Cot-1 DNA is placental DNA and yields variable DNA fragments of 50–300 bp in length. Batch to batch variation makes Cot-1 DNA a highly variable element in the array procedure. For oligo aCGH, Cot-1 DNA is not necessary, since the oligonucleotides are designed to be repeat-free.

In 2004 the Wan Lam laboratory (BCCRC, Vancouver; Ishkanian *et al.* 2004), published the first sub-megabase resolution tiling-set (SMRT) array that contiguously covered the human genome in a tiling path manner. By using overlapping clones, the resolution of the array was increased beyond the size of a single BAC clone and gains and losses of regions as small as 40–80 kb are detectable. A major advantage of using a tiling-path array is in identifying small (gene level) gains and losses. The probability of missing a small genetic alteration is inversely proportional to the genome coverage or representation of the detection strategy. The tiling path array offers a much greater

probability of detecting small-sized alterations (e.g. 40 kb) than marker-based genomic arrays. The even distribution of markers throughout the genome, as opposed to having a large number of small loci clustered in selected regions, is a key consideration in improving the resolution of genomic scanning strategies. The first version of the (SMRT) tiling-path array was composed of 32,433 clones spotted in triplicate over two microarray slides (Ishkanian *et al.* 2004). In version 2 of the SMRT array the Lam lab has created the SMRT re-array (SMRTr). The SMRTr array contains a more selective set of clones representing 83% of the original collection, eliminating unnecessary redundancy while maintaining tiling path coverage. As a result, the 27,000 clones are spotted in duplicate, on a single slide, reducing the cost and time of analysis (Ylstra *et al.* 2006).

These various platforms have their own distinctive characteristics that can be an asset while discovering novel CNVs. When this study was initiated, the use of SMRT array with the use of BACs spotted in duplicate has the one of the best effective resolutions (comparable to the NimbleGen 385k and Agilent 244k arrays) while using a small amount of genomic DNA. Moreover, the SMRT array requires much less DNA, 400ng compared to 1-2 μ g. Due to these facts we have chosen to use this platform in our study. The SMRT micro array used in our lab contains 26,363 clones spotted in duplicate with an effective resolution of 50kb. Although the BAC array used in our lab has a weaker breakpoint resolution, they have been able to call 10X more CNVs than on similar resolution oligo arrays. This is due to the fact that our array contains non-unique sequence allowing for a larger number of possible CNVs to be found. The slides were printed by the Wan Lam lab at BC Cancer Agency Research Centre in Vancouver,

Canada (<http://arraycgh.ca/index.php>). Our lab is a beta-site for the Wan Lam lab, by use of the platform, skills and competency to begin CNV discovery. In addition, the quality control and bioinformatics used in our lab was internalized with the help of the Wan Lam lab. In Table 1. we show Real Time PCR validation of CNV calls made in our lab. The CNVs called in our study are thus high confidence in part by calibration with the experimental and analytical methods developed and stringently validated by our collaborators (Ishkanian *et al.* 2005, Wong *et al.* 2006).

In order to validate the methodology of our platform, multiple controls were performed. As seen in (Fig. 2.), the visualization of trisomy 21 in individual 4 is quite compelling. Essentially all BACs are shifted to the right, confirming the one copy gain of chromosome 21 in this individual.

CNV1 Chr4 133123563- 133299949 (UCSC May 2004)	Sample	315	323	735	751	766
	aCGH threshold score	-9.846	-9.951	-8.773	-7.857	-3.86
	relative change (qPCR v. unaffected sibling (767))	0.6445	0.7935	n/a	n/a	n/a
	relative change (qPCR v. pooled males)	0.6675	0.891	0.7935	0.8125	0.841
	Comments	Both references have same copy #; all samples show copy number loss				
CNV2 Chr4 8722950- 8847678 (UCSC May 2004)	Sample	323	735	770		
	aCGH threshold score	6.149	7.437	2.735		
	relative change (qPCR v. unaffected sibling (315))	1.2895	1.7015	1.1485		
	relative change (qPCR v. pooled males)	1.26	1.4475	n/a		
	Comments	Both references have same copy #; samples 323 & 735 show copy number gain				
CNV3 Chr17 33369043- 33427467 (UCSC May 2004)	Sample	315	323			
	aCGH threshold score	3.6244	-4.054			
	relative change (qPCR v. unaffected sibling (767))	0.8705	0.536			
	relative change (qPCR v. pooled males)	2.1435	1.0235			
	Comments	References have different copy #; 315 has more copies than 323				

Table 1. CNV validation by Real Time PCR. The father in this family transmitted two diseases to children from two mothers. CNV discovery was done by aCGH using a reference of DNA pooled from multiple individuals. In this Real Time PCR study ($\Delta\Delta$ Ct method, Livak and Schmittgen 2001), CNVs are validated between affected and unaffected siblings and compared to a second pooled reference. aCGH

threshold score is copy number change (Ratio) / standard deviation autosomal ($> \pm 2.85$ are called significant gain/loss) (Chen *et al.* 2008).

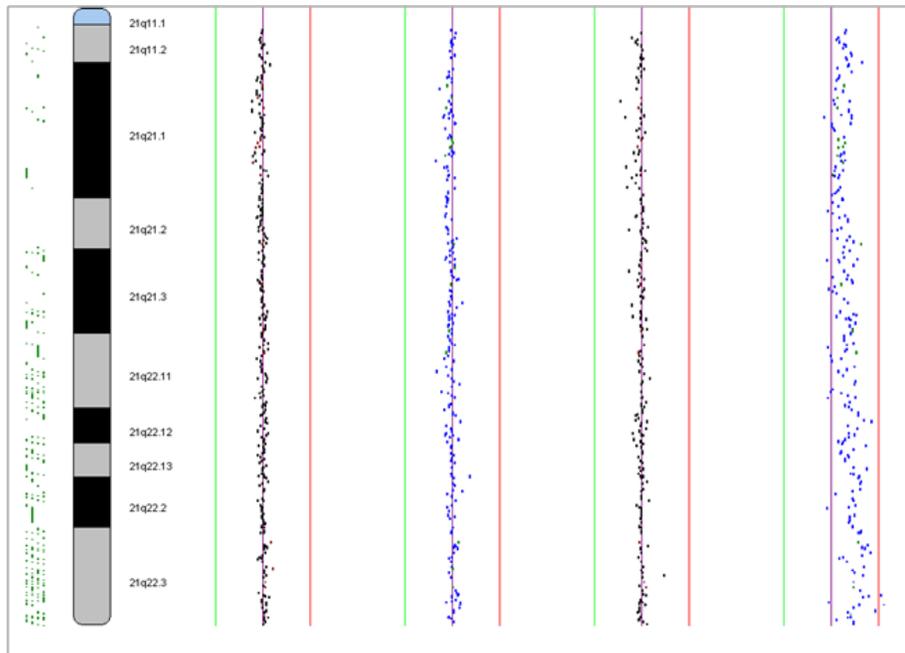


Figure 2. This example illustrates the methodology of the SMRT BAC array platform. Starting from the left the green spots represent genes which correspond to genes on the chromosome, then the banding of the long arm of chromosome 21. Finally the four panels represent four distinct individuals. The midline represents a 1/1 ratio of *cye3* and *cye5*, and any deviation to the left or right suggests a loss or gain in copy number respectively. The first three panels show normal individuals who have almost no copy number alterations. However, the fourth panel clearly demonstrates a one copy gain of the entire chromosome and validates trisomy 21.

Overview of Autism

Autism Spectrum Disorders (ASDs) are a group of childhood neurodevelopmental and neuropsychiatric disorders characterized by impairments in social interaction, verbal communication, and restricted and repetitive patterns of interests and behaviors (Wang *et al.* 2009). Diagnostic tools based on phenotype for ASD have been standardized for reliable diagnoses. This was accomplished by the release of the Autism Diagnostic Interview-Revised (ADI-R), a parents' or caregivers' questionnaire, and the Autism Diagnostic Observation Schedule-Generic (ADOS-G), a direct testing tool of the patients' current behavioral pattern. In addition, the diagnostic

criteria of DSM-IV have all helped to standardize the diagnosis of ASDs in the world. ASDs include a range of clinically defined conditions including autism, Asperger syndrome, Rett syndrome (RTT), childhood disintegrative disorder (CDD), and pervasive developmental disorder not otherwise specified (PDD-NOS).

ASD symptoms are recognized typically within the first 3 years of age with a lifelong persistence. Current prevalence estimates in the United States are 0.1–0.2% for autism and 0.6% for ASDs (Newschaffer *et al.* 2007). In addition, it has been found that ASDs are about four times more common in boys than girls, and at present around 1 in 150 children in the United States have a diagnosis of an ASD (Wang *et al.* 2009). It has also been estimated the ASDs collectively affect 60 in 10,000 individuals (Fombonne 2005).

In several studies it has been found that ASD susceptibility has a strong genetic component. For example there are much higher concordance rates of ASDs in monozygotic twins (92%) than dizygotic twins (10%) (Bailey *et al.* 1995). Autism has estimated heritability of greater than 90%. In addition, familial clustering largely is explained by genetic factors. Moreover, observations about the patterns of inheritance among families, including studies of twins, first-degree relatives, and more distantly related individuals, all suggest that ASDs are not caused by the action of a single gene transmitted in a simple dominant, recessive, or X-linked fashion. Indeed, the number of contributing loci, or genetic regions, has been estimated to be in the neighborhood of 15 (Risch *et al.* 1999; Gupta *et al.* 2006). Several diagnosable medical conditions show symptomology of autism spectrum disorders as well (e.g. fragile X syndrome, tuberous sclerosis complex, and neurofibromatosis). These cases account only for 10% of

patients with an autistic phenotype, while the majority of idiopathic ASDs may be based on specific gene regions which may be suggestive of an ASD. (Klauck *et al.* 2006).

There are only a handful of loci which have been found to be present in a larger population of Autism patients. The more prevalent ones are 16p11.2 which has only been found in 1% of cases (Weiss *et al.* 2008), maternal Duplication of 15q11 (Depienne, *et al.* 2009), and more recently 5p14.1 (Wang *et al.* 2009). The majority of the genetic associations involved with autism are found to be *de novo* variants or random mutation which occurs in the germ line DNA. Rare *de novo* copy number variants have been implicated in 7% of families with ASDs, but only in 1% of control families (Sebat *et al.* 2007). Due to the prevalence of CNVs found in Autism it is our main goal to conduct whole-genome screens of autistic patients to define regions with acknowledged susceptibility genes for autism for further fine mapping by association studies and to verify *de novo* or inherited variants. As previously stated, CNVs are a prevalent structural variation which has been implicated in many disorders by various studies. Presently there has been a large volume of research dedicated to understanding the pathophysiology of autism and other neurological and psychiatric conditions. To supplement this data, we are conducting CNV discovery to find novel and *de novo* variations in autistic patients.

Due to the commonness of this disorder there are reasons to suspect that common variation will be found to play a large role in the etiology of ASDs. Within the normal human genome, most of the variation in a given population is accounted for by polymorphisms that are present in more than 1% of individuals. These, by definition, are considered common alleles. It would be logical then to hypothesize that a disease that

is common would reflect this overall architecture. This would hold true as long as natural selection against these alleles did not play a significant role (Gupta *et al.* 2006).

Methods and Materials

DNA samples

Reference and sample genomic DNA samples were gathered under human studies protocol IRB06-00414 (Carlos Alvarez, Nationwide Children's Hospital). The reference DNA was from a normal individual who has no known disease or genetic abnormality. The blood samples were collected from autistic children and their parents, and provided to us by the Central Ohio Registry for Autism (CORA, which has its own IRB protocol), which was created by and continues to be headed by Dr. Gail Herman (Nationwide Children's Hospital). Blood was processed into purified DNA, and live lymphocyte B cells from the latter were EBV-immortalized (Epstein - Barr virus) and multiple aliquots were frozen in N₂ (liq). DNA was then extracted and purified from those cell lines to be used in our study.

The children involved in our study were clinically evaluated and diagnosed with autism (CORA). All these autistic children were negative for large genomic alterations using clinical cytogenetics (but had not been analyzed for submicroscopic alterations by aCGH (CORA)). All autistic children in this study are male, ranging from ages 4-18, and nine are of European ancestry and one individual is of African (Cameroon) ancestry.

DNA purity was determined by UV spectrophotometry (Nanodrop 1000). 260/280nm Absorbance ratios >1.8, and 260/230nm Absorbance ratios >2.0 were considered high quality. In addition DNA was analyzed by agarose gel electrophoresis to ensure that the DNA was of high molecular weight.

DNA labeling and array hybridization

Once the quality control on the DNA was complete 400ng of sample and reference DNA were labeled with Cyanine 3–dCTP and Cyanine 5–dCTP, respectively (Cy3/Cy5, Perkin Elmer Life Sciences). Klenow, enzyme, and a random octomer primer (Operon) were used in the random priming reaction. The enzyme reactions were incubated in the dark at 37°C for approximately 18 hours. DNA samples were purified from unincorporated label using spin columns (Microcon YM-30, Millipore). Purified samples were mixed with 100 µL of human Cot-1 DNA (purified highly repetitive-sequence DNA (Invitrogen)), which was used to block repetitive sequence. DNA was resuspended in 45 µl of DIG Easy hybridization solution (Roche). Quality control of DNA labeling was conducted by measuring dye incorporations (Nanodrop 1000). A threshold of 8pmol/µL was used to determine acceptable dye incorporations. The sample mixture was denatured at 85°C for 10 min., and repetitive sequences were blocked by hybridization at 45°C for 1 hour. The mixture was applied onto SMRT BAC arrays containing 26,363 clones spotted in duplicate on single slides (manufactured in the lab of Dr. Wan Lam, BCCRC, Vancouver; Wong *et al.* 2006). [The clones were selected from the SMRT clone set, to optimize tiling coverage of the genome; the clone list is available at the SMRT Array Web site (Ishkanian *et al.* 2004).] A cover slip was added (Fisher Scientific), and the slide was placed in a hybridization chamber along with 15µl of water for approximately 36 hours. Cover slips were then removed and arrays were washed three times for 5 min each with agitation in 0.1% saline sodium citrate (SSC) and 0.1X SDS at 45°C. Arrays were rinsed three times for 5 min each in 0.1X SSC at room temperature and were dried by being placed in 50mL (Falcon) tubes and

centrifuged at 700g for 3 min. Once completely dry the arrays were scanned on a Genepix 4100 scanner which individually scans at Cy3 and Cy5 optimal wavelengths of 595 and 685, respectively.

Visual inspection of the scans were first used to detect major problems, such as printing errors due to manufacturing defects, and large regions of noise. The fluorescence intensities at 595 and 685 nm were used for the quality control, normalization, and processing of the data. The data was exported to an Excel file where the stringent criteria were applied for CNV discovery. Using the same BAC array platform, and the same criteria, Wong *et al.* (2006) experimentally determined the false positive detection rate to be 54.7%. The main CNV calling criteria were the standard deviation (SD) and signal to noise ratios (SNR) of individual BACs, and the \log_2 spot ratio showing a difference between the DNA quantity in the subject and the reference. SeeGH-Norm software was used to individually normalize array-wide signal intensities. SeeGH software (Chi *et al.* 2004) was used to visualize all high confidence BACs on the human genome assembly (Mar. 2006 (hg18), a.k.a. NCBI Build 36.1). Once all CNVs were visualized, certain thresholds were used to focus our findings. High quality probe signals were determined using a (SNR) of 10 for our very high confidence calls denoted as Tier 1 CNVs and a lower stringency of an SNR ratio of 3 for a lower confidence denoted as Tier 2 CNVs. This filtered data was then exported again on to an Excel file where a standard deviation of the autosomal data was found for each CNV. Spots which had no duplicate were removed, and any BACs that had been found previously in 4% or more of normals by Wong *et al.* (2006) were removed from our data. That study examined 95 normal individuals and found a large number of CNVs that occur naturally

in the population. Notably, that study and ours were both biased for Western Europeans. Most importantly, the individuals from that study were used as our negative controls since that platform was identical to ours. The cut-off used by our lab to determine a called CNV was a standard deviation of -3.33 and 3.33. A loss would be called if the standard deviation of a BAC was -3.33 or less, and a gain would be called if the standard deviation of a BAC was 3.33 or more. The thresholds, quality control process, and the past expertise and guidance was provided by the Wan Lam lab (who created the platform and analysis tools).

Results

Array CGH was used to conduct CNV discovery in germ line DNA from 10 children with autism, and presently, in the parents of one of the children. DNA copy number of each sample was quantified by comparative hybridization with a single normal reference DNA. The array platform used was the SMRT BAC array, which has tiling resolution representation of the human genome at 1.5-fold coverage (Wan Lam, BCCRC, Vancouver). The false positive detection rate must be considered during CNV discovery. The CNVs found in the study are not only to be accurate if validation is confirmed. In addition, as reported in the Wong *et al.* paper, the false positive and negative detection rate were 54.7% and 45.3% respectively (Wong *et al.* 2006).

As a control population, we used SMRT BAC array data generated under the same conditions from 95 normal individuals (Wong *et al.* 2006). Both that study and ours were biased for Western European ancestry (and both also included individuals of African ancestry). Candidate CNVs were further analyzed by the use of the genome browser database at the University of California, Santa Cruz. The BAC-spanning

genomic coordinates of each called CNV region were entered into the browser and various supplementary data was accessed. In this way, we annotated the gene content of CNVs. We also used this database to determine whether previous studies had identified CNVs that coincided or overlapped with ours. Additionally, we determined whether CNV regions have segmental duplications (or low copy repeats) in the vicinity.

The data has been categorized to optimize the discovery of candidate autism associated CNVs. First CNVs were called based on the number of array elements altered, either being single or multiple BACs. In total, our study identified 805 high confidence single-BAC CNVs, and 390 lower confidence single BAC CNVs (Table 2.) Our study has an average of more than 80 CNVs per individual. A single-BAC CNV has only one array element which has been altered and therefore has a lower confidence than a multi-BAC CNV which has more than two altered BACs (Fig. 3). The total number of CNVs found in our study is lower than that of Wong *et al.* (2006), using the same platform. This is presumably due to aspects of DNA hybridization that could relate to DNA quality, (which affects probe labeling efficiency) or manual processing of array hybridizations. In addition, variability of lab conditions such as temperature, and levels of ozone (which degrade Cy5) may have also led to lower CNV calls. Due to the thresholds based on SD and SNR, we have kept similar standards as the Wong *et al.* paper and can therefore assume the same false-positive percentage.

Second, we identified the CNVs that were the most likely to be associated with disease: i.e., did not overlap regions of common CNV in normal individuals, affected genes previously implicated in autism, or were *de novo* mutations in a child (i.e., were not inherited from a parent) (*Appendices*, Table A1). We have called more than 20 high

confidence candidate autism CNVs per individual (Table 3). Of these CNVs 13 are tier 1 (SNR cutoff of 10; see above) multi-BAC, and 3 are tier 2 (SNR cutoff of 3) multi-BAC CNVs.

Total Number of CNVs

	# of Single BACs	# of Multi-BACs	average # of CNV per individual	# of gains per individual	# of losses per individual
Tier 1 (SNR 10)	805	37	81	47	33
Tier 2 (SNR 3)	390	9	39	22	17

Table 2. Total number of CNVs found in our study. Tier 1 are higher confidence CNVs which have a signal to noise ratio of 10 or greater. Tier 2 are lower confidence CNVs which have a signal to noise ratio of 3 or greater. The tier number determines a higher quality hybridization of slides.

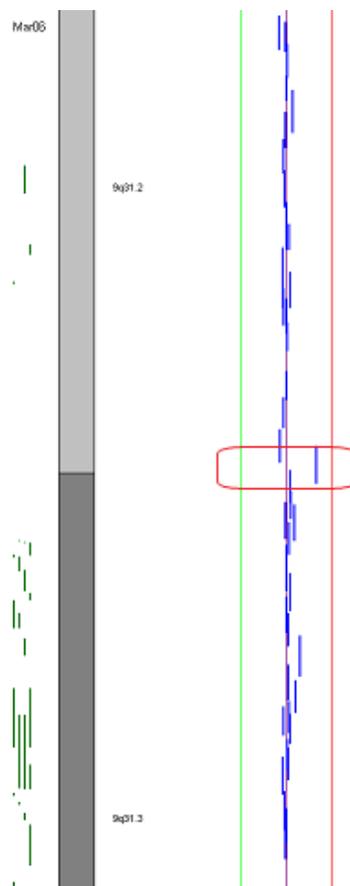


Figure 3. Single-BAC candidate CNV called in our study. Sample 137 shows a single-BAC CNV spanning 110156709-110355803 on chromosome 9q31.3. The BAC shifted to the right as seen by red outline represents a gain.

The candidate CNVs were prioritized according to two criteria. First we identified the highest possible confidence CNVs. The highest confidence CNV calls are those that are spanned by multiple BACs because they are detected by multiple array elements instead of one. In addition, they span a larger region of the genome and are thus likely to include more genes. The more genes involved in a copy number alteration, the higher likelihood it will have a phenotypic effect. Included in these were 37 high confidence and 9 lower confidence multi-BAC CNVs (Table 2). All of our high confidence multi-BAC CNVs follow our optimal thresholds of SD, correlating to being called a gain or loss, and SNR, correlating to their placement in the tier system.

In order to find high interest candidate CNVs possibly associated with autism, we filtered out CNVs that were detected in 4% or more of 95 normal individuals by aCGH with the same BAC array platform we used (Wong et al. 2006). The biochemical/genetic pathways and biological significance of genes potentially affected by CNV were screened for relevance to autism and brain development (using PubMed to screen the database of biomedical literature, and the Google search engine to screen the entire internet for published information). All CNVs were annotated manually by use of the UCSC genomic browser database (<http://genome.ucsc.edu/>). An example of a multi-BAC CNV has been visualized by SeeGH software, which allows us to compare the probes within and between individuals (Chi *et al.* 2004) (Fig. 4). The top 12 multi-BAC candidate autism CNVs are among the most promising in our study, for their CNV call confidence and for the large number of genes affected. (Appendices, Table A2).

Candidate Autism CNVs

	# of Single BACs	# of Multi-BACs	average # of CNV per individual	# of gains per individual	# of losses per individual
Tier 1 (SNR 10)	231	13	23	18	5
Tier 2 (SNR 3)	149	3	15	11	4

Table 3. CNVs that are high confidence (based on standard deviation, signal to noise ratio) and high interest (not being present in more than 4% of normal individuals which was found in another study of 95 normal individuals using the same platform (Wong *et al.* 2006)).

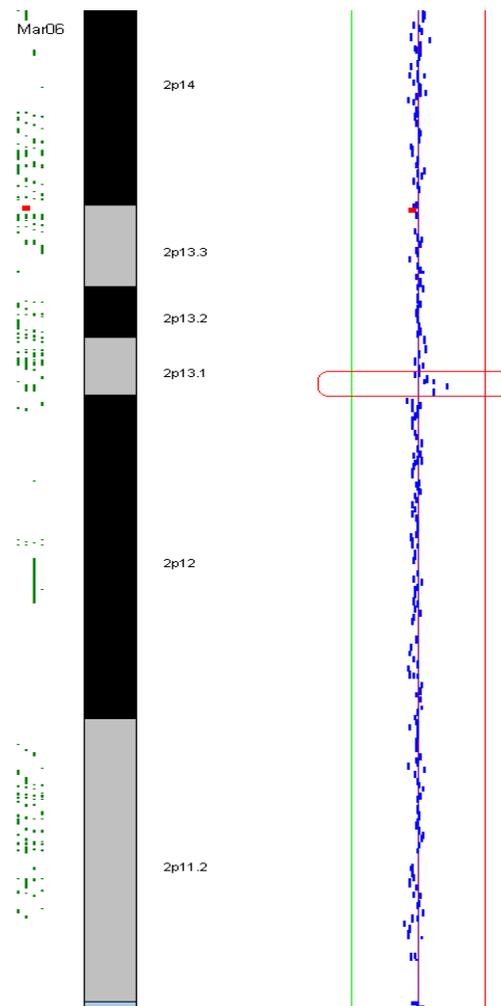


Figure 4. Multi BAC CNV found in our study. Sample A75 shows a multi-BAC CNV gain on chromosome 2p13.1 which is marked in red. Multi-BAC CNVs are particularly high confidence because they affect multiple probes and can be of higher clinical interest as they can affect more genes since they span a larger region of the genome.

We have separated our results into three forms: (1) rare or novel candidate CNVs that have not been found in previous studies, (2) CNVs that have been found in our study as well as larger studies also looking for CNV associations with autism, and (3) *de novo* or non- inherited CNVs.

Novel candidate autism CNVs have been found in our study. These CNVs are very interesting because they have rarely or never been seen in normals or other autistic children. The called CNV on chromosome 12 is a curious candidate CNV (Fig. 5). Sample A39 has a multi-BAC CNV gain in the 12q13 region, with three overlapping BACs shifted to the right. The CNV is not found in any other children and has not been found in other studies. In addition, the high SNR make this CNV very compelling. Many such high interest candidate CNVs are in non polymorphic regions and are not found in less than 4% of normals. The majority of these multi-BAC CNVs are found in gene rich regions, suggesting they are likely to have phenotypic effects (which may be associated with autism).

More than 25% of the total CNVs that have been called in our study have been found in current Autism CNV literature as well as the Autism Chromosome Rearrangement Database (ACRD). ACRD is an online database that has compiled the majority of autism CNVs from published studies such as *C. R. Marshall et al. 2008*, *L. Sebat et al. 2007*, and *Christian et al. 2008*. Specifically, 64/231 of tier 1 CNVs and 30/149 of tier 2 CNVs have been seen in other studies.

Due to the possibility that some CNVs may be common in normal individuals, but be present at higher rates (that are statistically significant) in autistic individuals, we also

created our own database of all CNVs previously implicated in published studies (studies were identified by PubMed searching, and the data was manually compiled and curated by myself and my advisor, Dr. Alvarez) (Appendices Table A3.)

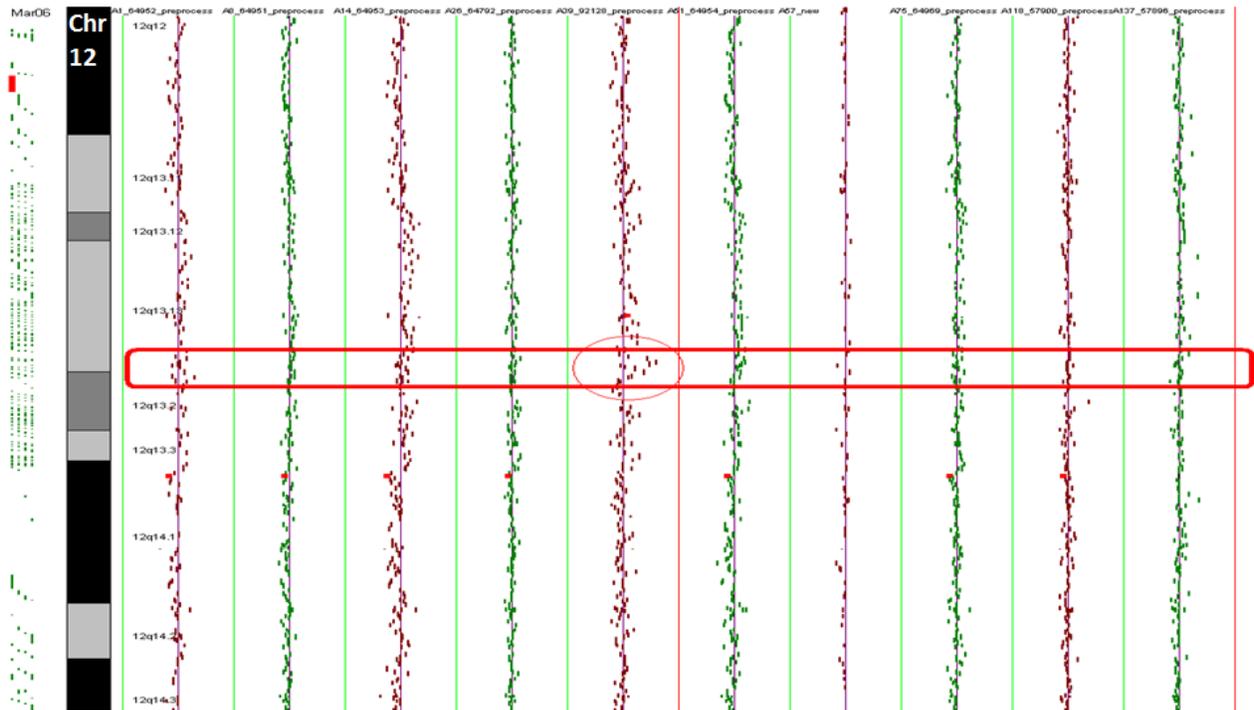


Figure 5. Close up of chr12q aCGH analysis representing the 10 children who took part in this study. The middle panel, sample A39 has a large multi-BAC CNV that is seen by its deviation from the midline to the right. This shows a gain in copy number in this gene region. In addition one can see that all of the other children have no deviation at the same loci.

We initiated additional analysis to understand the heritability of these autism associated gene regions. Currently we have conducted one triad study to identify *de novo* variants in an autistic individual. The triad A39, A40, and A41 represent the child, father, and mother respectively. By conducting aCGH analysis on the parents, similarities and difference between the parents and the child can be observed. As seen in (Fig. 6), neither the father nor mother has a significant shift from the midline for any of the BACs that are affected in the child. However, the child has a compelling and interesting multi-BAC CNV (i.e., affects several genes) in the same region. Some genes

that are found in this region are MTIF, MTM, MTF, MT3, BBS2, NUP93, MTE, MT1X, and MT1E. Most of these genes begin with “MT”; they are known as metallothionein genes. They are a cluster of genes involved with protection against metal toxicity, based on experimental findings (Egli *et al.* 2006). Of the others, BBS2 could be suggestive of autism because of the associated with Bardet-Biedl syndrome which has clinical presentation of renal dystrophy, and mental retardation. Since neither parent appears to have this CNV, we consider this to be a likely *de novo* variant.

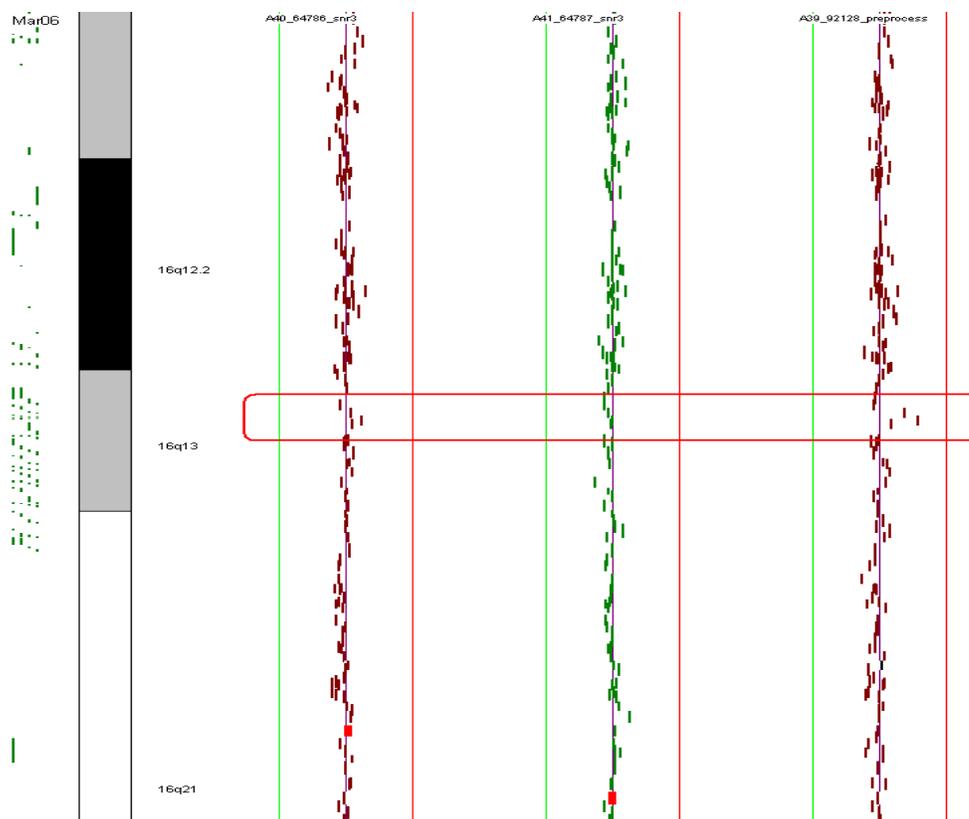


Figure 6. Possible *de novo* variant found in our study. From left to right, the first individual is the father, the second the mother, and the third the autistic child. The child has a multi-BAC CNV which is not found in either parent. This suggests this variant is a *de novo* mutation (i.e., not inherited from either parent).

Conclusion and Discussion

Our aCGH analysis of 10 autistic children yielded interesting findings. Once validated in follow-up studies, these results are likely to add to the current knowledge

about autism genetics and pathophysiology. In addition to identifying novel or rare CNVs, we have also found alterations in regions previously implicated in autism, and have identified multiple candidate *de novo* CNVs. Analyzing the genetic and biochemical pathways of the genes located in the CNV regions will help us to understand their significance. Although these CNVs need to be confirmed by other methods, we have designed our study to only detect high confidence variants. Our experimental and data analysis methodologies were calibrated to that of the Wan Lam lab.

The aCGH platform we used here has BAC array elements that span 150-200kb. These can generally detect CNVs at least 50kb in size, and can detect a subset of CNVs as small as 10kb (Kidd *et al.* 2006). We first focused on those high-confidence CNV calls which were spanned by multiple BACs. Not only were these calls detected by more than one array element, they were physically larger, and therefore more likely to be associated with disease. The more genes that are either duplicated or deleted, the higher the chance of phenotypic consequences (Table 4).

A large number of our candidate CNVs have also been found in other studies (see ACRD (<http://projects.tcag.ca/autism/>); (Appendix, Table A3)). One of these was identified by Glessner *et al.* (2009) in 33/2195 autistic patients, but only in 7/2519 normal healthy individuals (Fig. 7a). This CNV's association with autism was determined to have a P value of 3.57×10^{-6} and an odds ratio of 5.547. These values show the statistical significance of this CNV. As seen in Fig. 7b, the Glessner CNV clearly overlaps a region we identified in our sample A26. Another interesting aspect of this

CNV is the fact that it is upstream of the AK123120 gene that is expressed in the brain and thus consistent with a possible role in autism.

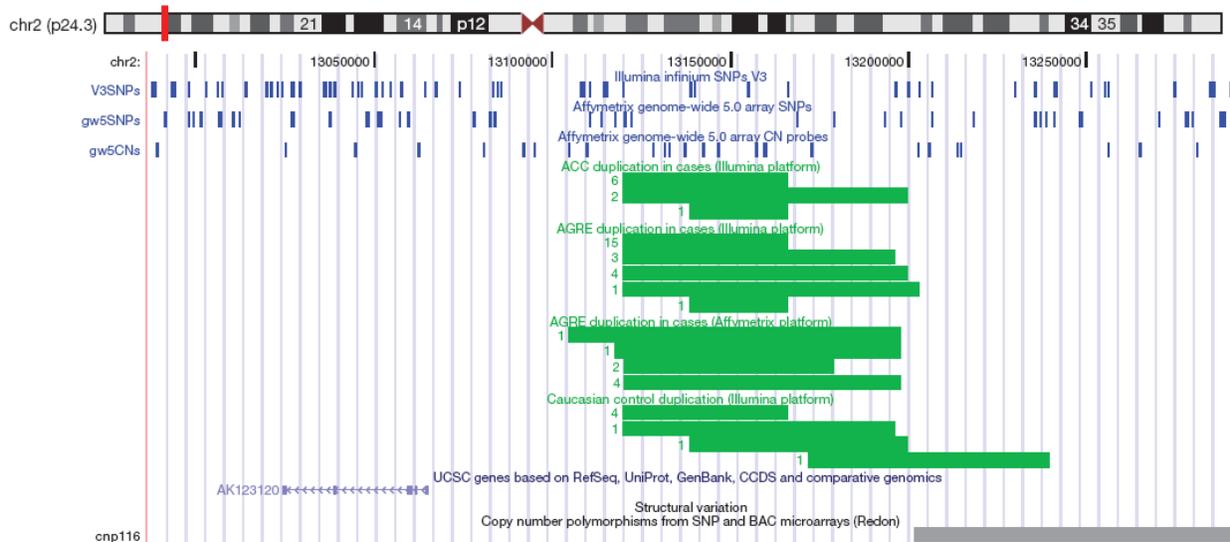


Figure 7a. A CNV identified by Glessner *et al.* (2009). The blue lines at the top indicate the SNP and copy number probe coverage. The green bars represent the BACs which span the CNV region of interest. The average breakpoint of the BAC is 13100000-13190000. This overlaps very well with a candidate CNV found in our study. This is an example of a CNV which has been called in our study along with larger studies involved with CNV discovery (Glessner *et al.* 2009).

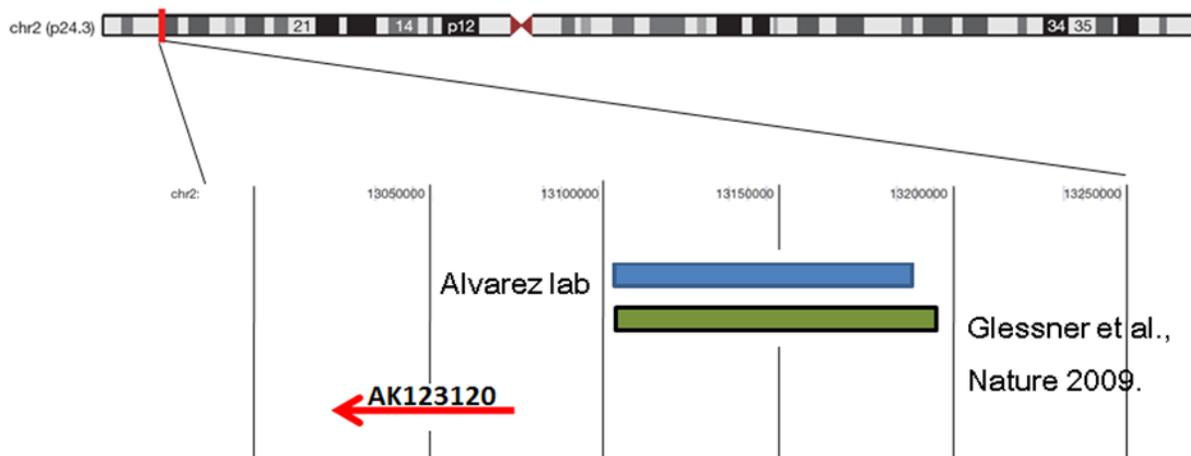


Figure 7b. The corresponding CNV to figure 7a. As one can see the blue line represents the CNV regions found in our study. The green line represents the CNV region found in the Glessner, Nature study. The close overlap shows a very compelling find. Since we have a CNV which corresponds to another published by another study examining over 1000 patients while our sample pool is only 10, our findings are very current and relevant.

In our study we have identified variations affecting many interesting genes that may be associated with the phenotypic expression of autism. One such gene, present in

a high-confidence CNV region, is FOXP4. FOXP4 is a member of the forkhead box family of transcription factors which are involved with development of the brain and other organs. FOXP2, which is known to be co-expressed with FOXP4, has been implicated in the development of language skills, indicating this CNV has a possible role in the characteristic traits of autism (Kato *et al.* 2004). Additional analysis is being done to find the genetic significance of the large number of other genes that have been found in CNV regions.

In our study we have found nearly 80 high-interest CNVs in 10 autistic children, and more than 20 high-confidence candidate autism CNVs have been called per individual. This number is much higher than the majority of published studies of autism CNV discovery. For example, the recent Glessner *et al.* (2009) article in Nature, found 15.5 CNV calls per individual. Our increased CNV call rate is due to the SMRT BAC array platform, which includes non-unique sequences –mainly those in regions of segmental duplications (or low copy repeats) –that have a 10-fold elevated rate of CNV. On the other hand, such non-unique sequences are not covered by oligonucleotide platforms because they cannot be uniquely mapped. Although the false negative rate of the platform used in our study is quite high, we still have found a large number of high-interest CNVs which may have been missed in others, but only a portion of those may actually be associated with autism.

The next steps are to finish the triad analysis on families which have interesting candidate CNVs present in the child, and then validate the findings to ensure they are true calls and that their breakpoints are also accurate. The validation will be done by fluorescence *in situ* hybridization (FISH), southern blotting, and Real Time PCR, all of

which give vital information that will enhance the findings so further examination can be conducted.

The use of the SMRT array platform has greatly increased the quantity (due to its ability to find a larger number of CNVs) and quality (by providing higher-confidence calls) of CNVs discovered using aCGH techniques. This robust and versatile platform has clear advantages over techniques used in other recent studies, including higher signal-to-noise ratios and better standard deviations than other platforms, and it requires no sample amplification or the complexity reduction that is necessary for the SNP array approach to CNV discovery. This study has supplied extensive data that will continue to be analyzed, validated, and annotated for autism relevance. We recommend the continued study of genome-wide analysis of structural variations, especially CNVs, to examine their role in ASD susceptibility, expression, and inheritability. This information can lead to clinical diagnostic testing to verify the origins of the disorder, and to the development of various therapeutic techniques to improve the lives of autistic individuals.

References

- Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E, Rutter M. Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med.* 1995 Jan;25(1):63-77. PubMed PMID: 7792363.
- Barber, J.C.K., Joyce, C.A., Collinson, M.N., Nicholson, J.C., Willatt, L.R., Dyson, H.M., Bateman, M.S., Green, A.J., Yates, J.R.W., and Dennis, N.R. 1998. Duplication of 8p23.1: A cytogenetic anomaly with no established clinical significance. *J. Med. Genet.* 35: 491–496.
- Bignell GR, Huang J, Greshock J, Watt S, Butler A, West S, Grigorova M, Jones KW, Wei W, Stratton MR, Futreal PA, Weber B, Shapero MH, Wooster R. High-resolution analysis of DNA copy number using oligonucleotide microarrays. *Genome Res.* 2004 Feb;14(2):287-95. PubMed PMID: 14762065; PubMed Central PMCID: PMC327104.
- Blouin MS, Parsons M, Lacaille V, Lotz S. Use of microsatellite loci to classify individuals by relatedness. *Mol Ecol.* 1996 Jun;5(3):393-401. PubMed PMID: 8688959.
- Chen WK, Swartz JD, Rush LJ, Alvarez CE. Mapping DNA structural variation in dogs. *Genome Res.* 2009 Mar;19(3):500-9. Epub 2008 Nov 17. PubMed PMID: 19015322; PubMed Central PMCID: PMC2661804.
- Chi B, DeLeeuw RJ, Coe BP, MacAulay C, Lam WL. SeeGH--a software tool for visualization of whole genome array comparative genomic hybridization data. *BMC Bioinformatics.* 2004 Feb 9;5:13. PubMed PMID: 15040819; PubMed Central PMCID:PMC373529.
- Conrad, D.F., Andrews, T.D., Carter, N.P., Hurles, M.E., and Pritchard, J.K. 2006. A high-resolution survey of deletion polymorphism in the human genome. *Nat. Genet.* **38**: 75–81
- Dakin EE, Avise JC. Microsatellite null alleles in parentage analysis. *Heredity.* 2004 Nov;93(5):504-9. Review. PubMed PMID: 15292911.
- Davies JJ, Wilson IM, Lam WL. Array CGH technologies and their applications to cancer genomes. *Chromosome Res.* 2005;13(3):237-48. Review. Erratum in: *Chromosome Res.* 2005;13(4):423. PubMed PMID: 15868418.
- Depienne C, Moreno-De-Luca D, Heron D, Bouteiller D, Gennetier A, Delorme R, Chaste P, Siffroi JP, Chantot-Bastarud S, Benyahia B, Trouillard O, Nygren G, Kopp S, Johansson M, Rastam M, Burglen L, Leguern E, Verloes A, Leboyer M, Brice A, Gillberg C, Betancur C. Screening for Genomic Rearrangements and Methylation Abnormalities of the 15q11-q13 Region in Autism Spectrum Disorders. *Biol Psychiatry.* 2009 Mar 17. [Epub ahead of print] PubMed PMID: 19278672.
- de Smith AJ, Walters RG, Froguel P, Blakemore AI. Human genes involved in copy number variation: mechanisms of origin, functional effects and implications for disease. *Cytogenet Genome Res.* 2008;123(1-4):17-26. Epub 2009 Mar 11. PubMed PMID: 19287135.
- Eenennaam AV. Genomics and Animal Agriculture. Animal and Biotechnology, UC Davis. 2008 Aug 20.
- Egli D, Domènech J, Selvaraj A, Balamurugan K, Hua H, Capdevila M, Georgiev O, Schaffner W, Atrian S. The four members of the Drosophila metallothionein family exhibit distinct yet overlapping roles in heavy metal homeostasis and detoxification. *Genes Cells.* 2006 Jun;11(6):647-58. PubMed PMID: 16716195.
- Eichler, E.E., Clark, R.A., and She, X. 2004. An assessment of the sequence gaps: Unfinished business in a finished human genome. *Nat. Rev. Genet.* **5**: 345–354.
- Engelen, J.J.M., Moog, U., Evers, J.L.H., Dassen, H., Albrechts, J.C.M., and Hamers, A.J.H. 2000. Duplication of chromosome region 8p23.1-p23.3: A benign variant? *Am. J. Med. Genet.* **91**: 18–21.

- Fombonne E. Epidemiology of autistic disorder and other pervasive developmental disorders. *J Clin Psychiatry*. 2005;66 Suppl 10:3-8. PubMed PMID: 16401144.
- Freeman JL, Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM, Aburatani H, Jones KW, Tyler-Smith C, Hurler ME, Carter NP, Scherer SW, Lee C. Copy number variation: new insights in genome diversity. *Genome Res*. 2006 Aug;16(8):949-61. Epub 2006 Jun 29. Review. PubMed PMID: 16809666.
- Glessner JT, Wang K, Cai G, Korvatska O, Kim CE, Wood S, Zhang H, Estes A, Brune CW, Bradfield JP, Imielinski M, Frackelton EC, Reichert J, Crawford EL, Munson J, Sleiman PM, Chiavacci R, Annaiah K, Thomas K, Hou C, Glaberson W, Flory J, Otieno F, Garris M, Soorya L, Klei L, Piven J, Meyer KJ, Anagnostou E, Sakurai T, Game RM, Rudd DS, Zurawiecki D, McDougle CJ, Davis LK, Miller J, Posey DJ, Michaels S, Kolevzon A, Silverman JM, Bernier R, Levy SE, Schultz RT, Dawson G, Owley T, McMahon WM, Wassink TH, Sweeney JA, Nurnberger JI, Coon H, Sutcliffe JS, Minshew NJ, Grant SF, Bucan M, Cook EH, Buxbaum JD, Devlin B, Schellenberg GD, Hakonarson H. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature*. 2009 Apr 28. [Epub ahead of print] PubMed PMID: 19404257.
- Gupta AR, State MW. Recent advances in the genetics of autism. *Biol Psychiatry*. 2007 Feb 15;61(4):429-37. Epub 2006 Sep 25. Review. PubMed PMID: 16996486.
- Hinds DA, Kloek AP, Jen M, Chen X, Frazer KA. Common deletions and SNPs are in linkage disequilibrium in the human genome. *Nat Genet*. 2006 Jan;38(1):82-5. Epub 2005 Dec 4. PubMed PMID: 16327809.
- Hodgson G, Hager JH, Volik S, Hariono S, Wernick M, Moore D, Nowak N, Albertson DG, Pinkel D, Collins C, Hanahan D, Gray JW. Genome scanning with array CGH delineates regional alterations in mouse islet carcinomas. *Nat Genet*. 2001 Dec;29(4):459-64. Erratum in: *Nat Genet* 2001 Dec;29(4):491. PubMed PMID: 11694878.
- Inoue, K. and Lupski, J.R. 2002. Molecular mechanisms for genomic disorders. *Annu. Rev. Genomics Hum. Genet.* 3: 199–242.
- Ishkanian AS, Malloff CA, Watson SK, DeLeeuw RJ, Chi B, Coe BP, Snijders A, Albertson DG, Pinkel D, Marra MA, Ling V, MacAulay C, Lam WL. A tiling resolution DNA microarray with complete coverage of the human genome. *Nat Genet*. 2004 Mar;36(3):299-303. Epub 2004 Feb 15. PubMed PMID: 14981516.
- Ji, Y., Eichler, E.E., Schwartz, S., and Nicholls, R.D. 2000. Structure of chromosomal duplicons and their role in mediating human genomic disorders. *Genome Res*. 10: 597–610.
- Katoh M, Katoh M. Human FOX gene family (Review). *Int J Oncol*. 2004 Nov;25(5):1495-500. Review. PubMed PMID: 15492844.
- Kidd JM, Cooper GM, Donahue WF, Hayden HS, Sampas N, Graves T, Hansen N, Teague B, Alkan C, Antonacci F, Haugen E, Zerr T, Yamada NA, Tsang P, Newman TL, Tüzün E, Cheng Z, Ebling HM, Tusneem N, David R, Gillett W, Phelps KA, Weaver M, Saranga D, Brand A, Tao W, Gustafson E, McKernan K, Chen L, Malig M, Smith JD, Korn JM, McCarroll SA, Altshuler DA, Peiffer DA, Dorschner M, Stamatoyannopoulos J, Schwartz D, Nickerson DA, Mullikin JC, Wilson RK, Bruhn L, Olson MV, Kaul R, Smith DR, Eichler EE. Mapping and sequencing of structural variation from eight human genomes. *Nature*. 2008 May 1;453(7191):56-64. PubMed PMID: 18451855; PubMed Central PMCID: PMC2424287.
- Klauck SM. Genetics of autism spectrum disorder. *Eur J Hum Genet*. 2006 Jun;14(6):714-20. Review. PubMed PMID: 16721407.
- Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, Simons JF, Kim PM, Palejev D, Carriero NJ, Du L, Taillon BE, Chen Z, Tanzer A, Saunders AC, Chi J, Yang F, Carter NP, Hurler ME, Weissman SM, Harkins TT, Gerstein MB, Egholm M, Snyder M. Paired-end mapping reveals extensive structural variation in the

human genome. *Science*. 2007 Oct 19;318(5849):420-6. Epub 2007 Sep 27. PubMed PMID: 17901297; PubMed Central PMCID: PMC2674581.

Lucito R, Healy J, Alexander J, Reiner A, Esposito D, Chi M, Rodgers L, Brady A, Sebat J, Troge J, West JA, Rostan S, Nguyen KC, Powers S, Ye KQ, Olshen A, Venkatraman E, Norton L, Wigler M. Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. *Genome Res*. 2003 Oct;13(10):2291-305. Epub 2003 Sep 15. PubMed PMID: 12975311; PubMed Central PMCID: PMC403708.

Lupski, J.R. 1998. Genomic disorders: Structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet*. 14: 417–422.

MacDonald ME, Barnes G, Srinidhi J, Duyao MP, Ambrose CM, Myers RH, Gray J, Conneally PM, Young A, Penney J, *et al*. Gametic but not somatic instability of CAG repeat length in Huntington's disease. *J Med Genet*. 1993 Dec;30(12):982-6. PubMed PMID: 8133508; PubMed Central PMCID: PMC1016628.

McCarroll, S.A., Hadnott, T.N., Perry, G.H., Sabeti, P., Zodi, M.C., Barrett, J., Dallaire, S., Gabriel, S.B., Lee, C., Daly, M.J., *et al*. 2006. Common deletion variants in the human genome. *Nat. Genet*. **38**: 86–92.

Newschaffer CJ, Croen LA, Daniels J, Giarelli E, Grether JK, Levy SE, Mandell DS, Miller LA, Pinto-Martin J, Reaven J, Reynolds AM, Rice CE, Schendel D, Windham GC. The epidemiology of autism spectrum disorders. *Annu Rev Public Health*. 2007;28:235-58. Review. PubMed PMID: 17367287.

Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, Brown PO. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet*. 1999 Sep;23(1):41-6. PubMed PMID: 10471496.

Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, González JR, Gratacòs M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. Global variation in copy number in the human genome. *Nature*. 2006 Nov 23;444(7118):444-54. PubMed PMID: 17122850; PubMed Central PMCID: PMC2669898.

Risch N, Spiker D, Lotspeich L, Nouri N, Hinds D, Hallmayer J, Kalaydjieva L, McCague P, Dimiceli S, Pitts T, Nguyen L, Yang J, Harper C, Thorpe D, Vermeer S, Young H, Hebert J, Lin A, Ferguson J, Chiotti C, Wiese-Slater S, Rogers T, Salmon B, Nicholas P, Petersen PB, Pingree C, McMahon W, Wong DL, Cavalli-Sforza LL, Kraemer HC, Myers RM. A genomic screen of autism: evidence for a multilocus etiology. *Am J Hum Genet*. 1999 Aug;65(2):493-507. PubMed PMID: 10417292; PubMed Central PMCID: PMC1377948.

Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J, Leotta A, Pai D, Zhang R, Lee YH, Hicks J, Spence SJ, Lee AT, Puura K, Lehtimäki T, Ledbetter D, Gregersen PK, Bregman J, Sutcliffe JS, Jobanputra V, Chung W, Warburton D, King MC, Skuse D, Geschwind DH, Gilliam TC, Ye K, Wigler M. Strong association of de novo copy number mutations with autism. *Science*. 2007 Apr 20;316(5823):445-9. Epub 2007 Mar 15. PubMed PMID: 17363630.

Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., Maner, S., Massa, H., Walker, M., Chi, M., *et al*. 2004. Large-scale copy number polymorphism in the human genome. *Science* **305**: 525–528.

Scherer SW, Cheung J, MacDonald JR, Osborne LR, Nakabayashi K, Herbrick JA, Carson AR, Parker-Katirae L, Skaug J, Khaja R, Zhang J, Hudek AK, Li M, Haddad M, Duggan GE, Fernandez BA, Kanematsu E, Gentles S, Christopoulos CC, Choufani S, Kwasnicka D, Zheng XH, Lai Z, Nusskern D, Zhang Q, Gu Z, Lu F, Zeesman S, Nowaczyk MJ, Teshima I, Chitayat D, Shuman C, Weksberg R, Zackai EH, Grebe TA, Cox SR, Kirkpatrick SJ, Rahman N, Friedman JM, Heng HH, Pelicci PG, Lo-Coco F, Belloni E, Shaffer LG, Pober B, Morton CC, Gusella JF, Bruns GA, Korf BR, Quade BJ, Ligon AH, Ferguson H, Higgins AW, Leach NT, Herrick SR, Lemyre E, Farra CG, Kim HG, Summers AM, Gripp KW, Roberts W, Szatmari P, Winsor EJ,

- Grzeschik KH, Teebi A, Minassian BA, Kere J, Armengol L, Pujana MA, Estivill X, Wilson MD, Koop BF, Tosi S, Moore GE, Boright AP, Zlotorynski E, Kerem B, Kroisel PM, Petek E, Oscier DG, Mould SJ, Döhner H, Döhner K, Rommens JM, Vincent JB, Venter JC, Li PW, Mural RJ, Adams MD, Tsui LC. Human chromosome 7: DNA sequence and biology. *Science*. 2003 May 2;300(5620):767-72. Epub 2003 Apr 10. PubMed PMID: 12690205.
- Scherer SW, Lee C, Birney E, Altshuler DM, Eichler EE, Carter NP, Hurles ME, Feuk L. Challenges and standards in integrating surveys of structural variation. *Nat Genet*. 2007 Jul;39(7 Suppl):S7-15. Review. PubMed PMID: 17597783.
- Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, Pertz LM, Clark RA, Schwartz S, Seagraves R, Oseroff VV, Albertson DG, Pinkel D, Eichler EE. Segmental duplications and copy-number variation in the human genome. *Am J Hum Genet*. 2005 Jul;77(1):78-88. Epub 2005 May 25. PubMed PMID: 15918152; PubMed Central PMCID: PMC1226196.
- Shaw, C.J. and Lupski, J.R. 2004. Implications of human genome architecture for rearrangement-based disorders: The genomic basis of disease. *Hum. Mol. Genet.* **13**: R57–R64.
- Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Döhner H, Cremer T, Lichter P. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer*. 1997 Dec;20(4):399-407. PubMed PMID: 9408757.
- Stankiewicz, P. and Lupski, J.R. 2002. Genome architecture, rearrangements and genomic disorders. *Trends Genet*. **18**: 74–82.
- The International HapMap Consortium. 2005. A haplotype map of the human genome. *Nature* **437**: 1299–1320.
- Thorisson GA, Stein LD. The SNP Consortium website: past, present and future. *Nucleic Acids Res*. 2003 Jan 1;31(1):124-7. PubMed PMID: 12519964; PubMed Central PMCID: PMC165499.
- Tuzun, E., Sharp, A.J., Bailey, J.A., Kaul, R., Morrison, V.A., Pertz, L.M., Haugen, E., Hayden, H., Albertson, D., Pinkel, D., *et al.* 2005. Fine-scale structural variation of the human genome. *Nat. Genet.* **37**: 727–732.
- Wang K, Zhang H, Ma D, Bucan M, Glessner JT, Abrahams BS, Salyakina D, Imielinski M, Bradfield JP, Sleiman PM, Kim CE, Hou C, Frackelton E, Chiavacci R, Takahashi N, Sakurai T, Rappaport E, Lajonchere CM, Munson J, Estes A, Korvatska O, Piven J, Sonnenblick LI, Alvarez Retuerto AI, Herman EI, Dong H, Hutman T, Sigman M, Ozonoff S, Klin A, Owley T, Sweeney JA, Brune CW, Cantor RM, Bernier R, Gilbert JR, Cuccaro ML, McMahon WM, Miller J, State MW, Wassink TH, Coon H, Levy SE, Schultz RT, Nurnberger JI, Haines JL, Sutcliffe JS, Cook EH, Minshew NJ, Buxbaum JD, Dawson G, Grant SF, Geschwind DH, Pericak-Vance MA, Schellenberg GD, Hakonarson H. Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature*. 2009 Apr 28. [Epub ahead of print] PubMed PMID: 19404256.
- Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, Saemundsen E, Stefansson H, Ferreira MA, Green T, Platt OS, Ruderfer DM, Walsh CA, Altshuler D, Chakravarti A, Tanzi RE, Stefansson K, Santangelo SL, Gusella JF, Sklar P, Wu BL, Daly MJ; Autism Consortium. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med*. 2008 Feb 14;358(7):667-75. Epub 2008 Jan 9. PubMed PMID: 18184952.
- Wong KK, deLeeuw RJ, Dosanjh NS, Kimm LR, Cheng Z, Horsman DE, MacAulay C, Ng RT, Brown CJ, Eichler EE, Lam WL. A comprehensive analysis of common copy-number variations in the human genome. *Am J Hum Genet*. 2007 Jan;80(1):91-104. Epub 2006 Dec 5. PubMed PMID: 17160897; PubMed Central PMCID: PMC1785303.

Ylstra B, van den Ijssel P, Carvalho B, Brakenhoff RH, Meijer GA. BAC to the future! or oligonucleotides: a perspective for micro array comparative genomic hybridization (array CGH). *Nucleic Acids Res.* 2006 Jan 26;34(2):445-50. Print 2006. Review. PubMed PMID: 16439806; PubMed Central PMCID: PMC1356528.

Zhao,X., Weir,B.A., LaFramboise,T., Lin,M., Beroukhim,R., Garraway,L., Beheshti,J., Lee,J.C., Naoki,K., Richards,W.G. *et al.* (2005) Homozygous deletions and chromosome amplifications in human lung carcinomas revealed by single nucleotide polymorphism array analysis. *Cancer Res.*, 65, 5561–5570.

Appendix

Appendix 1: Table A1. All Tier 1 and Tier 2 Candidate Autism CNVs

Unique ID	Clone_Name	Chr No.	Banding	BP StartPos (UCSC Mar 2006)	BP EndPos (UCSC Mar 2006)	Ratio (Normalized Cy3/Cy5 Log2 Ratio)	Standard Deviation	SNR Ch1 (595)	SNR Ch2 (685)	SDauto (standard deviation of the ratios of each probe from chr 1 22)	Thresh Score (Ratio/SD auto)	Sample (1 color / sample)	g/l
Tier 1													
RO32D18	RP11-465B22	1	p36.33	918163	1046699	0.239257	0.0398179	22.04	23.88	0.066748	3.584482	A26	gain
RO67D03	RP13-586E24	1	p36.32	2445100	2529860	0.269604	0.0375785	60.421	63.822	0.066748	4.039132	A26	gain
RO66F20	RP11-777P8	1	p36.13	16665329	16836250	0.3183	0.0291823	36.824	43.006	0.074813	4.254608	A8	gain
RO39F09	RP11-595I7	1	p36.11	24652247	24859224	0.299005	0.0381159	43.927	13.077	0.07167	4.171969	A75	gain
RO70O09	RP11-110H4	1	p36.11	25942591	26096516	0.238638	0.0114926	85.662	74.498	0.066748	3.575208	A26	gain
RO1F09	RP11-5J9	1	p34.3	36957071	37139825	-0.228811	0.0294411	59.623	42.789	0.066748	-3.42798	A26	loss
RO1B07	RP11-4M3	1	p33	46796925	46970566	0.258205	0.0256199	19.803	18.91	0.073771	3.500088	A137	gain
RO3P15	RP11-25A21	1	p32.1, p31.3	60807944	60980338	0.366584	0.0346807	14.537	12.389	0.073771	4.969216	A137	gain
RO42G20	RP11-627A12	1	p22.3	87569368	87753216	0.308185	0.0672437	13.552	13.867	0.073771	4.17759	A137	gain
RO26I24	RP11-361I6	1	p21.1	106247329	106413175	0.299415	0.0282921	39.587	30.651	0.066748	4.485752	A26	gain
RO1B14	RP11-7G12	1	q24.1	164123488	164287973	-0.237668	0.0029055	95.886	70.237	0.066748	-3.56068	A26	loss
RO20O17	RP11-267L3	1	q24.2	167222860	167381584	0.601437	0.0510008	15.779	21.489	0.073771	8.152757	A137	gain
RO27O02	RP11-380B22	1	q25.3	179466161	179671444	0.260274	0.0398688	25.499	27.352	0.073771	3.528134	A137	gain
RO58H09	CTD-2003B3	1	q32.1	204676883	204837914	0.257326	0.0132901	84.669	88.308	0.066748	3.855187	A26	gain
RO41A14	RP11-615M18	1	q32.2	205195912	205363199	0.488515	0.053383	13.854	15.353	0.073771	6.622047	A137	gain
RO50C17	RP11-718L9	1	q32.2	206299430	206471975	0.264449	0.0026008	15.212	12.054	0.073771	3.584728	A137	gain
RO17C23	RP11-196C8	2	p24.3	13032063	13188099	-0.238977	0.0139788	35.189	24.526	0.066748	-3.58029	A26	loss
RO50K16	RP11-723F23	2	p24.3	14978618	15161195	0.549819	0.0610438	12.692	16.756	0.073771	7.453051	A137	gain
RO36A02	RP11-526C24	2	p24.2	17558392	17595292	-0.257904	0.0307542	90.253	76.256	0.07167	-3.59849	A75	loss
RO34C02	RP11-482J4	2	p13.3	70684434	70893187	0.250953	0.0482021	18.036	20.897	0.073771	3.401784	A137	gain
RO2J19	RP11-15G4	2	p13.1	74988527	75142707	0.247251	0.0305937	50.426	54.109	0.07167	3.449853	A75	gain
RO12O17	RP11-124P5	2	p13.1	75115704	75274189	0.448689	0.0254438	22.177	19.261	0.07167	6.260486	A75	gain
RO50K22	RP11-723H1	2	p13.1	75226841	75425458	0.242051	0.0018498	38.945	46.245	0.07167	3.377299	A75	gain
RO6A04	RP11-58E9	2	p11.1	91654153	91685783	0.351565	5.23E-05	24.3	23.879	0.066748	5.267049	A26	gain
RO17I12	RP11-203K11	2	q11.1	95399978	95568811	0.464566	0.0002383	86.516	162.83	0.111086	4.182039	A39	gain
RO42O15	RP11-625O3	2	q11.2	95857272	96031199	-0.341205	0.0293874	16.539	16.305	0.100303	-3.40174	A51	loss
RO66P11	RP11-763O1	2	q11.2	96212625	96383203	0.350505	0.0059807	20.845	36.255	0.100303	3.494462	A51	gain
RO21H08	RP11-297B14	2	q13	113143567	113352583	0.279864	0.0194546	65.921	56.978	0.07167	3.904897	A75	gain
RO8C07	RP11-78E20	2	q14.2	121240000	121441961	-0.407064	0.0161383	64.844	99.175	0.111086	-3.6644	A39	loss
RO49H24	RP11-716M19	2	q14.2	121375188	121577515	0.403387	0.0208024	12.494	19.77	0.100303	4.021684	A51	gain
RO40E07	RP11-601F13	2	q22.1	138230775	138373820	-0.236377	0.0555022	14.119	12.67	0.06782	-3.48536	A118	loss
RO25K12	RP11-350N1	2	q22.1	140265043	140466338	0.321984	0.0300033	10.648	13.873	0.06782	4.747626	A118	gain
RO17I14	RP11-203K19	2	q31.1	176631931	176808639	0.298515	0.0255287	79.566	82.436	0.066748	4.472269	A26	gain
RO70O11	RP11-111C16	2	q31.1	177083977	177242722	0.292749	0.0007078	75.093	72.711	0.066748	4.385884	A26	gain
RO53G22	RP11-758O1	2	q31.1	177427021	177609208	0.273342	0.0136663	10.393	12.304	0.073771	3.705277	A137	gain
RO51C09	RP11-730M4	2	q31.3	180961085	181117649	0.296393	0.01393	15.189	16.221	0.073771	4.017744	A137	gain
RO25I10	RP11-355N23	2	q33.3	206900670	206969163	0.294401	0.0017416	64.761	56.097	0.066748	4.410634	A26	gain
RO40L08	RP11-611J23	2	q33.3, q34	208953051	209141331	0.319108	0.0113696	74.004	62.374	0.07167	4.452463	A75	gain
RO9N22	RP11-101G6	2	q36.1	221890614	222065341	-0.30663	0.0035702	18.414	12.658	0.066748	-4.59385	A26	loss
RO21D08	RP11-296C11	2	q36.2	225777102	225943830	0.257383	0.0116644	37.426	40.444	0.07167	3.591224	A75	gain
RO14G12	RP11-154A7	2	q37.1	232092877	232245500	0.247784	0.0653572	19.718	12.273	0.07167	3.45729	A75	gain
RO35M15	RP11-502F10	3	p22.1	40168920	40336928	0.270671	0.0225228	11.526	11.995	0.073771	3.669071	A137	gain
RO51N14	RP11-740P15	3	q22.1	43042214	43198446	0.336884	0.0356339	19.905	26.035	0.06782	4.967325	A118	gain
RO36D08	RP11-542F6	3	q25.1	134415378	134606591	0.318403	0.0585909	62.978	54.105	0.07167	4.442626	A75	gain
RO62P23	RP11-117K20	3	q26.31	159315572	159494764	0.27218	0.0254855	85.082	62.139	0.07167	3.797684	A75	gain
RO33O01	RP11-468N7	3	q27.1	173085106	173277471	0.299726	0.0652653	10.254	10.187	0.073771	4.062924	A137	gain
RO42J08	RP11-631P4	3	q29	185337137	185502784	0.24442	0.0008252	39.746	43.632	0.07167	3.410353	A75	gain
RO47L16	RP11-694E12	3	q29	193655933	193832580	-0.454193	0.0028617	109.59	82.12	0.07167	-6.33728	A75	loss
RO60N17	CTD-2235O1	3	q29	195414663	195511908	0.291416	0.042198	10.555	11.129	0.06782	4.296904	A118	gain
RO70O08	RP11-141L5	4	p16.3	819328	983918	0.32789	0.0011929	20.856	21.556	0.066748	4.912357	A26	gain
RO37F18	RP11-565M6	4	p16.3	1530742	1691087	0.222639	0.0349339	26.53	22.752	0.066748	3.335516	A26	gain

RO31O14	RP11-448F22	4	p16.3	2384072	2472983	0.332472	0.006866	23.859	26.812	0.073771	4.506812	A137	gain
RO31O13	RP11-444J4	4	p16.3	2475663	2659104	0.337627	0.0021786	19.321	25.107	0.073771	4.57669	A137	gain
RO30K14	RP11-431J16	4	p15.31	21912514	22118512	0.382154	0.042181	10.916	13.891	0.073771	5.180274	A137	gain
RO31P08	RP11-453M22	4	p14	37206944	37369549	0.283444	0.0598149	60.75	65.942	0.07167	3.954849	A75	gain
RO41N01	RP11-620D19	4	p11	48760365	48941412	-0.2838	0.0269542	127.71	87.773	0.066748	-4.25181	A26	loss
RO23K06	RP11-322A9	4	p11	48855123	48941378	-0.23954	0.0124458	102.81	77.167	0.066748	-3.58872	A26	loss
RO17B09	RP11-205G17	4	p11	49184918	49238843	-0.434054	0.0352295	38.012	11.879	0.112428	-3.86073	A1	loss
RO51E09	RP11-730P6	4	p11	49281935	49326349	-0.226318	0.0030547	128.64	87.012	0.066748	-3.39063	A26	loss
RO70L16	RP11-186C5	4	q13.3	75475766	75637578	-0.239452	0.0108392	45.775	43.097	0.07167	-3.34104	A75	loss
RO68K09	RP11-18022	4	q26	117414505	117570802	0.240949	0.0089449	70.636	68.062	0.066748	3.609831	A26	gain
RO27A12	RP11-375P10	4	q31.1	139960018	140140654	0.263714	0.0176508	56.051	59.119	0.066748	3.95089	A26	gain
RO4L03	RP11-37E18	4	q32.1	157999304	158162935	0.405522	0.0132561	13.9	13.703	0.073771	5.497038	A137	gain
RO45F04	RP11-663K16	4	q32.3	168924247	169095790	-0.333067	0.0223474	24.956	26.757	0.06782	-4.91104	A118	loss
RO45A04	RP11-656C20	5	p15.33	350722	447421	0.251496	0.0173114	64.233	58.454	0.066748	3.767843	A26	gain
RO14N24	RP11-161F13	5	p15.33	1562894	1724268	-0.255112	0.0135842	18.612	14.805	0.06782	-3.7616	A118	loss
RO55C23	RP11-778E14	5	p15.33	2592645	2774473	0.289887	0.0625217	21.958	25.165	0.073771	3.929552	A137	gain
RO70L14	RP11-186B21	5	q13.2	69527543	69653811	-0.30141	0.0005827	117.98	93.54	0.07167	-4.20553	A75	loss
RO58O16	CTD-2001M9	5	q13.2	70661551	70764108	-0.331258	0.0281294	57.908	43.987	0.07167	-4.62199	A75	loss
RO70G21	RP11-96H22	5	q13.3	74959279	75127937	0.262196	0.0008365	47.373	39.849	0.066748	3.928148	A26	gain
RO65J17	RP11-594A24	5	q13.3	75109899	75295154	0.414459	0.0038848	12.21	13.135	0.073771	5.618183	A137	gain
RO69C14	RP11-20M13	5	q14.3	82908893	83054341	0.32423	0.0950719	10.952	10.522	0.06782	4.780743	A118	gain
RO40O18	RP11-606H14	5	q33.2	153245135	153421108	0.343733	0.0157204	21.997	29.774	0.073771	4.65946	A137	gain
RO58G19	RP11-812K10	6	p25.3	71610	284699	0.26557	0.0093798	58.101	55.975	0.066748	3.978696	A26	gain
RO66I01	RP11-664H18	6	p25.1	5710060	5886212	-0.246554	0.091728	129.51	87.015	0.066748	-3.6938	A26	loss
RO70O21	RP11-113F21	6	p22.1	29148090	29304253	0.286076	0.0039669	21.004	19.488	0.066748	4.285911	A26	gain
RO58D24	CTD-2007P10	6	p21.33	30972604	31085380	0.423626	0.0591509	53.84	95.403	0.111086	3.813496	A39	gain
RO12D20	RP11-135J9	6	p21.33	31440461	31606959	0.258084	0.0278176	29.574	34.486	0.073771	3.498448	A137	gain
RO10D07	RP11-107C8	6	p21.31	34540001	34703255	0.249976	0.0578555	19.305	12.838	0.073771	3.38854	A137	gain
RO43I01	RP11-634F13	6	p21.1	41456894	41671891	0.2532961	0.091675	58.274	133.7	0.111086	4.797733	A39	gain
RO42C17	RP11-624F22	6	p21.1	41510587	41701770	0.431051	0.0025908	54.377	81.485	0.111086	3.880336	A39	gain
RO70I24	RP11-130K11	6	q14.3	86034664	86138958	-0.285843	0.0246773	13.995	11.292	0.06782	-4.21473	A118	loss
RO42O20	RP11-628G14	6	q15	87679748	87841029	0.399379	0.047176	27.246	34.081	0.073771	5.413767	A137	gain
RO50O10	RP11-723N6	6	q15	88886560	89038114	0.356687	0.0051315	13.218	19.318	0.073771	4.835057	A137	gain
RO40D08	RP11-610E23	6	q16.3	101611771	101775605	0.250149	0.0170286	67.66	52.704	0.07167	3.490289	A75	gain
RO64G13	RP11-346C16	6	q21	110941778	111130824	0.28682	0.0018873	11.036	13.805	0.073771	3.887978	A137	gain
RO61G21	CTD-2310D7	6	q22.1	113915311	114052447	-0.350276	0.0871679	14.866	11.191	0.06782	-5.16479	A118	loss
RO28L07	RP11-403C11	6	q22.1	114313227	114352792	-0.257576	0.0301249	83.226	69.687	0.066748	-3.85893	A26	loss
RO32K13	RP11-455L8	6	q23.2	131432645	131623624	0.358431	0.0160195	12.156	14.356	0.073771	4.858698	A137	gain
RO32K01	RP11-455I14	6	q25.2	155418033	155593336	0.262583	0.0910598	13.557	16.742	0.073771	3.559434	A137	gain
RO50J04	RP11-728F5	6	q27	170345513	170469123	0.262764	0.0238634	61.703	55.16	0.066748	3.936657	A26	gain
RO60F04	CTD-2260L17	7	p22.3	797562	898228	0.226066	0.0228289	42.207	43.148	0.066748	3.386858	A26	gain
RO31K13	RP11-442P13	7	p22.1	5007384	5179039	0.329093	0.0304544	23.405	28.441	0.073771	4.461008	A137	gain
RO61G10	CTD-2380A4	7	p21.2	15155609	15320915	0.242054	0.0106752	10.604	10.458	0.066748	3.626386	A26	gain
RO32K03	RP11-455J15	7	p21.1	16700151	16863920	0.318534	0.0127032	15.958	16.922	0.073771	4.317876	A137	gain
RO40E05	RP11-601F7	7	p15.3	24617413	24785003	0.271527	0.0003769	14.893	11.649	0.073771	3.680674	A137	gain
RO61O20	CTD-2010E11	7	p14.1	40782944	40818119	0.242927	0.0239009	39.819	40.455	0.066748	3.639465	A26	gain
RO36D20	RP11-542I1	7	p11.2	56616401	56796200	0.365986	0.0423444	95.806	72.534	0.07167	5.106544	A75	gain
RO63N18	RP11-325K1	7	q11.21	66036595	66209226	0.522774	0.0812204	12.339	14.063	0.073771	7.086443	A137	gain
RO35L04	RP11-519I5	7	q11.22	71722139	71877344	-0.2311	0.0976005	15.116	14.405	0.06782	-3.40755	A118	loss
RO68N06	RP11-818I12	7	q31.33	124755570	124943364	0.28215	0.0416231	17.913	17.156	0.066748	4.227093	A26	gain
RO22J13	RP11-313O20	7	q34	140553376	140720323	0.255915	0.0944334	14.379	18.6	0.073771	3.469046	A137	gain
RO13F09	RP11-143G2	7	q34	141730127	141884603	0.265932	0.0333132	29.643	34.038	0.074813	3.554623	A8	gain
RO52H15	RP11-749O22	7	q36.1	148697310	148888695	0.306651	0.0160294	52.567	50.532	0.066748	4.59416	A26	gain
RO30O05	RP11-428M3	7	q36.1	150572713	150800286	0.228269	0.0197092	45.52	44.686	0.066748	3.419863	A26	gain
RO3L16	RP11-26I6	7	q36.2	152632106	152791685	0.281292	0.0641551	12.535	16.133	0.073771	3.813043	A137	gain
RO59B10	CTD-2041J20	7	q36.3	157396437	157519058	0.26356	0.0291922	17.256	19.154	0.074813	3.522917	A8	gain
RO52G07	RP11-744K13	7	q36.3	157760321	157945126	0.49325	0.0090227	29.363	62.037	0.111086	4.440253	A39	gain
RO44L08	RP11-652N4	8	p23.2	4901876	5127620	0.240296	0.0615756	74.877	55.144	0.07167	3.352811	A75	gain
RO42A02	RP11-625P23	8	q12.1	58003559	58174053	0.249751	0.0119968	14.929	13.913	0.073771	3.38549	A137	gain
RO6J22	RP11-65D13	8	q21.11	77643686	77816008	0.269201	0.0592499	12.038	10.021	0.07167	3.756118	A75	gain
RO59H22	CTD-2085I3	8	q22.1	94655709	94761228	0.224219	0.0118865	96.433	104.99	0.066748	3.359187	A26	gain

RO50K04	RP11-723E18	8	q24.13	126155750	126356347	0.561082	0.0029083	13.207	17.029	0.073771	7.605726	A137	gain
RO35J04	RP11-518L1	8	q24.3	144584483	144718307	0.245826	0.0073539	15.161	16	0.066748	3.682897	A26	gain
RO57J07	RP11-806N14	9	p22.2	18221982	18449492	0.272444	0.0374392	25.771	25.428	0.07167	3.801367	A75	gain
RO41D22	RP11-621D23	9	p13.3	35715248	35951025	0.44273	0.0636283	87.364	194.47	0.111086	3.985471	A39	gain
RO22M24	RP11-312A20	9	p13.3	35825343	36043213	0.232816	0.0090786	53.073	54.533	0.066748	3.487985	A26	gain
RO26P18	RP11-370J5	9	p13.1	39073855	39224644	-0.244675	0.0318827	59.581	51.682	0.07167	-3.41391	A75	loss
RO17M12	RP11-204I2	9	p11.2	43583843	43746434	-0.255637	0.0269047	17.831	14.053	0.07167	-3.56686	A75	loss
RO8H01	RP11-88D3	9	q12	66405147	66476865	0.251042	0.0954679	35.263	40.091	0.073771	3.40299	A137	gain
RO23I05	RP11-318L1	9	q12	68164115	68282615	-0.387626	0.0059263	16.026	14.248	0.112428	-3.44777	A1	loss
RO8D23	RP11-87H9	9	q12	68221319	68447264	-0.376495	0.0375693	32.902	37.234	0.112428	-3.34877	A1	loss
RO59G20	CTD-2019O15	9	q21.32	84869245	85007158	0.363446	0.0463226	16.402	19.346	0.06782	5.35898	A118	gain
RO32C03	RP11-454G15	9	q31.3	110156709	110355803	0.678205	0.0187376	11.699	15.267	0.073771	9.193382	A137	gain
RO18E03	RP11-220D10	9	q31.3	110395342	110547877	0.350597	0.0375134	24.441	32.858	0.100303	3.495379	A51	gain
RO50K10	RP11-723F2	9	q31.3	111163247	111380118	0.310704	0.0454656	23.687	28.806	0.073771	4.211736	A137	gain
RO48C19	RP11-695L17	9	q34.3	138380703	138494631	0.222865	0.0281676	26.798	27.225	0.066748	3.338902	A26	gain
RO62H21	RP11-102I24	10	p13	13089369	13228260	0.251713	0.0154517	91.261	94.577	0.066748	3.771094	A26	gain
RO67J07	RP13-617N7	10	q11.1	42020851	42040383	0.223057	0.0110125	97.237	106.62	0.066748	3.341778	A26	gain
RO61K11	CTD-2318O3	10	q11.1	42020851	42040383	0.42606	0.0017692	127.98	314.81	0.111086	3.835407	A39	gain
RO58F11	RP11-818P1	10	q11.21	42987824	43202215	0.281415	0.0025307	10.327	11.606	0.074813	3.761579	A8	gain
RO4M13	RP11-30N1	10	q11.23	51121592	51195063	-0.289387	0.005939	17.055	13.701	0.06782	-4.26699	A118	loss
RO50C08	RP11-722F10	10	q23.1	84452403	84618004	0.251853	0.0120866	18.7	18.324	0.073771	3.413984	A137	gain
RO64F18	RP11-448I7	10	q26.3	134955684	135066765	0.247872	0.0041274	35.496	35.033	0.066748	3.713549	A26	gain
RO62A23	RP11-23J21	11	p15.4	3349519	3535987	-0.258074	0.0404571	59.242	41.311	0.07167	-3.60087	A75	loss
RO66P09	RP11-763N14	11	p15.1	16437283	16630905	0.261995	0.0106759	124.09	121.67	0.066748	3.925136	A26	gain
RO66P21	RP11-766K9	11	p14.3	23361050	23525759	0.262915	0.0372596	102.82	109.38	0.066748	3.93892	A26	gain
RO16P20	RP11-195I23	11	p11.2	44985558	45145067	0.298712	0.0206256	60.993	66.465	0.07167	4.167881	A75	gain
RO35H03	RP11-511O3	11	q12.2	59935908	59963112	-0.346224	0.0060762	122.3	96.389	0.066748	-5.18703	A26	loss
RO44E17	RP11-642F7	11	q13.1	65023037	65185629	0.270586	0.0029536	39.971	45.262	0.07167	3.775443	A75	loss
RO56C23	RP11-791I20	11	q13.4	74652389	74835071	0.263234	0.0774126	19.616	19.974	0.073771	3.568259	A137	gain
RO17P04	RP11-217K21	11	q13.5	76392326	76536143	0.247836	0.0317434	46.635	15.624	0.066748	3.71301	A26	gain
RO18B14	RP11-239K22	11	q21	95994439	96152262	0.290185	0.0924047	16.715	26.075	0.06782	4.278753	A118	gain
RO14C05	RP11-150H9	11	q23.3	118603789	118758082	0.230834	0.0169338	44.992	41.54	0.066748	3.458291	A26	gain
RO21F12	RP11-296M15	11	q25	133545363	133721889	0.271078	0.0195374	71.407	71.264	0.066748	4.061215	A26	gain
RO50N14	RP11-729M2	12	p13.33	2098197	2246866	0.307968	0.0250917	21.003	26.011	0.06782	4.540961	A118	gain
RO31K14	RP11-447J10	12	p13.31	5573581	5743850	0.312687	0.0919338	11.957	22.74	0.073771	4.238617	A137	gain
RO31C13	RP11-440N2	12	q12	36720939	36889175	0.405225	0.0549528	16.721	21.189	0.073771	5.493012	A137	gain
RO14P21	RP11-158I2	12	q13.13	52541438	52732799	0.458524	0.009104	79.974	128.2	0.111086	4.127649	A39	gain
RO32J03	RP11-462E2	12	q13.13	52695946	52866729	0.609011	0.0088091	114.9	273.57	0.111086	5.482338	A39	gain
RO27L23	RP11-383J7	12	q13.2	53041821	53231620	0.389835	0.0262153	79.703	103.39	0.111086	3.509308	A39	gain
RO3L04	RP11-26G8	12	q14.1	57636373	57812533	0.362373	0.0441885	18.937	20.302	0.073771	4.912133	A137	gain
RO4C09	RP11-27M6	12	q23.3	106831834	106997898	-0.32588	0.0318997	37.869	31.34	0.07167	-4.54695	A75	loss
RO1A21	RP11-1C11	12	q23.3	106899825	107062279	-0.332661	0.0360865	49.144	20.44	0.07167	-4.64157	A75	loss
RO60N16	CTD-2280A16	12	q24.13	112581835	112696534	0.316783	0.0179711	11.197	13.785	0.06782	4.670938	A118	loss
RO40L15	RP11-608N12	13	q12.13	25309664	25502697	0.404711	0.0517673	15.196	20.754	0.06782	5.967428	A118	gain
RO1B03	RP11-4K24	13	q14.11	41655648	41810459	0.308232	0.0097277	10.817	12.028	0.073771	4.178227	A137	gain
RO32C14	RP11-457M18	13	q14.11	42325670	42471777	0.316926	0.0148054	13.269	15.378	0.073771	4.296078	A137	gain
RO32K02	RP11-459A1	13	q22.2	74479268	74687974	0.357404	0.0224436	12.736	14.695	0.073771	4.844776	A137	gain
RO1I13	RP11-1M7	13	q22.2	75198429	75381103	0.535017	0.0674467	11.431	14.913	0.073771	7.252403	A137	gain
RO68D22	RP11-480F17	13	q22.3	76260854	76446840	0.312069	0.0027563	17.022	16.79	0.07167	4.354249	A75	gain
RO9P19	RP11-991I1	13	q34	112550821	112746773	0.284104	0.0371097	21.032	23.396	0.073771	3.851161	A137	gain
RO35N22	RP11-520H13	14	q23.3	64706822	64880159	0.264777	0.0223679	18.641	17.335	0.07167	3.694391	A75	gain
RO61E15	CTD-2305J20	14	q32.33	104788384	104893766	0.256025	0.0170491	33.843	38.265	0.07167	3.572276	A75	gain
RO67D17	RP13-594H12	14	q32.33	104822654	104944254	0.290204	0.0104206	46.644	49.238	0.066748	4.347756	A26	gain
RO40H07	RP11-607H20	15	q11.1	18273500	18473311	-0.251218	0.0055642	14.473	10.047	0.074813	-3.35795	A8	loss
RO44A01	RP11-641K15	15	q11.2	18781625	18946020	-0.248226	0.0027648	24.739	23.85	0.06782	-3.66007	A118	loss
RO28L19	RP11-403L7	15	q11.2	18898414	19048601	-0.385608	0.0093409	255.27	386.61	0.111086	-3.47126	A39	loss
RO37J13	RP11-561P13	15	q13.1	26135317	26310601	-0.245656	0.0483739	18.284	15.656	0.06782	-3.62218	A118	loss
RO61F22	RP11-6L23	15	q24.1	71985268	72141641	0.445803	0.0017855	27.509	52.143	0.111086	4.013134	A39	gain
RO54P21	RP11-775D1	15	q24.1	72244696	72412336	-0.379857	0.0284384	121.92	179.79	0.111086	-3.41949	A39	loss
RO34D17	RP11-486P18	16	p12.3	20442292	20598180	0.254507	0.0022627	26.95	31.016	0.06782	3.752684	A118	gain
RO42E14	RP11-626K17	16	p11.2	32287041	32440412	0.259538	0.0068257	115.8	119.75	0.066748	3.888326	A26	gain
RO1C23	RP11-1F10	16	q12.1	47811144	48013771	-0.233299	0.0091252	56.105	22.273	0.066748	-3.49522	A26	loss

RO40K14	RP11-605N4	16	q13	55095980	55251063	0.388252	0.0066532	68.105	163.08	0.111086	3.495058	A39	gain
RO24J03	RP11-343H19	16	q13	55211316	55369561	0.584413	0.0157119	44.675	141.42	0.111086	5.260906	A39	gain
RO57F06	RP11-809H20	17	p11.2	17534762	17706303	0.313564	0.0009701	34.221	36.099	0.066748	4.697729	A26	gain
RO13D12	RP11-146G23	17	q11.2	28624283	28648111	-0.259572	0.0029571	96.169	66.917	0.066748	-3.88884	A26	loss
RO29J09	RP11-420B16	17	q21.32	43866044	44065387	0.24495	0.0301362	49.254	43.499	0.066748	3.669773	A26	gain
RO40C11	RP11-600O7	17	q21.32	44003179	44186865	0.452488	0.0085277	58.569	140.63	0.111086	4.073313	A39	gain
RO31G13	RP11-441O7	17	q23.3	58730625	58892053	0.306991	0.0270136	15.733	20.688	0.073771	4.161405	A137	gain
RO67A21	RP11-802F16	17	q25.3	73485102	73659590	-0.261082	0.0710048	23.332	14.339	0.074813	-3.48979	A8	loss
RO56D22	RP11-798C7	17	q25.3	76834696	77076142	0.248783	0.0440796	13.981	22.257	0.07167	3.471229	A75	gain
RO67I10	RP13-538H2	18	q11.2	19017248	19059585	0.239763	0.0025887	42.967	48.331	0.07167	3.345375	A75	gain
RO45A21	RP11-654B8	18	q21.1	42715677	42800239	0.50121	0.0158717	83.348	90.309	0.066748	7.508989	A26	gain
RO31P20	RP11-453P5	18	q21.32	56301614	56482301	0.244348	0.0560184	50.517	55.016	0.07167	3.409348	A75	gain
RO49C12	RP11-711F2	18	q23	71337342	71520536	0.238207	0.0453517	10.429	11.517	0.06782	3.512341	A118	gain
RO69E17	RP11-7H17	18	q23	75216360	75402240	0.256705	0.0354918	54.956	54.595	0.066748	3.845883	A26	gain
RO60F06	CTD-2260M22	19	p13.3	1377765	1490288	0.242766	0.0050084	26.584	32.1	0.07167	3.387275	A75	gain
RO67F13	RP13-600E22	19	p13.3	3408988	3533861	0.24949	0.0098373	66.675	79.588	0.07167	3.481094	A75	gain
RO5G04	RP11-46L10	19	p13.12	14015002	14153121	0.25763	0.0176784	42.295	45.186	0.07167	3.59467	A75	gain
RO44I16	RP11-652K21	19	p13.12	14069923	14245943	0.253994	0.0347055	49.545	44.306	0.07167	3.543937	A75	gain
RO36H08	RP11-543F7	19	q12	33742778	33936080	0.331168	0.0272844	81.343	58.754	0.07167	4.620734	A75	gain
RO10E12	RP11-104J24	19	q12	34942164	35091892	0.223538	0.0275906	32.664	11.699	0.066748	3.348984	A26	gain
RO44H08	RP11-652C23	19	q12	35155612	35336561	0.261465	0.0206779	74.085	58.742	0.07167	3.648179	A75	gain
RO61E06	CTD-2344L3	19	q13.32	50576103	50664187	0.225805	0.0038332	50.43	46.282	0.066748	3.382948	A26	gain
RO48B12	RP11-705C4	19	q13.42	60715146	60850444	0.241451	0.0253611	26.324	26.441	0.07167	3.368927	A75	gain
RO57M10	RP11-804M21	20	p13	1895714	2066809	-0.225321	0.0721362	83.454	56.92	0.066748	-3.3757	A26	loss
RO50D09	RP11-724J12	20	q11.21	30432283	30631518	0.230552	0.0162274	50.36	58.786	0.066748	3.454066	A26	gain
RO46G09	RP11-667G22	20	q11.22	31699917	31807042	0.234802	0.0086402	56.248	52.975	0.066748	3.517738	A26	gain
RO27A22	RP11-376B8	20	q11.22	33077333	33284163	-0.240186	0.0118327	17.418	13.416	0.06782	-3.54152	A118	loss
RO40L24	RP11-612A10	20	q11.22	33095140	33146318	0.266599	0.0203116	11.195	13.373	0.073771	3.613873	A137	gain
RO31O15	RP11-444K9	20	q13.13	47777796	47958064	0.448028	0.0844809	17.262	20.943	0.073771	6.073227	A137	gain
RO19D14	RP11-262B23	20	q13.2	53643329	53684580	0.239257	0.0042936	32.376	39.968	0.07167	3.338314	A75	gain
RO50C09	RP11-718I14	20	q13.33	60103776	60288903	0.224404	0.0118186	53.935	52.522	0.066748	3.361958	A26	gain
RO33J16	RP11-476I15	20	q13.33	62263781	62434320	0.26224	0.0019177	34.083	32.494	0.066748	3.928807	A26	gain
RO19F11	RP11-259G22	21	p11.1	10032184	10197104	-0.25157	0.0171311	96.007	72.421	0.07167	-3.51012	A75	loss
RO70F14	RP11-175M15	21	p11.1	10037729	10092616	0.246981	0.0129584	12.552	12.84	0.073771	3.347942	A137	gain
RO19P08	RP11-264C24	21	q21.1	17391475	17557175	0.250601	0.0384228	16.228	12.131	0.073771	3.397012	A137	gain
RO45B24	RP11-663E24	21	q21.3	28721529	28897123	-0.297887	0.0177427	22.795	19.221	0.06782	-4.39232	A118	loss
RO23D14	RP11-327M4	21	q22.3	44413607	44443805	0.261532	0.0176996	19.003	21.164	0.07167	3.649114	A75	gain
RO28K16	RP11-397E9	21	q22.3	44699015	44846952	0.447093	0.0508791	56.579	144.49	0.111086	4.024747	A39	gain
RO63G13	RP11-161A23	21	q22.3	45025735	45208462	0.237371	0.0313729	91.117	96.215	0.066748	3.556226	A26	gain
RO44H20	RP11-652F11	22	q21.1	18002150	18171498	0.291579	0.0207932	72.264	53.665	0.07167	4.068355	A75	gain
RO19B20	RP11-261O10	22	q21.2	24039097	24206958	0.260817	0.0099759	20.481	22.21	0.074813	3.486252	A8	gain
RO62B05	RP11-86H17	22	q21.3	28894288	29037473	0.259636	0.0290522	13.641	13.372	0.073771	3.519486	A137	gain
RO48P01	RP11-704C5	22	q22.13	37362475	37572224	0.262249	0.0248796	59.37	67.354	0.066748	3.928942	A26	gain
RO29J12	RP11-423E19	22	q22.3	49413876	49518700	0.241691	0.0292466	56.172	33.546	0.066748	3.620947	A26	gain
RO21H20	RP11-297F13	X	p11.23	48575765	48734088	0.349604	0.0088735	52.794	65.633	0.07167	4.877968	A75	gain
RO17B08	RP11-211H10	X	p11.23	49201123	49376616	-0.228992	0.0939533	17.517	14.119	0.06782	-3.37647	A118	loss
RO32D20	RP11-465B24	X	p11.21	56465528	56632629	0.259046	0.0620309	74.634	88.328	0.07167	3.614427	A75	gain
RO58D07	RP11-818I17	X	p11.1	56752841	56935688	-0.446649	0.0071086	30.331	19.819	0.06782	-6.5858	A118	loss
RO18M04	RP11-231J22	X	q22.3	105097484	105276081	-0.247856	0.0071793	32.716	25.318	0.07167	-3.45829	A75	loss
RO4A15	RP11-27E21	Y	p11.2	6053399	6213222	0.586923	0.084705	91.583	220.15	0.111086	5.283501	A39	gain
RO53L12	RP11-764M21	Y	p11.2	6507372	6664486	0.321714	0.0264861	62.714	66.877	0.066748	4.81983	A26	gain
RO65F22	RP11-622D24	Y	p11.2	6516741	6719180	0.234058	0.0366642	36.348	29.581	0.066748	3.506592	A26	gain
RO19M18	RP11-258E22	Y	p11.2	6540216	6691826	0.246632	0.0108286	70.828	69.215	0.066748	3.694972	A26	gain
RO10J07	RP11-108F14	Y	p11.2	7755631	7921510	-0.250488	0.0038544	40.62	30.877	0.07167	-3.49502	A75	loss
RO27A16	RP11-375P13	Y	p11.2	8465905	8635791	-0.287174	0.0194659	128.31	92.872	0.066748	-4.30236	A26	loss
RO31G02	RP11-446M21	Y	q11.221	15089236	15303617	0.285442	0.0438328	10.655	11.804	0.073771	3.869298	A137	gain
RO1B21	RP11-5C5	Y	q11.223	23920245	24077176	-0.292971	0.0138734	77.592	31.379	0.066748	-4.38921	A26	loss
Tier 2													
RO34A05	RP11-478I22	1	p36.23	7959002	8148278	0.282164	0.0671978	8.309	8.653	0.076899	3.66928	A137	gain
RO68A24	RP11-84P16	1	p36.11	27205702	27404075	0.299351	0.0198414	7.62	5.099	0.076899	3.892781	A137	gain
RO68M09	RP11-561A23	1	p33	47992238	48169933	0.298906	0.0707262	9.241	9.201	0.071866	4.159213	A118	gain
RO9O23	RP11-94O18	1	p33	49280862	49433096	0.247428	0.0945091	3.008	3.503	0.071866	3.442908	A118	gain

RO2M22	RP11-13M4	1	p31.3	63033731	63227884	-0.246048	0.0387297	4.549	3.099	0.071851	-3.42442	A75	loss
RO38H12	RP11-585M16	1	p31.1	74734998	74917065	0.504963	0.0274138	5.538	8.822	0.071866	7.026452	A118	gain
RO22N12	RP11-317F9	1	q23.1	156709465	156863209	0.274621	0.0775682	6.877	4.443	0.076899	3.571191	A137	gain
RO68M21	RP11-719L1	1	q31.3	193537539	193704669	0.430891	0.0171587	5.064	5.885	0.071866	5.995756	A118	gain
RO29B15	RP11-419B18	1	q31.3	195446765	195615132	0.269704	0.0397316	5.837	7.013	0.076899	3.50725	A137	gain
RO13G21	RP11-138M22	1	q42.13	228338069	228505163	0.521773	0.0348915	49.808	58.39	0.066893	7.799471	A26	gain
RO20J04	RP11-278O12	1	q42.3	232875925	233058172	0.259071	0.015055	5.784	4.822	0.076849	3.371169	A8	gain
RO68L18	RP11-488L18	1	q44	245339707	245542773	-0.267396	0.0339793	3.592	4.095	0.071866	-3.72076	A118	loss
RO61B09	CTD-2014D7	2	p16.1	57197583	57336716	0.370165	0.0298074	5.55	6.853	0.071866	5.150767	A118	gain
RO65N17	RP11-602P8	2	p12	75843805	76014895	0.390207	0.0110931	7.573	8.107	0.076899	5.074279	A137	gain
RO22D07	RP11-313B1	2	q21.1	130401477	130570107	0.539988	0.0963221	5.352	5.731	0.111152	4.858104	A39	gain
RO54G16	RP11-769N11	2	q21.1	130801038	130941683	0.379786	0.0721277	110.16	291.78	0.111152	3.416817	A39	gain
RO2J05	RP11-15D9	2	q22.1	139452943	139619513	-0.293557	0.0115449	6.298	4.227	0.076849	-3.81992	A8	loss
RO43C13	RP11-633H24	2	q24.1	154895951	155104293	0.252524	0.0343442	4.399	6.016	0.071866	3.513817	A118	gain
RO50I03	RP11-719O1	2	q24.3	165653881	165842181	0.27515	0.0594719	9.813	10.675	0.076899	3.57807	A137	gain
RO32K05	RP11-455J20	2	q32.2	190268101	190449901	0.315058	0.0092207	5.281	5.871	0.076899	4.097036	A137	gain
RO5E07	RP11-44A23	2	q34	212401435	212568601	-0.274036	0.0718385	11.777	6.499	0.076849	-3.5659	A8	loss
RO57F17	RP11-806C2	3	p26.3	1841954	2043596	0.240232	0.0079839	8.316	12.329	0.071866	3.342777	A118	gain
RO44C07	RP11-642A2	3	p24.3	21962388	22134044	0.422388	0.0169458	8.143	8.936	0.076899	5.492763	A137	gain
RO60I18	CTD-2185K4	3	p11.1	90525740	90584932	-0.24723	0.0648275	6.083	5.258	0.071866	-3.44015	A118	loss
RO43N10	RP11-141F4	3	q11.2	99591113	99756754	0.302473	0.0800014	6.077	5.741	0.071851	4.209726	A75	gain
RO3E02	RP11-20N7	3	q13.11	107633524	107818033	0.269103	0.0236845	8.401	8.459	0.076899	3.499434	A137	gain
RO65N05	RP11-599H1	3	q13.12	109384850	109586860	0.450016	0.0970822	8.18	8.958	0.076899	5.85204	A137	gain
RO68E12	RP11-539H11	3	q13.12	109435013	109636045	0.287275	0.0051343	6.524	4.181	0.076899	3.735744	A137	gain
RO50J03	RP11-725I19	3	q24	148478234	148676766	0.32309	0.0748734	6.35	8.755	0.071866	4.495728	A118	gain
RO27E10	RP11-378E6	3	q24	149152070	149322873	-0.260601	0.0161051	5.636	4.4	0.076849	-3.39108	A8	loss
RO66M20	RP11-718K10	3	q25.2	154255436	154412488	0.260832	0.0192821	6.535	7.868	0.076899	3.391878	A137	gain
RO45K17	RP11-655G22	3	q28	193133121	193305756	0.264123	0.0447648	16.052	8.214	0.076899	3.434674	A137	gain
RO59J10	CTD-209O84	3	q29	194228217	194348803	0.906815	0.0213221	4.026	6.548	0.076899	11.79229	A137	gain
RO68A21	RP11-72I14	4	p16.3	2716373	2926405	0.412427	0.0956369	3.812	3.715	0.076899	5.36323	A137	gain
RO10B01	RP11-106H18	4	q13.3	72039825	72197229	-0.276666	0.0546742	5.111	5.205	0.071866	-3.84975	A118	loss
RO63N22	RP11-326O22	4	q13.3	72541949	72732136	0.393061	0.0421344	5.277	5.845	0.076899	5.113393	A137	gain
RO2B03	RP11-14A23	4	q21.23	86301196	86469048	0.355453	0.0407534	8.115	8.915	0.076899	4.622336	A137	gain
RO61H24	RP11-9H11	4	q26	115629364	115809774	0.364207	0.0365235	4.498	5.521	0.071866	5.067862	A118	gain
RO4I01	RP11-29M9	4	q28.3	134450123	134636219	-0.257497	0.0380494	8.058	6.269	0.076899	-3.34851	A137	loss
RO2D21	RP11-14K14	4	q35.2	190673393	190811644	0.398252	0.0870498	113.57	202.53	0.111152	3.582949	A39	gain
RO59C05	CTD-2011O21	4	q35.2	191039899	191138136	-0.371553	0.0119558	50.382	28.003	0.111027	-3.34651	A1	loss
RO9L14	RP11-101B14	5	q11.2	56361220	56544061	0.260312	0.0144271	9.004	12.895	0.076899	3.385116	A137	gain
RO2B04	RP11-15O24	5	q11.2	58487753	58646102	0.26282	0.0058421	5.667	5.776	0.076899	3.41773	A137	gain
RO66F05	RP11-734L8	5	q21.1	99194930	99380487	0.284348	0.0069671	8.292	8.876	0.076899	3.697681	A137	gain
RO59H12	CTD-2075G19	5	q21.2	103890873	104011391	0.319472	0.0169826	6.677	7.662	0.071866	4.445384	A118	gain
RO62P24	RP11-144H10	5	q23.1	120001738	120179752	0.405011	0.0413233	6.166	6.076	0.071866	5.635641	A118	gain
RO56F03	RP11-795P7	5	q23.2	123122540	123337265	0.434802	0.0284611	8.257	12.073	0.071866	6.050177	A118	gain
RO71D12	RP11-281H9	5	q23.3	128933631	129091165	0.443518	0.0229032	6.557	3.874	0.071866	6.171458	A118	gain
RO17P21	RP11-211F5	5	q33.1	149230059	149401295	0.230369	0.0210011	6.96	8.655	0.066893	3.443843	A26	gain
RO50J13	RP11-725K6	5	q33.2	152452352	152621841	0.336881	0.0315645	4.329	9.449	0.071866	4.687627	A118	gain
RO66P24	RP11-798A10	5	q35.2	175185697	175351684	0.485902	0.0227787	11.089	5.603	0.071866	6.761222	A118	gain
RO35G22	RP11-506L4	5	q35.3	179650299	179831215	0.287939	0.0414556	10.536	9.282	0.071851	4.007446	A75	gain
RO42K13	RP11-625D22	6	p23	14164567	14370368	0.313224	0.0407591	8.809	11.924	0.071866	4.358445	A118	gain
RO66C01	RP11-650L9	6	p11.1	58483534	58667977	0.31482	0.0857721	5.217	8.314	0.076899	4.093941	A137	gain
RO9A13	RP11-92I12	6	q16.3	103442960	103604028	-0.242568	0.0182229	4.97	4.239	0.071866	-3.37528	A118	loss
RO20F08	RP11-278D12	6	q22.1	113973438	114122489	-0.453926	0.0664879	12.549	8.891	0.071866	-6.31628	A118	loss
RO63N10	RP11-324H19	6	q22.31	125778192	125951756	0.36266	0.0680322	4.833	4.766	0.076899	4.716056	A137	gain
RO42F09	RP11-629G5	7	p21.1	18699485	18843722	0.272669	0.0120229	8.796	11.119	0.071866	3.794131	A118	gain
RO66F15	RP11-737K4	7	p14.1	38317139	38487572	0.270778	0.0836083	5.876	7.195	0.071866	3.767818	A118	gain
RO59D24	CTD-2057A11	7	p14.1	39094056	39250068	0.343873	0.0083736	7.782	9.9	0.071866	4.784919	A118	gain
RO63D12	RP11-299O9	7	q21.11	84194243	84365508	0.261287	0.0786911	5.634	4.704	0.071866	3.635753	A118	gain
RO67I21	RP11-1279M16	7	q21.13	88320492	88498710	0.46968	0.0461967	4.25	4.237	0.076899	6.107752	A137	gain
RO8D11	RP11-86H8	7	q31.1	110936386	111097114	-0.47501	0.0092985	16.874	9.816	0.071866	-6.60966	A118	loss
RO21I14	RP11-288L5	7	q32.3	130338884	130529807	-0.267959	0.0036932	9.555	5.364	0.076849	-3.48682	A8	loss

RO32M12	RP11-459H7	7	q36.3	155707994	155932827	0.353235	0.0052071	22.606	7.246	0.076899	4.593493	A137	gain
RO67L21	RP13-631H19	8	p12	37399409	37535863	0.239449	0.0686205	16.19	9.359	0.071866	3.331882	A118	gain
RO64N10	RP11-465K16	8	p11.21	40778858	40972561	0.273926	0.0186655	8.474	5.456	0.071851	3.812417	A75	gain
RO44F10	RP11-651O22	8	p11.21	43028640	43224677	0.359587	0.0478506	4.669	3.746	0.071851	5.004621	A75	gain
RO40B22	RP11-610C11	8	q13.1	67209150	67436697	0.313038	0.0193316	5.438	3.974	0.071851	4.356766	A75	gain
RO58L24	CTD-2010C15	8	q21.11	77209526	77359732	0.444097	0.0590123	10.213	7.881	0.100216	4.431398	A51	gain
RO58N16	CTD-2010G21	8	q23.1	107583259	107713603	0.390155	0.009087	6.122	7.438	0.076899	5.073603	A137	gain
RO1A15	RP11-1A23	8	q24.13	124780127	124954329	-0.268898	0.0656775	8.283	6.38	0.076899	-3.49677	A137	loss
RO18P15	RP11-238O2	8	q24.22	132286383	132447254	0.29836	0.0102269	3.467	3.717	0.076899	3.879894	A137	gain
RO70B03	RP11-144K8	8	q24.3	142872355	143036068	0.501523	0.0945861	7.75	10.518	0.076899	6.52184	A137	gain
RO52N16	RP11-754E16	8	q24.3	143213865	143301529	0.41669	0.0497216	4.364	5.156	0.100216	4.157919	A51	gain
RO33G13	RP11-467K20	9	p21.3	21965732	22150198	0.444524	0.0467249	7.135	9.093	0.076899	5.780621	A137	gain
RO68F06	RP11-603E21	9	p21.1	29426014	29599173	0.331981	0.0067543	5.38	4.587	0.076899	3.417104	A137	gain
RO7F14	RP11-76C9	9	p21.1	31283693	31430872	-0.244388	0.0733835	5.46	4.681	0.071866	-3.40061	A118	loss
RO16D22	RP11-192I12	9	p13.3	35317928	35497319	0.666228	0.0970702	4.261	6.645	0.111152	5.993846	A39	gain
RO8N11	RP11-89I3	9	p12	40664093	40859472	-0.277965	0.0075816	16.648	8.281	0.076849	-3.61703	A8	loss
RO5I19	RP11-44K23	9	p11.2	43454734	43631985	-0.398667	0.0153272	8.137	5.699	0.076849	-5.18767	A8	loss
RO24P08	RP11-346L5	9	q12	68280728	68468579	-0.440136	0.0335968	5.995	10.256	0.131551	-3.34574	A14	loss
RO70F06	RP11-175G16	9	q22.1	89988405	90172770	0.24585	0.0199616	3.71	4.867	0.071866	3.42095	A118	gain
RO33C01	RP11-467B11	9	q31.1	101917303	102098845	0.28466	0.0549768	7.779	10.707	0.076899	3.701739	A137	gain
RO68A12	RP11-44G6	9	q31.2	107666667	107813656	0.303278	0.0027068	7.452	4.063	0.076899	3.943848	A137	gain
RO66D15	RP11-730A19	10	p13	13062307	13253628	-0.305214	0.007694	6.934	8.734	0.071866	-4.24699	A118	loss
RO60F20	CTD-2265D13	10	p11.22	31619726	31669689	-0.47732	0.0131501	6.411	8.046	0.131551	-3.6284	A14	loss
RO4I02	RP11-33I16	10	p11.22	31711322	31868133	-0.27203	0.0293789	14.3	7.336	0.076849	-3.5398	A8	loss
RO66H23	RP11-745A18	10	q11.23	53025431	53219133	0.247577	0.0688736	7.317	6.843	0.071866	3.444981	A118	gain
RO1I02	RP11-3D5	10	q21.1	57052410	57216834	-0.258206	0.0313411	5.922	3.949	0.076899	-3.35773	A137	loss
RO63J22	RP11-318J23	10	q21.3	65711959	65891133	0.583614	0.0814764	4.356	5.302	0.076899	7.589357	A137	gain
RO49F21	RP11-714J11	10	q23.33	95288778	95447597	0.24062	0.0474172	9.407	10.491	0.071866	3.348176	A118	gain
RO44G21	RP11-642K1	10	q26.13	125774542	125856402	0.313314	0.0276415	8.847	9.466	0.076899	4.074357	A137	gain
RO65D12	RP11-614E16	11	p14.1	27482159	27659272	0.325842	0.0414669	3.046	3.03	0.071866	4.534022	A118	gain
RO36C21	RP11-522B1	11	q13.4	74035539	74186161	0.266201	0.0328281	7.116	17.275	0.076899	3.461697	A137	gain
RO10B19	RP11-106P8	11	q14.1	79463518	79616021	-0.253873	0.0677924	6.739	6.443	0.071851	-3.53333	A75	loss
RO50I14	RP11-728G23	11	q14.1	84557237	84734682	0.337978	0.0340642	7.285	9.13	0.071866	4.702891	A118	gain
RO4E19	RP11-29E15	11	q14.1	84858148	85007763	-0.258203	0.0171912	11.884	6.614	0.076849	-3.35987	A8	loss
RO4A06	RP11-31E14	11	q22.3	106116314	106270764	-0.261308	0.05727	6.467	4.033	0.076849	-3.40028	A8	loss
RO10P15	RP11-109E10	11	q24.1	122629237	122798552	0.332965	0.0714934	4.373	3.928	0.076849	4.332717	A8	gain
RO50N15	RP11-726G1	12	p13.31	9478300	9654291	0.422425	0.0787759	8.693	12.036	0.071866	5.877953	A118	gain
RO6A15	RP11-55E6	12	p12.3	17693010	17816827	-0.292601	0.0250252	3.969	3.054	0.076899	-3.805	A137	loss
RO59N10	CTD-2103G15	12	q13.13	48916511	49097669	0.349693	0.0791161	7.767	8.691	0.076899	4.547432	A137	gain
RO18G23	RP11-222A15	12	q13.2	54118477	54287319	0.398468	0.0811836	5.489	4.834	0.071866	5.544597	A118	gain
RO69L10	RP11-78L11	12	q14.2	61570376	61745520	0.2413	0.0481929	20.234	5.343	0.071851	3.358339	A75	gain
RO9I16	RP11-96C17	12	q21.1	72035686	72183994	-0.312471	0.0349198	10.521	5.898	0.076849	-4.06604	A8	loss
RO68I21	CTD-2259B14	12	q21.1	72078504	72180449	0.302768	0.0510871	5.711	5.457	0.076899	3.937216	A137	gain
RO50L19	RP11-726C4	12	q21.32	85201787	85406009	0.312927	0.04517	5.227	6.116	0.076849	4.071972	A8	gain
RO13E22	RP11-140J20	12	q23.1	96283836	96447914	0.299916	0.0904121	4.876	4.411	0.076899	3.900129	A137	gain
RO69B04	RP11-63B18	12	q23.1	96437047	96592806	0.471741	0.0005402	40.47	72.191	0.111152	4.244107	A39	gain
RO50I15	RP11-719P9	12	q23.1	99888331	100086721	0.392805	0.0268227	8.689	11.255	0.076899	5.108064	A137	gain
RO7M07	RP11-68I1	13	q13.2	34485139	34666538	-0.29899	0.0186556	7.648	4.721	0.076849	-3.89062	A8	loss
RO13E01	RP11-138D23	13	q21.32	66378445	66542967	-0.299936	0.0473627	4.869	3.511	0.076899	-3.90039	A137	loss
RO13N12	RP11-148O11	13	q21.32	67175227	67343618	0.303854	0.0822882	5.954	3.438	0.076899	3.951339	A137	gain
RO65L17	RP11-597C19	13	q21.33	67840459	68017187	-0.263205	0.0086522	7.072	4.626	0.076849	-3.42496	A8	loss
RO64B06	RP11-440M10	13	q31.1	82700870	82860135	0.688722	0.0846591	4.812	6.025	0.076899	8.956189	A137	gain
RO67G17	RP11-811J13	13	q32.1	96593476	96807474	-0.425103	0.0511132	7.924	6.5	0.071866	-5.91522	A118	loss
RO57F05	RP11-805P2	14	q12	26206655	26423170	0.252798	0.0313015	5.888	8.876	0.071866	3.51763	A118	gain
RO42G08	RP11-626P14	14	q12	26268979	26451739	0.497712	0.0497846	4.253	3.937	0.076899	6.472282	A137	gain
RO20O05	RP11-267E16	14	q13.1	33137499	33295974	0.850474	0.0564773	8.45	16.106	0.076899	11.05962	A137	gain
RO53F05	RP11-761B1	14	q21.1	39255957	39395014	0.246268	0.0361119	7.179	8.673	0.071866	3.426766	A118	gain
RO6D18	RP11-64F23	14	q21.2	41577260	41758651	0.202027	0.0055642	4.731	5.796	0.071866	3.342429	A118	gain
RO1A11	RP11-1A16	14	q21.3	45849446	46005115	-0.26412	0.0560403	11.353	9.676	0.076899	-3.43464	A137	loss
RO42O01	RP11-625K13	14	q23.1	58416882	58604908	0.269958	0.0615515	6.349	7.971	0.071866	3.756408	A118	gain
RO4A02	RP11-31E3	14	q24.3	73034707	73200904	-0.269286	0.0745114	8.889	6.07	0.076899	-3.50181	A137	loss
RO22O10	RP11-312F4	15	q21.3	54444378	54598110	0.658435	0.0623159	5.97	9.103	0.071866	9.161982	A118	gain

RO5O21	RP11-45J10	15	q22.31	61579923	61736764	0.310979	0.0158349	12.833	8.957	0.066893	4.648902	A26	gain
RO29M16	RP11-418F16	15	q25.3	83235499	83414337	0.266466	0.0275984	15.926	7.079	0.076899	3.465143	A137	gain
RO1A24	RP11-2M20	16	p12.3	17074705	17250823	-0.267167	0.0743339	14.906	3.56	0.076899	-3.47426	A137	loss
RO32O01	RP11-456C20	17	q22	50530395	50751652	0.257666	0.0690553	9.363	11.737	0.076899	3.350707	A137	gain
RO69O17	RP11-17H21	17	q23.3	59684680	59845552	0.274758	0.0071057	10.52	9.025	0.076849	3.575297	A8	gain
RO70I09	RP11-98G11	17	q25.3	75393657	75565478	0.258249	0.030605	7.365	7.729	0.076899	3.358288	A137	gain
RO69N15	RP11-59I11	18	p11.32	37518	178873	0.906524	0.0297812	6.392	11.332	0.076899	11.7885	A137	gain
RO68A08	RP11-20I15	19	q13.33	57123599	57310836	0.27021	0.0255294	8.219	7.879	0.076899	3.51383	A137	gain
RO54P16	RP11-777H15	20	q11.21	30235601	30446679	0.476557	0.0044265	3.586	4.502	0.111027	4.292262	A1	gain
RO1J08	RP11-8J1	20	q13.13	48087051	48213089	-0.48188	0.0700382	14.208	13.568	0.131551	-3.66307	A14	loss
RO22F12	RP11-316F6	21	q21.2	25340087	25515908	0.288355	0.0839074	8.124	7.539	0.076899	3.749789	A137	gain
RO61I08	CTD-2385G3	21	q22.3	44518520	44620833	0.418591	0.0408149	3.547	3.976	0.076899	5.443387	A137	gain
RO68O02	RP11-513N1	X	p21.3	27774072	27944947	0.276639	0.0693283	5.563	5.874	0.076849	3.599774	A8	gain
RO68B18	RP11-183D14	X	p21.3	28368505	28540790	0.427304	0.0317378	7.309	6.696	0.076899	5.556691	A137	gain
RO61J17	CTD-2183E12	X	p11.4	42299586	42425357	0.381155	0.0093833	8.247	8.666	0.076899	4.956566	A137	gain
RO67L07	RP13-626I16	X	p11.21	56286352	56436154	-0.27009	0.0026177	4.906	3.875	0.071866	-3.75824	A118	loss
RO58H12	CTD-2008K20	X	q21.32	92605050	92702456	0.327184	0.0500391	5.759	6.821	0.071866	4.552695	A118	gain
RO61P12	RP11-19M10	X	q27.3	145275725	145451069	0.268799	0.0327518	4.392	4.661	0.071866	3.740281	A118	gain
RO15B01	RP11-166O18	Y	q11.221	17496617	17617587	-0.262797	0.0397974	6.544	5.919	0.071866	-3.65676	A118	loss
RO1F03	RP11-5I7	Y	q11.222	20889976	21063742	-0.313962	0.0168836	7	6.363	0.071866	-4.36871	A118	loss
RO58C02	RP11-814F19	Y	q11.223	25156032	25333309	0.53047	0.0022069	6.159	14.612	0.131551	4.032428	A14	gain
RO29I22	RP11-418A1	Y	q11.23	25546517	25733298	0.466481	0.0321437	3.275	4.214	0.131551	3.546009	A14	gain
RO50D01	RP11-724G24	Y	q11.23	25678272	25845689	-0.37497	0.0400576	9.152	8.175	0.100216	-3.74162	A51	loss

Legend	
Sample ID	
A1	Grey
A8	Black
A14	White
A26	Purple
A39	Blue
A51	Green
A75	Yellow
A118	Orange
A137	Red

Tier 1
Thresh score (SD/SNR) <-3.33 or >3.33
SNR of BAC ≥10
Filtered out all CNVs present in >1 individual in 95 normals (same platform; Wan Lam)
Tier 2
Thresh score <-3.33 or >3.33
SNR of BAC ≥3
Filtered out all CNVs present in >1 individual in 95 normals (same platform; Wan Lam)

Appendix 2: Table A2. High Confidence Autism Associated Multi-BAC CNVs

Chr No.	Banding	BP StartPos (UCSC Mar 2006)	BP EndPos (UCSC Mar 2006)	Ratio (Normalized Cy3/Cy5 Log2 Ratio)	Standard Deviation	SNR Ch1 (595)	SNR Ch2 (685)	S _{Dauto} (standard deviation of the ratios of each probe from chr 1-22)	Thresh Score (Ratio/S _{Dauto})	Sample (1 color / sample)	g/l
2	p13.1	74988527	75142707	0.2473	0.030594	50.426	54.109	0.07167	3.4499	A75	gain
2	p13.1	75115704	75274189	0.4487	0.025444	22.177	19.261	0.07167	6.2605	A75	gain
2	p13.1	75226841	75425458	0.2421	0.00185	38.945	46.245	0.07167	3.3773	A75	gain
2	q11.1	95399978	95568811	0.4646	0.000238	86.516	162.83	0.111086	4.182	A39	gain
2	q11.2	95857272	96031199	-0.3412	0.029387	16.539	16.305	0.100303	-3.402	A51	loss
2	q11.2	96212625	96383203	0.3505	0.005981	20.845	36.255	0.100303	3.4945	A51	gain
2	q31.1	176631931	176808639	0.2985	0.025529	79.566	82.436	0.066748	4.4723	A26	gain
2	q31.1	177083977	177242722	0.2927	0.000708	75.093	72.711	0.066748	4.3859	A26	gain
2	q31.1	177427021	177609208	0.2733	0.013666	10.393	12.304	0.073771	3.7053	A137	gain
4	p16.3	2384072	2472983	0.3325	0.006866	23.859	26.812	0.073771	4.5068	A137	gain
4	p16.3	2475663	2659104	0.3376	0.002179	19.321	25.107	0.073771	4.5767	A137	gain
4	p11	48760365	48941412	-0.2838	0.026954	127.71	87.773	0.066748	-4.252	A26	loss
4	p11	48855123	48941378	-0.2395	0.012446	102.81	77.167	0.066748	-3.589	A26	loss
4	p11	49184918	49238843	-0.4341	0.03523	38.012	11.879	0.112428	-3.861	A1	loss
4	p11	49281935	49326349	-0.2263	0.003055	128.64	87.012	0.066748	-3.391	A26	loss
6	p21.1	41456894	41671891	0.533	0.091675	58.274	133.7	0.111086	4.7977	A39	gain
6	p21.1	41510587	41701770	0.4311	0.002591	54.377	81.485	0.111086	3.8803	A39	gain
9	q12	68164115	68282615	-0.3876	0.005926	16.026	14.248	0.112428	-3.448	A1	loss
9	q12	68221319	68447264	-0.3765	0.037569	32.902	37.234	0.112428	-3.349	A1	loss
12	q13.13	52541438	52732799	0.4585	0.009104	79.974	128.2	0.111086	4.1276	A39	gain
12	q13.13	52695946	52866729	0.609	0.008809	114.9	273.57	0.111086	5.4823	A39	gain
12	q13.2	53041821	53231620	0.3898	0.026215	79.703	103.39	0.111086	3.5093	A39	gain
12	q23.3	106831834	106997898	-0.3259	0.0319	37.869	31.34	0.07167	-4.547	A75	loss
12	q23.3	106899825	107062279	-0.3327	0.036087	49.144	20.44	0.07167	-4.642	A75	loss

15	q24.1	71985268	72141641	0.4458	0.001785	27.509	52.143	0.111086	4.0131	A39	gain
15	q24.1	72244696	72412336	-0.3799	0.028438	121.92	179.79	0.111086	-3.419	A39	loss
16	q13	55095980	55251063	0.3883	0.006653	68.105	163.08	0.111086	3.4951	A39	gain
16	q13	55211316	55369561	0.5844	0.015712	44.675	141.42	0.111086	5.2609	A39	gain
19	p13.12	14015002	14153121	0.2576	0.017678	42.295	45.186	0.07167	3.5947	A75	gain
19	p13.12	14069923	14245943	0.254	0.034706	49.545	44.306	0.07167	3.5439	A75	gain

Appendix 3: Table A3. CNVs Found in ACRD and Our Study

Clone Name	Chr No.	BP Start (UCSC Mar 2006)	BP End (UCSC Mar 2006)	Ratio (Normalized Cy3/Cy5 Log2 Ratio)	Standard Deviation	SNR Ch1 (595)	SNR Ch2 (685)	SDauto (standard deviation of the ratios of each probe from chr 1-22)	Thresh Score (Ratio/SD auto)	Sample (1 color / sample)
Tier 1										
RP11-465B22	1	918163	1046699	0.239257	0.039818	22.04	23.88	0.066748	3.584482	A26
RP11-25A21	1	60807944	60980338	0.366584	0.034681	14.54	12.39	0.073771	4.969216	A137
RP11-212D5	1	150227817	150404751	0.373318	0.06219	61.24	134	0.111086	3.360622	A39
CTD-2003B3	1	204676883	204837914	0.257326	0.01329	84.67	88.31	0.066748	3.855187	A26
RP11-196C8	2	13032063	13188099	-0.238977	0.013979	35.19	24.53	0.066748	-3.58029	A26
RP11-723F23	2	14978618	15161195	0.549819	0.061044	12.69	16.76	0.073771	7.453051	A137
RP11-526C24	2	17558392	17595292	-0.257904	0.030754	90.25	76.26	0.07167	-3.59849	A75
RP11-117K20	3	159315572	159494764	0.27218	0.025486	85.08	62.14	0.07167	3.797684	A75
RP11-694E12	3	193655933	193832580	-0.454193	0.002862	109.6	82.12	0.07167	-6.33728	A75
RP11-141L5	4	819328	983918	0.32789	0.001193	20.86	21.56	0.066748	4.912357	A26
RP11-565M6	4	1530742	1691087	0.222639	0.034934	26.53	22.75	0.066748	3.335516	A26
RP11-448F22	4	2384072	2472983	0.332472	0.006866	23.86	26.81	0.073771	4.506812	A137
RP11-444J4	4	2475663	2659104	0.337627	0.002179	19.32	25.11	0.073771	4.57669	A137
RP11-431J16	4	21912514	22118512	0.382154	0.042181	10.92	13.89	0.073771	5.180274	A137
RP11-134M4	4	90847963	91018998	-0.550377	0.009356	44.23	51.05	0.111086	-4.95451	A39
RP11-659P17	4	97390836	97583629	0.484092	0.020737	97.36	195.2	0.111086	4.357813	A39
RP11-637I24	4	164173596	164206678	0.44381	0.038487	34.38	83.38	0.111086	3.995193	A39
RP11-14K14	4	190673393	190811644	0.398252	0.08705	113.6	202.5	0.111086	3.585078	A39
RP11-656C20	5	350722	447421	0.251496	0.017311	64.23	58.45	0.066748	3.767843	A26
RP11-161F13	5	1562894	1724268	-0.255112	0.013584	18.61	14.81	0.06782	-3.7616	A118
RP11-778E14	5	2592645	2774473	0.289887	0.062522	21.96	25.17	0.073771	3.929552	A137
RP11-99E23	5	174208465	174386441	-0.428273	0.083377	44.96	38.91	0.111086	-3.85533	A39
RP11-26H18	6	7545199	7735839	0.353082	0.023407	28.15	29.93	0.100303	3.520154	A51
RP11-346C16	6	110941778	111130824	0.28682	0.001887	11.04	13.81	0.073771	3.887978	A137
CTD-2310D7	6	113915311	114052447	-0.350276	0.087168	14.87	11.19	0.06782	-5.16479	A118
RP11-403C11	6	114313227	114352792	-0.257576	0.030125	83.23	69.69	0.066748	-3.85893	A26
CTD-2260L17	7	797562	898228	0.226066	0.022829	42.21	43.15	0.066748	3.386858	A26
CTD-2380A4	7	15155609	15320915	0.242054	0.010675	10.6	10.46	0.066748	3.626386	A26
RP11-455J15	7	16700151	16863920	0.318534	0.012703	15.96	16.92	0.073771	4.317876	A137
RP11-818I12	7	124755570	124943364	0.28215	0.041623	17.91	17.16	0.066748	4.227093	A26
RP11-674O21	8	3049068	3251231	0.429	0.015461	48.48	116.5	0.111086	3.861873	A39
RP11-652N4	8	4901876	5127620	0.240296	0.061576	74.88	55.14	0.07167	3.352811	A75
RP11-102I24	10	13089369	13228260	0.251713	0.015452	91.26	94.58	0.066748	3.771094	A26
RP11-818P1	10	42987824	43202215	0.281415	0.002531	10.33	11.61	0.074813	3.761579	A8
RP11-30N1	10	51121592	51195063	-0.289387	0.005939	17.06	13.7	0.06782	-4.26699	A118

RP11-296M15	11	133545363	133721889	0.271078	0.019537	71.41	71.26	0.066748	4.061215	A26
RP11-440N2	12	36720939	36889175	0.405225	0.054953	16.72	21.19	0.073771	5.493012	A137
RP11-26G8	12	57636373	57812533	0.362373	0.044189	18.94	20.3	0.073771	4.912133	A137
RP13-552N19	12	130837437	130869768	0.403183	0.086379	10.94	28.74	0.111086	3.629467	A39
RP11-563G5	13	18014607	18190654	0.49662	0.020139	111.7	110	0.111086	4.47059	A39
RP11-51H10	13	113492713	113658045	0.489403	0.042209	42.72	92.72	0.111086	4.405623	A39
RP11-520H13	14	64706822	64880159	0.264777	0.022368	18.64	17.34	0.07167	3.694391	A75
RP11-218L8	14	86490679	86613015	-0.387777	0.024405	26.08	29.42	0.112428	-3.44911	A1
RP11-607H20	15	18273500	18473311	-0.251218	0.005564	14.47	10.05	0.074813	-3.35795	A8
RP11-641K15	15	18781625	18946020	-0.248226	0.002765	24.74	23.85	0.06782	-3.66007	A118
RP11-403L7	15	18898414	19048601	-0.385608	0.009341	255.3	386.6	0.111086	-3.47126	A39
RP11-768D20	15	21886511	22055073	-0.38372	0.022012	18.81	20.56	0.112428	-3.41303	A1
CTD-2315M12	15	24567294	24683390	0.434072	0.004535	37.98	33.56	0.111086	3.907531	A39
RP11-561P13	15	26135317	26310601	-0.245656	0.048374	18.28	15.66	0.06782	-3.62218	A118
RP11-6L23	15	71985268	72141641	0.445803	0.001785	27.51	52.14	0.111086	4.013134	A39
RP11-775D1	15	72244696	72412336	-0.379857	0.028438	121.9	179.8	0.111086	-3.41949	A39
RP11-605N4	16	55095980	55251063	0.388252	0.006653	68.11	163.1	0.111086	3.495058	A39
RP11-343H19	16	55211316	55369561	0.584413	0.015712	44.68	141.4	0.111086	5.260906	A39
RP11-798C7	17	76834696	77076142	0.248783	0.04408	13.98	22.26	0.07167	3.471229	A75
RP11-453P5	18	56301614	56482301	0.244348	0.056018	50.52	55.02	0.07167	3.409348	A75
RP11-711F2	18	71337342	71520536	0.238207	0.045352	10.43	11.52	0.06782	3.512341	A118
RP11-703H17	18	74556232	74771957	0.383839	0.051755	34.16	81.72	0.111086	3.455332	A39
RP11-7H17	18	75216360	75402240	0.256705	0.035492	54.96	54.6	0.066748	3.845883	A26
RP11-543F7	19	33742778	33936080	0.331168	0.027284	81.34	58.75	0.07167	4.620734	A75
RP11-705C4	19	60715146	60850444	0.241451	0.025361	26.32	26.44	0.07167	3.368927	A75
RP11-476I15	20	62263781	62434320	0.26224	0.001918	34.08	32.49	0.066748	3.928807	A26
RP11-652F11	22	18002150	18171498	0.291579	0.020793	72.26	53.67	0.07167	4.068355	A75
RP11-799D2	22	46893547	47064651	0.33546	0.000556	29.09	31.35	0.100303	3.344466	A51
RP11-423E19	22	49413876	49518700	0.241691	0.029247	56.17	33.55	0.066748	3.620947	A26
RP11-297F13	X	48575765	48734088	0.349604	0.008873	52.79	65.63	0.07167	4.877968	A75
RP11-211H10	X	49201123	49376616	-0.228992	0.093953	17.52	14.12	0.06782	-3.37647	A118
Tier 2										
RP11-201O16	1	241113402	241143060	-0.377712	0.003692	29.14	13.37	0.111027	-3.40198	A1
RP11-633H24	2	154895951	155104293	0.252524	0.034344	4.399	6.016	0.071866	3.513817	A118
RP11-719O1	2	165653881	165842181	0.27515	0.059472	9.813	10.68	0.076899	3.57807	A137
RP11-806C2	3	1841954	2043596	0.240232	0.007984	8.316	12.33	0.071866	3.342777	A118
RP11-655G22	3	193133121	193305756	0.264123	0.044765	16.05	8.214	0.076899	3.434674	A137
RP11-721I4	4	2716373	2926405	0.412427	0.095637	3.812	3.715	0.076899	5.36323	A137
RP11-18O22	4	117414505	117570802	0.240949	0.008945	70.64	68.06	0.066893	3.602006	A26
CTD-2011O21	4	191039899	191138136	-0.371553	0.011956	50.38	28	0.111027	-3.34651	A1
RP11-263G17	5	39200686	39246329	-0.40495	0.026786	3.098	3.662	0.111027	-3.64731	A1
CTD-2075G19	5	103890873	104011391	0.319472	0.016983	6.677	7.662	0.071866	4.445384	A118
RP11-144H10	5	120001738	120179752	0.405011	0.041323	6.166	6.076	0.071866	5.635641	A118
RP11-281H9	5	128933631	129091165	0.443518	0.022903	6.557	3.874	0.071866	6.171458	A118
RP11-625D22	6	14164567	14370368	0.313224	0.040759	8.809	11.92	0.071866	4.358445	A118
RP11-278D12	6	113973438	114122489	-0.453926	0.066488	12.55	8.891	0.071866	-6.31628	A118

RP11-324H19	6	125778192	125951756	0.36266	0.068032	4.833	4.766	0.076899	4.716056	A137
RP11-299O9	7	84194243	84365508	0.261287	0.078691	5.634	4.704	0.071866	3.635753	A118
RP11-1279M16	7	88320492	88498710	0.46968	0.046197	4.25	4.237	0.076899	6.107752	A137
RP11-86H8	7	110936386	111097114	-0.47501	0.009298	16.87	9.816	0.071866	-6.60966	A118
RP11-651O22	8	43028640	43224677	0.359587	0.047851	4.669	3.746	0.071851	5.004621	A75
RP11-467K20	9	21965732	22150198	0.444524	0.046725	7.135	9.093	0.076899	5.780621	A137
RP11-730N17	9	23690948	23798886	-0.452101	0.026159	6.635	5.62	0.131551	-3.4367	A14
RP11-603E21	9	29426014	29599173	0.331981	0.006754	5.38	4.587	0.076899	4.317104	A137
RP11-76C9	9	31283693	31430872	-0.244388	0.073384	5.46	4.681	0.071866	-3.40061	A118
RP11-730A19	10	13062307	13253628	-0.305214	0.007694	6.934	8.734	0.071866	-4.24699	A118
RP11-745A18	10	53025431	53219133	0.247577	0.068874	7.317	6.843	0.071866	3.444981	A118
RP11-3D5	10	57052410	57216834	-0.258206	0.031341	5.922	3.949	0.076899	-3.35773	A137
RP11-318J23	10	65711959	65891133	0.583614	0.081476	4.356	5.302	0.076899	7.589357	A137
RP11-138D23	13	66378445	66542967	-0.299936	0.047363	4.869	3.511	0.076899	-3.90039	A137
RP11-745K2	14	75867497	76087405	0.435484	0.014216	8.298	14.74	0.100216	4.345454	A51
RP11-418F16	15	83235499	83414337	0.266466	0.027598	15.93	7.079	0.076899	3.465143	A137