Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____   _____
Morna Aiko Ikeda   Date

1

High Resolution X Chromosome Copy Number Variation in Autism

By

Morna A. Ikeda
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Genetics and Molecular Biology

_____
Stephen T. Warren, Ph.D.
Advisor

_____
David Cutler, Ph.D.
Committee Member

_____
Christa Lese Martin, Ph.D.
Committee Member

_____
Stephanie Sherman, Ph.D.
Committee Member

_____
Michael Zwick, Ph.D.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____ Date

High Resolution X Chromosome Copy Number Variation in Autism

By

Morna A. Ikeda
B.A., Northwestern University, 2000

Advisor: Stephen T. Warren, Ph.D.

An abstract of
a dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in
the Graduate Division of Biological and Biomedical Sciences
Genetics and Molecular Biology

2012

Abstract

High Resolution X Chromosome Copy Number Variation in Autism

By Morna A. Ikeda

The autism spectrum disorders (ASD) are a broadly defined set of developmental disorders that include autism and Asperger syndrome. Individuals with ASD are defined as having impairments in social interaction, deficiencies in communication, as well as restricted and stereotyped behaviors and interests. Leo Kanner first described autism in 1943, and subsequent twin and family studies have demonstrated a substantial genetic component underlying ASD. A marked increase in the prevalence of ASD has been noted in the last decade, and the most recent estimate suggests a prevalence of 1:88. A four- to ten-fold male preponderance of ASD suggests the existence of sex-specific risk alleles and the possibility of a recessive susceptibility locus on the X-chromosome. In the last eight years, copy number variation (CNV) has been appreciated as a rich source of both inherited and *de novo* human genomic variation.

Technological advances in array Comparative Genomic Hybridization (aCGH), the common microarray based assay used to assess copy number state, have enabled detection of increasingly smaller variants. We developed a custom high-density array consisting of 2.1 million oligonucleotide probes dedicated to the X-chromosome. Non-repetitive sequence is probed at a resolution of one probe per 50 bp or 96.8 megabases of unique sequence. Additionally, we further enhanced the stringency and thus rigor with which samples are interrogated by developing a hybridization and wash protocol for the Tecan HSPro 4800. The application of this machine standardized hybridization, wash, and dry conditions across all samples.

Using our custom X-chromosome CGH microarrays, we screened three cohorts for X-chromosome CNV. The first cohort was a series of 100 ASD males from the Autism Genetic Resource Exchange collection. Our second cohort of 64 ASD males was derived from the Simons Simplex Collection. Finally, our third cohort consisted of 100 males from a National Institute of Mental Health control population where individuals were selected as controls for the study of neuropsychiatric disease. In total, 164 ASD males and 100 non-ASD males were evaluated for X-chromosome CNV.

High Resolution X Chromosome Copy Number Variation in Autism

By

Morna A. Ikeda
B.A., Northwestern University, 2000

Advisor: Stephen T. Warren, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in
Graduate Division of Biological and Biomedical Sciences
Genetics and Molecular Biology

2012

Acknowledgements

I would like to first thank Stephen Warren, Dave Cutler, Christa Lese Martin, Stephanie Sherman and Mike Zwick for their constant support, guidance and encouragement through all the different stages of my thesis work and scientific growth. Truly, I left every encounter with a sense of forward progression in my experiments and/or critical thinking of the work at hand. With their gentle nudges, my mind and confidence in what I knew grew more than I ever imagined. Furthermore, it was through their efforts that I have learned to move beyond the answer at hand to the question it now posed.

I would also like to thank all those in the Department of Human Genetics. I received a tremendous amount of encouragement and supportive inquiry of my science from faculty, staff and students. But, their efforts to support and nourish me personally made DoHG a genuine resource and community for me. And, I would be completely remiss if I didn't acknowledge the incredible support and guidance from my lab mates: Reid Alisch, Steven Bray, Pankaj Chopra, Brad Coffee, Stephen Collins, Anne Dodd, Krayton Keith, Bryan Lynch, Tamika Malone, Julie Mowrey, Jen Mulle, Leila Myrick, Mika Nakamoto, Mike Sanotoro, Josh Suhl, Tao Wang, and Fuping Zhang. Of course, a special note of thanks to my peers, Adam Locke and Mike Santoro, for their efforts to better my scientific thoughts and work throughout my graduate career.

Finally, I would like to thank my family and friends from both within and outside of the Emory community. While some claim to never have understood how I spent my time and efforts, all were a constant source of encouragement and refuge when I sought balance in my life.

TABLE OF CONTENTS

I. INTRODUCTION

**II. OPTIMIZATION OF THE NIMBLEGEN ARRAY COMPARATIVE
GENOMIC HYBRIDAZATION PROTOCOL**

APPENDIX

TABLES

FIGURES

Chapter 1. Introduction

1. Introduction

a. Background and prevalence of the autism spectrum disorders

Autism is a developmental disorder characterized by deficiencies in verbal communication, impaired social interaction, and restricted or patterned behaviors and interests.[1] Leo Kanner first described autism or an "inborn autistic disturbance of affective contact" in 11 children in 1943.[2] While Kanner notes that there are similarities to childhood schizophrenia, he argues that these children affected with autistic disturbance are a distinct group and necessitate further study. With each update of the *Diagnostic and Statistical Manual of Mental Disorders (DSM)* put out by the American Psychiatric Association, the criteria for autistic diagnosis has continually expanded and the definition further refined.[1, 3, 4] The fourth edition of the *DSM* was first released in 1994 and later revised in 2000. In this edition, autism has been considered one end of a larger series of disorders collectively identified as autism spectrum disorder (ASD).[1, 5] Individuals with ASD have a range of verbal, social, and behavioral deficits that can individually define a syndrome but collectively are considered variable presentations of the same disorder. The umbrella designation of ASD includes pervasive developmental disorder, not otherwise specified (PDD-NOS), Asperger Syndrome, as well as autism. Autism is the most predominant and severe of the ASD.

With an expanded and more specific clinical definition for autism and related presentations introduced with the *DSM-III* in 1980, a marked increase in the diagnosis of autism and its related disorders was subsequently observed. [4, 6-8] Gillberg *et al* noted

that the rate of autism nearly doubled from 0.5 in 1000 to 1 in 1000 over a thirty-year

period. However, the increase seemed to begin prior to the release of the third edition.

Studies conducted from 1970-1997 reported the higher rate (1:1000) as compared to

those conducted between 1966 and 1970 (0.5:1000). [7] *DSM-IV* was introduced in 1994

and further clarified and organized autism and ASD under the 'Pervasive Developmental

Disorders' along with Childhood Disintegrative Disorder and Rett's Disorder. Despite

specific criteria by which autism and ASD are defined, they remain phenotypically

heterogeneous in their presentation. While any or all aspects of autism may manifest

before the age of three and a diagnosis possible at that time, this disorder is typically

diagnosed between four-five years of age. [9] Additionally, despite the *DSM-IV*

definitions in place for nearly 20 years, the prevalence of ASD has shown a clear increase

in recent years with a majority of studies estimating 1:110 children with ASD and the

2012 release of 2008 data the Autism and Developmental Disabilities Monitoring

Network estimating 1:88. [4, 9-12]

A consistent and notable feature of ASD is the preponderance of males affected by these

disorders. Despite the increasing prevalence estimates and various studies and

populations ascertained, the male to female sex ratio in autism has been repeatedly

estimated to be about 4:1. The expansion of diagnosis to include all ASD increases this

bias to 10:1. However, when stratified by intellectual disability, for IQ < 70, the male to

female ratio in autism approaches 1:1. [9, 13]

    b. A strong genetic component underlies ASD

Epidemiologic twin and family studies have demonstrated a strong genetic component to these disorders. Twin studies assessing disease concordance among monozygotic twins (MZ) have estimated rates of 47-96%, while dizygotic (DZ) twins have been found to be concordant at rates of 10-36%.[14-21] Consistent with the DZ concordance rate for ASD, the sibling risk ranges from 3-11% representing an almost ten-fold risk over that of the general population.[19, 22-25] These data suggest a highly heritable component to autism and ASD. In spite of the large number of genetic studies, to date nearly 70-85% of individuals affected with these disorders remain idiopathic.[26]

c. Genes and loci causal and implicated in ASD

Within the ASD, a genetic basis for disease can be identified in 15-27% percent of individuals affected.[27] Fragile X Syndrome has long been the most common diagnosis, accounting for nearly 0-8% of all individuals with ASD. A duplication or triplication event of maternal 15q11-13 has been shown to account for 1-3%, and co-morbidity with tuberous sclerosis has been reported to account for a similar rate of 1-4%. [26, 28] A growing body of literature has identified multiple candidate loci associated with ASD through association and linkage studies.[29-34] A recent review of the literature by Catalina Betancur has identified 103 genes and 44 genomic imbalances located throughout the genome as playing a role in ASD.[35] The distributive nature of these findings further substantiates the significant genetic heterogeneity underlying these disorders. Additionally the representation of multiple loci with structural changes in ASD suggests a significant mutational mechanism in ASD. While monogenic forms of ASD

clearly exist, genomic loci disrupted in ASD often bear multiple genes. This suggests that dosage sensitivity can also be a factor in susceptibility to these disorders.

The current understanding of the molecular basis underlying ASD comes from the accumulated knowledge gathered from linkage and genome wide association studies as well as cellular pathways disrupted by genomic imbalances. However, despite the genes and loci implicated in ASD, the vast majority of individuals affected with autism or ASD remain idiopathic.

2. Known genomic structural changes in ASD

Copy number variants (CNV) are lengths of sequence greater than one kilobase (kb) in size that can vary in copy number from the reference of two copies autosomally, one copy of chromosome X and one copy chromosome Y (males), or two copies chromosome X (females). CNV can be either pathogenic or benign. In the mid 2000s, submicroscopic CNV (less than one Mb in size) in autism was a relatively understudied mutational mechanism. Cytogenetic observations (resolution of five megabase (MB) though two to three Mb is possible) and BAC-based arrays (genome-wide: one Mb resolution, targeted tiling: ~150 kb) accounted for the known structural changes in ASD.[36] Increasing advances in oligonucleotide array technology beginning in the early 2000s enabled researchers to screen target regions at a much greater resolution than that afforded by cytogenetic or BAC-based analysis.

a. Knowledge of CNV as discerned from oligonucleotide-array based studies

Advances in oligonucleotide array technology in the last decade have enabled the remarkable discovery of large-scale copy number changes within normal individual genomes as well as across populations.[37] In 2004, two landmark papers were published characterizing the widespread occurrence of large-scale CNV within individuals as well as across different ethnic populations.[38, 39] These initial descriptions characterized changes 200 kilobases (kb) or larger in 75 individuals and estimated at least 11-12 CNV/person exist in the normal population. The Sebat and Iafrate studies motivated multiple research efforts to more fully characterize normal CNV in many different populations.

Since 2004, the search for copy number changes has undergone further advances in the ability to discriminate smaller sequence changes (increased resolution) as well as an expanded diversity of platforms from which to interrogate individual genomes. More probes per array, single nucleotide polymorphism (SNP) based arrays, paired-end sequencing, and most recently, next-generation sequencing (NGS) technologies have contributed to a greater understanding of genomic structure at greater and greater resolution.

Several characteristics seem to hold true with each new study and advance in technology. Copy number variation exists in normal populations, and copy number variant loci are shared across different ethnic populations.[40-51] Consistent with SNP studies of African populations, a greater diversity of CNV exists in African derived populations than in any other.[38, 40, 44, 46, 47, 50-54] Additionally, the majority of copy number changes

identified within the HapMap populations have suggested that the majority of CNV are inherited. [42, 43, 47, 51-53, 55]

While there is growing clarity about the pervasive nature of CNV, several aspects remain to be resolved about these copy number changes. It is unclear at what size distributions CNV exist as well as whether CNV are randomly dispersed or clustered throughout the genome. It has already been demonstrated that a significant set of CNV located genome-wide are mediated by non-allelic homologous recombination between flanking segmental duplications.[44, 46, 47, 52, 56-60] However, the full interrogation of CNV not associated with segmental duplications is incomplete. While many studies of genome wide CNV have been conducted throughout the mid-late 2000s, array platform, number and distribution of probes, study population, and CNV identification algorithms have varied greatly from study to study.[54, 61] For example, while one study might identify a locus as altered by a single, large variant another study might later identified the locus as harboring multiple, smaller variants. This example illustrates conflicting results that can arrive when using multiple array platforms and criteria by which CNV are identified. The array resolution (i.e., the number of probes interrogating a fixed length) and probe distribution (e.g., SNP-based arrays initially lacked even coverage throughout the genome and other arrays used a targeted rather than tiling based approach) are two variables that can easily contribute to differing interpretations of the same biological phenomena.[53, 54, 61, 62]

Additionally, characterization of CNV on the sex chromosomes has proceeded more slowly than autosomal loci. Searching for structural changes on the sex chromosomes is difficult due to the comparative nature of the technologies used to identify such changes. Using an array Comparative Genomic Hybridization (aCGH) or SNP-based platform, fluorescent intensity of a given marker in the tested sample is compared to that of a normal reference at the same marker. Where intensity values are the same, the locus is interpreted to have the same copy number in both samples. A relative increase in fluorescence of the tested sample compared to the reference suggests an increase in copy number in the test relative to reference, while a decrease in test fluorescence relative to the reference suggests a decrease in copy number in the test relative to the reference. In a sex-mismatched comparison, much of the entirety of either sex chromosome in the test will appear shifted relative to the reference at a statistically significant level. [63] Mining for additional copy number changes (i.e., identifying segments of the sex chromosome that are deleted or duplicated relative to flanking sequence) in such data is quite difficult, and often, studies will not include the sex chromosomes for this reason. Consistent with this notion, the Database of Genomic Variants has proportionally fewer CNV and indels reported in the sex chromosomes than in the autosomes.

The Database of Genomic Variants (DGV) is a collection of sequences reported as copy number variant. [64] (projects.tcag.ca/variation) CNV and indels (changes in copy number sequence from 100 base pairs to 1 kb) in this database are considered to be polymorphic rather than mutant, although this database is not well curated and therefore prone to some error. In a brief assessment of reported genomic variation, we observed

two features specific to the sex chromosomes. First, of the 42 studies reporting CNV and indels in DGV, the proportion of references for all autosomes is quite similar at a mean percentage of 90%, but this rate declines for the sex chromosomes (mean: 55%). (Figure 1.1.a) Second, when we asses the proportion of all calls reported for each chromosome, we observe a similar proportion as the percent of the genome (length of the chromosome) that the chromosome represents for the autosomes. Specifically, the mean difference between the percentage of CNV and indels reported for the autosomes and the percent size of the autosome is 0.15%. However, these similar proportions deviate when accounting for the sex chromosomes. Both chromosome X and Y show a greater mean difference (-1.68%) between CNV reported in the DGV and the percentage of the genome these chromosomes represent. (Figure 1.1.b) Several reasons may explain why there are fewer CNV reported for chromosome X and Y. First, it may be that CNV formation on these chromosomes occurs less frequently than their autosomal counterparts. Second, while the sex chromosomes may be equally susceptible to such rearrangements, these changes may not be well tolerated by the organism and thus more heavily selected against.[52, 53, 55] Finally, a non-biological explanation may simply extend from the fact that fewer studies have evaluated these chromosomes for copy number changes as described above.

b. CNV in the ASD

Large rearrangements identifiable by cytogenetic and FISH based assays have long played a role in the identification of genomic disorders.[26, 65-73] With increasing resolution made possible by technical advances, normal variation is being cataloged and

curated in individuals and diverse populations worldwide, and similarly, multiple studies

are also being conducted in effort to ascertain a more comprehensive description for the

role of genomic structural change in disease as well.[74] Included in these efforts have

been multiple exploratory studies of genome wide copy number variation in populations

affected with the autism spectrum disorders. Out of these efforts, over 40 loci have been

implicated and identified as causal in ASD.[35] Structural changes have been identified

on most autosomes as well as the X chromosome. Imbalances can range anywhere in

magnitude from cytogenetically visible changes (e.g., five Mb disruptions) to less than

300 kb, and no ethnic/race-specific loci have been identified.

In 2007, Sebat and colleagues were the first to report increased *de novo* rates of CNV in

ASD relative to those rates observed in normal populations.[75] Interestingly, a higher

proportion of females to males affected with ASD in the *de novo* group (male:female

ratio: 1.8:1 *de novo*, 5:1 overall sample) were observed.[75] This reduction in the skewed

sex ratio suggested to the authors that these *de novo* CNV were more penetrant than

inherited CNV and likely contributed to disease in a sex independent fashion. Since this

initial report, additional groups studying other ASD populations have found similar

increases in *de novo* rates of CNV in their cases relative to controls.[66, 68, 70, 72, 76,

77]

3. A role for the X chromosome in ASD

As noted previously, a conspicuous feature of autism is the significantly larger proportion

of males diagnosed with autism or ASD. Males continually outnumber females four to

one within the autistic diagnosis, while expansion to include all ASD extends this bias to as high as 10 to 1.[5, 10, 11, 78] This extreme skewing of the sex ratio is consistent with a pattern of X-linked inheritance. Under such a model, a genetic lesion on the X chromosome would be unveiled in the male, hemizygous state. [5]

a. A hemizygous X chromosome as a susceptibility to ASD

Typically, studies of the hemizygous X chromosome occur in affected males. However, women affected with Turner Syndrome (45, X) are similarly susceptible to changes on the X chromosome. One study of women with TS observed distinct behavioral and cognitive differences between women who inherited the paternal versus maternal X chromosome. Specifically, the authors suggested the existence of an imprinted locus on the X chromosome where females with TS who maintained the paternal X chromosome appeared to be more socially adjusted than those who had inherited the maternal, imprinted X chromosome.[79] The authors suggested this locus may help to moderate social adaptability. Additionally, in an unselected sample of women with TS, five of 150 were identified as having autism as well as TS. All individuals with autism also had intact maternal X chromosomes.[80] These data suggest that a locus on the X chromosome may serve to play a role in autism susceptibility.

b. Sex-specific risk loci exist in ASD

Another compelling role for the X chromosome arises from the results of a 2004 linkage study conducted by Stone *et al* suggesting sex-specific risk alleles exist in autism. In this study, multiplex families were stratified by the presence of an affected female child,

thereby splitting the families into 'male only' or 'female containing' groups. Significant linkage peaks were obtained within both sets, but the identified loci did not overlap suggesting different etiologies were responsible for disease in the two different groups. [81] While 17q11 was implicated in the male stratified sample set, their findings do not exclude the X chromosome from playing a role in ASD as well.

Additionally, in 2006, Gauthier *et al* identified two markers (DXS6789 and DXS8043) located at Xq21.33 and Xq27.3 to be significantly associated with males affected with ASD in a French Canadian cohort (FC). This finding is remarkable for the fact that a strong founder effect exists within the FC population. [82] While additional studies of autism susceptibility loci in this population reported in 2010 and 2011 did not identify any further signal at or near DXS6789 and DXS8043, these follow-up studies also included an expanded, non-FC population. This diversified sample population may have diluted the original signal identified by Gauthier and colleagues. [83, 84] In another distinct study, Vincent and colleagues identified a modest signal at Xq27 in multiplex families with autism and related phenotypes.[85] In combination, these data suggest that a subset of individuals affected with ASD may bear a lesion on the X chromosome that underlies their autistic phenotype.

c. Skewed X-inactivation is increased in autistic females

Talebizadeh *et al* found significantly increased skewed X inactivation patterns (> 80:20%) in females with autism (33%) as compared to females unaffected by this disorder. Furthermore, the mothers of daughters affected with autism show an increased

rate of skewed X-inactivation.[86] This is striking in light of analysis conducted by

Amos-Landgraf and colleagues. The X inactivation patterns for over 1,000

phenotypically unaffected females were determined. These authors found that only 8.8%

of this normal female population showed skewing of greater than 80:20%. [87] Showing

a nearly 4 fold increase in skewed X-inactivation, the findings of Talebizadeh and

colleagues evidences the potential for a disrupted locus on the active X chromosome.

Additionally, Thomas *et al* identified a possible critical region for autism at Xp22.3.

Three of eight females bearing a deletion at this locus were also diagnosed with autism.

Skewed inactivation of the intact chromosome led the authors to suggest the loss of

genetic material at this loci (including the genes *SS, CDPX, STS, KAL,* and *MLS*) explains

the autism in the three affected females.[88] Both studies suggest that those X

chromosomes that remain active also harbor a genetic lesion that either significantly

contributes to or increases susceptibility to ASD.

4. Conclusion

    a. Summary of autism, CNV in disease, and the X chromosome in ASD

    ASD are a phenotypically heterogeneous set of developmental disorders and are

very common with a prevalence of 1:88. While family and twin studies have firmly

established a genetic component underlying susceptibility, monogenic findings as well as

genomic imbalances have implicated multiple regions throughout the genome as

functioning in ASD etiology. Cytogenetic analysis has long played a role in identifying

novel loci involved with disease. Recent technological advances have enabled the

detection of genomic imbalance on a finer scale. As CNV discovery in normal

populations is progressing, so is the identification of CNV in disease.

b. Identification of novel ASD loci: Hypothesis and Experimental design

Given the strong genetic character of the ASD and as well as the striking male bias, we

hypothesized that high susceptibility gene(s) to ASD may reside on the X chromosome.

When disrupted, an autistic locus on the X chromosome would put a 46, XY male at

tremendous risk for disease. To this end, we sought to capitalize on the recent advances

in oligonucleotide array technology to explore the nature of fine structural variation on

the X chromosome of individuals with autism. In addition, our findings would expand the

current description of CNV on chromosome X.

c. Summary of our experiments and samples studied

Our initial experimental strategy was to screen 100 unrelated males with autism from the

Autism Genetic Resource Exchange (AGRE, research.agre.org) multiplex cohort by high-

resolution array Comparative Genomic Hybridization (aCGH). Those CNV identified and

validated from this sample set would then be genotyped in two additional sample sets.

First, an additional 200 unrelated males with autism as well as all the presumably

unaffected and unrelated fathers (n=300) would be genotyped to assess population

frequency. Secondly, because families selected from the AGRE cohort were multiplex,

mothers, concordant siblings, and discordant siblings would be genotyped to assess

whether a CNV was inherited and/or tracked with autistic disorder within the family.

As the study progressed, two more sample populations became available. The first

population to become available was the Simons Simplex Collection (SSC, sfari.org). This

is a collection of families with one child affected with autism or ASD (simplex). We

screened 64 unrelated males affected with autism from this cohort for CNV on

chromosome X. Given that the SSC is a collection of simplex families, we anticipated

that probands from this cohort may harbor more *de novo* structural changes involved with

ASD than probands from the multiplex families from the AGRE cohort.

The second new population we screened by chromosome X high-resolution aCGH was

the National Institute of Mental Health (NIMH) Human Genetics Initiative (HGI) control

population. This population is largely comprised of adult males and females who self-

report no known neuropsychiatric disorders (proportion of adults: ~3900/4300). We

chose to use this cohort in two ways. First, we selected 100 males over the age of 18 to be

used to determine if our autistic samples had a larger burden of CNV than this unaffected

population. Secondly, a total number of 1,500 adult males over the age of 18 would be

CNV genotyped to estimate general population frequency of any CNV identified.

In total, 164 unrelated males with ASD were selected for fine-scale structural analysis of

the X chromosome. Validated CNV were to be re-genotyped in the 164 as well as an

additional 200 unrelated males affected with ASD to estimate CNV frequency in an

autistic population (total n = 364). One hundred males unaffected with ASD were

screened by aCGH (NIMH), and a total of 1,800 unaffected males (300 unrelated fathers

from AGRE and 1,500 males from NIMH) were genotyped to estimate CNV frequency in an unaffected population. (Table 1.1)

High-resolution aCGH analysis of the X chromosome first began with our use of the chromosome X, 385K CGH array from NimbleGen (now Roche NimbleGen). Tiled across the X chromosome, 385,000 oligonucleotide probes targeting repeat masked X chromosome sequence were synthesized in a single array. We processed and ran four samples from the AGRE cohort before the opportunity to use the 2.1M CGH arrays arose. The new 'High Density' CGH arrays were comprised of 2.1 million probes tiled along the X chromosome. By shifting our array platform to the 2.1M CGH arrays, our probe coverage increased from an average of four probes/kb to 20 probes/kb. Fifty samples from the AGRE cohort were run on the 2.1M aCGH arrays.

Concurrent with the sample processing and array hybridization were our early efforts of validating array identified CNV. By the time we had processed the first 50 AGRE samples on the HD aCGH arrays, a high false positive rate prompted us to explore aCGH protocol changes that might increase our capture of true CNV and reduce the false CNV call rate. Once an optimized protocol was established, we processed 100 AGRE samples as well as the 64 SSC samples using the new protocol on the HD CGH arrays.

Array behavior appeared to shift with purchases of new arrays in the latter end of our 2.1M aCGH experiments with the samples from the autistic populations. Discussions with Roche NimbleGen confirmed our suspicions of altered array synthesis chemistry. It

was decided that the 100 NIMH control samples for 2.1M aCGH screening would be processed and hybridized following the manufacturer's protocol and not the optimized one we had developed.

In summary, we conducted four independent studies of two populations using two CGH arrays and two different protocols. Three different approaches were taken to interrogate individuals with autism. We first began with a brief analysis of four individuals from the AGRE cohort and screened the X chromosome for CNV using the 385k CGH array and the NimbleGen protocol. Next, we increased our array resolution and screened 50 individuals from AGRE using the 2.1M CGH array and the NimbleGen protocol. Then, our optimized protocol was used to screen 100 individuals from AGRE and 64 individuals from SSC on the 2.1M CGH array. Finally, our studies of the NIMH normal controls were conducted on the 2.1M CGH array using NimbleGen's protocol for the 100 males selected for study.

5. References

1.   Association, A.P., *Diagnostic and Statistical Manual of Mental Disorders*. 4th - Text Revision ed. Vol. Four. 2000, Washington, DC: R.R Donnelly & Sons Company.
2.   Kanner, L., *Autistic Disturbances of Affective Contact.* Nervous Child, 1943. **2**: p. 217-50.
3.   Association, A.P., *Diagnostic and Statistical Manual of Mental Disorders*. 2nd ed. Vol. Two. 1968, Washington, DC.
4.   Saracino, J., et al., *Diagnostic and Assessment Issues in Autism Surveillance and Prevalence.* Journal of developmental and physical disabilities, 2010. **22**: p. 317-330.
5.   Folstein, S.E. and B. Rosen-Sheidley, *Genetics of autism: complex aetiology for a heterogeneous disorder.* Nat Rev Genet, 2001. **2**(12): p. 943-55.
6.   Mayes, R. and A.V. Horwitz, *DSM-III and the revolution in the classification of mental illness.* J Hist Behav Sci, 2005. **41**(3): p. 249-67.
7.   Gillberg, C. and L. Wing, *Autism: not an extremely rare disorder.* Acta Psychiatr Scand, 1999. **99**(6): p. 399-406.
8.   Keyes, K.M., et al., *Cohort effects explain the increase in autism diagnosis among children born from 1992 to 2003 in California.* Int J Epidemiol, 2011.
9.   *Prevalence of autism spectrum disorders - Autism and Developmental Disabilities Monitoring Network, United States, 2006.* MMWR Surveill Summ, 2009. **58**(10): p. 1-20.
10.   Boyle, C.A., et al., *Trends in the prevalence of developmental disabilities in US children, 1997-2008.* Pediatrics, 2011. **127**(6): p. 1034-42.
11.   Kogan, M.D., et al., *Prevalence of parent-reported diagnosis of autism spectrum disorder among children in the US, 2007.* Pediatrics, 2009. **124**(5): p. 1395-403.
12.   Szatmari, P., et al., *Sex differences in repetitive stereotyped behaviors in autism: implications for genetic liability.* Am J Med Genet B Neuropsychiatr Genet, 2012. **159B**(1): p. 5-12.
13.   Ritvo, E.R., et al., *Concordance for the syndrome of autism in 40 pairs of afflicted twins.* Am J Psychiatry, 1985. **142**(1): p. 74-7.
14.   Steffenburg, S., et al., *A twin study of autism in Denmark, Finland, Iceland, Norway and Sweden.* J Child Psychol Psychiatry, 1989. **30**(3): p. 405-16.
15.   Bailey, A., et al., *Autism as a strongly genetic disorder: evidence from a British twin study.* Psychol Med, 1995. **25**(1): p. 63-77.
16.   Lichtenstein, P., et al., *The genetics of autism spectrum disorders and related neuropsychiatric disorders in childhood.* Am J Psychiatry, 2010. **167**(11): p. 1357-63.
17.   Ronald, A., et al., *Genetic heterogeneity between the three components of the autism spectrum: a twin study.* J Am Acad Child Adolesc Psychiatry, 2006. **45**(6): p. 691-9.
18.   Hoekstra, R.A., et al., *Heritability of autistic traits in the general population.* Arch Pediatr Adolesc Med, 2007. **161**(4): p. 372-7.

19.   Ronald, A. and R.A. Hoekstra, *Autism spectrum disorders and autistic traits: a decade of new twin studies.* Am J Med Genet B Neuropsychiatr Genet, 2011. **156B**(3): p. 255-74.

20.   Hallmayer, J., et al., *Genetic heritability and shared environmental factors among twin pairs with autism.* Arch Gen Psychiatry, 2011. **68**(11): p. 1095-102.

21.   Constantino, J.N., et al., *Sibling recurrence and the genetic epidemiology of autism.* Am J Psychiatry, 2010. **167**(11): p. 1349-56.

22.   Ritvo, E.R., et al., *The UCLA-University of Utah epidemiologic survey of autism: recurrence risk estimates and genetic counseling.* Am J Psychiatry, 1989. **146**(8): p. 1032-6.

23.   Bolton, P., et al., *A case-control family history study of autism.* J Child Psychol Psychiatry, 1994. **35**(5): p. 877-900.

24.   Sumi, S., et al., *Sibling risk of pervasive developmental disorder estimated by means of an epidemiologic survey in Nagoya, Japan.* J Hum Genet, 2006. **51**(6): p. 518-22.

25.   Martin, C.L. and D.H. Ledbetter, *Autism and cytogenetic abnormalities: solving autism one chromosome at a time.* Curr Psychiatry Rep, 2007. **9**(2): p. 141-7.

26.   Cooper, G.M., et al., *A copy number variation morbidity map of developmental delay.* Nat Genet, 2011. **43**(9): p. 838-46.

27.   Zafeiriou, D.I., A. Ververi, and E. Vargiami, *Childhood autism and associated comorbidities.* Brain Dev, 2007. **29**(5): p. 257-72.

28.   Consortium, I.M.G.S.o.A., *A full genome screen for autism with evidence for linkage to a region on chromosome 7q.* Hum Mol Genet, 1998. **7**(3): p. 571-8.

29.   Philippe, A., et al., *Genome-wide scan for autism susceptibility genes. Paris Autism Research International Sibpair Study.* Hum Mol Genet, 1999. **8**(5): p. 805-12.

30.   Risch, N., et al., *A genomic screen of autism: evidence for a multilocus etiology.* Am J Hum Genet, 1999. **65**(2): p. 493-507.

31.   Liu, J., et al., *A genomewide screen for autism susceptibility loci.* Am J Hum Genet, 2001. **69**(2): p. 327-40.

32.   Molloy, C.A., M. Keddache, and L.J. Martin, *Evidence for linkage on 21q and 7q in a subset of autism characterized by developmental regression.* Mol Psychiatry, 2005. **10**(8): p. 741-6.

33.   Szatmari, P., et al., *Mapping autism risk loci using genetic linkage and chromosomal rearrangements.* Nat Genet, 2007. **39**(3): p. 319-28.

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

35.   Lockwood, W.W., et al., *Recent advances in array comparative genomic hybridization technologies and their applications in human genetics.* Eur J Hum Genet, 2006. **14**(2): p. 139-48.

36.   Feuk, L., A.R. Carson, and S.W. Scherer, *Structural variation in the human genome.* Nat Rev Genet, 2006. **7**(2): p. 85-97.

37.   Sebat, J., et al., *Large-scale copy number polymorphism in the human genome.* Science, 2004. **305**(5683): p. 525-8.

38. Iafrate, A.J., et al., *Detection of large-scale variation in the human genome.* Nat Genet, 2004. **36**(9): p. 949-51.

39. Sharp, A.J., et al., *Segmental duplications and copy-number variation in the human genome.* Am J Hum Genet, 2005. **77**(1): p. 78-88.

40. Tuzun, E., et al., *Fine-scale structural variation of the human genome.* Nat Genet, 2005. **37**(7): p. 727-32.

41. Hinds, D.A., et al., *Common deletions and SNPs are in linkage disequilibrium in the human genome.* Nat Genet, 2006. **38**(1): p. 82-5.

42. McCarroll, S.A., et al., *Common deletion polymorphisms in the human genome.* Nat Genet, 2006. **38**(1): p. 86-92.

43. Freeman, J.L., et al., *Copy number variation: new insights in genome diversity.* Genome Res, 2006. **16**(8): p. 949-61.

44. Redon, R., et al., *Global variation in copy number in the human genome.* Nature, 2006. **444**(7118): p. 444-54.

45. de Stahl, T.D., et al., *Profiling of copy number variations (CNVs) in healthy individuals from three ethnic groups using a human genome 32 K BAC-clone-based array.* Hum Mutat, 2007.

46. Korbel, J.O., et al., *Paired-end mapping reveals extensive structural variation in the human genome.* Science, 2007. **318**(5849): p. 420-6.

47. Wong, K.K., et al., *A comprehensive analysis of common copy-number variations in the human genome.* Am J Hum Genet, 2007. **80**(1): p. 91-104.

48. de Smith, A.J., et al., *Array CGH analysis of copy number variation identifies 1284 new genes variant in healthy white males: implications for association studies of complex diseases.* Hum Mol Genet, 2007. **16**(23): p. 2783-94.

49. Kidd, J.M., et al., *Mapping and sequencing of structural variation from eight human genomes.* Nature, 2008. **453**(7191): p. 56-64.

50. McCarroll, S.A., et al., *Integrated detection and population-genetic analysis of SNPs and copy number variation.* Nat Genet, 2008. **40**(10): p. 1166-74.

51. Conrad, D.F., et al., *A high-resolution survey of deletion polymorphism in the human genome.* Nat Genet, 2006. **38**(1): p. 75-81.

52. Conrad, D.F. and M.E. Hurles, *The population genetics of structural variation.* Nat Genet, 2007. **39**(7 Suppl): p. S30-6.

53. Perry, G.H., et al., *The fine-scale and complex architecture of human copy-number variation.* Am J Hum Genet, 2008. **82**(3): p. 685-95.

54. Locke, D.P., et al., *Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome.* Am J Hum Genet, 2006. **79**(2): p. 275-90.

55. Bailey, J.A., et al., *Segmental duplications: organization and impact within the current human genome project assembly.* Genome Res, 2001. **11**(6): p. 1005-17.

56. Sharp, A.J., et al., *Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome.* Nat Genet, 2006. **38**(9): p. 1038-42.

57. Eichler, E.E., *Recent duplication, domain accretion and the dynamic mutation of the human genome.* Trends Genet, 2001. **17**(11): p. 661-9.

58. Bailey, J.A., J.M. Kidd, and E.E. Eichler, *Human copy number polymorphic genes.* Cytogenet Genome Res, 2008. **123**(1-4): p. 234-43.

59.   Itsara, A., et al., *Population analysis of large copy number variants and hotspots of human genetic disease.* Am J Hum Genet, 2009. **84**(2): p. 148-61.

60.   Scherer, S.W., et al., *Challenges and standards in integrating surveys of structural variation.* Nat Genet, 2007. **39**(7 Suppl): p. S7-15.

61.   Cooper, G.M., et al., *Systematic assessment of copy number variant detection via genome-wide SNP genotyping.* Nat Genet, 2008. **40**(10): p. 1199-203.

62.   Yatsenko, S.A., et al., *Microarray-based comparative genomic hybridization using sex-matched reference DNA provides greater sensitivity for detection of sex chromosome imbalances than array-comparative genomic hybridization with sex-mismatched reference DNA.* J Mol Diagn, 2009. **11**(3): p. 226-37.

63.   Iafrate, A.J., et al. *Database of Genomic Variants.* 2004 Nov 02, 2010; Available from: projects.tcag.ca/variation/.

64.   Marshall, C.R. and S.W. Scherer, *Detection and characterization of copy number variation in autism spectrum disorder.* Methods Mol Biol, 2012. **838**: p. 115-35.

65.   Girirajan, S., et al., *Relative burden of large CNVs on a range of neurodevelopmental phenotypes.* PLoS Genet, 2011. **7**(11): p. e1002334.

66.   Vaags, A.K., et al., *Rare deletions at the neurexin 3 locus in autism spectrum disorder.* Am J Hum Genet, 2012. **90**(1): p. 133-41.

67.   Sanders, S.J., et al., *Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism.* Neuron, 2011. **70**(5): p. 863-85.

68.   Scherer, S.W., et al., *Human chromosome 7: DNA sequence and biology.* Science, 2003. **300**(5620): p. 767-72.

69.   Jacquemont, M.L., et al., *Array-based comparative genomic hybridisation identifies high frequency of cryptic chromosomal rearrangements in patients with syndromic autism spectrum disorders.* J Med Genet, 2006. **43**(11): p. 843-9.

70.   Romano, C., et al., *3rd International Meeting on Cryptic Chromosomal Rearrangements in Mental Retardation and Autism.* Eur J Hum Genet, 2007. **15**(10): p. 1098-101.

71.   Marshall, C.R., et al., *Structural variation of chromosomes in autism spectrum disorder.* Am J Hum Genet, 2008. **82**(2): p. 477-88.

72.   Fantes, J.A., et al., *Organisation of the pericentromeric region of chromosome 15: at least four partial gene copies are amplified in patients with a proximal duplication of 15q.* J Med Genet, 2002. **39**(3): p. 170-7.

73.   O'Donovan, M.C., G. Kirov, and M.J. Owen, *Phenotypic variations on the theme of CNVs.* Nat Genet, 2008. **40**(12): p. 1392-3.

74.   Sebat, J., et al., *Strong association of de novo copy number mutations with autism.* Science, 2007. **316**(5823): p. 445-9.

75.   Cho, S.C., et al., *Copy number variations associated with idiopathic autism identified by whole-genome microarray-based comparative genomic hybridization.* Psychiatr Genet, 2009. **19**(4): p. 177-85.

76.   Fernandez, B.A., et al., *Phenotypic spectrum associated with de novo and inherited deletions and duplications at 16p11.2 in individuals ascertained for diagnosis of autism spectrum disorder.* J Med Genet, 2010. **47**(3): p. 195-203.

77.   Yeargin-Allsopp, M., et al., *Prevalence of autism in a US metropolitan area.* Jama, 2003. **289**(1): p. 49-55.

78. Skuse, D.H., et al., *Evidence from Turner's syndrome of an imprinted X-linked locus affecting cognitive function.* Nature, 1997. **387**(6634): p. 705-8.

79. Creswell, C. and D. Skuse, *Autism in Association with Turner Syndrome: Genetic Implications for Male Vulnerability to Pervasive Developmental Disorders.* Neurocase, 1999. **5**: p. 511-518.

80. Stone, J.L., et al., *Evidence for sex-specific risk alleles in autism spectrum disorder.* Am J Hum Genet, 2004. **75**(6): p. 1117-23.

81. Gauthier, J., et al., *Autism spectrum disorders associated with X chromosome markers in French-Canadian males.* Mol Psychiatry, 2006. **11**(2): p. 206-13.

82. Noor, A., et al., *Disruption at the PTCHD1 Locus on Xp22.11 in Autism spectrum disorder and intellectual disability.* Sci Transl Med, 2010. **2**(49): p. 49ra68.

83. Piton, A., et al., *Systematic resequencing of X-chromosome synaptic genes in autism spectrum disorder and schizophrenia.* Mol Psychiatry, 2011. **16**(8): p. 867-80.

84. Vincent, J.B., et al., *Genetic linkage analysis of the X chromosome in autism, with emphasis on the fragile X region.* Psychiatr Genet, 2005. **15**(2): p. 83-90.

85. Talebizadeh, Z., et al., *Brief report: non-random X chromosome inactivation in females with autism.* J Autism Dev Disord, 2005. **35**(5): p. 675-81.

86. Amos-Landgraf, J.M., et al., *X chromosome-inactivation patterns of 1,005 phenotypically unaffected females.* Am J Hum Genet, 2006. **79**(3): p. 493-9.

87. Thomas, N.S., et al., *Xp deletions associated with autism in three females.* Hum Genet, 1999. **104**(1): p. 43-8.

| | NimbleGen Protocol | | Emory Protocol | |
Sample	385K aCGH	2.1M aCGH	2.1M aCGH	CNV genotyping
AGRE - autistic proband	4	50	100	300
AGRE - unaffected father	-	-	-	300
SSC - autistic proband	-	-	64	64
NIMH - unaffected control	-	100	-	1,500

Table 1.1 – CNV detection and genotyping in autistic and unaffected populations.
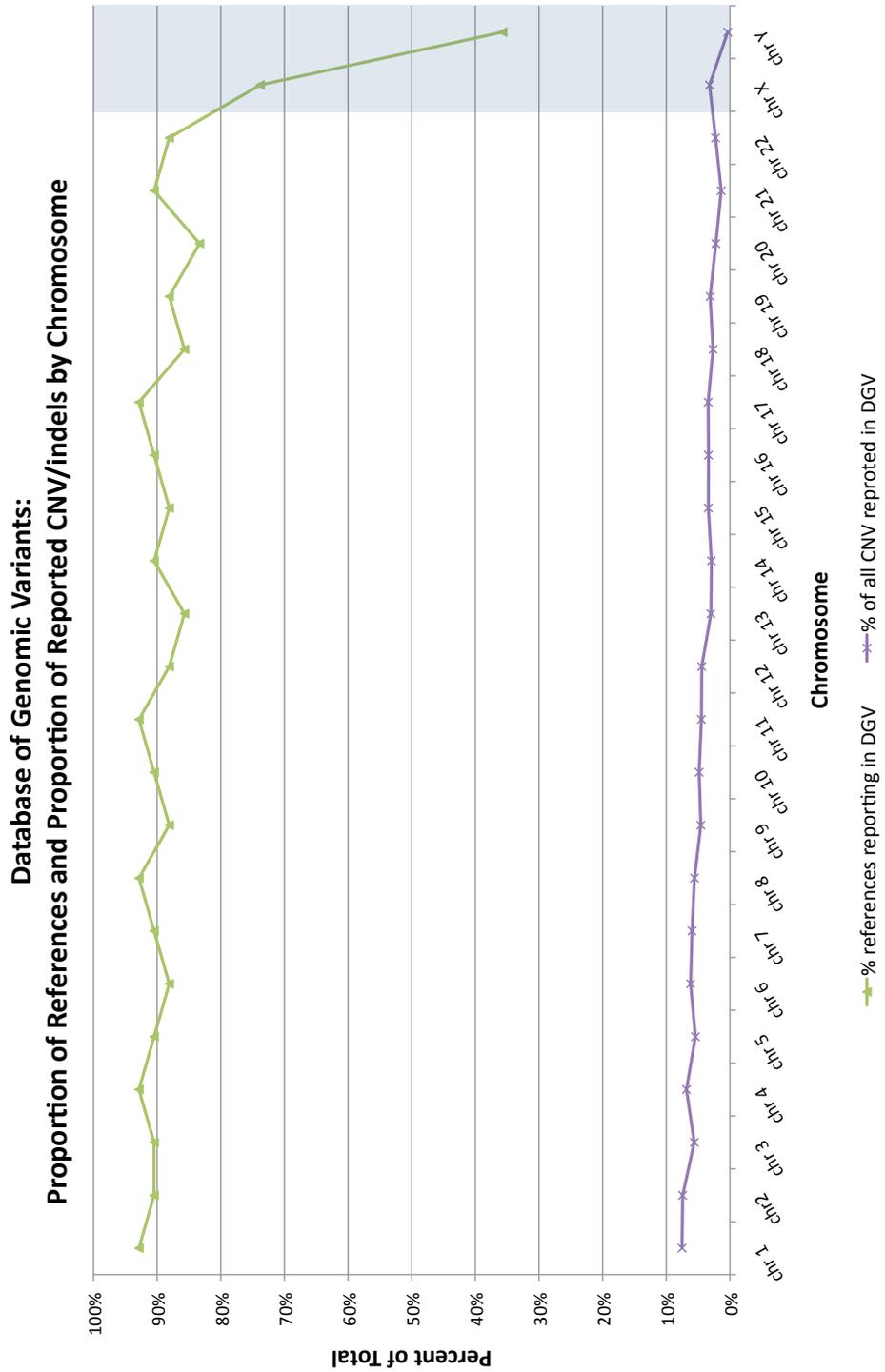
Figure 1.1.a – Proportion of References and CNV or indels reported in the DGV. Proportion of all references reporting CNV or indel in the DGV (Hg18) as well as the proportion of all CNV or indel in the DGV plotted by chromosome.
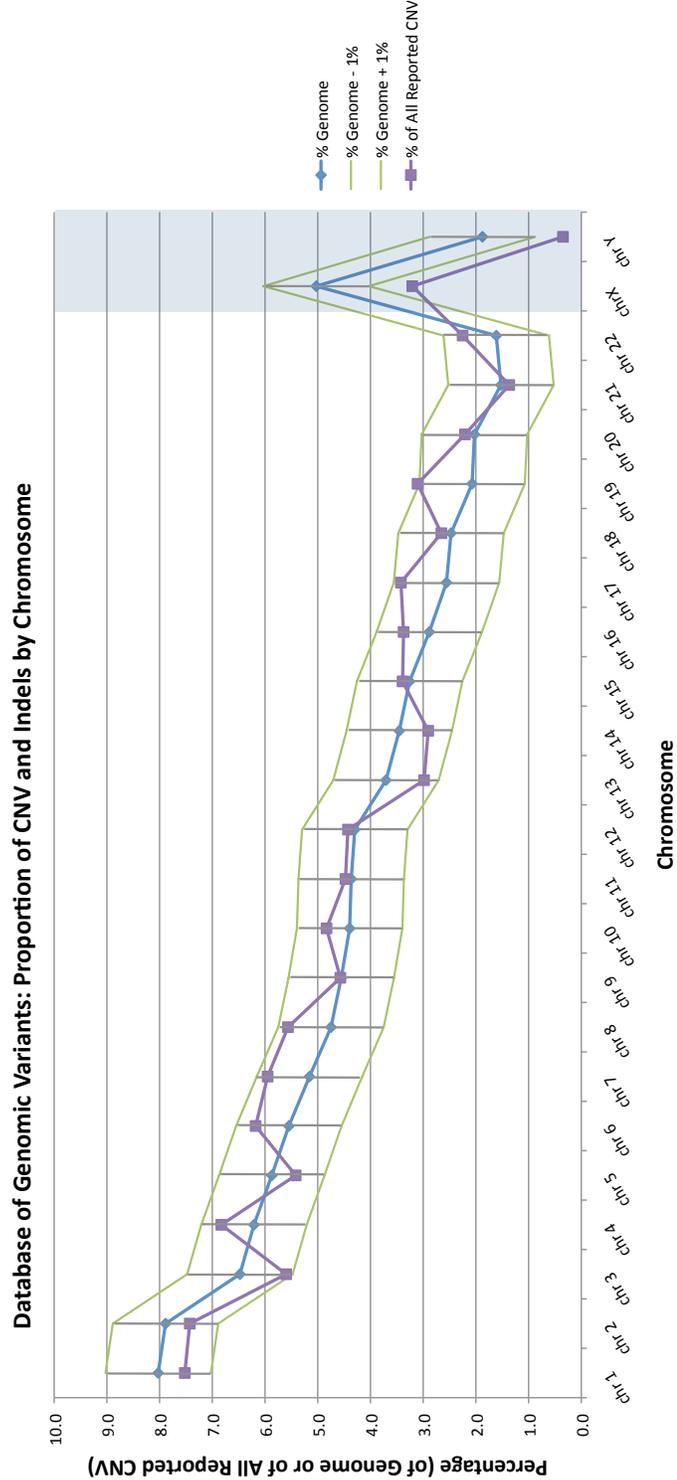
Figure 1.1.b - Proportion of CNV or indels reported in the DGV by chromosome.

Chapter 2. Optimization of the NimbleGen array Comparative Genomic Hybridization (aCGH) Protocol

1. Background

a. Properties of Oligonucleotide Microarray Data

While oligonucleotide CGH arrays enable a higher resolution of copy number variant (CNV) detection than BAC-based arrays, it is recognized that multiple oligonucleotide probes are required to 'call' a CNV with the same confidence as a single BAC probe due to increased noise in oligonucleotide based arrays. [36, 46] BAC-based arrays have a higher signal to noise ratio than oligonucleotide-based arrays enabling CNV detection by fewer probes.[36, 51, 89, 90] Reported false positive rates (or the identification of CNV that fail to validate) for oligonucleotide arrays tend to range from 5-17%, rates have been reported as high as 32-66%. [38, 44, 45, 49, 52, 55, 91-96] Using the genome-wide NimbleGen high-density array (2.1M) with the resolution of one probe/35 kilobases (kb), Itsara and colleagues reported a false call rate of 23%. However, for those variants greater than 100 kb, the false call rate was reduced to 15%. [60]

Additionally, Locke *et al* reported a high false negative call rates (true CNV that fail to be identified) for oligonucleotide-based arrays. Using a BAC-based array to identify CNV in 269 HapMap individuals, the authors employed a custom 385K NimbleGen CGH array dedicated to the same regions as the BAC array to validate their findings in nine individuals. They found that 66% of CNV were identified by both arrays. Forty of 63 calls not made by the oligonucleotide array were called by other platforms in other

studies. The authors concluded the failure to capture all calls by both arrays was due to the high false negative (34%) of the oligonucleotide array. [55]

Our initial analysis of 50 samples hybridized to the 2.1M CGH array, resulted in an unusually high false call rate of 65%. Additionally, our false negative calls (those CNV not called despite being real) were high. We base this evaluation on two validated CNV, a deletion and duplication, in our reference. An 826 bp deletion in the reference (duplication in test samples) failed to be called in 39/48 (81%) individuals, and a 293 bp duplication (deletion in test samples) failed to be called in 21/48 (44%).

A number of strategies exist to identify CNV that are likely real (true calls) rather than a technical artifact. These include stringent probe selection to reduce cross hybridization, dye-swap experiments to identify CNV that are likely real and not random noise-generated, self-vs-self hybridizations to determine a platform specific false call rate, or comparison of CNV calls for one sample across multiple platforms to determine false discovery rate for a given platform again to increase fidelity of calls.[61] While these strategies are all possible, they are also quite labor and resource intensive. Additionally, more difficult to assess are the loss of CNV to false negative calls. However, an ideal platform for CNV identification would be able to accurately call CNV that exist (increase true calls, decrease false negatives) while minimizing noise or other artifacts that result in false calls.

> *b. Specifications for the CGH arrays (385K, 2.1M)*

Our study of the fine scale nature of copy number variation on the X chromosome began

with the 385K Comparative Genomic Hybridization array (385K aCGH) from

NimbleGen (now Roche NimbleGen). We used a custom design array that tiled 385,000

probes along the X chromosome with an average intermarker distance of 270 bp or four

probes/kb. However, early in our screen of the Autism Genetic Resource Exchange

(AGRE, research.agre.org) cohort, a high density CGH array (2.1M aCGH) for the X

chromosome became available. Again, NimbleGen designed a custom array dedicated to

the X chromosome that now tiled 2,100,000 probes among three subarrays (A01, A02,

and A03; approximately 714,000 probes each) synthesized on a single glass slide. Our

intermarker distance was greatly shortened to 50 bp or 20 probes/kb.

c. General description of the manufacturer's protocol

Oligonucleotide based aCGH sample processing and hybridization are largely similar

across different platforms. For different aCGH protocols, sample amplification is often a

variation of a randomly primed single-strand displacement reaction[97], and Cy3 or Cy5

fluorophores are conjugated to either the randomer primers or a subset of nucleotides that

are then incorporated during extension. Hybridization is mediated during sample

agitation over the array area for an extended period of time depending on probe density

and at a platform specific temperature. Array washing is either manual or automated.

The NimbleGen protocols for sample processing and array hybridization of the 385K and

2.1M CGH arrays are quite similar. Depending on the density of array, one to two

micrograms of DNA are sonicated to 200 – 2,000 bp. This material is then concurrently

amplified and labeled using Cy3 or Cy5 5' fluorinated random nonamers in a two hour

Klenow extension reaction. The Klenow polymerase fragment does not have exonuclease

activity, and as the polymerase extends along the template, it will displace rather than

degrade any double stranded product downstream of its starting position.[97] DNA being

tested for CNV is typically labeled with the Cy3 fluorophore, and the reference DNA

typically labeled with the Cy5 fluorophore. The amplified and labeled product is then

isolated, and a spectrophotometer is used to estimate amount of labeled product as well as

the amount of fluorescence per unit DNA (Specific Activity, picomol/microgram).

Equal amounts (15 micrograms (μg) for 385K, 30 μg for 2.1M) of the labeled test and

reference samples are then mixed together in a hybridization solution and applied to the

microarrays. Disposable 'lids' are applied to each array creating a hybridization chamber

over the array area, sample is injected into the chamber, and sealing of the airports at

either end of the chamber create a closed system. The lid-array assembly is then loaded

onto a BioMicro MAUI mixing station where a preprogrammed setting controls the

frequency of agitation by which an internal airbladder shifts the hybridization solution

over the array area. Both array and lid assembly sit are held at 42° Celsius during

hybridization. The 385K arrays are hybridized for 16-20 hours and the 2.1M arrays for

48-60 hours. Once hybridization is complete, the lid-array assembly is separated, the

arrays are manually rinsed in three increasingly stringent washes, and then arrays are

spun dry. Scanning for either array is done on the Axon 4000B dual laser scanner at five

micron resolution. At this resolution, there are approximately 10 pixels/feature for the

385K array (feature size: 16 um x 16 um), and 7 pixels/feature for the 2.1M array (feature size: 13 um x 13 um).

d. Experimental steps that were altered

In effort to reduce our observed false positive rates and false negative rates, we explored ways to make the NimbleGen protocol more stringent. We hypothesized that a large underlying component to these relatively high rates may be cross hybridization of non-specific DNA. We considered how we might incorporate a Tecan array hybridization and wash station (HSPro 4800) into NimbleGen's array processing protocol. This instrument was already utilized by the Agilent aCGH platform with great success and would allow a more dynamic hybridization program as well as automated washing and drying. Additionally, rubber gaskets ringing the array area of the hybridization chamber created the closed system as opposed to the BioMicro lids. The lid system is based on an adhesive seal, and, on occasion, this seal would leak during hybridization. Typically, array data was not recoverable from arrays that had 'leaked'. The Tecan hybridization chamber and gasket sealing system could minimize array loss. Additionally, automation would provide more consistent processing conditions across experiments and batches. By minimizing the degree and nature of manual processing, we sought to reduce noise introduced by any manual procedures, and increased stringency in hybridization would reduce signal due to non-specific hybridization.

2. Optimized Steps

a. Experimental Protocol Changes and quality controls instituted

Reference used and Ozone-free space used to process arrays

Our protocol optimization is based on the 2006 version of the manufacturer's protocol. At that time, the original protocol implemented the use of sonication to normalize genomic DNA size across samples; however, the manufacturer no longer considered sonication necessary in subsequent, updated versions of their protocol. Nonetheless, we continued to sonicate all test and reference DNA prior to labeling and amplification to allow for a direct comparison of further downstream changes that we would make to the manufacturer's protocol. Additionally, once sonicated sizes were confirmed by gel electrophoresis, all sonicated reference DNAs were then pooled together before labeling and amplification. Pooling created a normalized reference source of sonicated DNAs for each batch of processed samples.

High ozone and humidity levels have been identified as strong contributors to Cy5 signal degradation. [98, 99] To reduce ambient exposure of labeled samples and hybridized arrays to ozone and humidity, a separate room was retrofitted for sample processing, array hybridization and wash, and array scanning. This room was maintained under positive pressure, a dehumidifier continually operated to reduce humidity, and a continually operated ozone scrubber removed ozone from ambient air. A highly sensitivite ozone detector installed in the room typically registered 8-10 parts per billion (ppb) or less. While this level might seem somewhat elevated, ozone levels inside the room were greatly reduced as compared to those outside of the room (ranging as high as 100+ ppb in peak ozone season), ozone levels were maintained at a constant level throughout high and low ambient ozone season, and the measurement error of the reader

was five ppb suggesting the actual ozone level may have been on the order of three to five ppb. All steps in sample and array processing described below were conducted in this environment.

Labeling and Amplification

Further downstream in the aCGH protocol, the stringency of sample hybridization and washing were increased to minimize non-specific hybridizations. Initial runs indicated that raw fluorescent signal was too close to background noise to make identification of copy number changes possible. While this decrease in signal suggests an increase in the stringency of hybridization kinetics, it was clear that we needed to increase overall sample fluorescence in order to proceed with the more stringent application.

Labeling of the test and reference DNAs with fluorescent molecules Cy3 and Cy5 respectively occurs during the amplification step. Traditional aCGH protocols employ one of two labeling strategies that feature either a 5'-fluorinated primer to end label the amplified material, or a nucleotide conjugated with Cy3 or Cy5 is incorporated during extension. To maximize these labeling potentials, we combined both of these methods. (Figure 2.1) One drawback of backbone labeling with a Cy3 or a Cy5 fluoraphor is that the fluoraphor is bulky. Lee *et al* hypothesized that an observed intensity bias of Cy3 over Cy5 stemmed from the difference in size of the two fluorophors contributed to a bias in the fluorinated nucleotide incorporation.[100] The hypothesis is predicated on the assumption that the Klenow fragment has difficulty incorporating the fluorinated nucleotide. This is substantiated by our own observations that samples labeled with a

44

fluorinated nucleotide result in a shorter fragment population than those labeled with unmodified nucleotides.

To circumvent any bias in labeling efficiency and to maintain a similar population length of amplified molecules as seen in the manufacturer's protocol, we employed an indirect labeling method that utilizes a uradine base conjugated with an aminohexylacrylamido (aha) tag.[101] Amine reactive Cy3 or Cy5 dyes are then added to the DNA once amplification is complete. The reactive fluors label the backbone of the amplified fragments. The aha tag is considerably smaller than the Cy3 or Cy5 fluoraphor, which improves the Klenow fragment's ability to incorporate the modified nucleotides and allows for larger fragments to be amplified, generating a labeled product similar in size to the manufacturer's protocol 5' end labeled size.[100, 101] Utilizing the same sized tag in both the test and reference DNA enabled a systematic incorporation rate of aha-modified nucleotides for either the sample or reference. Remarkably, by using both the 5' labeled primer and the backbone labeling strategy, we were able to increase the overall fluorescence of the hybridized samples over 2.23 fold with the 5' end labeling strategy yielding an average 30 pmol/ug and the combined strategy yielding an average 67 pmol/ug in initial tests. With this gain in fluorescence per amplified unit, we could continue to hybridize the same amounts of amplified material as recommended by the manufacturer without losing signal in a more stringent hybridization and wash environment.

Array Hybridization and Wash

The manufacturer's hybridization protocol utilizes the BioMicro MAUI hybridization system and a manual wash protocol. The MAUI system uses disposable, one-time use hybridization chambers that result in a low-volume and closed hybridization environment. Airbladders within each chamber mix hybridization liquid over the array surface and frequency of agitation is hard-coded into the hybridization station with four frequency-of-agitation options available on the MAUI 4-bay system (two options on the 12-bay system). Manual washing inherently introduces variation in frequency, force and duration of array exposure to wash solutions and the environment. This design makes it nearly impossible to ensure each array is treated identically.

Thus, we automated the hybridization and wash steps of the protocol by utilizing the Tecan 4800Pro Hybridization station. This stand-alone system uniformly hybridizes, washes, and dries microarrays with constant agitation frequencies, force, and time duration. Additionally, an Active Bubble Suppression system is incorporated into the fluid dynamics of the machine to prevent bubble formation over array areas. Moreover, the Tecan system comes equipped with a program designed specifically for Agilent CGH arrays. With assistance from the technical support staff from Tecan, we custom developed a hybridization, wash, and dry program for the Roche NimbleGen arrays. The hybridization annealing temperature was increased from 42°C to 52°C to select for a decrease in non-specific hybridization and false calls. Hybridization force and frequency were elevated, and hybridization time was increased from 48 hours to 60 hours as Dai *et al* have reported that longer hybridization times reduce noise in array data.[102] Also, additional washes with the least stringent buffer were added in effort to remove any

precipitates that may have formed during the long hybridization time. Sterile nitrogen gas was used to dry arrays. Arrays were then held in the closed, nitrogen filled environment until ready for scanning to protect array signal from possible ozone degradation.

b. Computational Protocol Changes and Quality Control Steps Instituted

<u>*Array Scanning*</u>

We developed an array scanning protocol meant to maintain as consistent a set of scanning conditions and parameters as possible for all arrays scanned. To start, all arrays were scanned on the same GenePix 4000B scanner with the same resolution setting (five micron). The scanner was calibrated once a month or more depending on overall usage. The GenePix 4000B is a dual laser scanner (532 nanometers (nm) and 635nm) that can scan for both Cy3 and Cy5 fluors using both lasers concurrently.

Using the 'Histogram' function hard coded into the scanner software, pixel intensities are plotted against their log-transformed frequencies. Each channel, 532 nm (green color, Cy3) and 635 nm (red color, Cy5) are plotted together. The 'Count Ratio' measurement is the ratio of the sum of fluorescent signal from each channel with an ideal measurement of one.[103] Scanned images were evaluated by their Count Ratio (CR) value and the final placement of the high-intensity 'tails' of the intensity histogram. The photomultiplier (PMT) gain settings were manually adjusted to generate CR values of 1 +/- 0.15 to provide the greatest overlap of total fluorescence, and the high intensity-tails targeted to the 1e-5 to 1e-4 range. (Figure 2.2)

We used a Roche NimbleGen custom microarray comprised of approximately two million oligonucleotide probes tiled along the X chromosome. This first step in our computational optimization was the development of a bioinformatic strategy that would determine whether any of the probes behaved inconsistently. We hypothesized that removal of such probes would purge a portion of false positive variants from our data. [104]

Multiple studies of immobilized probe sequence selection and kinetic behavior have been previously conducted in an effort to identify those properties of a probe that enable it to hybridize reliably across experiments. The parameters that have been evaluated and found to play a significant role in the fidelity of specific oligonucleotide hybridization include probe GC content, melting temperature, sequence homopolymers, stem loop structure, self-folding, and sequence similarity.[104-114] In a recent work, Mulle *et al.* found that probe melting temperature, single nucleotide polymorphisms (SNPs), and homocytosine motifs affected oligonucleotide array data quality. [104] These authors identified probes having irregular patterns of hybridization by evaluating individual probe log2 values their variance across nineteen arrays.

Using the data generated by the fifty AGRE samples run by the NimbleGen protocol on the 2.1M arrays, we evaluated the log2 probe variance in this data set. For each of the three subarrays (A01, A02, and A03), the log2 values for the 714,456 experimental probes on each subarray were collected, and variance for each probe determined. As with

the findings made by Mulle *et al*, we similarly identified a subset of probes with a

relatively high variance as compared to the remainder of probes on the subarray. These

probes with the high variance were defined as "poorly" behaving if their variance was

greater than 0.175. The value of 0.175 was chosen based on the comparative behavior of

probe variance within a subarray. We determined the variance for each probes, sorted the

probes by their variance, and plotted all probes from the subarray by their ranking (x-

axis) and variance (y-axis). (Figure 2.3a-c) The cutoff value of 0.175 was subjectively

chosen because in all three subarrays, this value appeared to remove the majority of high-

variance probes. Removal of these increased variance probes resulted in the total loss of

122,544 experimental probes (5.72% of all experimental probes). (Table 2.1)

After deviant probes were removed from analysis, the intensity distribution from each

array was assessed using an MA plot ('M' is the intensity ratio or $\log2(R) - \log2(G)$, and

'A' is the average intensity or $1/2 (\log2(R) + \log2(G))$). An MA plot is an intensity-

dependent ratio of raw microarray data plotted against the total average intensities.

(Figure 2.4.a) MA plots were evaluated on several parameters. Acceptable plots exhibited

an oval shape symmetric about the y=0 line, extended within x=9 to x=16 values, and

overall height bounded by y=+/-1. Additionally, a fitted linear regression model was

applied to each subarray where the M values were regressed on the A values. The

outputted coefficients of intercept and slope were plotted on each MA plot, and the slope

recorded as a further assessment of array performance.

Data exhibiting non-symmetric shapes about the y=0 axis, demonstrating a more condensed body (typically, 2-3 units along the x-axis), showing dense scatter beyond +/-1 on the y-axis, and/or having a regressed slope deviating greatly from '0' were removed from further analysis and samples re-run. (Figure 2.4.b-c) Each subarray was evaluated by these criteria, and a passing subarray was then assessed for copy number changes.

Identification and selection of high confidence CNV

The manufacturer's segment analysis software, NimbleScan (v2.4) was used to assess copy number change in subarrays that passed the MA plot quality control step. The output file, 'segtable', is a text report that returns the entire evaluated region (i.e., in our study, the chromosome X from bases 2,709,520 to 154,583,236; Build Hg18) as defined by its log2 values. With a majority of the chromosome near log2=0, this extensive list of segments or potential CNV is quite cumbersome to go filter by hand. In order to extract relevant and plausible segments for further analysis/validation, we developed an automated selection strategy to refine our candidate validation list.

We evaluated segments that had a log2 value greater (more positive or more negative depending on the shift) than one standard deviation (SD) from the mean of all log2 values of the respective subarray. Further inspection of our data showed larger CNV as represented by multiple segments shifted in the same direction. To identify the most parsimonious set of segments for validation, we merged all segments that ended within three kb of each other and shifted in the same direction (i.e., all log2s were either greater than one SD or less than one SD from the mean). These merged variants were then

defined as having the most extreme 5' and 3' boundaries of the originally identified segments, and the log2 value assigned to the merged variant was an average of the original log2s weighted by the number of probes within the original segment. (Figure 2.5) We also developed an exclusionary strategy that removed segments identified by fewer than nine probes per kb. And, we excluded individuals with a number of variants greater than three standard deviations from the mean number of variants called for all individuals. Together, these selection and exclusion strategies provided us with a list of loci that we further analyzed and validated for their contribution to disease status.

3. Results

a. Results from Experimental Changes

While reference selection and sonication contributed to overall experimental strategy and consistency throughout processing of all samples, the development of a labeling protocol that increased Specific Activity (fluorescence per amplified DNA unit) was critical for the implementation of the more stringent hybridization and wash parameters introduced by the use of the Tecan 4800Pro. Using the same samples to label and amplify by the NimbleGen and our optimized protocol, we increased the SA values from 29.7 pmol/μg to 67.3 pmol/μg.

The Tecan 4800Pro has excelled in its ability to hybridize, wash and dry Agilent arrays. Unfortunately, despite all efforts to refine the final hybridization and wash protocol for the NimbleGen arrays, the hybridization, wash and dry parameters do not have nearly the high degree of performance or reliability as seen for Agilent and other array platforms.

Specifically, areas of saturated fluorescence that mimicked a 'streaking' or 'drip' -like pattern from the liquid ports located at the top and bottom of the array slide were frequently observed. (Figure 2.6) We hypothesized that residual hybridization solution and/or sample may be 'sticking' in the capillary tubes of the hybridization chambers. But, despite rigorous purging and washing of the hybridization chambers and liquid handling lines of the machine after hybridizing arrays, the issue persisted. An alternative hypothesis that may explain our observations is that component(s) of the hybridization solution are not compatible with the increased annealing temperature and/or the hybridization chamber and tubing materials used in the Tecan system. While it may be possible to develop a hybridization solution that is compatible with both NimbleGen High Density microarrays and the Tecan HS4800Pro hybridization station, for purposes of time and effort efficiency, we did not pursue such an option.

b. Results from Computational Changes

We identified 122,544 experimental probes (5.72% of all experimental probes) as having a log2 variance > 0.175 across 50 arrays. We evaluated the two probe sets (greater or less than 0.175 log2 variance) by probe length, GC content, AG (purine) content and melting temperature (T_m). Probe length, GC content, Ag content, and T_m were significantly different for those probes having a log2 variance greater than 0.174 versus those less than 0.175. (Table 2.2, Figure 2.7) These findings are consistent with previous reports. [104, 110-112, 114]

We then assessed the 'poorly' behaving experimental probes by position across the X chromosome. These probes showed a similar positional distribution as the remaining probes, indicating that they were randomly distributed along the chromosome and interspersed among probes that behave more consistently. Removal of the 'poorly' behaving probes did not result in significant gaps in probe coverage of the X-chromosome, and further analysis without them will remain largely comprehensive.

   c. Results from the Sum of All Optimized (Experimental and Computational) Steps
We then sought to evaluate how our modifications both experimentally and computationally compared to values generated following NimbleGen's protocol. First, we assessed the overall performance of the arrays by their average log2 values from each protocol. The expected average log2 value for each probe is '0'. We found that the optimized technical protocol produced average log2 values that were closer to the expected value of '0,' irrespective of subarray. However, these optimized data were also 'noisier' as seen in their slightly larger standard deviations. (Table 2.3)

Next, we evaluated the CNV identified by both protocols for the same 30 samples for which there was data on both platforms. Using the Segment Filtering strategy described above, the original protocol identified 302 total variants with an average 10.1 CNV/chrX while the optimized protocol identified 174 total variants with an average 5.8 CNV/chrX. This reduction in number of calls is significantly different with $p < 0.001069$. The total number of variant bases identified increased with the optimized protocol because the average size of calls made using the optimized protocol were 2.7 times as large as those

made by the original. (Table 2.4, Figure 2.8) Another feature distinguishing the CNV

identified from the two protocols was the distribution of deletion and duplication events

identified. NimbleGen's protocol resulted in the identification of 184 deletions and 118

duplications (1.6 deletions:duplications) while the optimized protocol identified 44

deletions and 130 duplications (0.34 deletions:duplications). The five-fold decrease in

the deletion to duplication ratio was significantly different at $p < 1.36 \times 10^{-13}$.

The average size of calls made by the optimized protocol was nearly three times that of

the original. CNV identified by NimbleGen's protocol averaged 2,681 bases (SD =

5,673), while the average size of the optimized protocol was 7,264 bases (SD = 30,914).

While the optimized protocol resulted in larger CNV on average, this shift was

statistically significant when CNV size was evaluated by all variant (all CNV: $p < 0.05$).

When stratified by nature of the CNV, statistical significance was lost (deletions: $p <$

0.16, duplications: $p < 0.16$). (Table 2.5, Figure 2.9)

We also evaluated the probe coverage of CNV identified by each protocol. The entire

experimental probe set is comprised of 2,020,823 probes covering base positions

2,709,520 to 154,583,236 (build Hg18) on chromosome X, and after repeat masking,

average intermarker distance is 50 basepairs. This translates to about 20 probes/kb, and

we were willing tolerate copy number variants called by as low as nine probes/kb. Using

the same number of 'good' probes for copy number analysis, the optimized protocol

returns a slightly diminished probe coverage per variant than that of the original protocol

(18.0 probes/kb and 20.4 probes/kb respectively). This decrease in coverage is significant

for all CNV: $p < 5.24 \times 10^{-10}$, for deletions: $p < 2.23 \times 10^{-12}$, for duplications: $p < 2.23 \times 10^{-12}$. (Table 2.5, Figure 2.9)

GC content of all variants identified by both protocol was also evaluated and found to be significantly different. Variants identified by the original protocol tended to have a lower GC content than those identified by the optimized strategy (original: 0.45 (SD=0.08), optimized: 0.56 (SD=0.13)). This difference was significant for all CNV, deletions only, and duplications only. (all CNV: $p < 6.97 \times 10^{-21}$, deletions: $p < 0.03$, duplications: $p < 0.03$). (Table 2.5, Figure 2.9)

As a final measure of protocol performance, we evaluated by several metrics the true and false call rates for CNV called within the same samples from each protocol. First, the optimized protocol was successful in reducing our false negative call rate. For the 30 samples run on both protocols, 33 CNV are known to be real. The NimbleGen protocol failed to call 11 (33%) of these loci, and the false negative (true CNV that fail to be called) rate for the optimized protocol dropped this number to 7 or 21%. And the true positive call rate (true CNV that are called) increased with implementation of the optimized protocol. The true positive call rate in the NimbleGen protocol was 22/62 (35%) for PCR validated loci, and the optimized protocol rate increased for a different set of 62 loci to 26 (42%). However, these two rates were statistically indistinguishable with $p < 0.58$). (Table 2.6, Figure 2.10)

We also evaluated what proportion of true and false calls were made as stratified by GC content. Remarkably, for both protocols, we found GC content to be a strong indicator for a variant's true or false call identity within each protocol. All false positives, irrespective of copy number state (deletion or duplication), showed an elevated GC content over those loci that validated. For both the original and optimized protocols, the average GC content was significantly different between true and false calls with the original protocol: $p < 5.43E-06$ and optimized $p < 2.20E-16$.

While both protocols showed a significant difference in the GC content of the true and false calls, the optimized protocol enables better discrimination of true calls when evaluated by their GC content. For both protocols, a GC content of 56% or less accounted for all true calls while minimizing the number of false calls. Specifically, in the original protocol, variants with a GC content less than or equal to 56% accounted for 26/26 or 100% of all true calls and 59/68 or 87% of all false calls. While in the optimized protocol, variants with GC content equal to or less than 56% still accounted for 52/52 or 100% of all true calls, but the proportion of false calls dropped tremendously to 59/116 or 51% of all false calls. The proportion of false calls dropped by 41% under this threshold. (Figure 2.11a-b)

Additionally, we evaluated the GC content of the X chromosome by 1,000 bp bins. The increasing proportion of chromosome X is plotted by GC content on these two plots (Fig 2.11.a-b) as well. This feature in addition to our optimized call rate data suggests several conclusions. First, a significant proportion of the X chromosome remains assayable at

both the GC content cut offs of 56% and 47% (98.2% and 88.3% of chromosome X remain). Second, the true calls from the optimized protocol mimic a distribution similar to the chromosome on a whole. This should be expected if we were truly 'sampling' real loci for CNV on the X chromosome. And finally, using our optimized protocol, we can assay nearly all (98.2%) of the X chromosome for CNV with a GC content of 56% while eliminating the majority of false calls.

4. Summary

Overall, much effort was expended to develop an aCGH protocol and analysis strategy that would identify copy number variant calls with greater fidelity than the existing protocol. While issues remain to be resolved, our current version of optimized sample processing and array analysis outperforms the original protocol in its ability to better discriminate false from true calls when using GC content as an indicator.

E. References

1. Lockwood, W.W., et al., *Recent advances in array comparative genomic hybridization technologies and their applications in human genetics.* Eur J Hum Genet, 2006. **14**(2): p. 139-48.

2. de Stahl, T.D., et al., *Profiling of copy number variations (CNVs) in healthy individuals from three ethnic groups using a human genome 32 K BAC-clone-based array.* Hum Mutat, 2007.

3. McCarroll, S.A., et al., *Integrated detection and population-genetic analysis of SNPs and copy number variation.* Nat Genet, 2008. **40**(10): p. 1166-74.

4. Zhang, Z.F., et al., *Detection of submicroscopic constitutional chromosome aberrations in clinical diagnostics: a validation of the practical performance of different array platforms.* Eur J Hum Genet, 2008. **16**(7): p. 786-92.

5. Carter, N.P., *Methods and strategies for analyzing copy number variation using DNA microarrays.* Nat Genet, 2007. **39**(7 Suppl): p. S16-21.

6. Sebat, J., et al., *Large-scale copy number polymorphism in the human genome.* Science, 2004. **305**(5683): p. 525-8.

7. Conrad, D.F., et al., *A high-resolution survey of deletion polymorphism in the human genome.* Nat Genet, 2006. **38**(1): p. 75-81.

8. Freeman, J.L., et al., *Copy number variation: new insights in genome diversity.* Genome Res, 2006. **16**(8): p. 949-61.

9. Redon, R., et al., *Global variation in copy number in the human genome.* Nature, 2006. **444**(7118): p. 444-54.

10. de Smith, A.J., et al., *Array CGH analysis of copy number variation identifies 1284 new genes variant in healthy white males: implications for association studies of complex diseases.* Hum Mol Genet, 2007. **16**(23): p. 2783-94.

11. Baross, A., et al., *Assessment of algorithms for high throughput detection of genomic copy number variation in oligonucleotide microarray data.* BMC Bioinformatics, 2007. **8**: p. 368.

12. Korbel, J.O., et al., *Systematic prediction and validation of breakpoints associated with copy-number variants in the human genome.* Proc Natl Acad Sci U S A, 2007. **104**(24): p. 10110-5.

13. Locke, D.P., et al., *Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome.* Am J Hum Genet, 2006. **79**(2): p. 275-90.

14. Conrad, D.F., et al., *Origins and functional impact of copy number variation in the human genome.* Nature, 2010. **464**(7289): p. 704-12.

15. Tucker, T., et al., *Comparison of genome-wide array genomic hybridization platforms for the detection of copy number variants in idiopathic mental retardation.* BMC Med Genomics, 2011. **4**: p. 25.

16. Greenway, S.C., et al., *De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot.* Nat Genet, 2009. **41**(8): p. 931-5.

17. Malhotra, D., et al., *High frequencies of de novo CNVs in bipolar disorder and schizophrenia.* Neuron, 2011. **72**(6): p. 951-63.

18. Itsara, A., et al., *Population analysis of large copy number variants and hotspots of human genetic disease.* Am J Hum Genet, 2009. **84**(2): p. 148-61.

19. Scherer, S.W., et al., *Challenges and standards in integrating surveys of structural variation.* Nat Genet, 2007. **39**(7 Suppl): p. S7-15.

20. Lage, J.M., et al., *Whole genome analysis of genetic alterations in small DNA samples using hyperbranched strand displacement amplification and array-CGH.* Genome Res, 2003. **13**(2): p. 294-307.

21. Fare, T.L., et al., *Effects of atmospheric ozone on microarray data quality.* Anal Chem, 2003. **75**(17): p. 4672-5.

22. Branham, W.S., et al., *Elimination of laboratory ozone leads to a dramatic improvement in the reproducibility of microarray gene expression measurements.* BMC Biotechnol, 2007. **7**: p. 8.

23. Lee, M., J.M. Trent, and M.L. Bittner, *Optimization of oligonucleotide microarray fabricated by spotting 65-mer.* Anal Biochem, 2007. **368**(1): p. 61-9.

24. t Hoen, P.A., et al., *Fluorescent labelling of cRNA for microarray applications.* Nucleic Acids Res, 2003. **31**(5): p. e20.

25. Dai, H., et al., *Use of hybridization kinetics for differentiating specific from non-specific binding to oligonucleotide microarrays.* Nucleic Acids Res, 2002. **30**(16): p. e86.

26. Leung, Y.F. and D. Cavalieri, *Fundamentals of cDNA microarray data analysis.* Trends Genet, 2003. **19**(11): p. 649-59.

27. Mulle, J.G., et al., *Empirical evaluation of oligonucleotide probe selection for DNA microarrays.* PLoS One, 2010. **5**(3): p. e9921.

28. Mahajan, S., et al., *Oligonucleotide microarrays with stem-loop probes: enhancing the hybridization of nucleic acids for sensitive analysis.* Bioorg Med Chem Lett, 2008. **18**(12): p. 3585-8.

29. Leparc, G.G., et al., *Model-based probe set optimization for high-performance microarrays.* Nucleic Acids Res, 2009. **37**(3): p. e18.

30. Pozhitkov, A.E., D. Tautz, and P.A. Noble, *Oligonucleotide microarrays: widely applied--poorly understood.* Brief Funct Genomic Proteomic, 2007. **6**(2): p. 141-8.

31. Sharp, A.J., et al., *Optimal design of oligonucleotide microarrays for measurement of DNA copy-number.* Hum Mol Genet, 2007. **16**(22): p. 2770-9.

32. Xia, X.Q., et al., *Evaluating oligonucleotide properties for DNA microarray probe design.* Nucleic Acids Res, 2010. **38**(11): p. e121.

33. Cutler, D.J., et al., *High-throughput variation detection and genotyping using microarrays.* Genome Res, 2001. **11**(11): p. 1913-25.

34. Kane, M.D., et al., *Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays.* Nucleic Acids Res, 2000. **28**(22): p. 4552-7.

35. Flibotte, S. and D.G. Moerman, *Experimental analysis of oligonucleotide microarray design criteria to detect deletions by comparative genomic hybridization.* BMC Genomics, 2008. **9**: p. 497.

36. Patel, V.C., et al., *Microarray oligonucleotide probe designer (MOPeD): A web service.* Open Access Bioinformatics, 2010. **2**(2010): p. 145-155.

37. Tulpan, D., et al., *Thermodynamically based DNA strand design.* Nucleic Acids Res, 2005. **33**(15): p. 4951-64.

Subarray	Variance Cut-Off	# of probes (%)
All	-	122,540 (5.72%)
A01	> 0.175	46,266 (6.48%)
A02	> 0.175	38,865 (5.44%)
A03	> 0.175	37,410 (5.24%)

Table 2.1 – Number of high variance probes.

	Good Probes	Poor Probes	p value
N (% Total)	1,908,871 (94.5%)	122,544 (5.72%)	
Mean Length in bp (SD)	56 (4.6)	55 (4.3)	< 6.97E-49
Mean AG Content (SD)	50.1 (10.0)	50.7 (10.6)	< 2.20E-16
Mean GC Content (SD)	39.0 (8.4)	37.4 (9.0)	< 2.20E-16
Mean Tm in °C (SD)	75.6 (3.0)	74.4 (3.3)	< 2.20E-16

Table 2.2 - 'Poor' and Well Behaving Probes Evaluated Across Four Parameters

	# of probes	n (samples)	Mean Log2 (SD)	Median Log2
Subarray A01	668,189			
NimbleGen		48	0.00116 (0.26)	0.003
Optimized		75	-0.00002 (0.34)	0.002
Subarray A02	675,590			
NimbleGen		48	0.00232 (0.26)	0.005
Optimized		81	-0.00002 (0.34)	0.002
Subarray A03	677,044			
NimbleGen		48	0.00207 (0.28)	0.005
Optimized		80	-0.00001 (0.34)	0.001

Table 2.3 – Probe Behavior by NimbleGen and Optimized Protocols

N (samples)	Total CNV*	Total Variant Bases (% ChrX)	CNVs/ChrX	Mean # Variant Bases/ChrX (% ChrX)
30	302	809,723 (0.52%)	10.1	2,696 (0.0017%)
30	**174****	1,271,198 (0.82%)	5.8	11,059 (0.0071%)
n	**Total Deletions**	**Total Deleted Bases (% ChrX)**	**Deletions / ChrX**	**Mean # Deleted Bases/ChrX (% ChrX)**
30	184	376,981 (0.24%)	6.1	2,161 (0.0014%)
22	44	344,113 (0.22%)	2.0	11,787 (0.0076%)
n	**Total Duplications**	**Total Duplicated Bases (% ChrX)**	**Duplications / ChrX**	**Mean # Duplicated Bases/ChrX (% ChrX)**
28	118	432,742 (0.28%)	4.2	3,054 (0.0020%)
26	130	925,612 (0.60%)	5.0	6,751 (0.0044%)

*Counts do not include confirmed reference variants.

** Wilcoxon rank sum, p-value < 0.001069

Table 2.4 - CNV identified by the NimbleGen and Optimized protocols.
Both protocols were applied to the same thirty samples.

	N (samples)	Total CNV*	Mean Size in bp (SD)	Median Size in bp	Mean # Probes/kb (SD)	% GC (SD)
NimbleGen	30	302	2,681 (5,673)	1,178	20.4 (3.7)	45 (8)
Optimized	30	**174****	7,264 (30,914)	1,700	**18.0** (4.0)**	**56** (13)**
Deletions	n	Total	Mean Size in bp (SD)	Median Size in bp	Mean # Probes/kb (SD)	% GC (SD)
NimbleGen	30	184	2,049 (3,854)	1,100	21.0 (3.0)	44 (5)
Optimized	22	44	7,821 (26,891)	1,710	**15.2** (4.0)**	**41** (7)**
Duplications	n	Total	Mean Size in bp (SD)	Median Size in bp	Mean # Probes/kb (SD)	% GC (SD)
NimbleGen	28	118	3,667 (7,615)	1,532	19.4 (4.3)	47 (11)
Optimized	26	130	7,120 (32,367)	1,703	**18.9** (3.6)**	**61** (10)**

*Counts do not include confirmed reference variants.
** Optimized significantly different from NimbleGen.

Table 2.5 - CNV characteristics by the NimbleGen and Optimized protocols.
Both protocols were applied to the same thirty samples.

	Total CNV (% of Category)	Mean Size in bp (SD)	Median Size in bp	Mean Probes/kb (SD)	Median Probes/kb	Mean GC Content (SD)	Median GC Content
ALL CNV							
TRUE CALLS							
NimbleGen	22 (35%)	4,884 (7,113)	1,442	19.3 (3.2)	19.8	42% (6)	41%
Optimized	26 (42%)	18,392 (67,005)	2,148	**15.3 (3.7)***	15.7	**41% (5)**	39%
FALSE CALLS							
NimbleGen	40 (65%)	2,176 (2,119)	1,293	20.5 (1.8)	21.1	50% (8)	48%
Optimized	36 (58%)	**6,852 (29,366)***	1,935	**19.5 (2.9)***	20.0	58% (10)	61%
DELETIONS							
TRUE CALLS							
NimbleGen	20 (40%)	3,625 (6,129)	1,272	19.5 (3.1)	20.4	42% (6)	41%
Optimized	23 (79%)	4,302 (6,254)	1,955	15.3 (3.8)	15.5	41% (5)	39%
FALSE CALLS							
NimbleGen	30 (60%)	2,169 (2,429)	1,138	20.6 (2.0)	21.3	47% (5)	48%
Optimized	6 (21%)	31,169 (71,970)	675	17.9 (4.9)	16.9	43% (14)	40%
DUPLICATIONS							
TRUE CALLS							
NimbleGen	2 (17%)	17,471 (508)	17,471	16.9 (3.2)	16.9	41% (3)	41%
Optimized	3 (9%)	126,413 (189,487)	33,261	15.1 (3.7)	15.9	38% (1)	39%
FALSE CALLS							
NimbleGen	10 (83%)	2,198 (664)	2,140	20.4 (1.2)	20.8	59% (9)	61%
Optimized	30 (90%)	1,988 (943)	2,108	19.8 (2.3)	20.2	61% (4)	61%

Table 2.6 – True and False positive loci by the NimbleGen and Optimized protocols.
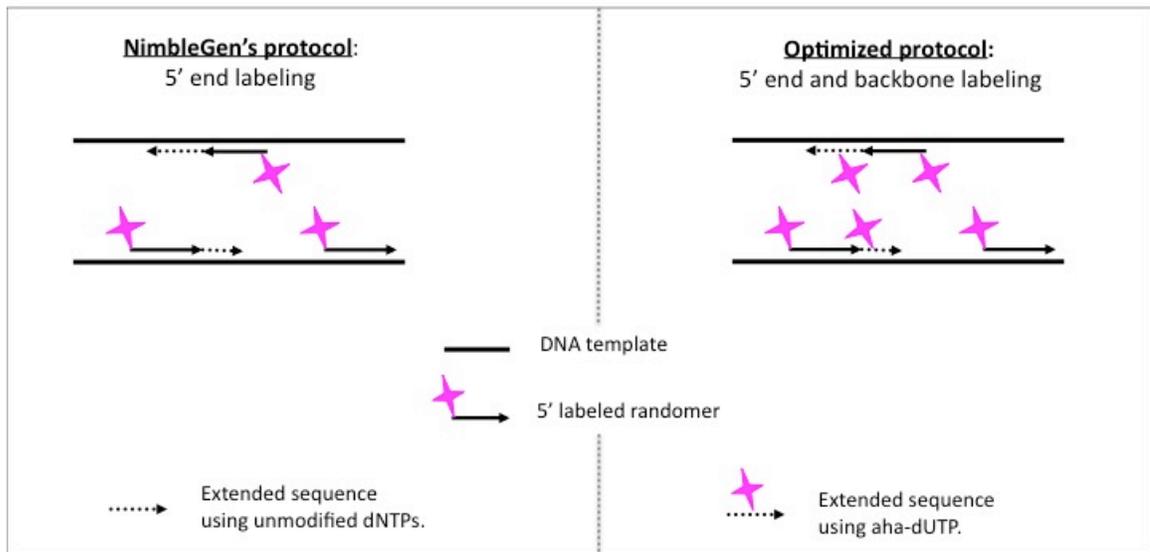
Figure 2.1 – NimbleGen and Optimized DNA labeling strategy.
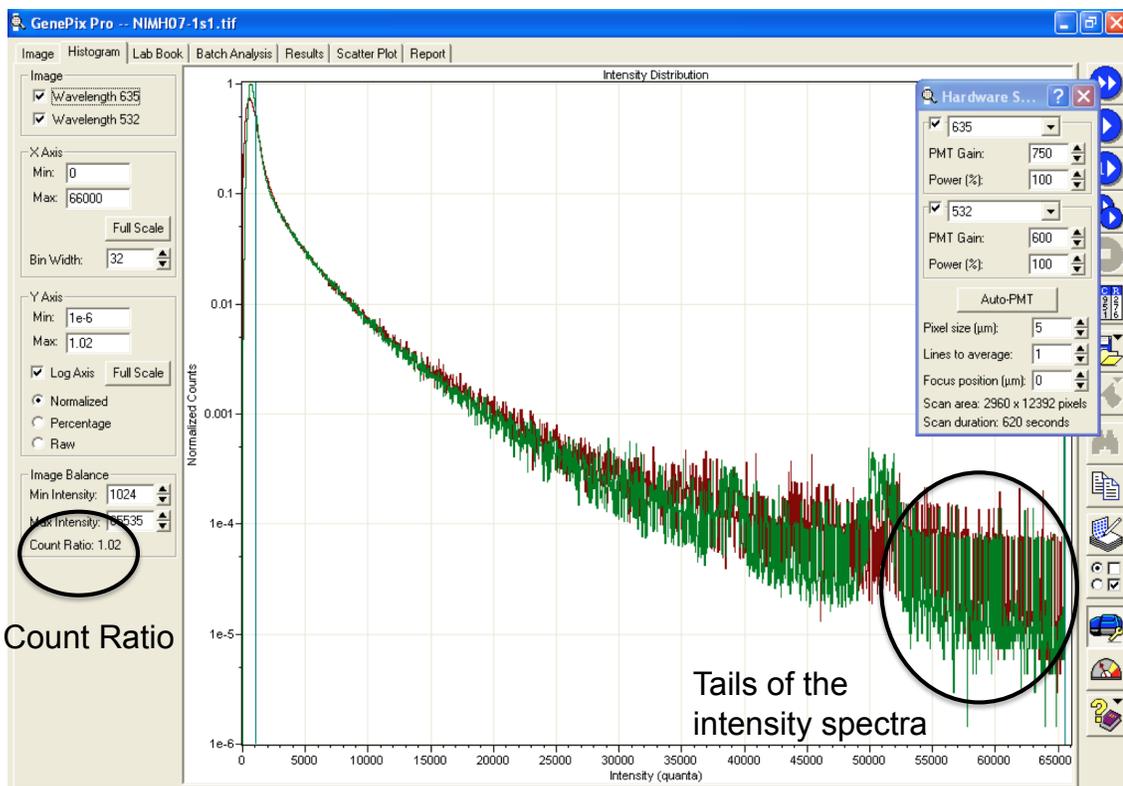
Figure 2.2 – Scanning parameters as informed by the Intensity Distribution Histogram

Log2 Variance of Subarray A01 Probes

Cutoff Variance: > 0.175
Probes Eliminated: 46,266 (6.48%)

0.175

Probe Log2 Variance
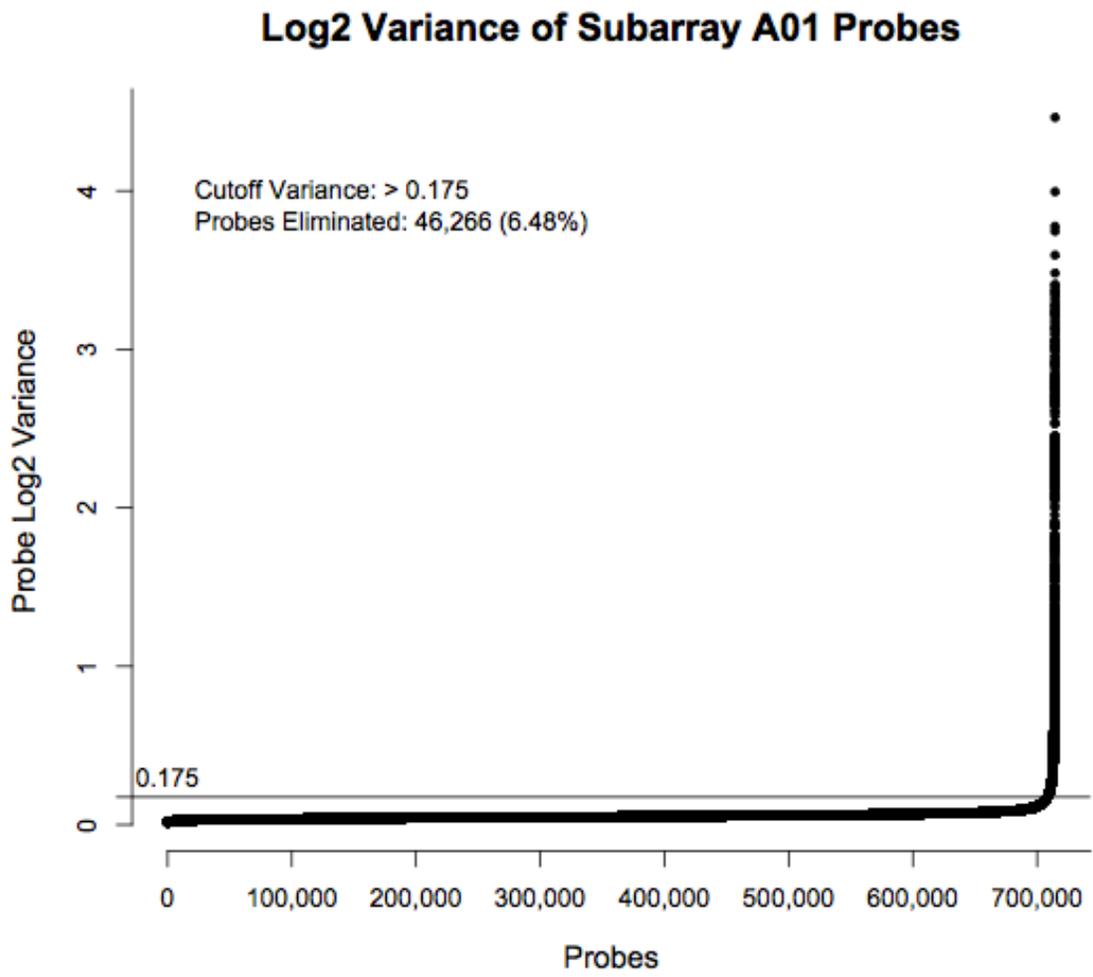
Probes

Figure 2.3.a – Variance Analysis of Subarray A01 Probes

Log2 Variance of Subarray A02 Probes

Cutoff Variance: > 0.175
Probes Eliminated: 38,865 (5.44%)

0.175

Probe Log2 Variance
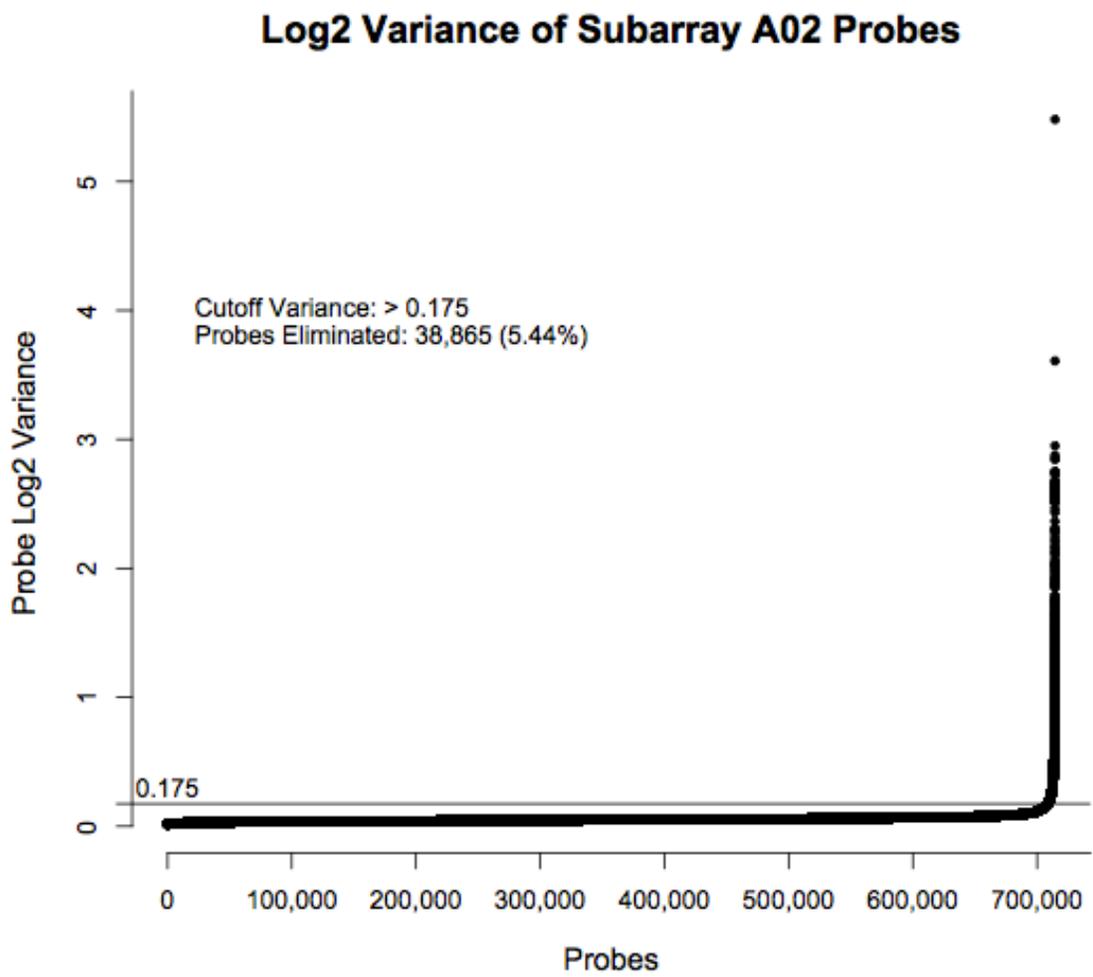
Probes

Figure 2.3.b – Variance Analysis of Subarray A02 Probes

Log2 Variance of Subarray A03 Probes

Cutoff Variance: > 0.175

Probes Eliminated: 37,410 (5.24%)

0.175

Probe Log2 Variance
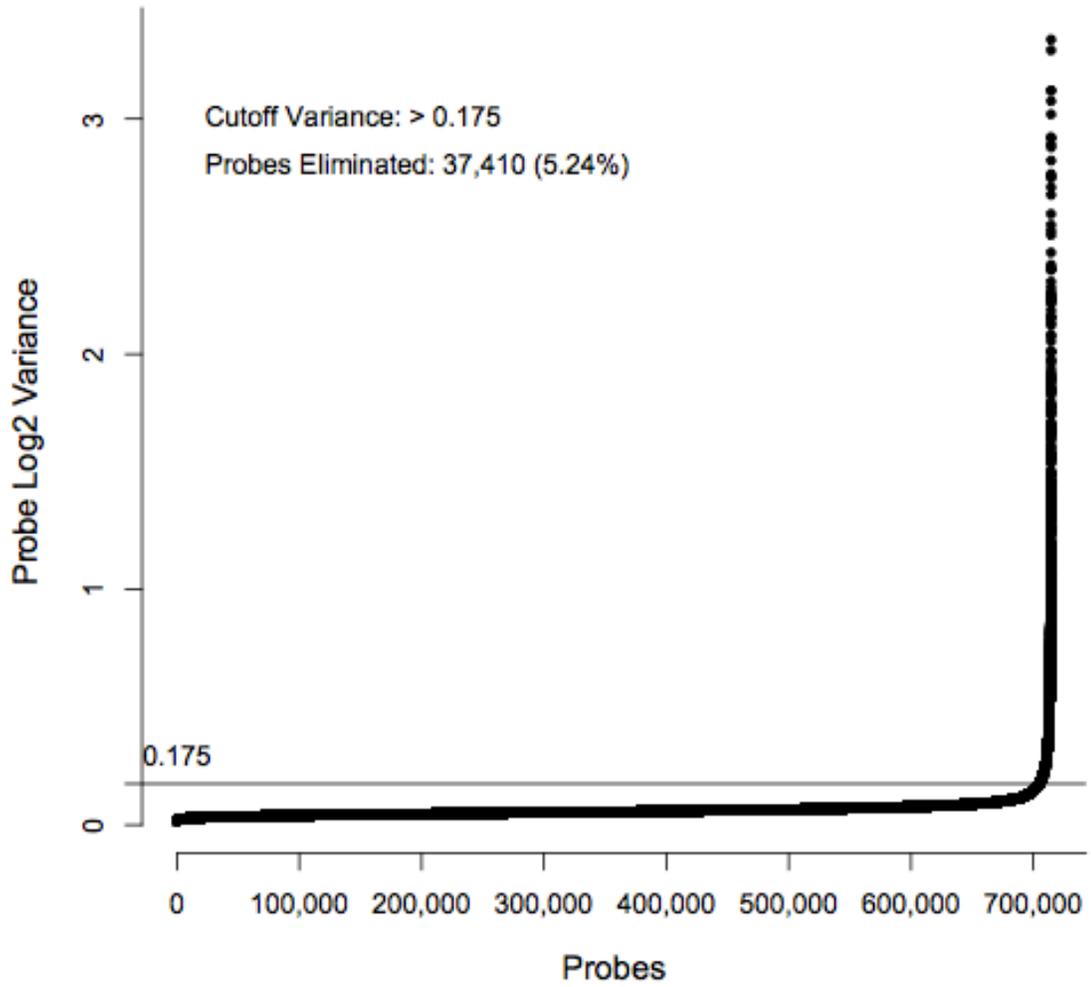
Probes

Figure 2.3.c – Variance Analysis of Subarray A03 Probes

Acceptable MA Plot

slope of fitted line:
0.000807095001647984
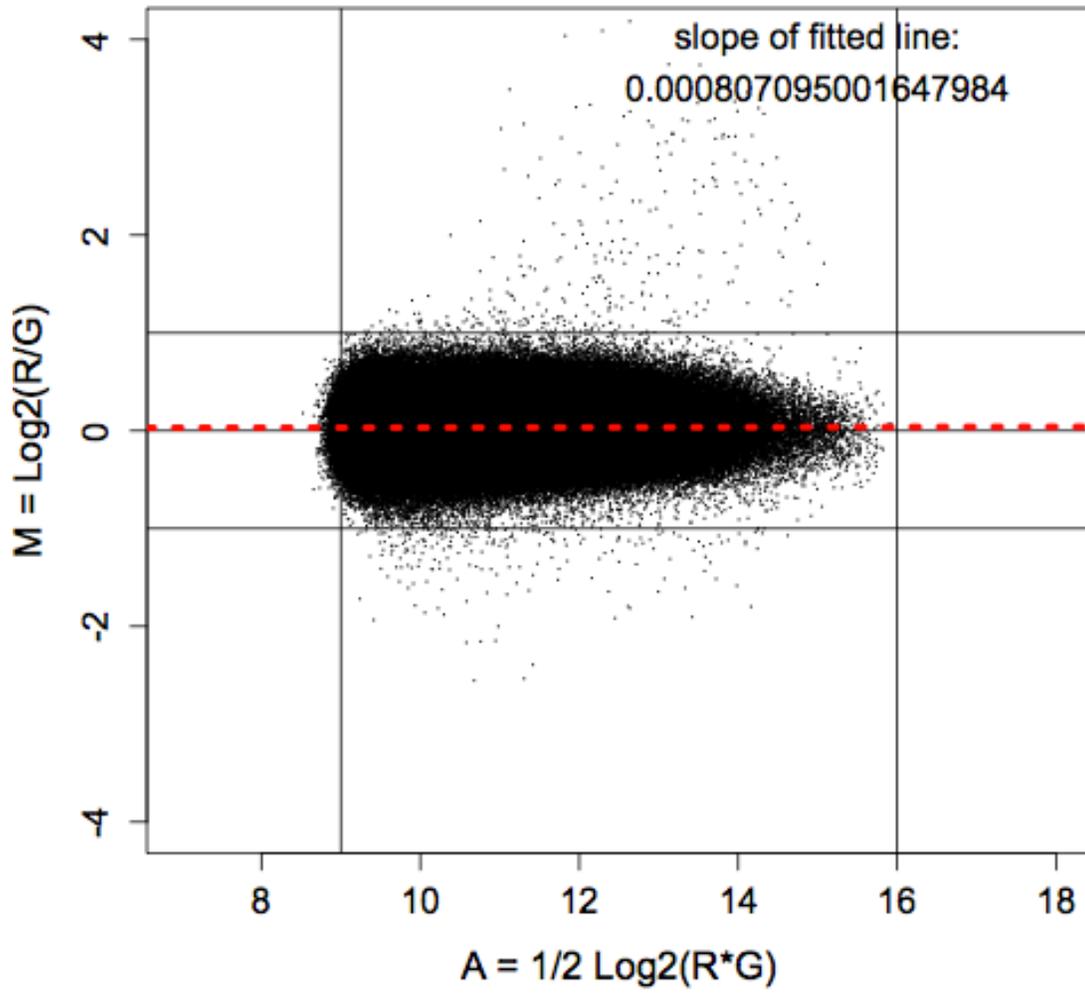
Figure 2.4.a - Acceptable MA Plot

Poor MA Plot

slope of fitted line:
-0.000907393571643005

M = Log2(R/G)

A = 1/2 Log2(R*G)
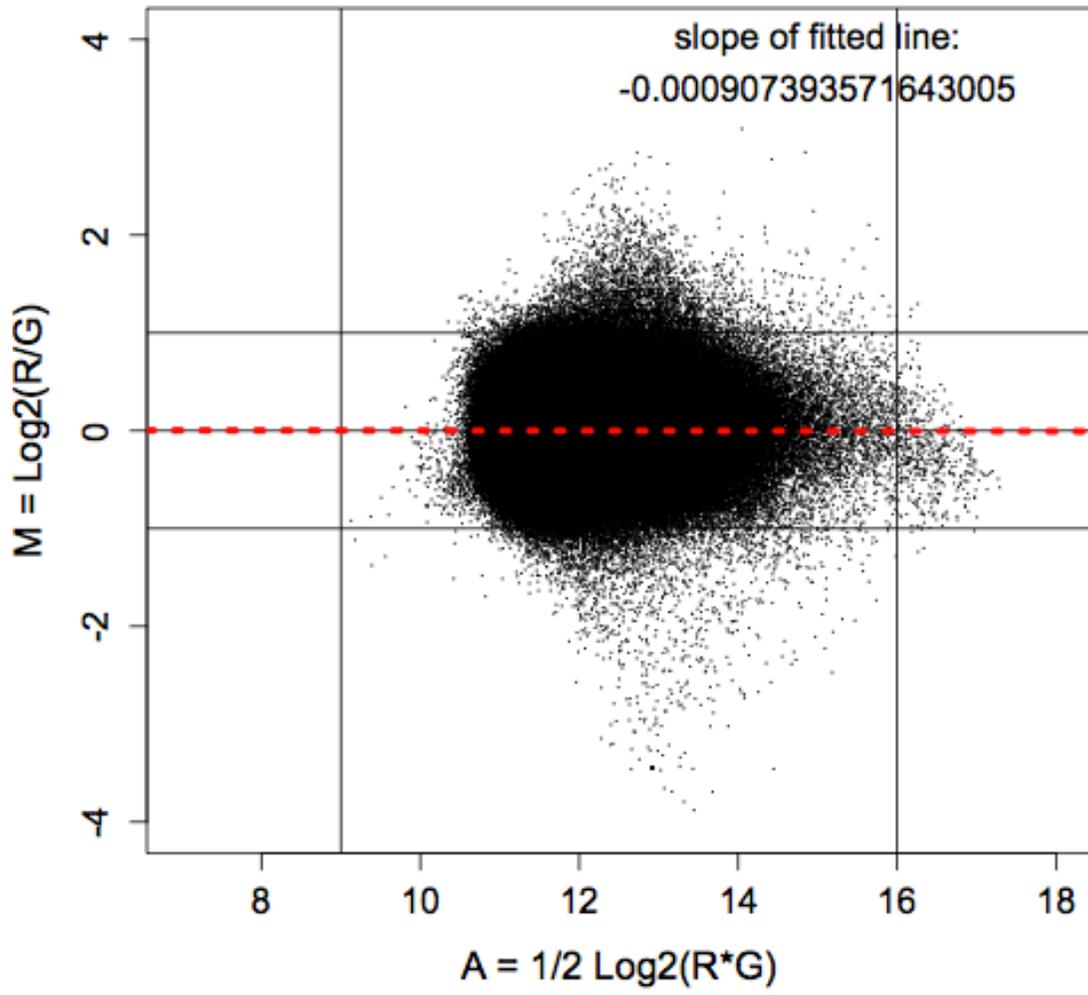
Figure 2.4.b - Poor MA Plot #1
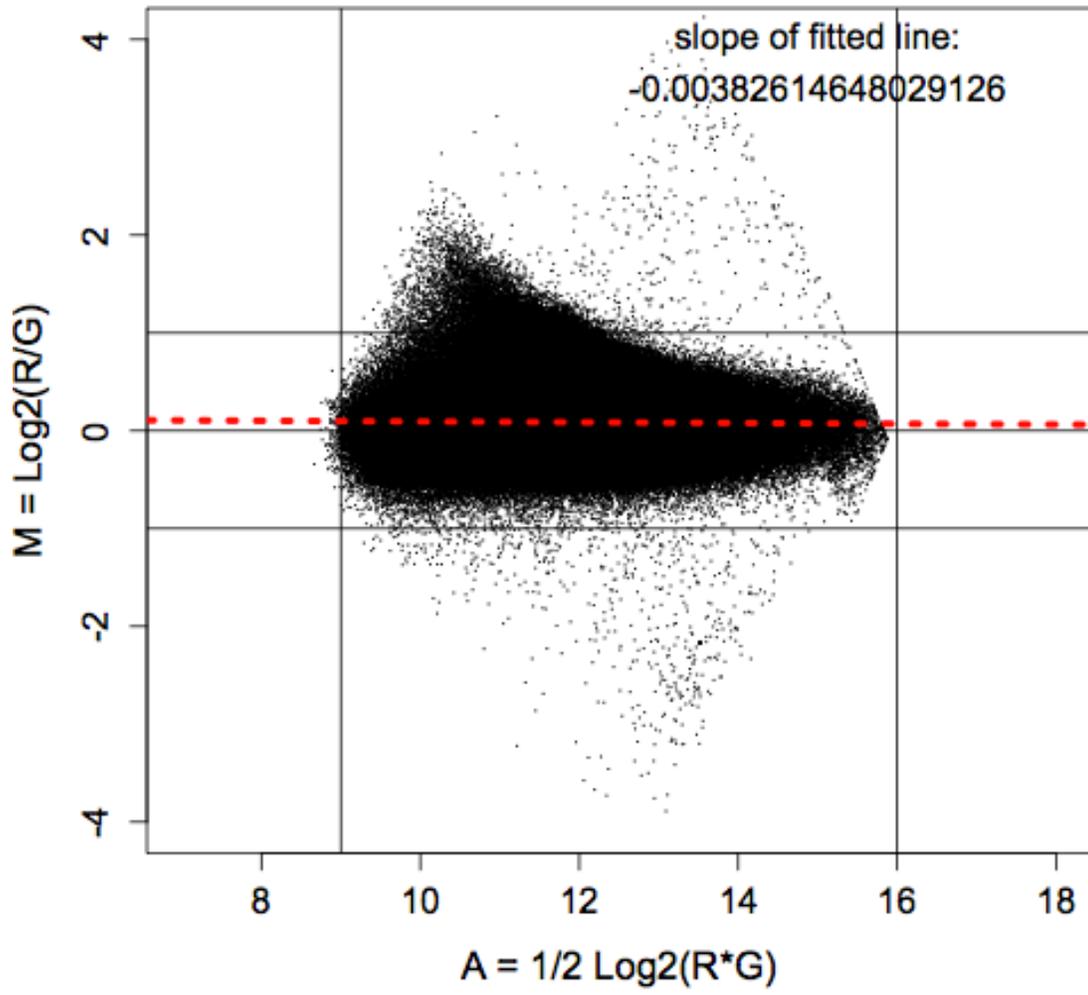
Figure 2.4.c - Poor MA Plot #2
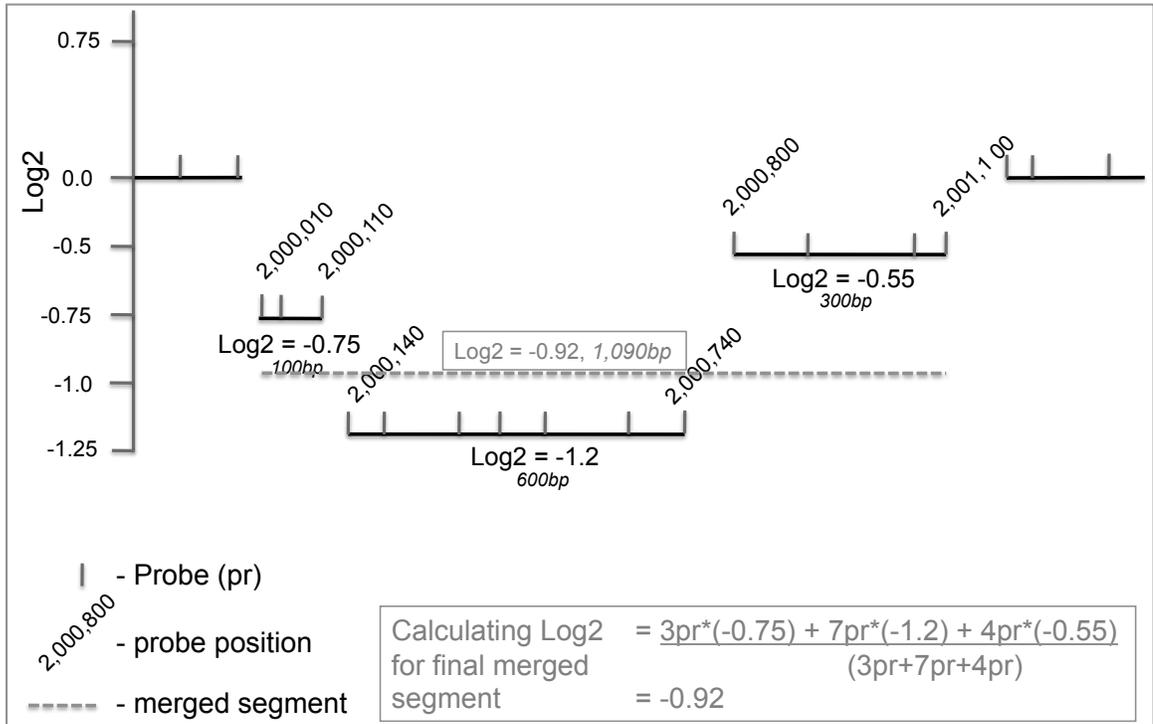
Figure 2.5 - Hypothetical example of merging multiple segments representing a single, deleted locus.

NimbleGen
Protocol

Optimized Protocol
– 'Good' array

Optimized Protocol
– 'Drip' array

CR = 1.02

CR = 0.94
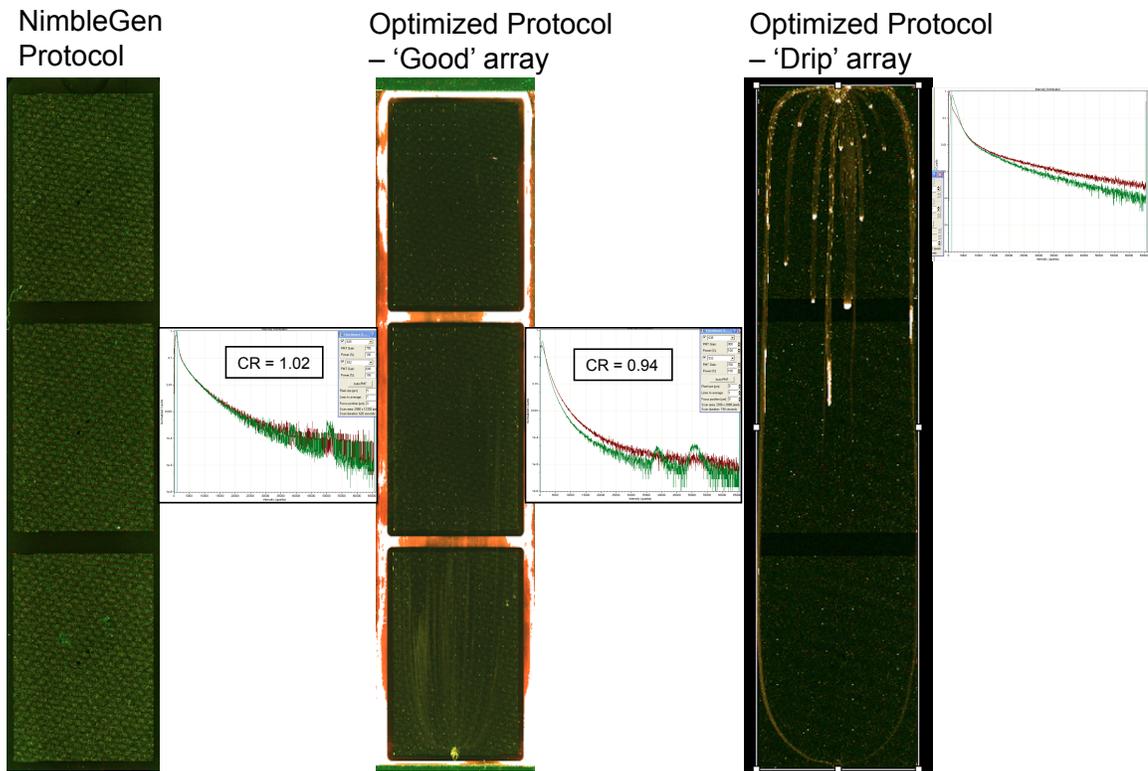
Figure 2.6 – Array performance by NimbleGen and Optimized protocols.

Probe Length: Good and Poor Probes
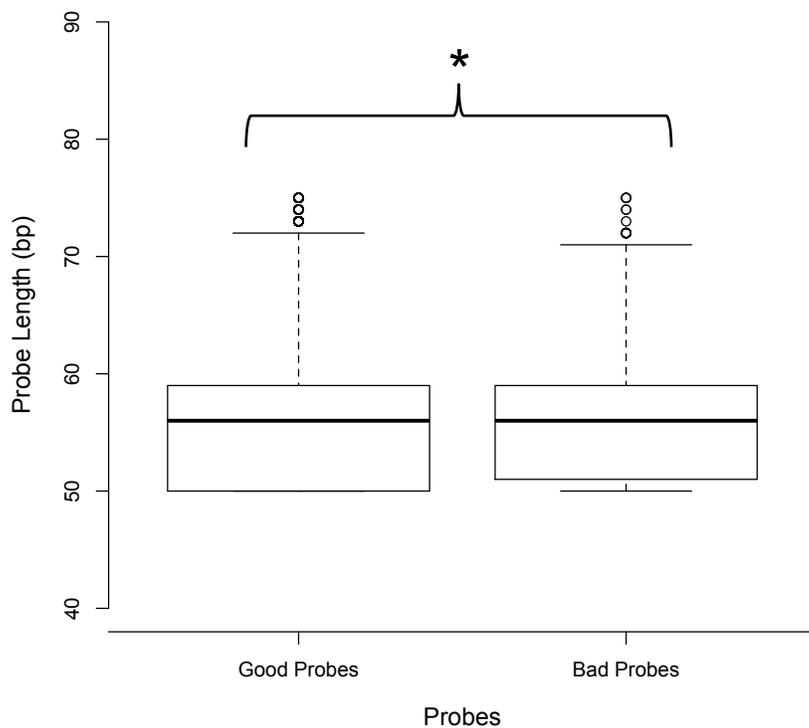
Figure 2.7.a – Boxplots of Good and Bad probes by Length.
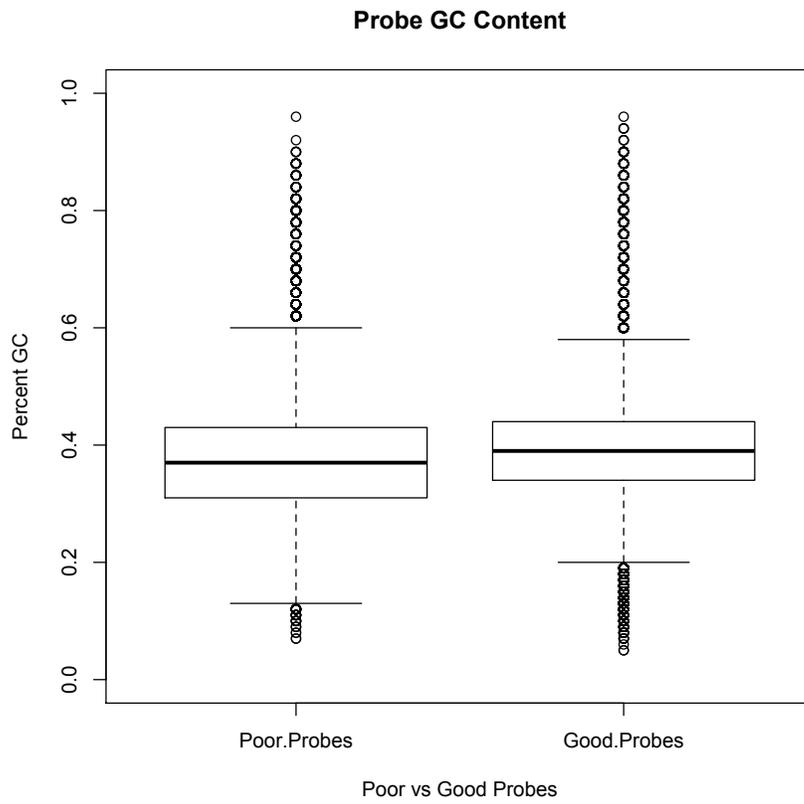
Probe GC Content

Figure 2.7.b – Boxplots of Good and Bad probes by GC Content.

C. Probe AG Content: Good and Poor Probes
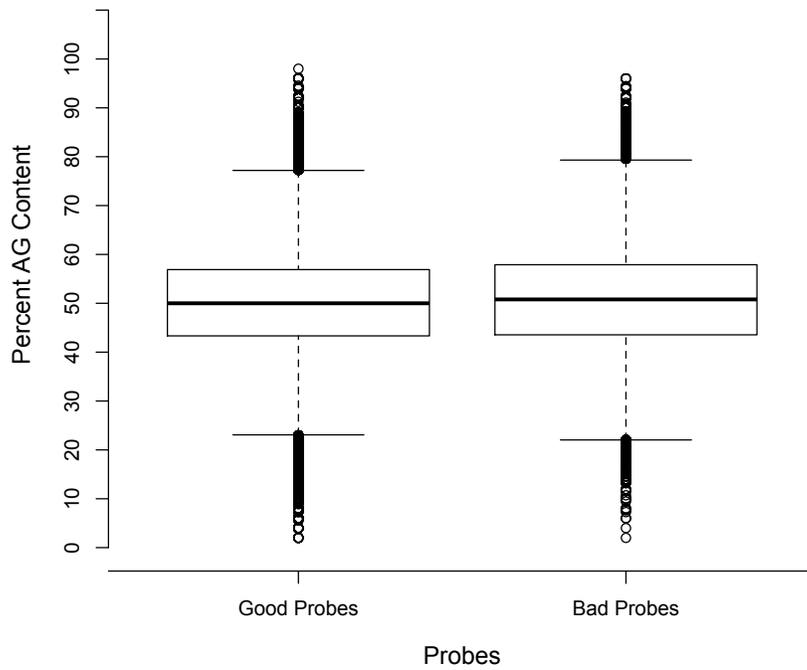
Figure 2.7.c – Boxplots of Good and Bad probes by AG Content.

Probe Tm: Good and Poor Probes
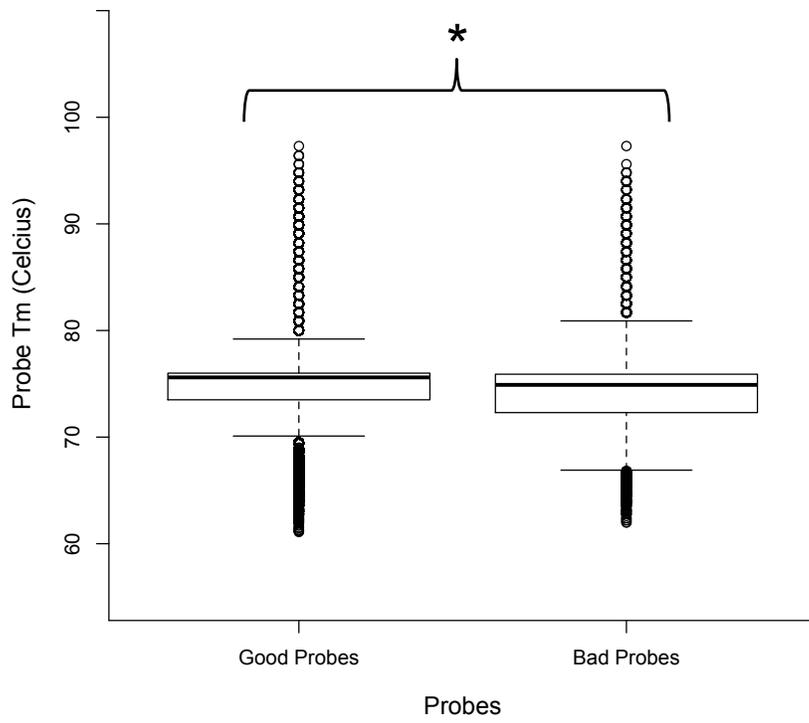
Figure 2.7.d – Boxplots of Good and Bad probes by melting temperature.

B. Deletions

Optimized:
44 deletions

NimbleGen:
184
deletions

20 shared loci
All 20 called deleted by NimbleGen
19/20 called deleted by Optimized

C. Duplications

Optimized
∴
130 dup.

NimbleGen
∴
118 dup.

22 shared loci
21/22 called duplicated by NimbleGen
All 22 called duplicated by Optimized

A. All calls

Optimized:
174 calls

NimbleGen:
302 calls

41 variants called
by both protocols

Figure 2.8 – CNV call overlap of NimbleGen and Optimized protocols.
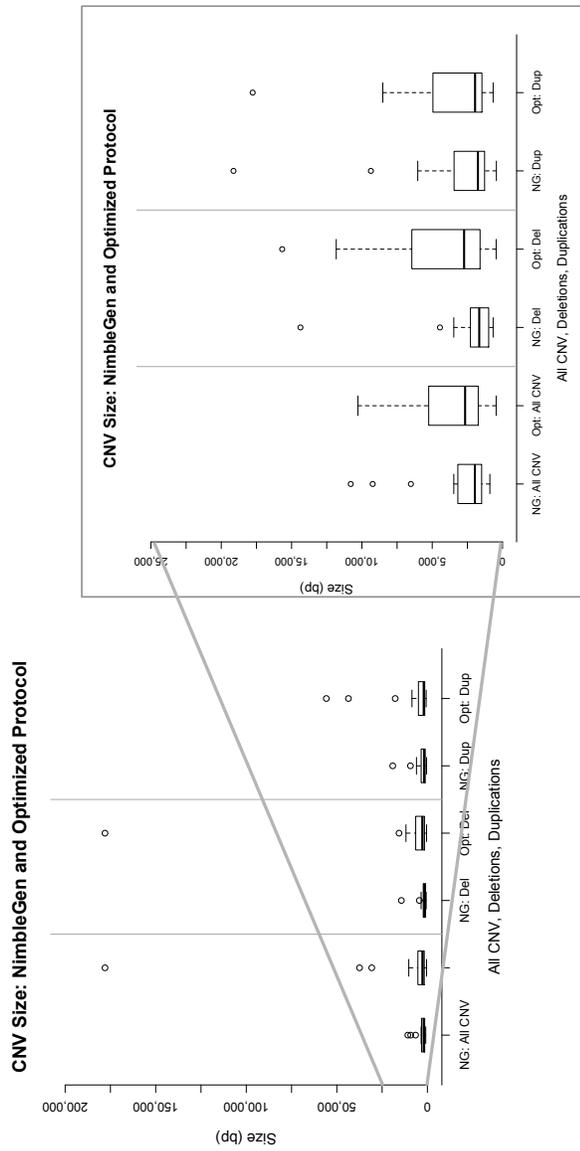
Fig. 2.9.a – CNV size distributions by the NimbleGen and Optimized protocols
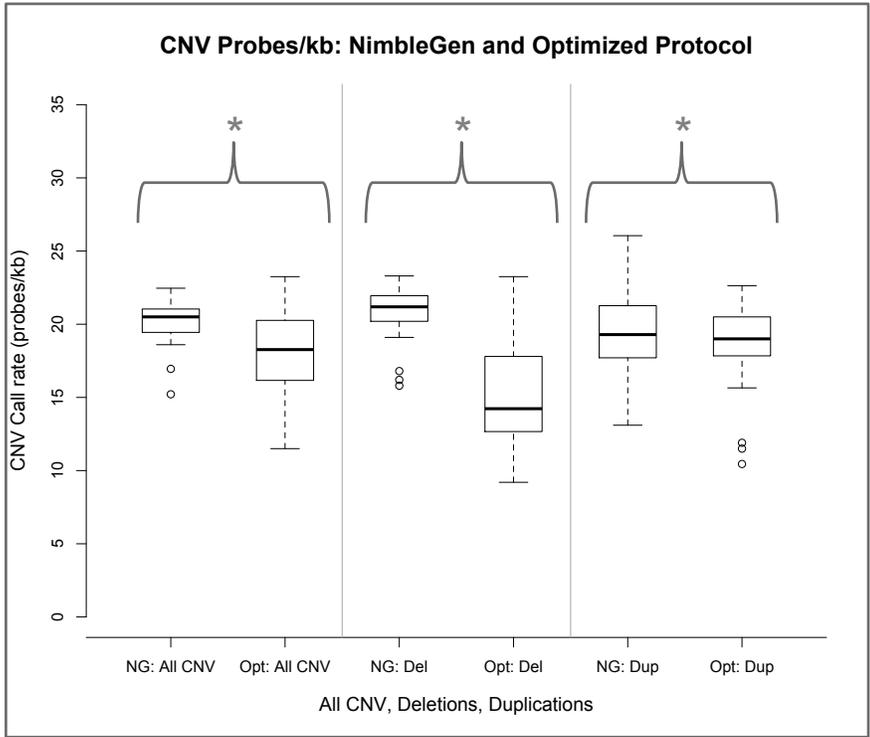
Fig. 2.9.b – CNV Probes/kb by the NimbleGen and Optimized protocols
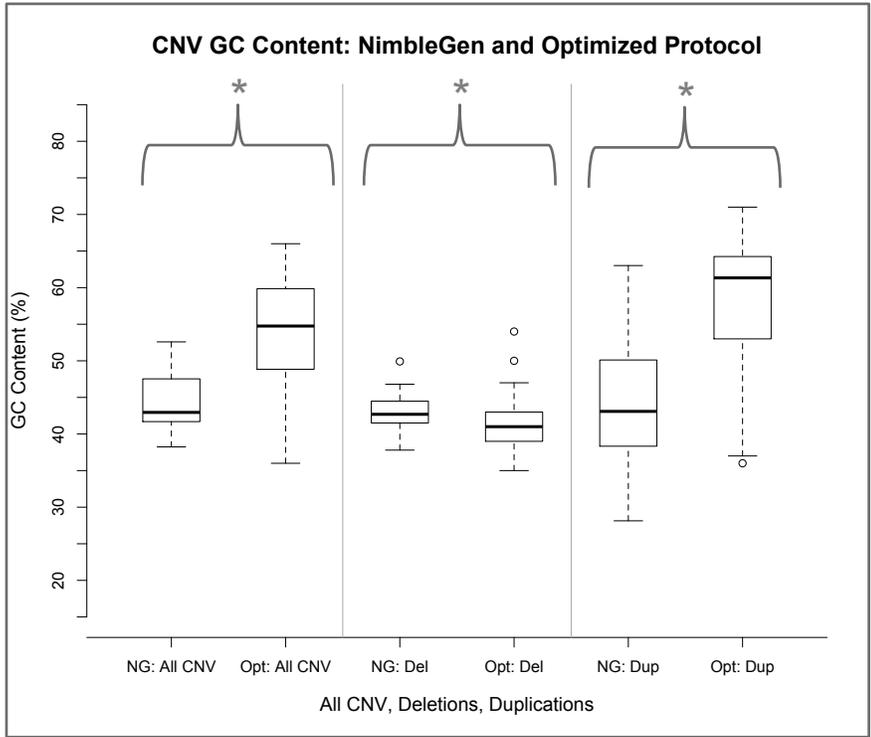
Fig. 2.9.c – CNV GC content by the NimbleGen and Optimized protocols

A. All calls – True Calls

15 TP loci by both protocols

NimbleGen: 22/62 TP

Optimized: 26/62 TP

B. All calls – False Calls

NimbleGen: 40/62 FP

Optimized: 36/62 FP

7 FP loci by both protocols

TP: True Positive Calls
FP: False Positive Calls
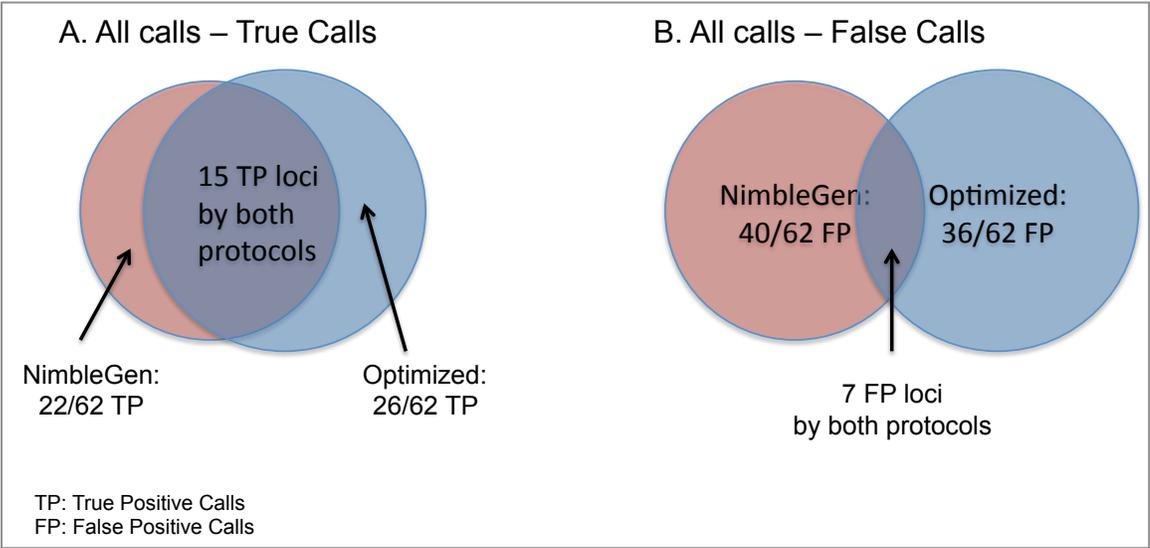
Figure 2.10 – Overlap of true and false CNV calls between protocols
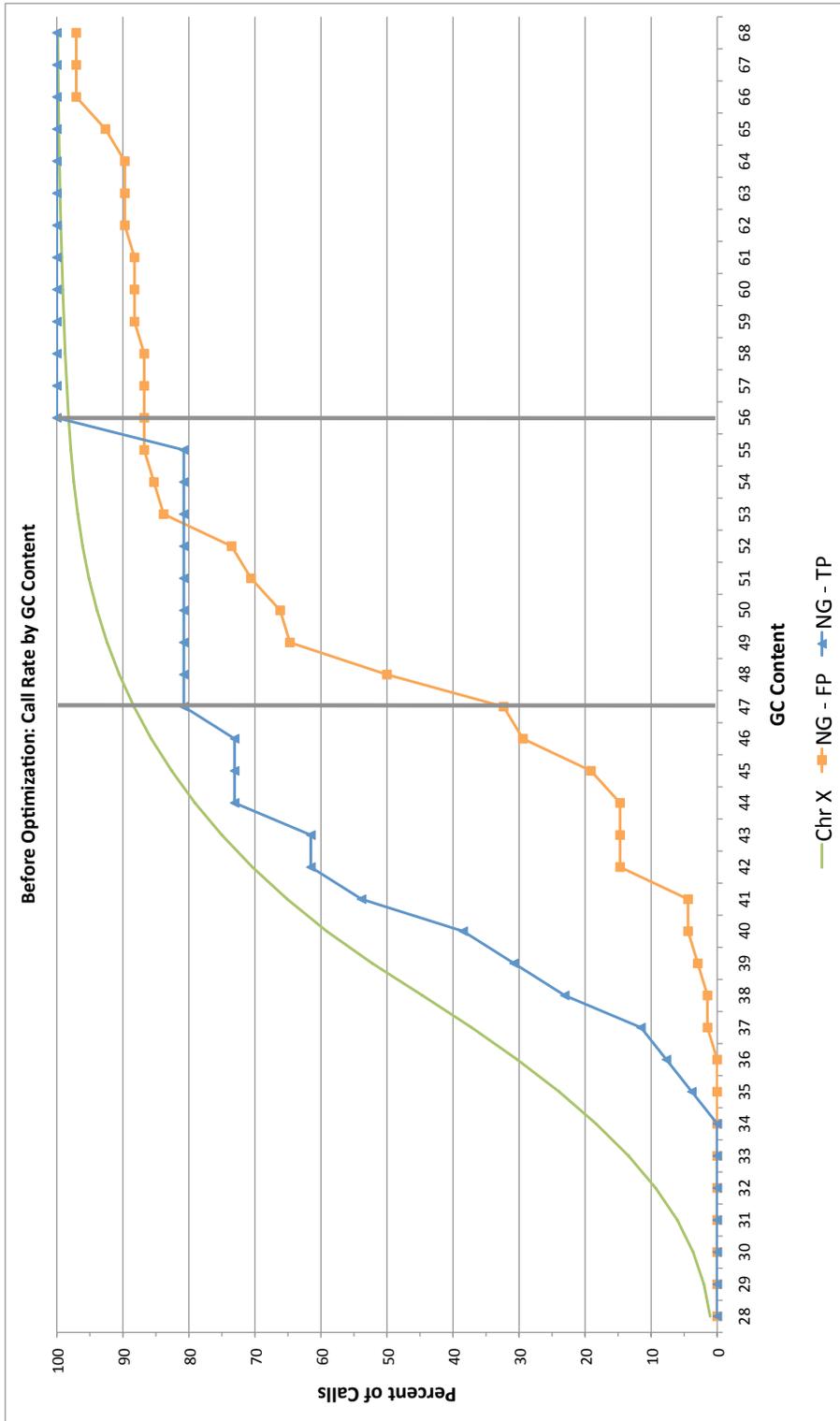
Figure 2.11.a – NimbleGen Protocol: Proportion of true and false calls plotted by GC
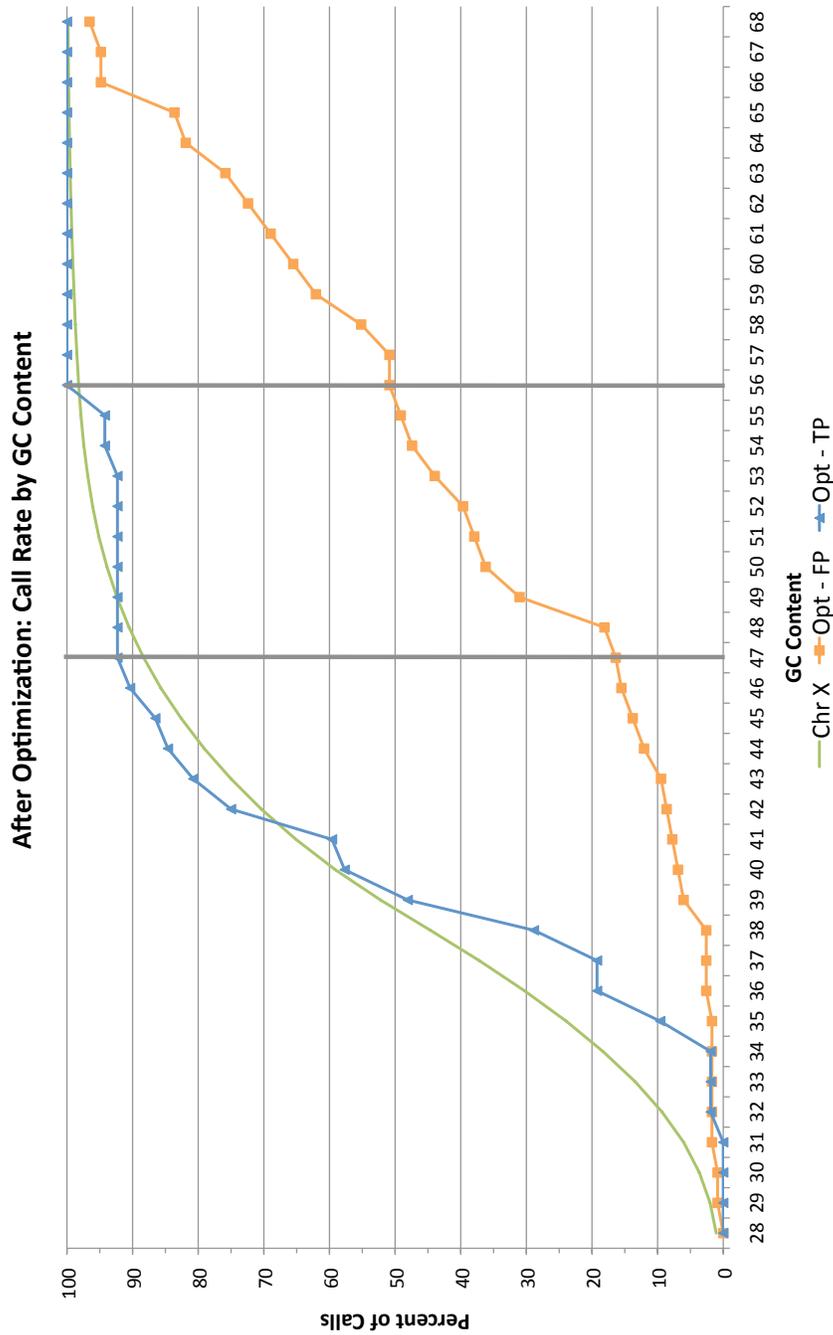
Content

Figure 2.11.b - After Optimization: Proportion of true and false calls plotted by GC

Content

Chapter 3. X-Chromsome Copy Number Variation and Breakpoint Analysis

1. Introduction to Copy Number Variation on the X Chromosome

We hypothesized that the X chromosome bears susceptibility loci for the autism spectrum

disorders (ASD). With over 40 genomic structural changes located throughout the

genome already identified as playing a role in ASD, we speculated that smaller changes

(less than one megabase) could similarly disrupt autism susceptibility loci on the X

chromosome. Capitalizing on advances in oligonucleotide-array technology that enable

identification of fine-scale structural variation (500 base pairs (bp) in size or greater), we

developed a study design to screen the X chromosome at high resolution in males with

ASD. We studied several populations of samples using two array platforms and two array

protocols to identify copy number variation on the X chromosome. This chapter describes

our copy number findings, validation efforts, functional studies, and analysis of structural

breaks.

Over the course of this project, we generated copy number variant (CNV, sequence that

varies in copy number from the reference one kb in size or greater, can be benign or

pathogenic) and indel (sequence that varies in copy number from the reference 100bp-

1kb in size, can be benign or pathogenic) data from four different combinations of

samples, two different array densities, and two different array-processing protocols . The

first, and most brief, screened four samples from the Autism Genetic Resource Exchange

(AGRE, multiplex families) by the 385K Comparative Genomic Hybridization (CGH)

array following the manufacturer's (NimbleGen) protocol. This array probes the X

chromosome with 385,000 oligonucleotide probes resulting in an intermarker distance of approximately one probe ever 200 bp (four probes/kilobase (kb)). Next, the platform was upgraded to the 2.1M CGH array harnessing 2,100,000 million probes to interrogate the X chromosome. Data from 50 AGRE samples were generated. Early validation efforts suggested a high false call rate (65%), so we developed an optimized protocol in effort to reduce our false positive and false negative calls. Our third set of CNV and indel was derived from 100 AGRE samples and 64 samples from the Simons Simplex Collection (SSC, simplex families) screened by the 2.1M CGH array using our optimized protocol. Finally, we assayed 100 male controls from the National Institute of Mental Health (NIMH) control population ascertained for the Human Genetics Initiative by the 2.1M CGH array using NimbleGen's protocol.

2. Copy Number Changes Identified by array CGH

Before discussing the specifics of our findings, there are two characteristics to appreciate about our data. First, we describe our CNV data set in two different ways. The first is a summary of all CNV identified within a group, and the second is a summary of 'distinct' CNV characteristics. Studies of CNV will often make a distinction between 'singletons' (those loci found only in one individual and no others) and 'copy number variant regions' (CNVR, a locus that is found to be variant in more than one individual and if boundaries are not shared, percent overlap is defined by the authors). We identified both singletons and CNVR in our CNV screens of the X chromosome, and we assessed the nature, size, and gene content of those sequences. However, we defined a 'distinct locus' as sequence found to be copy number variant in one or more individuals. This grouping of singleton

and CNVR allowed us to describe and characterize all copy number variant loci on the X chromosome as a whole. We defined our CNVR as having upstream and/or downstream breaks ending within three kb of one another or if one variant completely encompasses another; most often, ends were within one kb.

A second item to note is our characterization of CNV size. While the mean and standard deviation (SD) for the size of all and distinct CNV were calculated, this estimate of central tendency can be misleading. The majority of our data are quite small, but the presence of large variants within data sets inflates both the mean and SD estimates. These will be listed in the summary tables, but discussions of CNV or indel size will be based on the median values.

a. Study One: Four individuals from AGRE on the 385K array using the NimbleGen protocol

Using the 385K CGH arrays and following NimbleGen's protocol, we identified 16 CNV in four samples from the AGRE cohort (4.0 variants/person). The median size for these CNV was 3,263 bp. There were half as many deletions as duplications (0.45:1, five deletions, 11 duplications), and only one variant, a 149 kilobase (kb) duplication, involved a gene, *melanoma antigen family A, 8* (*MAGEA8*). (Table 3.1.a, Appendix: Table A.1)

The 16 CNV represented nine distinct loci in these four individuals. The median size of the distinct loci was 4,007 bp. Deletions and duplications were nearly equal in

representation with four deleted-only and four duplicated-only loci. No loci showed both

a deletion and a duplication event. All CNV in this data set overlap at least four or more

variants from the Database of Genomic Variants (DGV). In fact, 90 reported CNV and

indels intersect with our 16 CNV. (Table 3.1.b, Appendix: Table A.1)

b. Study Two: Fifty individuals from AGRE on the 2.1M array using the

NimbleGen protocol

Our next effort at screening the X chromosome for CNV and indels utilized the 2.1M

CGH array (20 probes/kb) to evaluate 50 AGRE samples. We were able to generate data

from 48 of these individuals, and 581 CNV and indels were identified (12.1

variants/person). Deletions accounted for 65% (379 CNV) of all CNV identified and

duplications the remaining 35% (202 CNV and indels). The deletion:duplication ratio

was 1.88:1. The median size for all CNV and indels was 1,080 bp. Overall, deleted loci

were smaller than duplicated loci with the median deleted size being 1,005 bp and the

median duplicated size 1,419 bp. Two hundred-sixty CNV and indels overlapped coding

or intragenic sequence. Proportional to the number of deletions in the data set overall,

deletions in coding or intragenic sequence accounted for 63% of all such CNV and indels

and duplications the remaining 37%. Many CNV and indels in our data set overlapped

with reported variants in the DGV. Of the 581 CNV and indels we identified, 420

overlapped with 509 previously reported variants while the remaining 161 CNV and

indels did not intersect with any reported DGV variants. (Table 3.1.a, Appendix: Table

A.2)

The 581 CNV and indel identified in these 50 AGRE samples by the NimbleGen protocol represented 282 distinct loci. There were 153 deleted-only loci (63%), 82 duplicated-only loci (34%), and seven loci showing both deletions and duplications (3%). For all categories of distinct loci, the median size was about 1,000 bp (all distinct: 1,152 bp; deletion-only: 994 bp; duplication-only: 1,443 bp; deleted and duplicated loci: 1,355 bp). A little less than half (112/282 distinct CNV loci, 46%) of the distinct variant loci overlapped coding or intragenic sequence. Fifty deleted-only loci (21%), 50 duplicated-only loci (21%) and 10 deleted and duplicated loci (4% of all distinct) involved a gene structure. (Table 3.1.b, Appendix: Table A.2)

c. Study Three: 100 affected males from AGRE and 64 affected males from SSC on the 2.1M array using the Optimized protocol

Using the more stringent protocol we developed and the 2.1M aCGH, we screened 164 samples from the AGRE (n=100) and SSC (n=64) cohorts for CNV on chromosome X. Of these samples, 149 (91%) passed array and data quality control measures, and 517 CNV were identified (3.5 variants/person). Nearly equal representation of deletions and duplications were observed (deletion:duplication = 0.88:1, 242 deletions (46.8%), 275 duplications (53.2%)). Unlike the CNV and indels identified by the NimbleGen protocol, under the conditions of our altered protocol, most CNV and indels in our data set are nearly the same size irrespective of state (median size all: 1,719 bp: median size deletions: 1,694 bp; median size duplications: 1,724). Sixty-one percent (313 CNV) of all CNV and indels overlapped genes with 108 deletions (21%) and 205 duplications (40%) involving coding or intragenic sequence. (Table 3.1.a, Appendix: Table A.3)

The 517 CNV and indels identified in our 164 individuals represent 244 distinct variant loci. Interestingly, similar rates of deleted-only CNV and duplicated-only CNV were observed. 118 deleted-only variants (48% of distinct), 106 duplicated-only (43% of distinct), and 20 both deleted and duplicated variants (8% of distinct) were found. Median size values suggest all distinct loci are about the same size at about 2,000 bp. The median size for all distinct loci is 1,951 bp, for deletion only: 1,608 bp, for duplication-only: 2,063 bp, and deleted and duplicated: 4,203 bp. About half (128 distinct CNV, 53%) of these distinct CNV intersect coding or intragenic sequence with duplications accounting for the majority (73 distinct CNV, 30%), and deleted-only (41 distinct CNV, 17%) and deleted and duplicated (14 distinct CNV, 6%) the remainder. (Table 3.1.b, Appendix: Table A.3**)**

d. Study Four: 100 unaffected males from NIMH on the 2.1M array using the NimbleGen protocol

Due to our agreement with Roche NimbleGen, we ran the 100 NIMH control samples following the NimbleGen protocol on the 2.1M arrays. Before running all 100, we first ran two samples in a proof-of-performance in our hands for Roche NimbleGen before proceeding with the remaining samples. Our analysis of 102 male samples yielded data from 101 arrays and identified 1,231 variants (12.2 variants/sample). Distinctly different from the two previous large studies of the AGRE samples, duplications comprised the majority of CNV (775 CNV, 63%) over deletions (456 CNV, 37%), and the deletion:duplication ratio was 0.59. Additionally, the observed CNV sizes in this data set

were larger, on average, than those observed in our earlier findings. The median size of all CNV is 2,015 bp, for deletions 933 bp, and for duplications 4,062 bp. Nearly half (634/1231 or 54%) of the CNV identified overlapped with 272 different genes. Consistent with an increase in observed duplications, 409 duplications overlapped with a gene while 225 deletions intersected a gene. Nearly three-fourths (851 of 1,231 or 69% CNV and indels) overlap with reported variants from the DGV. (Table 3.1.a, Appendix: Table A.4)

The 1,231 CNV identified in the NIMH control population represented 462 distinct regions where 155 distinct regions (34%) were deleted, 285 distinct regions (62%) were duplicated, and 22 distinct loci (5%) were deleted and duplicated. The median sizes for the distinct loci were more similar to those previously observed with 2,221 bp, 1,375 bp, 2,935 bp, and 6,435 bp for all, deleted-only, duplicated-only, and deleted and duplicated only sequences. Again, about half (216 or 49%) of the distinct loci overlapped with coding or intragenic sequence. Duplication-only events comprised most of these loci at 136, 60 deletion-only events and 20 deleted and duplicated loci overlapping gene structures. Of the 426 distinct loci, 241 overlapped with 664 CNV and indels reported in DGV. (Table 3.1.b, Appendix: Table A.4)

Emory CNV Overlap with Reports from the Database of Genomic Variants

While the sex chromosomes do not bear similar rates of reported copy number events as compared to the autosomes (Figure 3.1) in the Database of Genomic Variants (DGV, Hg18), 75% of those studies reporting variants in this database have reported structural

changes for the X chromosome. Currently, 3,268 CNV and indels from over 4,600

individuals (male and female) screened have been reported for chromosome X. These

data suggest that the X chromosome in the normal population has less than 0.7 copy

number changes on chromosome X per person. The exact value is difficult to calculate as

not all CNV and indels are identified as having been discovered in a male or female. We

observed a five-fold increase (3.5 CNV/chromosome X) among those samples derived

from the AGRE and SSC cohorts and a 17-fold increase (12.2 CNV/chromosome X) for

those samples derived from the NIMH cohort. The median size of deletions, duplications

or inversions reported on chromosome X in DGV is 2,013 bp. These reports are similar

to what we observed in the NIMH samples (median size: 2,105bp), but somewhat larger

than our observations of the AGRE and SSC cohorts (median size: 1,719 bp). The pattern

of size distribution, where a majority of calls are 'small' and larger CNV represent a

minority of calls, is similar when we compare the DGV and our own findings. However,

the majority of CNV identified among the AGRE and SSC samples, or the NIMH

samples are significantly smaller than those reported variants in DGV. Figure 3.2a-e

3. Validated Copy Number Changes

Our first pass selection of CNV for validation studies targeted possible candidate genes.

Our strategy was to select structurally altered genes with 1) a known disease association

(e.g., X-linked Mental Retardation), 2) a known molecular function that could plausibly

be involved with autistic disorder, or 3) found to be expressed in the brain. In our screen

of the AGRE and SSC cohorts, we identified 152 genes that intersected our array

identified CNV and indels. The most compelling candidates genes were derived from the

disease-associated; we found 16 (*ARHGEF9, ARX, CUL4B, DLG3, DMD, FTSJ1, JARID1C/KDM5C, KDM6A/UTX, KIAA2022, MAOA, PAK3, PHF6, RPS6KA3, SLC6A8, SYP, ZNF41*) to be disrupted. Eleven of these 16 were selected for their role in non-syndromic X linked Mental Retardation (*ARX, CUL4B, DLG3, FTSJ1, JARID1C/KDM5C, KIAA2022, PAK3, RPS6KA3, SLC6A8, SYP, and ZNF41*). We identified seven genes (*BCAP31, FRMPD4/KIAA0316, GABRQ, GPR173, TCEAL2, TMEM47, TREX2, HAUS7*) with a compelling cellular role such as functioning in neuronal processes (*FRMPD4* and *GABRQ*). And 50 genes (*ARHGEF9, ARX, BCOR, BGN, CUL4B, DDX26B, DDX3X, DLG3, DMD, ELK1, F8A1, FAM122B, FGF13, FOXO4, GPRASP2, GPRASP2, GRIPAP1, H2AFB3, HEPH, IDH3G, IQSEC2, KIAA2022, LAMP2, LOC550643, MID1IP1, MTMR1, NAP1L2, OTUD5, PCYT1B, PDZD4, PGRMC1, PHF6, PPP1R3F, PRKX, RBBP7, SH3KBP1, SHROOM2, SLC38A5, SLC6A8, SLITRK4, SSR4, STAG2, STS, SYP, TBL1X, TCEAL2, TMEM164, TMEM47, UBA1, USP9X*) were identified as being expressed in the brain. Of these 50, eight were identified from the disease-associated category and two from the cellular-function category.

a. Characteristics of validation assays developed for CNV identified in AGRE and SSC cohorts

Using PCR based strategies to capture deletion and duplication junction breakpoints, we ultimately developed validation assays for 105 distinct loci representing 179 CNV in the AGRE and SSC cohorts. Specifically, for each working assay, an amplicon of an expected variant size would confirm or refute the validity of the array identified copy

number change. For deletions, the expected amplicon would be smaller in size relative to intact sequece (i.e., reference and control DNAs); for duplications, the expected amplicon would be larger in size than single copy, reference or control DNAs, or a junction amplicon generated where no amplicon is expected from a single copy. It is important to note that this PCR based strategy for validation of duplications assumes that the duplicated sequence is located in tandem with the original. While most duplicated sequence appears to occur in tandem [46, 55], it is possible for the 'duplicated' sequence to exist elsewhere in the genome.

For these 105 distinct loci, 39 loci (37%) representing 58 CNV validated while the remaining 66 loci (63%) representing 121 CNV proved to be false positive. A majority of the true calls were deletion events (35 or 90%) while a somewhat more even distribution of deletions and duplications were found to be false (26 deleted (39%) and 35 duplicated (53%), the remaining five deleted or duplicated (8%)). The false call rate (failure to validate an array identified variant sequence) for deletions was 26 or 43% and for duplications was 35 or 90%. Additionally, those variants that validated tended to be larger with a median size of 2,343 bp versus 1,370 bp for variants that did not validate. (Table 3.2)

   b. Inheritance and Population Data for validated CNV

We assessed 33 of the 39 validated loci for maternal inheritance, the remaining six were not tested for inheritance. Of these, 32 CNV or indel was also identified in the mother. A 14 kb deletion was the only *de novo* structural change identified. This locus does not

overlap any genes but has a few semi-conserved sequences and is located 174 kb downstream of *gastrin-releasing peptide receptor* (*GRPR*, found to be disrupted in an individual with autism).

Genotyping in an extended autistic population (200 additional, unrelated AGRE samples) as well normal control populations (300 unrelated fathers from the AGRE cohort, 1500 males from the NIMH cohort) was begun for those CNV that validated. Two CNV, AU016.3 and AU122.1 were completely genotyped on all autistic samples (n = 364, 300 AGRE and 64 SSC) and control samples (300 AGRE fathers, 1,500 NIMH males). The first CNV, AU016.3, is a deletion approximately eight kb in size located at chrX: 16,249,363 - 16,266,863 (Hg18); sequencing is not complete and exact coordinates not yet determined. Our observed frequencies are 2/364 (0.5%) in our autistic samples and 2/1800 (0.1%) in our controls. These frequencies are not statistically different with $p < 0.2684$. Our second CNV, AU122.1, is the *GRIA3* 561 bp promoter deletion at chrX: 122,143,696 – 122,144,257 (Hg18). We observed 6/364 (1.6%) and 20/1799 (1.1%) for our autistic and normal populations. Unfortunately, these frequencies were also not statistically different with $p < 0.5531$. Additionally, genotyping has begun for seven additional CNV (AU015.2, AU033.95, AU065.4, AU082.1, AU096.5, AU113.2, and AU122.1). (Table 3.3)

c. Candidate genes that have been validated and remain to be validated

As noted previously, we initially identified 17 loci as having compelling molecular and disease related functions that prioritized these loci for validation. Eight of the 17 loci had

validation assays developed for them though seven of the eight proved to be false calls.

The remaining nine loci are *ARX, GABRQ, KDM6A/UTX, KIAA2022, MAOA, SYP,*

TMEM47, and *ZNF41*. Four genes, *GABRQ, MAOA, SYP,* and *ZNF41* have the

compelling feature of having GC content less than 53%. Based on our validation rates, a

greater than 50% chance remains that each will be a false call, but the potential exists that

these variants will validate true. (Figure 3.4)

The only validated CNV located in a gene was a deletion event in the first intron of the

FERM and PDZ domain containing 4 gene (*FRMPD4)*. This gene was first identified

from a large protein cDNA library of human brain. *Lee et* al demonstrated the gene

product's involvement with the maintenance of excitatory synaptic transmission as well

as its interactions with PSD95, a postsynaptic density protein.[115] Our 2.5 kb deletion

lies 277 kb downstream of the first exon and about 80 kb upstream of the second exon

within chrX: 12,344,184 – 12,346,734. One hypothesis might suggest that the deletion

creates an alternative splice site, and breakpoint sequencing would further substantiate

this possibility. While this deletion event could have removed an enhancer or promoter

for this gene, the deleted sequence is not significantly conserved but this does not

disprove the notion that a human specific modifier could be lost. (Figure 3.4)

Additional candidate CNV for validation were identified by their proximity but to coding

or intragenic sequence. While these CNV do not overlap exonic or intronic sequence,

they may arise upstream or downstream of gene. These locations might hold modifier

sequence that could affect (enhance or repress) gene expression, and gain or loss of such sequence may perturb expression levels enough to result in an autistic phenotype.

We identified a 560 bp deletion 1,541 bp upstream of the first coding exon of *glutamate receptor, ionotropic, AMPA3 (GRIA3)*. Array CGH identified two individuals as being deleted at this locus. This gene is interesting as a candidate for ASD because of its causal role in X-linked Mental Retardation (OMIM: 305915, MRX94) and its molecular role as an AMPA-responsive glutamate receptor, the predominant excitatory neurotransmitter receptors in the human brain.[116] We hypothesized a potential role in ASD as mis-regulation of protein expression for this gene product ultimately disturbs neuronal function. The 560 bp deletion removes relatively un-conserved sequence, but does not negate the possibility of human specific regulatory sequences in and overlapping the deleted sequence. (Figure 3.5) Confirmation PCR validated these deletion events, and breakpoint sequencing revealed the two individuals shared the same breakpoints. The di-nucleotides 'AG' and 'GT' are found at the upstream and downstream breakpoints suggesting a non-homologous end joining mechanism mediated a double strand break.

We conducted functional studies of this upstream promoter region by a luciferase assay. The intact and deleted sequences were placed upstream of a luciferase promoter construct, transfected into N2A cells, and luciferase expression from protein lysates was measured. We found an increase in luciferase expression when the reporter was modified by the deleted locus. (Figure 3.6) However, genotyping for the deletion in our expanded autistic cohort (300 AGRE samples) and a normal sample population (1,500 normal

white males and the 300 normal fathers) demonstrated similar frequencies in both populations (p < 0.2247).

4. Analysis of Junction Sequence

a. Validated CNV with breakpoint sequencing

CNV identified and sequenced from the autistic cohorts screened at Emory

We bi-directionally sequenced 27 validated copy number variant loci. These sequenced variants are comprised of 26 deletions and one duplication event. Eight were found to be variant in multiple individuals, and all breaks in these shared loci were the same suggesting that at some point the variant may have been inherited from a common ancestor. We further assessed the sequenced breakpoints for homology at both ends, insertion of sequence, and repeat content at the points of junction. Breakpoints, or the ends of the copy number variant, have been evaluated for the presence of homologous sequence in effort to discern a mechanism by which the structure was resolved. We found 26 junction fragments showing homology at the immediate ends of the upstream and downstream breaks. Seventeen of these homologous sequences were four bp or smaller, and two of the 17 had small insertions (one bp and four bp) at the breakpoint junction. The microhomology of four bp or less as well as the small insertions are suggestive of non-homologous end joining (NHEJ) to repair a double strand break during replication.[117-120] Five of the 27 showed homologies ranging from five to eight bp fall within the range of a non-classical NHEJ (previously known as microhomology mediated end-joining (MMEJ, typified by microhomologies ranging from 5-25 bp).[119-121] Additionally, *Alu* elements and a 200 bp sequence were found in both the upstream and

downstream breaks of three deletions suggesting non-allelic homologous recombination between the homologous sequences at each end may have mediated the deletion of intervening sequence.[119-122] (Table 3.4)

Microhomologies, repeat elements and other sequence-based motifs have been found at or around breakpoints, and their presence is an indicator for how a structural event may have been resolved.[93, 117-121, 123-128] While much effort went into the identification and collection of our set of sequenced breakpoints, our data set remains relatively small. By using data from other studies, additional sequenced breaks can be used to ask broader questions of the occurrence of microhomology as well as evaluate junctions for the enrichment of motif sequences. We searched the literature for sequenced breakpoints or junctions and identified 12 articles from 2002-2010.[49, 92, 93, 129-137]. (Table 3.5)

We collected 681 deletions or duplications having nucleotide resolution of their junctions. (Appendix: Table A.4) The deletions and duplications were identified by various methods and are distributed throughout the genome from diverse populations. (Figure 3.7.a-b) In all data sets, duplications are much larger than deletions. While most of the data are derived from deletions with only 17 duplications (3%) represented, this is similar to what we identified in our data where our single duplication represented 4% of our set. (Table 3.4, Table 3.6) The data generated by Conrad et al is split out from the rest of the reference as it bears the great similarity to our own data set. Conrad's study utilized high-density arrays from NimbleGen as well. Their study conducted a screen of

fine scale variation of the human genome using the HapMap population. However, no calls were made for the X chromosome from their study.

Nearly 80% (536) of the literature sequences were evaluated for homology at the breakpoints. About one third (200 or 37%) of the evaluated sequences showed no homology at both ends of the variant, and the remaining two thirds (337 or 63%) had some degree of homology. Of the 337 structural changes showing homology at their breaks, 74% (249/337) of the homology was four bp or less suggesting NHEJ as the major mechanism of repair to double strand breaks. The remaining structural changes were largely composed of breakpoints having microhomology ranging with the 5-25 bp in size that is indicative of a MMEJ mechanism (24%) of double strand break resolution. The remaining 2% ranged as high as 200 bp, and such relatively long stretches of homology suggest the non-allelic homologous recombination mechanism of repair. Additionally, 102 of the 563 evaluated breaks (18%) showed an insertion at the breakpoint, and 64% (65/z) of these insertions were seven bp or less in size suggesting some sort of slippage during repair or participating in NHEJ.[54] (Table 3.7, Figure 3.8.a-c)

Using all sequenced breakpoints (the 27 from Emory and 681 from the literature), we evaluated single base insertion and deletion events and found no bias for any particular nucleotide for either ($p < 0.61$, and $p < 0.73$ respectively).

b. Analysis of sequence motifs at breakpoint junctions

Various sequence motifs have been previously identified as being significantly associated with structural break resolution. We identified 39 such motifs from six different categories including 1) Chi-like sites, 2) hotspots for deletions, insertions, and indels, 3) DNA polymerase specific sequences, 4) Eukaryotic Transcription Regulation, 5) Mobile Elements, and 6) Recombination sites. [123, 129, 133, 138-148] (Table 3.8) We evaluated the frequency of these motifs at the breakpoints of both the sequenced data set and the randomly generated data sets. The goal was to assess whether these motifs were enriched in the sequenced breakpoint data set over that of the random set.

c. Development and evaluation of randomly chosen breakpoint sequences

We wanted to ask the question whether the frequency of previously reported motif sequences were increased at the breakpoints of CNV relative to random spots in the genome. To do so, we first identified a set of 565 'real' CNV with well-characterized breakpoints. Those CNV consisted of the 27 first described here, together with 305 described by Conrad et al, and the remaining from 11 published sources. For each of these 565 'real' CNV, we created 1,000 'random' CNV. Each random CNV was the same size as the corresponding real CNV, and located on the same chromosome, but with its start location uniformly distributed across the chromosome. The percent 'GC' and percent 'N' (undetermined nucleotide composition) of each 'random' set was similar to the 'GC' and 'N' content of the corresponding 'real' CNV. (Figure 3.9.a-b) 98% of the 1,000 random sets differed from the 'real' variants in their mean GC content at the breaks by less than 10%, and all but three data sets had the same mean 'N' content as the 'real'

values. This resulted in 565,000 randomly positioned 'variants' to be used to compare the frequency of observed characteristics versus a randomly selected population.

For each pair of breakpoints (real or random), we defined 50 bp windows on either side of each breakpoint in both the upstream and downstream breaks. (Figure 3.10) We considered all four windows (('A', 'B', 'C' and 'D') as representative of susceptibility to structural resolution. In other words, we hypothesized that the sequences located in A-D were more susceptible/amenable to resolve a structural event (the product being a deletion or duplication) than a randomly chosen sequence. This susceptibility/amenability to structural resolution may occur in windows inside (B, C) or outside (A, D) or all (A-D) of the CNV breakpoints. We evaluated all possibilites (All windows, inside, or outside) in 565 of the sequenced breakpoints ('real' sequences) and 565,000 random windows ('random' sequences) were evaluated for the frequency of the 39 motifs previously ascribed to breakpoint resolution. We identified eight motif sequences that were overrepresented in our data set as compared to the randomly generated data set. These sequences (Human hypervariable minisatellite core sequence, DNA polymerase α frameshift hotspots, DNA polymerase β frameshift hotspots, AT-rich signal, Curved DNA signal, Topoisomerase II consensus cleavage site (Drosophila), Murine LTR recombination hotspot , Nonamer recombination signal) did not cluster in any one category, and only one motif, Topoisomerase II consensus cleavage site (Drosophila), was significant in all three window types (all, inside and outside). Table 9

5. Summary

We conducted four different studies using three different sample populations, two different array platforms, and two different array hybridization protocols. In total, the X chromosome of 302 samples was screened by high-resolution aCGH. We identified 2,300 CNV of which 1,357 (59%) were deletions and 1,263 (55%) were duplications. We attempted to validate 331 loci, and 146 successful genotyping assays were developed. Of these 146 assays, 100 identified falsely called CNV loci, 44 identified truly called CNV loci, and two confirmed CNV in the reference sample. Of the validated loci, a 560 bp deletion 1,541 bp upstream of the *GRIA3* locus was confirmed. Functional studies in Neuro2A cells were developed to assay the effect of this region with and without the deleted sequence on expression of a luciferase gene. Loss of this sequence resulted in an increase in expression. Genotyping for this deletion in 364 individuals with autism and 1,799 normal individuals did not reveal enrichment for this deletion in the autistic population.

We characterized a subset of our confirmed CNV breakpoint junctions. In addition to identifying microhomology, insertions at the breakpoint, and repeats associated at the breaks, we also expanded our analysis to include an assessment of motif occurrence. We identified 39 sequence motifs that have been previously reported as being enriched at breakpoint junctions. We combined our breakpoints with 565 additional breakpoints with nucleotide resolution from the literature. Using our and the literature sequences, we identified eight different sequence motifs that were enriched within 50 bp windows of the breakpoints.

6. References

1. Locke, D.P., et al., *Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome.* Am J Hum Genet, 2006. **79**(2): p. 275-90.
2. de Stahl, T.D., et al., *Profiling of copy number variations (CNVs) in healthy individuals from three ethnic groups using a human genome 32 K BAC-clone-based array.* Hum Mutat, 2007.
3. Lee, H.W., et al., *Preso, a novel PSD-95-interacting FERM and PDZ domain protein that regulates dendritic spine morphogenesis.* J Neurosci, 2008. **28**(53): p. 14546-56.
4. Medicine, M.-N.I.o.G. *Online Mendelian Inheritance in Man*. 1987; Available from: omim.org/.
5. Kidd, J.M., et al., *A human genome structural variation sequencing resource reveals insights into mutational mechanisms.* Cell, 2010. **143**(5): p. 837-47.
6. Liu, P., C.M.B. Carvalho, and J.R. Lupski, *Mechanisms for recurrent and complex human genomic rearrangements.* 2012.
7. Carvalho, C., et al., *Evidence for disease penetrance relating to CNV size: Pelizaeus-Merzbacher disease and manifesting carriers with a familial 11 Mb duplication at Xq22.* Clin Genet, 2011.
8. Luo, Y., et al., *Diverse mutational mechanisms cause pathogenic subtelomeric rearrangements.* Hum Mol Genet, 2011. **20**(19): p. 3769-78.
9. Chen, J.M., et al., *Genomic rearrangements in inherited disease and cancer.* Semin Cancer Biol, 2010. **20**(4): p. 222-33.
10. Hastings, P.J., et al., *Mechanisms of change in gene copy number.* Nat Rev Genet, 2009. **10**(8): p. 551-64.
11. Conrad, D.F., et al., *Origins and functional impact of copy number variation in the human genome.* Nature, 2010. **464**(7289): p. 704-12.
12. Ball, E.V., et al., *Microdeletions and microinsertions causing human genetic disease: common mechanisms of mutagenesis and the role of local DNA sequence complexity.* Hum Mutat, 2005. **26**(3): p. 205-13.
13. Colnaghi, R., et al., *The consequences of structural genomic alterations in humans: Genomic Disorders, genomic instability and cancer.* Semin Cell Dev Biol, 2011. **22**(8): p. 875-85.
14. Koumbaris, G., et al., *FoSTeS, MMBIR and NAHR at the human proximal Xp region and the mechanisms of human Xq isochromosome formation.* Hum Mol Genet, 2011. **20**(10): p. 1925-36.
15. Liu, P., et al., *Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements.* Cell, 2011. **146**(6): p. 889-903.
16. Mills, R.E., et al., *Mapping copy number variation by population-scale genome sequencing.* Nature, 2011. **470**(7332): p. 59-65.
17. Myers, S., et al., *A common sequence motif associated with recombination hot spots and genome instability in humans.* Nat Genet, 2008. **40**(9): p. 1124-9.
18. de Smith, A.J., et al., *Array CGH analysis of copy number variation identifies 1284 new genes variant in healthy white males: implications for association studies of complex diseases.* Hum Mol Genet, 2007. **16**(23): p. 2783-94.

19.   Nichol Edamura, K. and C.E. Pearson, *DNA methylation and replication: implications for the "deletion hotspot" region of FMR1.* Hum Genet, 2005. **118**(2): p. 301-4.

20.   Korbel, J.O., et al., *Systematic prediction and validation of breakpoints associated with copy-number variants in the human genome.* Proc Natl Acad Sci U S A, 2007. **104**(24): p. 10110-5.

21.   Goldmann, R., et al., *Genomic characterization of large rearrangements of the LDLR gene in Czech patients with familial hypercholesterolemia.* BMC Med Genet, 2010. **11**: p. 115.

22.   Kim, P.M., et al., *Analysis of copy number variants and segmental duplications in the human genome: Evidence for a change in the process of formation in recent evolutionary history.* Genome Res, 2008. **18**(12): p. 1865-74.

23.   Lam, H.Y., et al., *Nucleotide-resolution analysis of structural variants using BreakSeq and a breakpoint library.* Nat Biotechnol, 2010. **28**(1): p. 47-55.

24.   Nobile, C., et al., *Analysis of 22 deletion breakpoints in dystrophin intron 49.* Hum Genet, 2002. **110**(5): p. 418-21.

25.   Park, H., et al., *Discovery of common Asian copy number variants using integrated high-resolution array CGH and massively parallel DNA sequencing.* Nat Genet, 2010. **42**(5): p. 400-5.

26.   Vissers, L.E., et al., *Rare pathogenic microdeletions and tandem duplications are microhomology-mediated and stimulated by local genomic architecture.* Hum Mol Genet, 2009. **18**(19): p. 3579-93.

27.   Woodward, K.J., et al., *Heterogeneous duplications in patients with Pelizaeus-Merzbacher disease suggest a mechanism of coupled homologous and nonhomologous recombination.* Am J Hum Genet, 2005. **77**(6): p. 966-87.

28.   Zhang, F., et al., *Mechanisms for nonrecurrent genomic rearrangements associated with CMT1A or HNPP: rare CNVs as a cause for missing heritability.* Am J Hum Genet, 2010. **86**(6): p. 892-903.

29.   Perry, G.H., et al., *The fine-scale and complex architecture of human copy-number variation.* Am J Hum Genet, 2008. **82**(3): p. 685-95.

30.   Calabretta, B., et al., *Genome instability in a region of human DNA enriched in Alu repeat sequences.* Nature, 1982. **296**(5854): p. 219-25.

31.   Chen, S.J., et al., *Ph1+bcr- acute leukemias: implication of Alu sequences in a chromosomal translocation occurring in the new cluster region within the BCR gene.* Oncogene, 1989. **4**(2): p. 195-202.

32.   Rudiger, N.S., N. Gregersen, and M.C. Kielland-Brandt, *One short well conserved region of Alu-sequences is involved in human gene rearrangements and has homology with prokaryotic chi.* Nucleic Acids Res, 1995. **23**(2): p. 256-60.

33.   Dohoney, K.M. and J. Gelles, *Chi-sequence recognition and DNA translocation by single RecBCD helicase/nuclease molecules.* Nature, 2001. **409**(6818): p. 370-4.

34.   Bacolla, A., et al., *Breakpoints of gross deletions coincide with non-B DNA conformations.* Proc Natl Acad Sci U S A, 2004. **101**(39): p. 14162-7.

35.   Lapidot, A., N. Baran, and H. Manor, *(dT-dC)n and (dG-dA)n tracts arrest single stranded DNA replication in vitro.* Nucleic Acids Res, 1989. **17**(3): p. 883-900.

36.   Abeysinghe, S.S., et al., *Translocation and gross deletion breakpoints in human inherited disease and cancer I: Nucleotide composition and recombination-associated motifs.* Hum Mutat, 2003. **22**(3): p. 229-44.

37.   Chuzhanova, N., et al., *Translocation and gross deletion breakpoints in human inherited disease and cancer II: Potential involvement of repetitive sequence elements in secondary structure formation between DNA ends.* Hum Mutat, 2003. **22**(3): p. 245-51.

38.   Chuzhanova, N., et al., *Gene conversion causing human inherited disease: evidence for involvement of non-B-DNA-forming sequences and recombination-promoting motifs in DNA breakage and repair.* Hum Mutat, 2009. **30**(8): p. 1189-98.

39.   Wu, T.C. and M. Lichten, *Meiosis-induced double-strand break sites determined by yeast chromatin structure.* Science, 1994. **263**(5146): p. 515-8.

40.   Aoki, K., et al., *A novel gene, Translin, encodes a recombination hotspot binding protein associated with chromosomal translocations.* Nat Genet, 1995. **10**(2): p. 167-74.

	# of Samples	# of CNVs (% all CNVs)	# of CNVs/Sample (SD)	Mean size in bp (SD)	Median size in bp	# of CNV which overlap genes (% all CNV)
AGRE - 385K, NimbleGen protocol						
ALL	4	16	4.0 (2.9)	13,415 (36,584)	3,263	1 (6.3%)
Deletions	-	5 (31%)	1.3 (1.5)	3,961 (4,239)	1,227	0 (0%)
Duplications	-	11 (69%)	2.8 (1.5)	17,712 (43,993)	3,539	1 (6.3%)
AGRE - 2.1M, NimbleGen protocol						
ALL	48	581	12.1	2,955 (21,279)	1,080	260
Deletions	-	379 (65.2%)	7.9	1,685 (2,974)	1,005	164 (63.1%)
Duplications	-	202 (34.8%)	4.2	5,340 (35,795)	1,419	96 (36.9%)
AGRE + SSC - 2.1M, Emory protocol						
ALL	149	517	3.5 (3.8)	7,447 (28,307)	1,719	313 (60.5%)
Deletions	-	242 (46.8%)	1.6	5,347 (16,628)	1,694	108 (20.9%)
Duplications	-	275 (53.2%)	1.9	9,282 (35,435)	1,724	205 (39.7%)
NIMH - 2.1M, NimbleGen protocol						
ALL	101	1186	11.7 (10.1)	26,606 (101,842)	2,104	634 (53.5%)
Deletions	-	413 (34.8%)	4.1	6,522 (79,825)	1,043	225 (19.0%)
Duplications	-	773 (65.2%)	7.7	37,336 (110,402)	4,091	409 (34.5%)

bp: base pairs, CNV: Copy Number Variant, SD: Standard Deviation

Table 3.1.a All Copy Number Variants Identified by high-density aCGH

	# of Samples	# of CNVs (% all CNVs)	# of Distinct CNV loci (% ofAll distinct)	Mean size in bp (SD)	Median size in bp	# of Distinct Loci which overlap genes (% all distinct)
AGRE - 385K, NimbleGen protocol						
ALL	**4**	**16**	**9**	**21,437 (48,395)**	**4,007**	**1 (6.3%)**
Deletions	-	5 (31%)	4 (44%)	4,645 (4,566)	3,514	0 (0%)
Duplications	-	11 (69%)	5 (56%)	34,872 (64,505)	4,007	1 (6.3%)
Deleted & Dupliated	-	-	-	-	-	-
AGRE - 2.1M, NimbleGen protocol						
ALL	**48**	**581**	**282**	**5,073 (32,850)**	**1,152**	**112 (46%)**
Deletions	-	379 (65.2%)	153 (63%)	2,171 (4,318)	994	52 (21%)
Duplications	-	202 (34.8%)	82 (34%)	10,677 (55,916)	1,443	50 (21%)
Deleted & Dupliated	-	-	7 (2.9%)	2,876 (3,538)	1,355	10 (4%)
AGRE + SSC - 2.1M, Emory protocol						
ALL	**149**	**517**	**244**	**10,890 (37,869)**	**1,951**	**128 (53%)**
Deletions	-	242 (46.8%)	118 (48%)	4,929 (10,902)	1,608	41 (17%)
Duplications	-	275 (53.2%)	106 (43%)	13,725 (45,971)	2,063	73 (30%)
Deleted & Dupliated	-	-	20 (8.2%)	31,034 (72,368)	4,203	14 (6%)
NIMH - 2.1M, NimbleGen protocol						
ALL	**101**	**1186**	**426**	**18,900 (107,802)**	**1,945**	**206 (48%)**
Deletions	-	413 (34.8%)	151 (35%)	15,351 (131,806)	1,375	58 (14%)
Duplications	-	773 (65.2%)	252 (59%)	20,910 (96,095)	2,407	128 (30%)
Deleted & Dupliated	-	-	23 (5.4%)	20,174 (23,647)	9,118	20 (4.7%)

Table 3.1.b Distinct Copy Number Variants Identified by high-density aCGH

	# of CNV	# of Unique Loci	Mean Size of All CNV (bp)	Median Size of All CNV (bp)	Mean GC Content of All CNV (%)
True Positive	**58**	**39**	**11,451**	**2,343**	**40.6%**
Deletion	54	35	11,778	2,243	40.6%
Duplication	4	4	7,036	7,141	39.9%
False Positive	**121**	**66**	**3,254**	**1,370**	**53.3%**
Deletion	59	26	1,481	1,205	54.0%
Duplication	62	35	4,940	1,439	52.6%
Deleted & Duplicated	-	5	3,165	4,114	52.4%

bp: base pairs, CNV: Copy Number Variant

Table 3.2 – Characteristics of validated copy number variants.

Locus	chrX Start	chrX Stop	Size (bp)	State	Autistic Frequency	Control Frequency
AU015.2	15,161,742	15,197,116	35,374	duplication	1/299	0/300
AU033.95	33,952,914	33,982,773	29,859	deletion	2/364	0/300
AU065.4	65,381,486	65,415,746	34,260	duplication	1/300	0/300
AU082.1	82,085,921	82,094,944	9,023	deletion	pooled DNA genotyped	?
AU096.5	96,493,413	96,495,398	1,985	deletion	pooled DNA genotyped	?
AU113.2	113,234,594	113,241,627	7,033	deletion	pooled DNA genotyped	?
AU120.4	120,416,100	120,419,999	3,899	deletion	pooled DNA genotyped	?

Table 3.3 – CNV genotyping in progress for autistic and unaffected populations.

ID	State	Start*	End*	Homology at Junction	Insertions
1	deletion	24,021,195	24,026,627	Alu element	-
2	deletion	30,048,860	30,056,098	TT	-
3	deletion	30,257,687	30,258,519	AG	A
4	deletion	33,952,914	33,982,773	AGGT	-
5	deletion	38,271,449	38,272,275	AAAAT	CT
6	deletion	44,264,761	44,266,313	Alu element	-
7	deletion	58,256,036	58,256,488	AGGCATTCTAATGATAGAGACACCTGTGGTGA	-
8	deletion	64,002,062	64,014,239	ACACT	-
9	duplication	65,381,485	65,415,746	C	-
10	deletion	80,303,149	80,304,088	ATA	-
11	deletion	82,085,921	82,094,944	CAGAG	-
12	deletion	87,545,357	87,547,591	ATTA	-
13	deletion	96,493,413	96,495,398	AG	-
14	deletion	97,842,279	97,843,597	ATT	-
15	deletion	103,289,074	103,289,688	ATTGCCCT	-
16	deletion	103,295,675	103,296,655	CATT	-
17	deletion	111,753,107	111,753,796	AA	-
18	deletion	113,234,594	113,241,627	GGC	-
19	deletion	116,588,176	116,591,267	GCT	CTTC
20	deletion	120,416,100	120,416,999	-	AATCAA
21	deletion	122,143,696	122,144,257	AG	-
22	deletion	131,767,115	131,769,226	GCC	-
23	deletion	133,692,157	133,693,471	AA	-
24	deletion	143,436,372	143,445,447	ATATCC	-
25	deletion	145,207,078	145,208,683	ATTT	-
26**	deletion	149,678,591	149,679,508	TCTG	-
27	deletion	150,457,729	150,462,994	~220 nt	-

*If breaks not clearly identifiable, the average position based on earliest and last possible base is shown.
**5 nt deleted 37 bp upstream of break

Table 3.4 - Validated and bidirectionally sequenced CNV breakpoints identified in our study.

Study	Year	Population Phenotype/Disease	Samples in CNV Analysis	CGH analysis Platform	# of CNVs Identified	Sequence Analysis Platform	# of Sequenced Breaks
Emory	2010	Autistic, male	142	**NimbleGen HD aCGH** ChrX	517	Sanger sequencing	27
Conrad	2010	Normal, HapMap, male/female	450	**NimbleGen HD aCGH** genome-wide	11,700	454 sequencing	324
deSmith	2008	Normal, French Caucasian males	50	**Agilent aCGH** 1) 185K genome-wide 2) 244K targeted	2,208	Sanger sequencing	20
Edamura	2005	Fragile X Syndrome	n/a	**various**	16	Sanger sequencing	16
Goldman	2010	Familial hypercholesterolemia	1447	**Multiple Ligation Probe Amplification**	37	Sanger sequencing	8
Kim	2008	Normal	Published Sources	**Published Sources:** Segmental Duplication Db, Human Copy Variation Consortium, Fosmid Paired-End Sequencing, Genome Comparison	NA	454 sequencing	67
Korbel	2007	Cat-Eye Syndrome, Emanuel syndrome, Dup22 syndrome, 22q11 Deletion syndrome, Digeorge syndrome	10	**Nimblegen aCGH** 385K, targeted	>400	Sanger sequencing	167
Lam	2010	CEPH male, Published Sources	Published Sources	**Published Sources:** Sanger sequencing, fosmid-paired-end-sequencing, next-gen sequencing, high-resolution aCGH	NA	Sanger sequencing	12
Nobile	2002	DMD/BMD	22		22	Sanger sequencing	4
Park	2010	Normal, HapMap, female	30	**Agilent HD aCGH** 1) 24M, genome-wide 2) 185K, targeted	20,099	Sanger sequencing	42
Vissers	2009	Mutiple Congenital Anomalies (MCA), MR with or without MCA, Epilepsy or Autism	38	**NimbleGen HD aCGH** 1) 385K, 2.1M genome-wide 2) two 385K targeted	38	Sanger sequencing	38
Woodward	2005	Pelizaeus-Merzbacher Disease	59	**FISH, multiplex PCR**	59	Sanger sequencing	12
Zhang	2010	Charcot-Marie-Tooth disease Type 1A, hereditary neuropathy with liability to pressure palsy	21	**Agilent aCGH** 15K, targeted	17	Sanger sequencing	9

Table 3.5 – Summary of articles identified in the literature.

	# of CNVs (% all CNVs)	Mean size in bp (SD)	Median size in bp
Emory			
ALL	**27**	**5,270 (8,377)**	**1,605**
Deletions	26 (96%)	4,155 (6,169)	1,579
Duplications	1 (4%)	34,261 (n/a)	34,261
Conrad			
ALL	**305**	**6,623 (19,157)**	**2,534**
Deletions	302 (99%)	6,592 (19,245)	2,447
Duplications	3 (1%)	9,707 (5,396)	9,235
11 Other Literature References			
ALL	**375**	**202,066 (993,372)**	**6,049**
Deletions	361 (96%)	177,580 (979,610)	6,035
Duplications	14 (4%)	833,461 (1,169,881)	491,597

Table 3.6 – Characteristics of breakpoint sequenced CNV from the literature and our studies.

A. Deletion and Duplication Proportions

	# CNV	Deletions (% of set)	Duplications (% of set)
All	708	690 (97%)	18 (3%)
Emory	27	26 (96%)	1 (4%)
Conrad	305	302 (99%)	3 (1%)
Literature*	376	362 (96%)	14 (4%)

* Literature less Conrad data

B. Homology characteristics

	# CNV	No homology at Break (% evaluated)	Microhomology ≤ 4 bp (% evaluted)	Microhomology 5-25 bp (% evaluted)	Homology >25 bp (% evaluted)
All	708	200/563 (36%)	266/563 (47%)	84/563 (15%)	13/563 (2%)
Emory	27	1/27 (4%)	17/27 (63%)	5/27 (19%)	4/27 (15%)
Conrad	305	115 (38%)	161 (53%)	27 (9%)	2 (1%)
Literature*^	376	85/232 (37%)	87/232 (38%)	53/232 (23%)	7/232 (3%)

* Literature less Conrad data

^not all sequences were evaluated for homologies or insertions

C. Insertion characteristics

	# CNV	No insertion at Break* (% evaluated)	Insertion at Break* (% evaluated)	Insertions* ≤ 7 bp (% evaluted)
All	708	461/563 (82%)	102/563 (18%)	69/563 (12%)
Emory	27	23/27 (85%)	4/27 (15%)	4/27 (15%)
Conrad	305	215 (70%)	90 (30%)	55 (18%)
Literature*	376	224/232 (97%)	7/232 (3%)	5/232 (2%)

* Literature less Conrad data

^not all sequences were evaluated for homologies or insertions

Table 3.7 - Characteristics of junction breaks from our study and the literature.

Site Name	Motif Sequence(s)	Reference
X– and X–like Sites		
Chi-sequence	GCTGGTGG	Abeysinghe, Dohoney
Chi-sequence truncated	TGGTGG	Chuzhanova
Human Fra(X) breakpoint cluster	CGGCGG	Chuzhanova
Human minisatellite conserved sequence/X-like element	GCWGGWGG	Abeysinghe, Kvikstad
Human hypervariable minisatellite core sequence	GGGCAGGANG, GGAGGTGGGCAGGARG	Abeysinghe, Kvikstad
Human hypervariable minisatellite recombination sequence	AGAGGTGGGCAGGTGG	Abeysinghe, Kvikstad
Deletion, Insertion, and Indel Hotspots		
Deletion hotspot consensus sequence	TGRRKM, YYTG	Abeysinghe, Kvikstad
Hamster deletion hotspot	TGGAG	Chuzhanova
Hamster and human APRT deletion hotspot	TTCTTTC	Chuzhanova
Indel hotspot	GTAAT	Kvikstad
Indel Super-hotspot motifs	CCCAG, TTCWCCCC, CCACCA, GGGACA, GCCCCG, AGCTG, CCATCT, GGAGAA	Chuzhanova
Insertion hotspots	ATMMGCC, TACCRC	Kvikstad
DNA Polymerase Pause or Frameshift Hotspots		
DNA polymerase α frameshift hotspots	TCCCCC, CTGGCG	Abeysinghe, Kvikstad
DNA polymerase β frameshift hotspots	ACCCWR, TTTT	Abeysinghe, Kvikstad
DNA polymerase α/β frameshift hotspots	TGGNGT, ACCCCA	Abeysinghe, Kvikstad
DNA polymerase α pause site core sequence	GAG, ACG, GCS	Abeysinghe, Kvikstad
DNA polymerase arrest site	WGGAG	Abeysinghe, Kvikstad
Eukaryotic Transcriptional Regulation		
AT-rich signal	WWWWWW	Singh
Curved DNA signal	AAAAN7AAAAN7AAAA, TTTTN7TTTTN7TTTT, TTTAAA	Singh
Kinked DNA signal	TAN3TGN3CA, TAN3CAN3TG, TGN3TAN3CA, TGN3CAN3TA, CAN3TAN3TG, CAN3TGN3TA	Singh
ORI signal	ATTA, ATTTA, ATTTTA	Singh
TG-rich signal	TGTTTTG, TGTTTTTTG, TTTTGGGG	Singh
Topoisomerase Cleavage Sites		
Topoisomerase I consensus cleavage site (Vaccinia)	YCCTT	Abeysinghe, Kvikstad
Topoisomerase I consensus cleavage site (vertebrate/plant)	CAT, CTY, GTY, RAT	Abeysinghe, Kvikstad
Topoisomerase II consensus cleavage site (Drosophila)	GTNWAYATTNATNNR	Singh
Topoisomerase II consensus cleavage site (vertebrate)	RNYNNCNNGYNGKTNYNY	Abeysinghe, Kvikstad, Singh
Mobile Elements		
Alu core element	CCTGTAATCCCAGCACTTTGGGAGGC	Abeysinghe, Rudinger
Mariner transposon-like element (3' end)	GAAAATGAAGCTATTTACCCAGGA	Abeysinghe, Kvikstad
Recombination Sites		
Autonomously replicated sequence	WRTTTATTTAW	Chuzhanova
Chinese hamster scaffold attachment site	AATAAAYAAA	Chuzhanova
Classical meiotic recombination hotspot	CCTCCCT	Chuzhanova
Hamster and human APRT deletion hotspot	TTCTTTC	Chuzhanova
Human hypervariable minisatellite recombination sequence	AGAGGTGGGCAGGTGG	Abeysinghe, Kvikstad
Indel Super-hotspot motifs	CCCAG, TTCWCCCC, CCACCA, GGGACA, GCCCCG, AGCTG, CCATCT, GGAGAA	Chuzhanova
Meiotic recombination hotspot	CCTCCCCT	Chuzhanova
Murine LTR recombination hotspot	TGGAAATCC	Abeysinghe, Kvikstad
Murine MHC recombination hotspot	CAGRCAGR	Abeysinghe, Kvikstad
Murine parvovirus recombination hotspot	CTWTTY	Abeysinghe, Kvikstad
Translin target sites	ATGCAG, GCCCWSSW	Abeysinghe, Aoki, Kvikstad
V(D)J Recombination Signals		
Heptamer recombination signal	CACAGTG	Abeysinghe, Kvikstad
Nonamer recombination signal	ACAAAAACC	Abeysinghe, Kvikstad
Immunoglobulin heavy chain class switch repeats	GAGCT, GGGCT, GGGGT, TGGGG, TGAGC	Abeysinghe, Kvikstad

Table 3.8. - Motifs previously associated with genomic rearrangement.

Site Name	p-values		
X– and X–like Sites	**ALL**	**INSIDE**	**OUTSIDE**
Chi-sequence	1	1	1
Chi-sequence truncated	1	1	1
Human Fra(X) breakpoint cluster	1	1	1
Human minisatellite conserved sequence/X-like element	1	1	1
Human hypervariable minisatellite core sequence	1	1	**8.09E-185**
Human hypervariable minisatellite recombination sequence	*	*	*
Deletion, Insertion, and Indel Hotspots			
Deletion hotspot consensus sequence	1	0.25	1
Hamster deletion hotspot	1	1	1
Hamster and human APRT deletion hotspot	1	1	1
Indel hotspot	1	1	1
Indel Super-hotspot motifs	1	1	1
Insertion hotspots	1	1	1
DNA Polymerase Pause or Frameshift Hotspots			
DNA polymerase α frameshift hotspots	1	1	**7.98E-02**
DNA polymerase β frameshift hotspots	**9.44E-03**	**2.80E-04**	1
DNA polymerase α/β frameshift hotspots	1	1	1
DNA polymerase α pause site core sequence	1	1	1
DNA polymerase arrest site	1	1	1
Eukaryotic Transcriptional Regulation			
AT-rich signal	0.52	**2.14E-02**	1
Curved DNA signal	0.52	**3.85E-02**	1
Kinked DNA signal	1	1	1
ORI signal	1	1	1
TG-rich signal	1	1	1
Topoisomerase Cleavage Sites			
Topoisomerase I consensus cleavage site (Vaccinia)	1	1	1
Topoisomerase I consensus cleavage site (vertebrate/plant)	1	1	1
Topoisomerase II consensus cleavage site (Drosophila)	**7.43E-280**	**5.66E-136**	**1.11E-146**
Topoisomerase II consensus cleavage site (vertebrate)	1	**0**	1
Mobile Elements			
Alu core element	1	1	1
Mariner transposon-like element (3' end)	*	*	*
Recombination Sites			
Autonomously replicated sequence	1	1	1
Chinese hamster scaffold attachment site	1	1	1
Classical meiotic recombination hotspot	1	1	1
Hamster and human APRT deletion hotspot	1	1	1
Human hypervariable minisatellite recombination sequence	*	*	*
Indel Super-hotspot motifs	1	1	1
Meiotic recombination hotspot	1	1	1
Murine LTR recombination hotspot	1	1	**5.28E-100**
Murine MHC recombination hotspot	1	1	1
Murine parvovirus recombination hotspot	1	1	1
Translin target sites	1	1	1
V(D)J Recombination Signals			
Heptamer recombination signal	1	1	1
Nonamer recombination signal	**0**	**1.79E-274**	**6.83E-282**
Immunoglobulin heavy chain class switch repeats	1	1	1

* motif not observed in literature and random sequences

Table 3.9 –Motifs with literature frequencies significantly different from random.
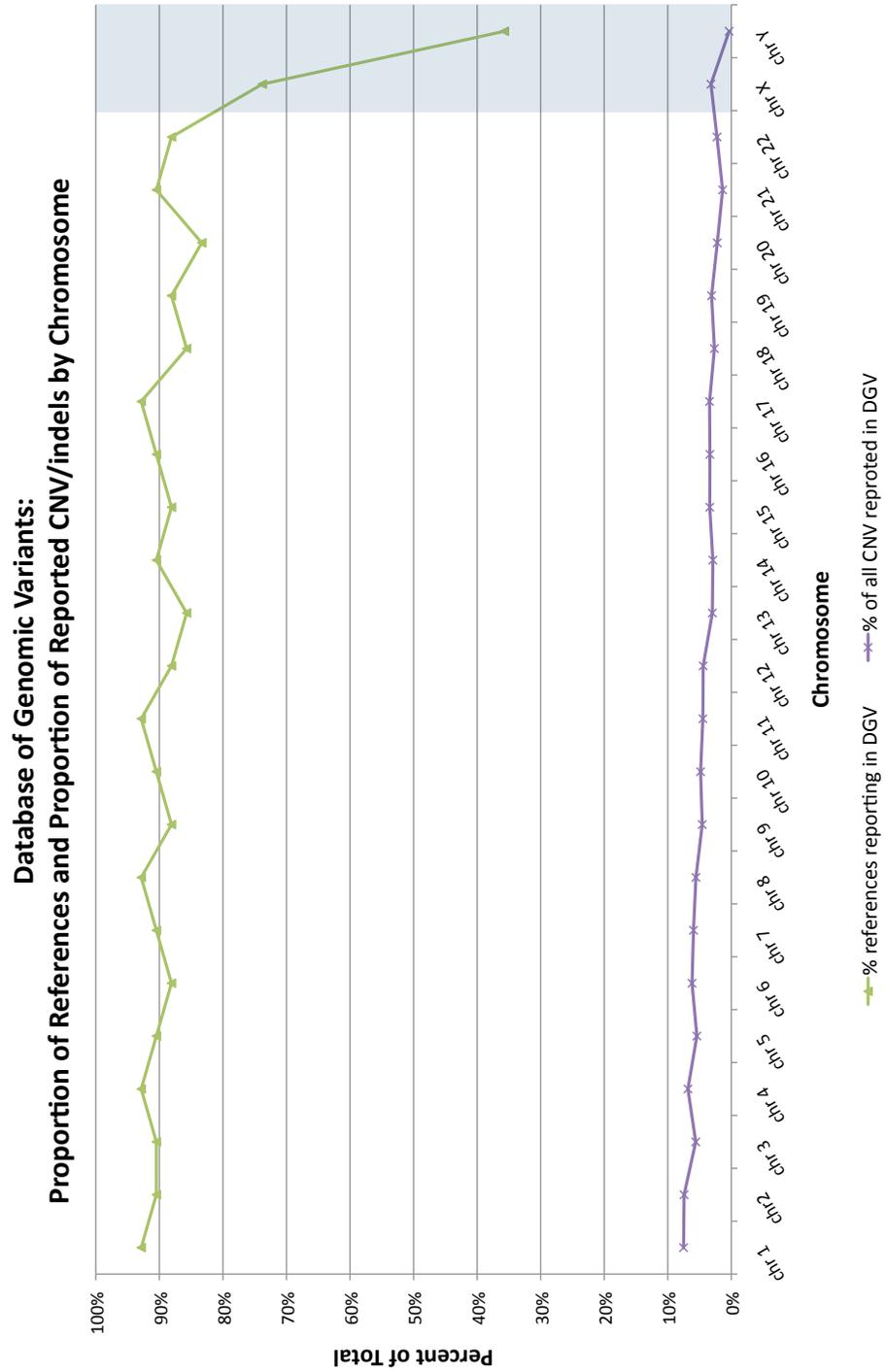
Figure 3.1 - Proportion of References and CNV or indels reported in the DGV.

Database of Genomic Variants: CNV and Indel for Chr X

3,268 CNV & indel
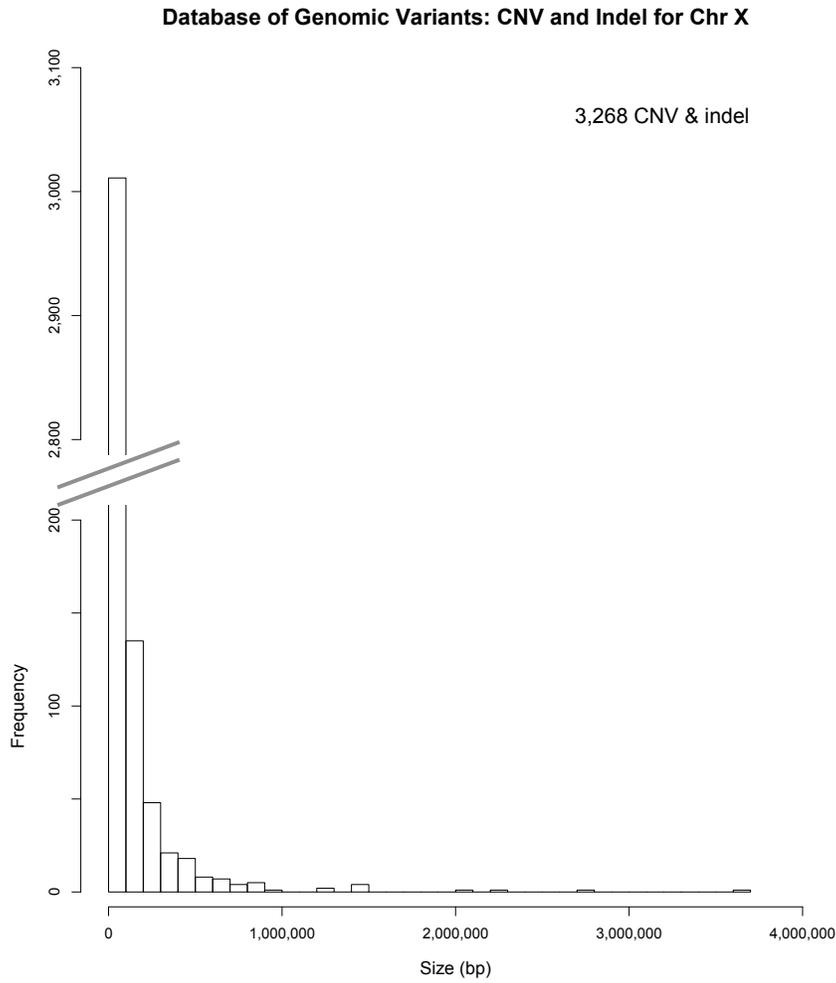
Figure 3.2.a – Size distribution of CNV and Indels from the Database of Genomic

Variants.

Conrad (2009): CNV and Indel for all autosomes

11,080 CNV & indel
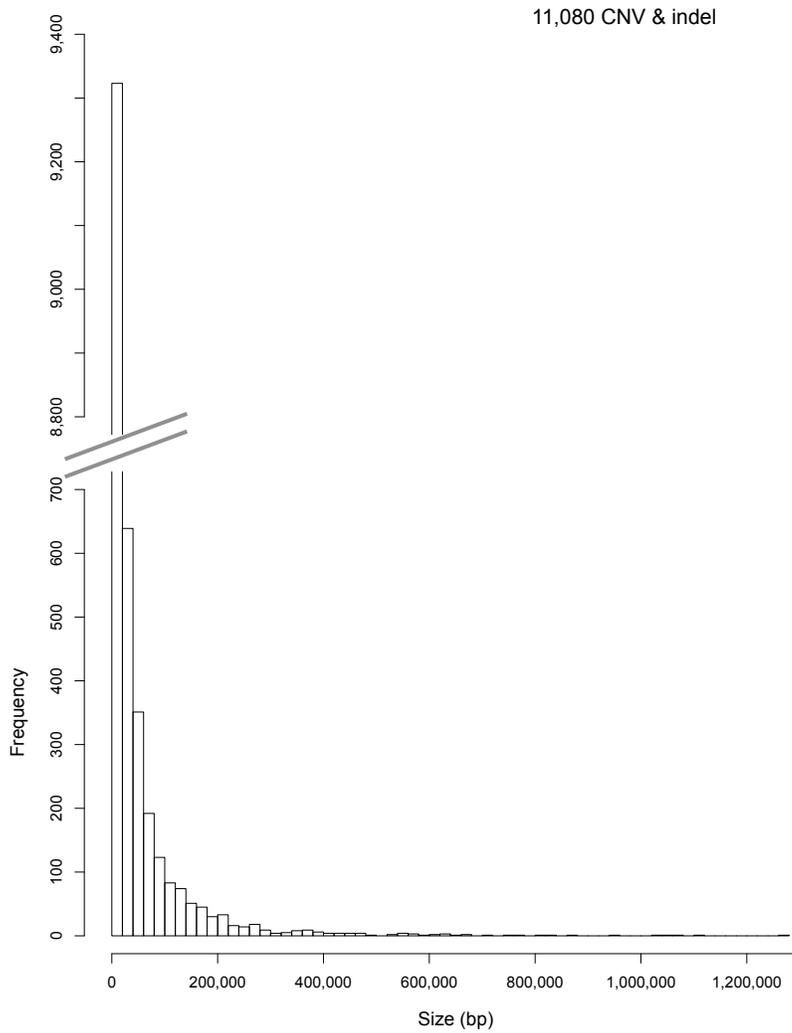
Figure 3.2b – Size distribution of CNV and indels from all autosomes reported by Conrad et al.

50 AGRE by NimbleGen protocol: CNV and Indel on Chr X
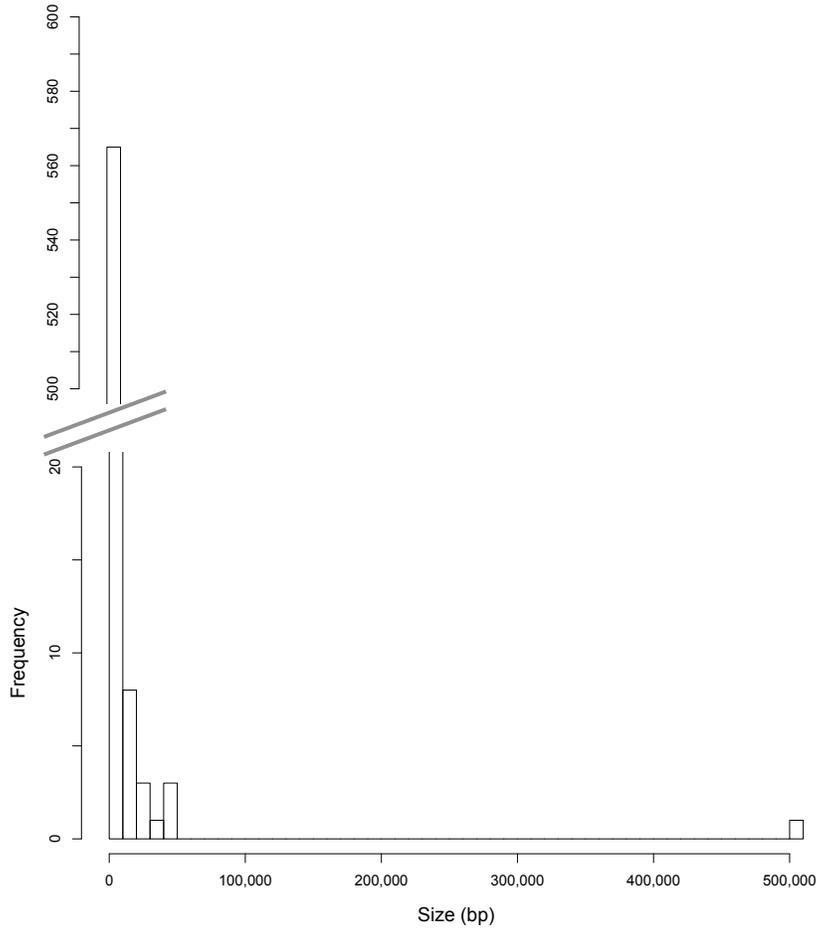
Figure 3.2.c – Size distribution of CNV and indels identified in the AGRE cohort folloing

the NimbleGen protocol.

164 AGRE & SSC by Optimized protocol: CNV and Indel on Chr X
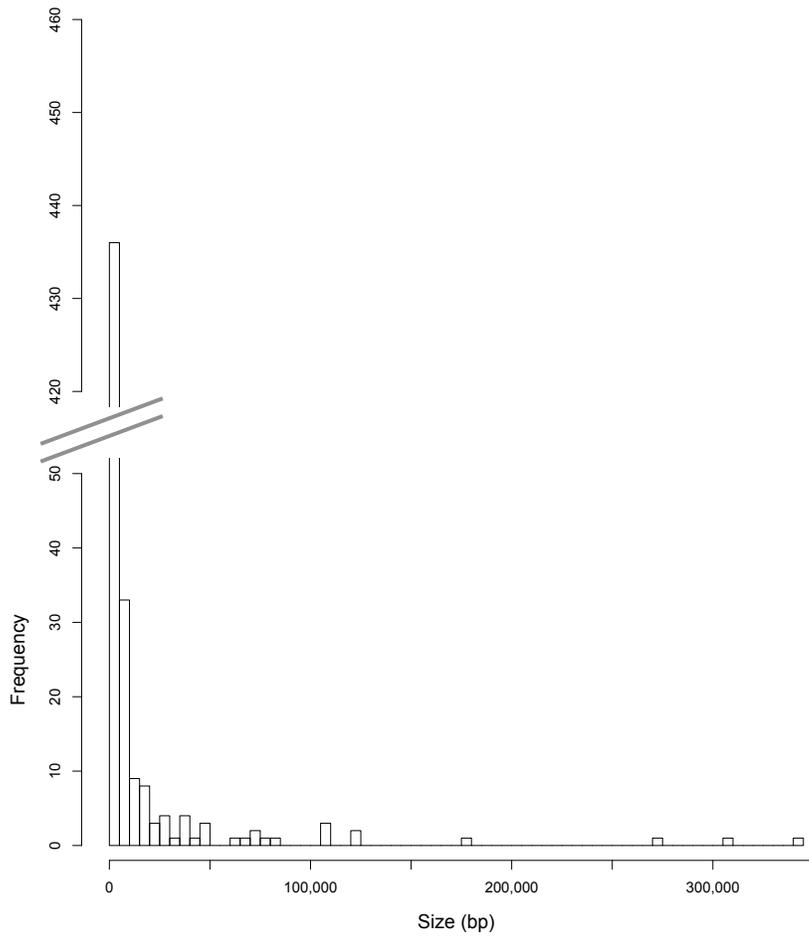
Figure 3.2.d – Size distribution of CNV and indels from the AGRE cohort following the

Optimized protocol.

100 NIMH by NimbleGen protocol: CNV and Indel on Chr X
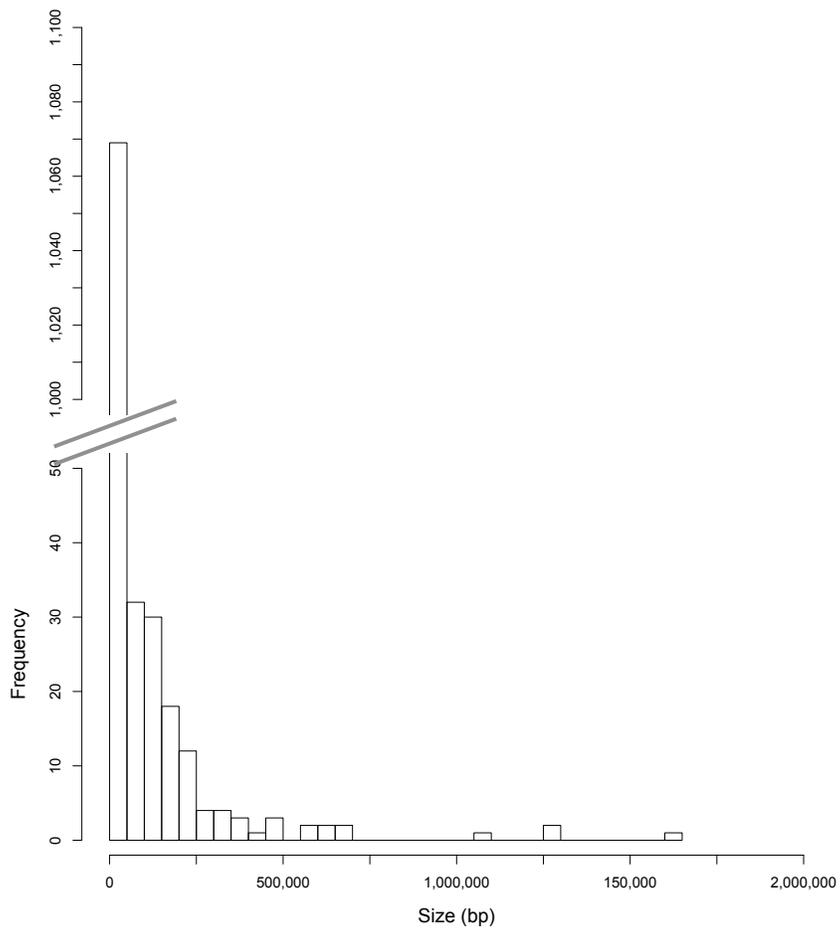
Figure 3.2.e – Size distribution of CNV and indels from the NIMH cohort following the NimbleGen protocol.
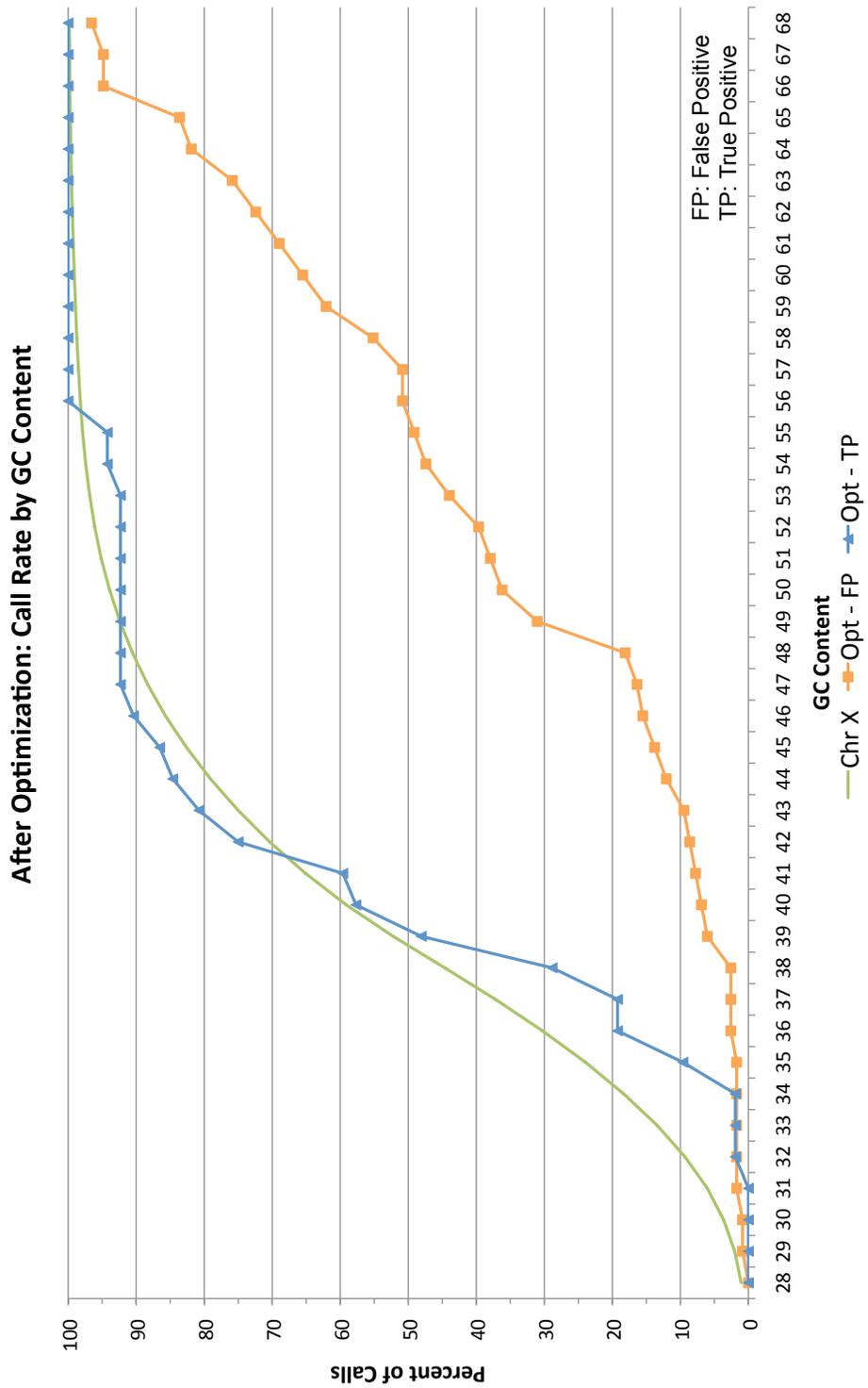
Figure 3.3 - Proportion of true and false calls plotted by GC Content
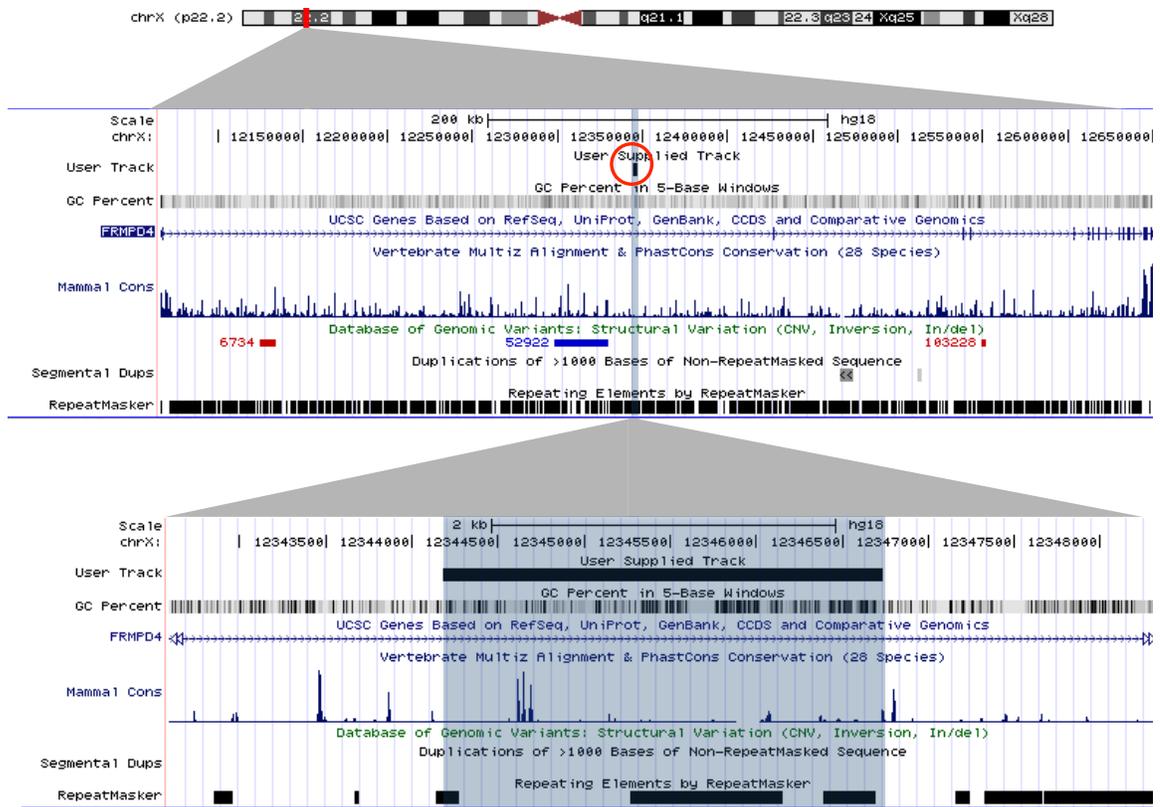
Figure 3.4 – 2.5kb intragenic deletion of FRMPD4.

Figure 3.5 – 561 bp deletion 1.5 kb upstream of GRIA3.
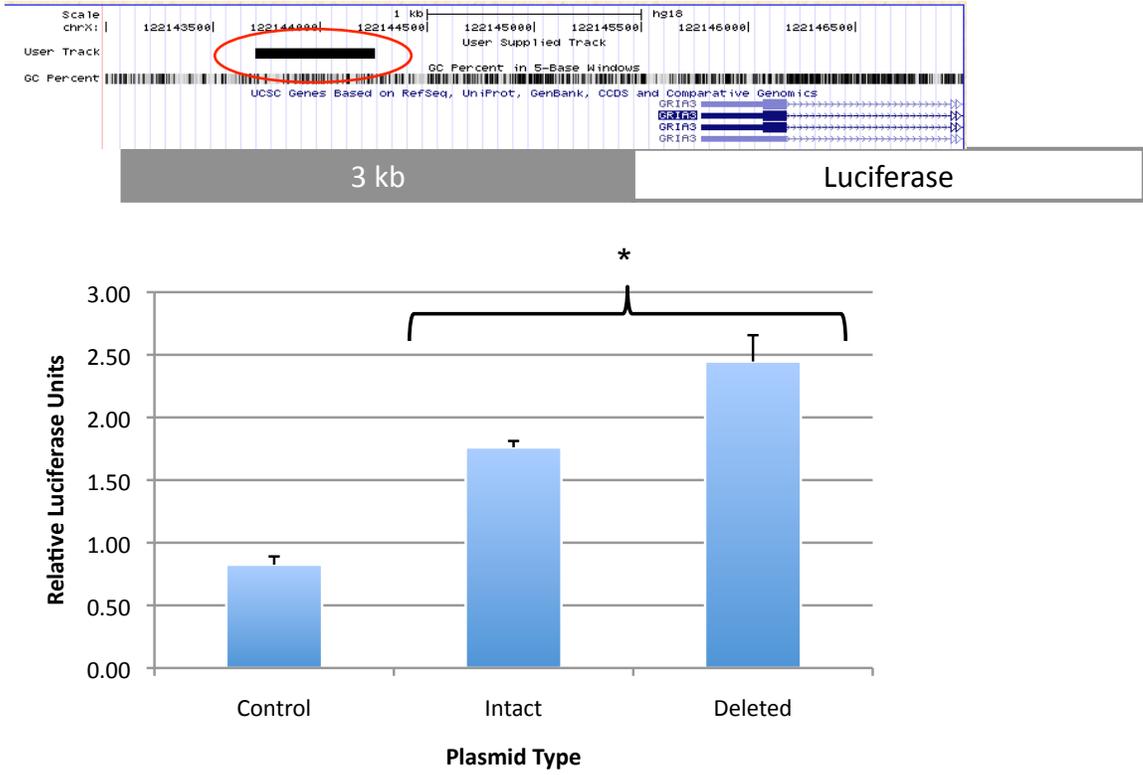
Figure 3.6 – The *GRIA3* promoter deletion has increased activity.
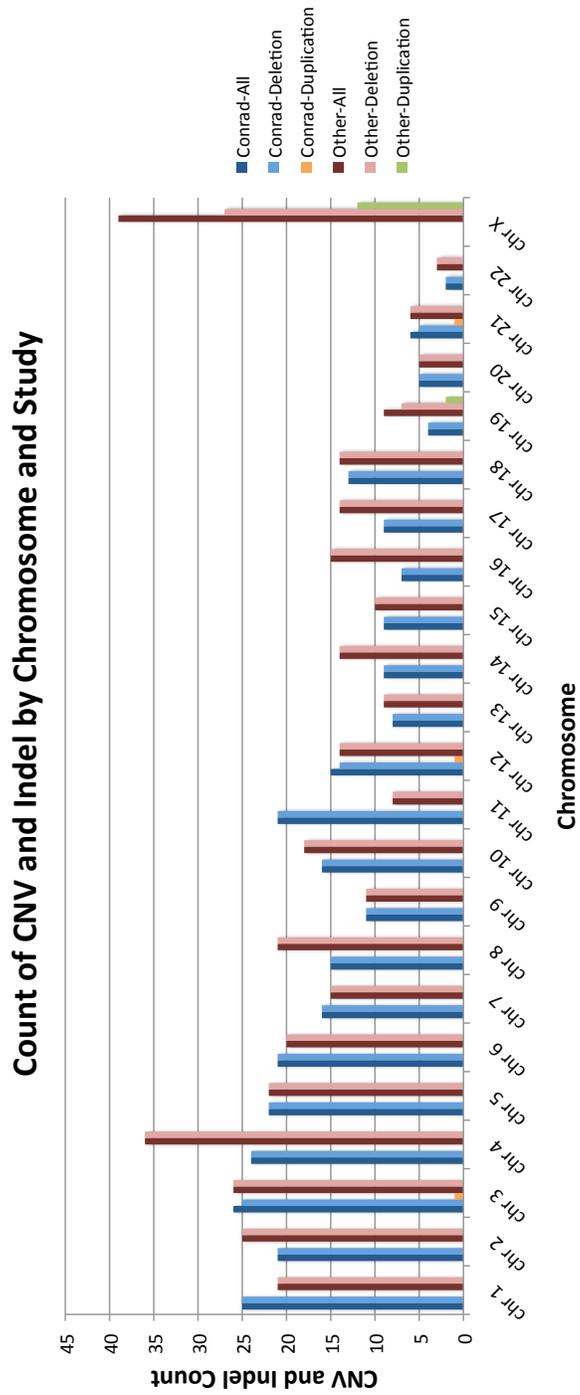
Figure 3.7.a – Literature breakpoint sequence distribution throughout the genome.
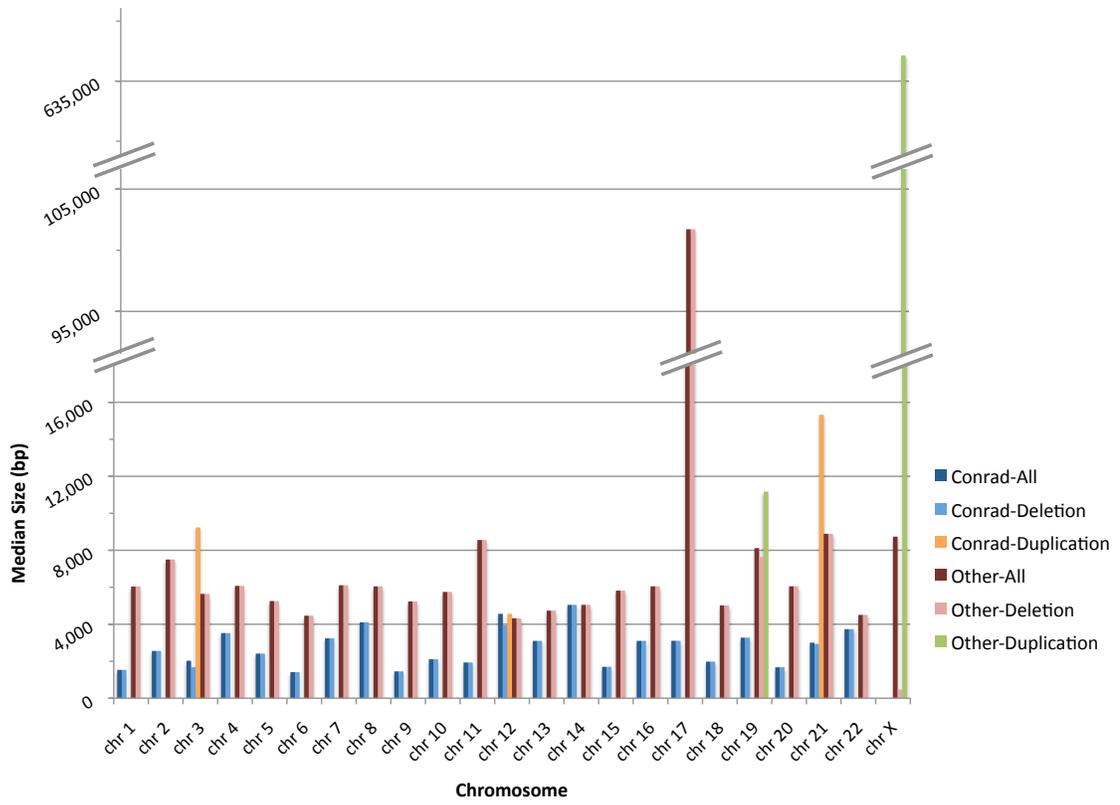
Figure 3.7.b – Median size of literature breakpoint sequence by chromosome.
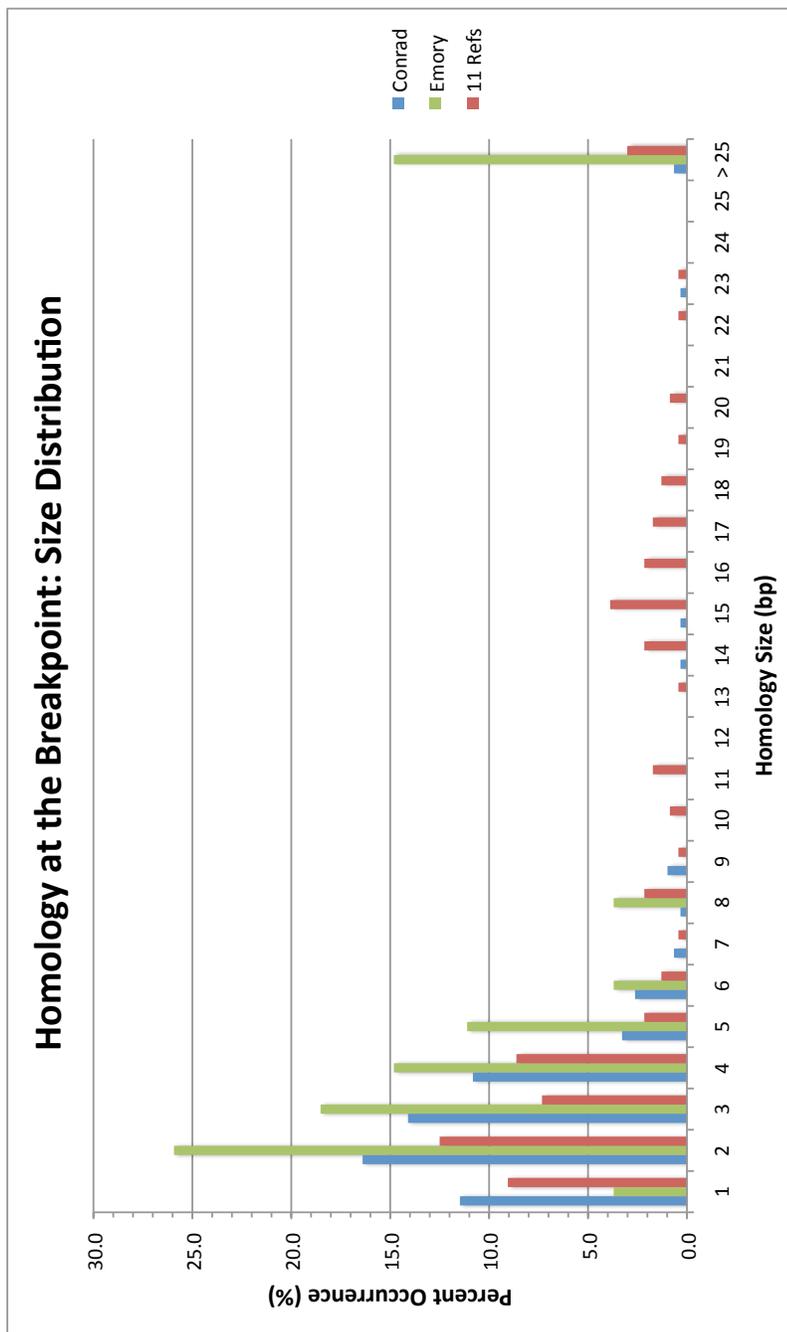
Figre 3.8.a - Size distribution of homologies at breakpoints.
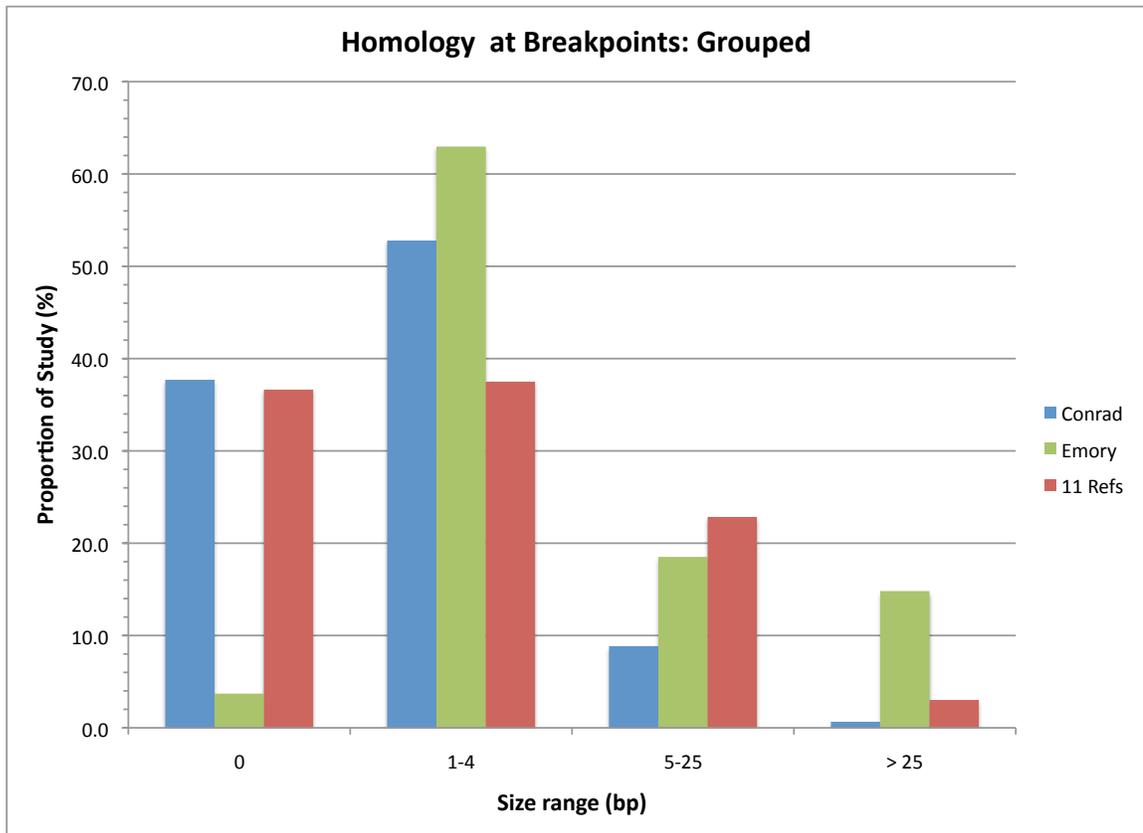
Figure 3.8.b - Grouped size distribution of homologies at breakpoints.
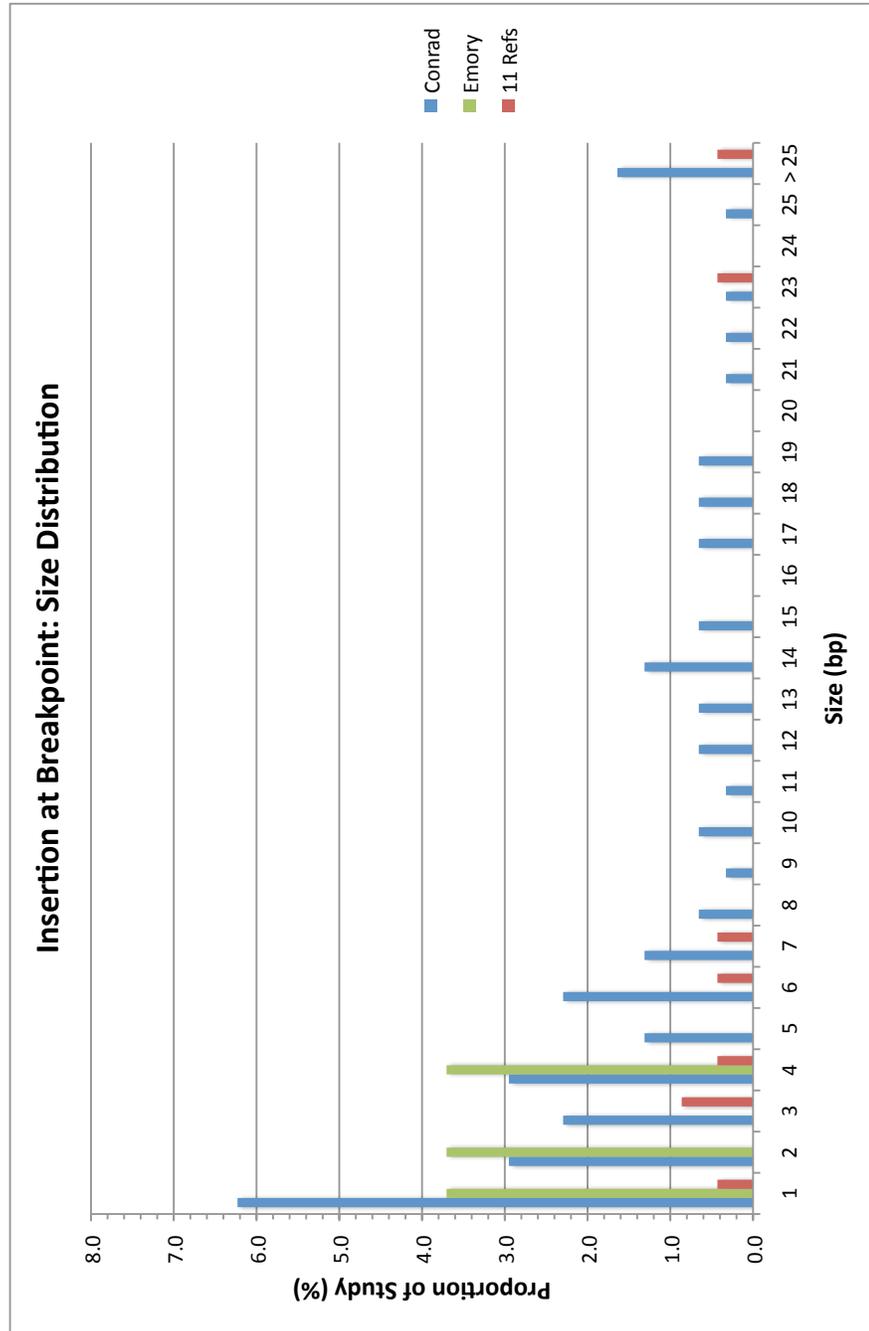
Figure 3.8.c - Size distribution of insertions at breakpoints.
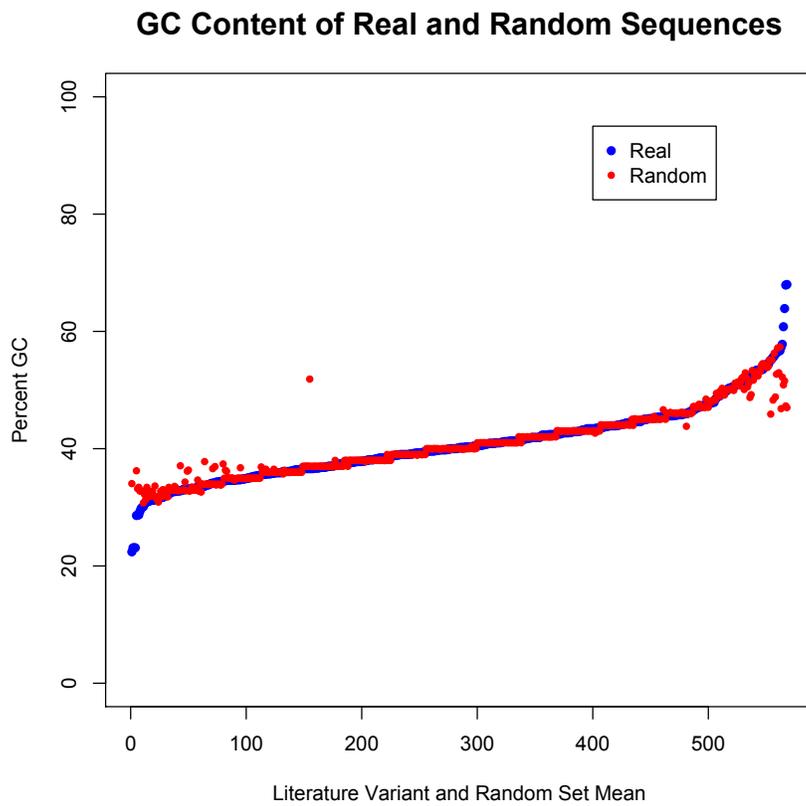
Figure 3.9.a – Average GC content of 1,000 Random Sequences and Real Sequence

N Content of Real and Random Sequences
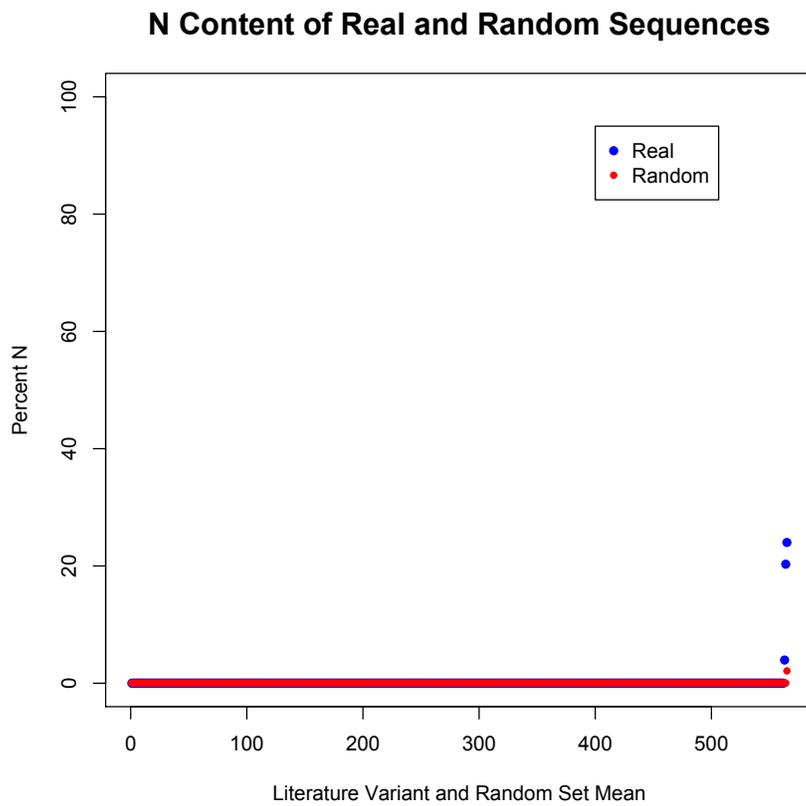
Figure 3.9.b – Average N content of 1,000 Random Sequences and Real Sequence.

50 bp — A
50 bp — B
50 bp — C
50 bp — D

All windows: A + B + C + D
Outside windows: A + D
Inside windows: B + C

deletion
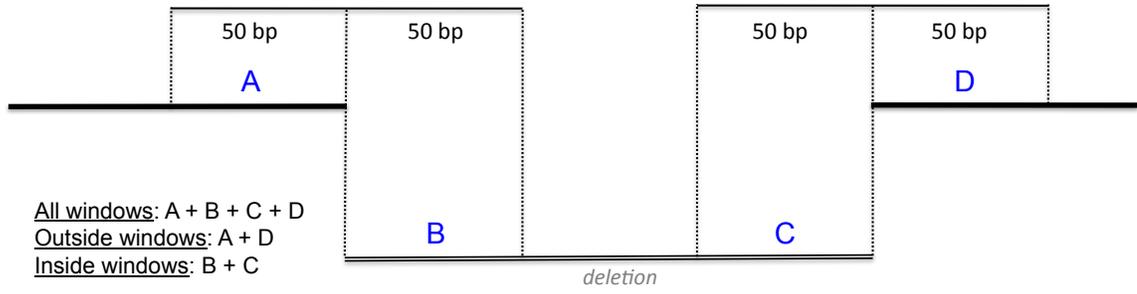
Figure 3.10 – Schematic for breakpoint junction analysis.

Chapter 4. Conclusion

With a prevalence currently estimated as being 1:88 [1], identifying the genetic components underlying the autism spectrum disorders (ASD) would benefit those individuals affected with these disorders as well as their caretakers. While understanding the molecular etiology of ASD can provide incite into unaffected behavior and cognitive functioning, this understanding could also focus the development of future drug and therapeutic strategies in the management of these disorders. A male bias of 4:1 in ASD [2-4] suggests a role for an X-linked susceptibility locus. Furthermore, skewed inactivation in females affected with autism and ASD as well as linkage for sex-specific risk alleles have further implicated the X chromosome.[5-11] These lines of evidence have led us to hypothesize that gene(s) residing on the X chromosome play a causal role in the susceptibility to ASD. To investigate this hypothesis, we developed an experimental strategy to screen copy number variation (CNV) on the X chromosome and assess the role of CNV in these disorders. This chapter discusses the methods and tools that we developed to identify high confidence CNV in autistic and normal cohorts, CNV identification and subsequent validation studies conducted on our findings, and finally conclusions that can be drawn and potential strategies to move forward in the investigation of the genetic underpinnings to autism.

1. The Optimization of array Comparative Genomic Hybridization and CNV Identification

We utilized a custom high density Comparative Genomic Hybridization array (aCGH) dedicated to the X chromosome to interrogate the size and distribution of CNV in our sample populations. During the course of our experiments, we developed a protocol that optimized the quality of the output from these arrays. Our efforts included increasing the fluorescence of labeled sample for greater signal discrimination, increasing the stringency of array hybridization and washing by moving to an automated platform (Tecan HSPro4800), establishing quality control measures of array data (fluorescent Count Ratio limits, MA plot evaluation), removal of unreliable probes, and developing an algorithm for identifying a parsimonious set of array-identified copy number variants (segment filtering). Under the conditions of the new protocol, restricting analysis to CNV with a GC content of less than 56% resulted in a drop from 86.8% of our false calls to 50.9% while still identifying all true calls. A caveat to limiting CNV validation to those CNV less than 56% GC content is reflected in the nature of genes and their elevated GC content. Most of the X chromosome (98%) has a GC content of 56% or less. However, it may be prudent to explore what percentage of genes exist within the remaining 2%. Validation studies require expenditures of effort and resources. However, if the goal is to identify changes in sequence as an underlying cause to disease, it may be warranted to continue efforts of validation for those CNV with a GC content greater than 56% and involving a gene structure.

Additionally, because our PCR based strategy assumes a tandem orientation of duplicated sequence, it is possible that the false call rate is inflated by our inability to confirm non-

tandem duplications. Overall, our optimized protocol reduces the false call rate for deletions from 60% of all deletions to 21% of all deletions irrespective of GC content.

While our false call rate may seem 'high', it is not inconsistent with reports in the literature. [12, 13] Greenway *et al* reported a false call rate of 66% for CNV greater than 20 kilobases (kb). Their validation strategy utilized the multiplex ligation-dependent probe amplification (MLPA), so those duplication events not in tandem would be positively identified.[12] Given our CNV data are much smaller (median size is approximately two kb), we observe a similar false call rate as for larger CNV.

2. Copy Number Variation in Autistic and Normal Populations

In total, four studies of 320 samples on two different aCGH resolutions by two different protocols were conducted. Our autistic populations were derived from the Autism Genetic Resource Exchange (AGRE, multiplex) cohort and the Simons Simplex Collection (SSC, simplex), and our normal DNA came from the National Institute of Mental Health (NIMH) control population ascertained for the Human Genetics Initiative. We used the 385K (385,000 oligonucleotide probes) and the 2.1M (2,100,000 oligonucleotide probes) CGH arrays from Roche NimbleGen, and processed both arrays by the manufacturer's recommended protocol as well as developing an optimized protocol for the 2.1M array.

a. Chromosome X CNV in individuals with autism

We identified 2,300 CNV and developed 146 validation assays. The majority of assays identified a false call (false call: 100 or 68%, true call: 46 or 32%). However, the caveat remains that our validation strategy assumed tandemly duplicated sequence. The possibility remains that a subset of those validation assays that failed to capture a junction fragment from a duplication call may simply be a failure of validation strategy not a failure of the array.

A genotyping assay was developed to characterize the effect of a deletion in identified in two individuals with autism on the immediate downstream gene, *GRIA3*, the gene identified as causing X-linked Mental Retardation (OMIM: 305915, MRX94).[14] While an increase due to loss of sequence was observed, genotyping in an expanded population with autism (n=364) and a normal population (1,799) revealed enrichment for this deletion in the population with autism at a statistically non-significant frequency ($p < 0.5531$).

Two validated loci remain compelling for further follow-up studies. First, familial and case and control genotyping for a 2.5 kb intronic deletion of the *FERM and PDZ domain containing* 4 gene (*FRMPD4*) remains to be conducted. This gene is highly compelling as a candidate gene due its regulatory function of excitatory synaptic transmission and its demonstrated interaction with postsynaptic density protein, PSD95. Secondly, a *de novo* deletion (AU016.3) about 175 kb downstream of *gastrin-releasing peptide receptor* (*GRPR)* has been identified. *GRPR* has been found to be disrupted in an individual with autism. While this CNV does not overlap with any gene structure its proximity to *GRPR*

and loss of sequence could effect expression. Again, genotyping of this variant locus in families and cases and controls remains to be conducted.

The remainder of our validated set of copy number changes is not immediately suggestive of a role in ASD. However, further research remains that might indeed unveil one or more of the variants identified as a susceptibility locus. Assessment of genotype frequency in both ASD and normal populations would help to prioritize variants for further characterization. Notably, those variants found to be enriched in the population of males with ASD over that of controls may loci involved with the manifestation of ASD.

While further validation and molecular study of our CNV data remains, we can conclude a few points from the data set that we have. Before discussing the specifics of our data, we wish to clarify our understanding and confidence in the larger CNV identified in our data set. Our experience is based on X chromosome CNV studies of two individuals manifesting developmental delay that were consistent with loss or disruption of the *fragile X mental retardation* 1 gene (*FMR1*), the gene underlying Fragile X Syndrome, the most common form of inherited intellectual disability. The first study was of a patient known to have a partial deletion of the *FMR1* gene.[15] This deletion began upstream of the first coding exon and ended before the eighth exon in *FMR1*. The breakpoints had not yet been determined, and this sample served as an ideal positive control for our CGH studies. We successfully identified the deletion boundaries by the 385K CGH array, and breakpoint sequencing determined the full size and location of the deletion (chrX: 146,703,942 - 146,820,448 (Hg18); 116,506 bp). The second study arose from Brad

Coffee's suspicion of a mosaic deletion involving *FMR1*.[16] Using the 2.1M CGH

arrays, we identified a 1,013,394 bp deletion that extended from chrX: 147,047,696 -

146,047,696 (Hg18) and included the *FMR1* gene as well as the next downstream gene,

FMR1NB. We use these two studies of CNV on the X chromosome as a measure of our

ability to accurately identify CNV greater 100 kb in size. To further support this

confidence threshold, we did identify and validate a 350 kb duplication in an AGRE

sample. First pass breakpoint sequencing suggests this duplication event is 345,617 bp in

size.

One of the first conclusions we can make from our data is regarding our large CNV.

From the CNV data collected from the 164 AGRE and SSC samples, we identified but

did not validate three deletions and six duplications greater than 100 kb in size. We

believe that these CNV are real and that all variants larger than 100kb in size from all

samples successfully screened were identified. We point out several significant

observations in relation to testing our original hypothesis. While the large duplication and

deletion variants (greater than 100kb) are likely to be real, we did not develop a

genotyping assay for these CNV. Thus, we were unable to compare the frequency data

for such structural changes among males with ASD and those of normal controls (NIMH

neuropsychiatric control samples and unaffected fathers of our autistic cohorts). Due to

their size, it is not unreasonable to hypothesize that these structural disruptions (three

deletions, six duplications) may contribute to ASD in many ways including but not

restricted to altering expression of dosage sensitive genes, altering usage of enhancer or

repressor loci, abrogating or creating species specific loci integral to gene and/or genomic function in cell processes (e.g., Matrix Assembly Regions).[17]

The largest variant identified was a 344 kb, and we can say with high confidence that there were no additional variants as large or larger within our data set. This suggests that large(r) structural changes on the X chromosome do not play a role in males with ASD. In retrospect, this may not be so surprising as one might expect that such changes in a hemizygous context (46, XY) may result in a more extreme phenotype than ASD and thus not be observed in the AGRE or SSC sample set.

A second point of discussion of our data is regarding the other end of the spectrum: the small CNV that comprise the majority of our data set. It is well known that oligonucleotide CGH arrays exhibit 'noisier' log2 values than their BAC-based predecessors and require more probes per variant to make a reliable call.[18-21] As described earlier, the 'streaking' sometimes generated by the optimized protocol likely contributed to more log2-based noise in our data as well. These two factors likely contributed to the high number of calls in the five kb and under range (84% of all calls). Unlike the variants greater than 100kb, we have less confidence in calls of this size.

However, we also found that the GC content of a given variant is a strong indicator for whether an aCGH identified CNV is likely a false or true positive call. Following either NimbleGen's protocol or our own optimized protocol, we found that all true positive calls made thus far had a GC content of 56% or less. While the optimized protocol did

improve our overall ability to identify true variants under this threshold, over half of all

calls with a GC content less than 56% are still likely to be false positives (i.e., at GC

cutoff of 56%: 59 false calls, 52 true calls). However, a GC cutoff of 56% also captures

nearly the entirety of the X chromosome (98%). A more stringent percent GC cutoff,

48%, may be warranted. This threshold shifts our true to false call proportion to 21 false

calls and 48 true calls while capturing 91% of chromosome X.

With these caveats in mind, 265 CNV less than five kb in size remain to be validated

from the AGRE and SSC study, and 121 (46%) of these have a GC content less than or

equal to 56%. While there exists a less than random chance that these variants will

validate, the possibility still remains that CNV at any or all of these loci may play a role

in autism susceptibility. Additionally, three of the 20 genes, *GABRQ, SYP,* and *ZNF4,*

overlap with this data set are from our original list of candidates.

Finally, our breakpoint junction analysis suggests that fewer motifs play a role in

resolving structural events then previously implicated. We tested 565 breakpoints for an

enrichment of 39 motifs previously identified as associating with structural breaks. The

eight motifs were found to be enriched in these breaks as compared to a randomly

generated data set (Human hypervariable minisatellite core sequence, DNA polymerase α

frameshift hotspots, DNA polymerase β frameshift hotspots, AT-rich signal, Curved

DNA signal, Topoisomerase II consensus cleavage site (Drosophila), Murine LTR

recombination hotspot, Nonamer recombination signal). We expected to identify more

motifs associated with the breakpoints as the 39 motifs selected had been previously

implicated in playing a role in resolving structural events. However, after multiple test correction, only these eight continued to remain significant. While it is possible other motifs are also involved in the breakpoint resolution of sequence gains and losses, these appear to play a more prominent role.

b. CNV identification in 102 normal males

Using the X-chromosome 2.1M CGH array and following NimbleGen's protocol, we were able to collect data from 101 of 102 individuals. We identified 1,231 CNV covering (12.2 CNV/person) with a median CNV size of 2,015 bp.

While it is difficult to discern much from the NIMH data without having validated any of these loci, there are a few conclusions that can be drawn. First, the size distribution of CNV identified within the NIMH population is consistent with those reported in the literature (and curated and stored in Database of Genomic Variants (DGV)), as well as the distribution observed in the AGRE samples processed by the NimbleGen. Both the NIMH cohort and the 50 AGRE samples by this method demonstrated a CNV rate of 12.2 and 12.1 CNV/person respectively. The median sizes for both sets were somewhat similar with the NIMH median size of 2, 015 bp and the AGRE median size of 1,080 bp. The NIMH samples had slightly larger median sizes but distributions were quite similar to both the AGRE samples as well as those in the DGV.[22]

Secondly, the NIMH data show a marked increase of duplication events over deletion events as compared to the prior studies in our AGRE cohort by both NimbleGen's

protocol and the optimized protocol. Following NimbleGen's protocol, the

deletion:duplication ratio has shifted from 1.88 in the AGRE samples (2.1M, NimbleGen)

to 0.53 in the NIMH samples (2.1M, NimbleGen). This 3.5 fold change may be a

function of biology in that the AGRE samples manifest more deletion events than the

NIMH samples, it may be a function of sample source in how cell lines were maintained

and DNA extracted, or, this shift may be due to the change in chemistry of array

synthesis and processing reagents. While it is not unreasonable to suggest that the

genome is more tolerant to increases in copy number rather than loss of sequence, reports

in the literature are conflicting over the observed deletion:duplication rates. [23-26]

Validation studies of the NIMH CNV data set would greatly increase our understanding

of the reliability of CNV called by the NimbleGen protocol. These variants may similarly

show stratification of true and false call rates by GC content. This metric would then help

in the interpretation of those loci that remain unvalidated, and a more accurate biological

understanding of CNV in this sample population might be determined.

3. Why might we have not found loci involved with ASD?

a. The technological limitations of array CGH

Array Comparative Genomic Hybridization (aCGH) has been used with success in

multiple studies, on different manufactured platforms, interrogating different regions of

the genome, and using different sources of DNA. It thus seemed reasonable to use the X

chromosome as our target region of interest to interrogate our cell line derived DNAs

from the AGRE and Simons cohorts.

In 2006, NimbleGen, now Roche NimbleGen, developed a novel method in which oligomer probes were synthesized by a photo-mediated chemistry on a glass substrate. This mask-less approach allowed for the dynamic building and synthesis of an array in a more cost-effective fashion. Any oligonucleotide-based array could be synthesized by this method, including arrays for Comparative Genomic Hybridization. As my thesis involved the identification of structural changes on the X chromosome, we chose to use NimbleGen's cost effective and flexible array platform for this purpose. NimbleGen agreed to train me in every step of their aCGH protocol at their labs in Madison, WI. My training started from the first step of assessing sample integrity and proceeded through to array analysis of copy number variation. As our relationship and commitment to the NimbleGen CGH technology progressed, we were also permitted early access to their new custom High Density arrays in which 2.1 million probes could be designed for any region(s) of interest. We shifted my fine scale exploration of X chromosome structural changes to an even finer resolution when we moved our platform design from the 385,000 probes of the 385K platform to the 2.1 million probes dedicated to the X chromosome.

Early validation efforts of our custom 2.1M CGH arrays indicated a high false positive rate (65%) that spurred us on to explore methods by which we could reduce this rate and increase the integrity of calls made by this platform. During the course of our own experimentation with NimbelGen's aCGH protocol, NimbleGen was conducting optimization efforts as well. Within several years time, we observed a shift in array behavior as compared to the earlier arrays we had been using. Specifically, when we

applied the same optimized protocol that we had previously developed to arrays

manufactured at a later date (a year or more), we found that the 'newer' arrays

underperformed (i.e., overall reduced signal across arrays, reduced call rates, increase in

noise) as compared to those on which we had conducted our earlier optimization

experiments. Discussions with NimbleGen technical and marketing representatives

revealed that they had indeed modified the array synthesis chemistries.

Concurrent with our optimization work, we also observed several revisions of

NimbleGen's recommended sample processing protocol. While the newer protocol

versions were more streamlined and user friendly, they were also less transparent.

NimbleGen may have been using the same chemical component and ratios as in earlier

versions, but it was not clear whether the noticeable change in array response was wholly

due to the improved array synthesis chemistry, to any changes in the sample processing

reagents, or a combination of both.

Our final study of 102 individuals from the NIMH control population used the latest

2.1M CGH arrays and was conducted using the latest reagent formulas following

NimbleGen's protocol. As discussed previously, we identified 1,231 variants with a

median size of 2,104 bp in size. While some of these variants may validate, we anticipate

a great many more are falsely called based on the number and size of variants identified

in this population. While these data have not been validated, given our experience with

this platform, we are inclined to believe that NimbleGen's improvements to array

manufacture and sample processing have not minimized the false call rates we had

previously observed.

One of our first efforts to increase array performance was to evaluate kinetics of the

probes on the custom arrays. Using the log2 values as a marker for probe behavior, we

identified a small subset of probes on the HD arrays as having an abnormally high

variance. Specifically, 5% of all experimental probes on the array showed a variance

greater than 0.175 across 50 arrays. While we were able to empirically determine that this

subset of probes was likely contributing noise and not signal to our analysis, it is possible

that the overall probe selection algorithm utilized by NimbleGen needed further

optimization. Perhaps, with less reliance on NimbleGen and increased effort on our part

to evaluate probes before array synthesis, we may have been able to develop a set of

probes that were better adapted to the unique characteristics of the X chromosome (e.g.,

overall repeat content, GC rich regions).

In the application of aCGH, oligonucleotide arrays are known to be much noisier than

BAC arrays. [18-20] While oligomer based arrays can interrogate regions on a much finer

scale than BAC-based arrays, more probes (typically, 10-20 oligonucleotide probes or

more versus 1-3 BAC probes) are required to call variants with statistical confidence.[27]

Additionally, it is possible that aCGH tiling arrays such as our high-density arrays,

irrespective of manufacturer (ex: Agilent, Illumina, Roche NimbleGen), are inherently

more noisy (i.e., increased standard deviation of log2 values across all probes) than their

reduced coverage predecessors. As probe density increases so does overall 'noise' of the

experiment. While it has been noted that genome wide aCGH evaluation has been used with great success, variants identified in such studies were orders of magnitude larger (e.g., 50 kb - 200kb) than the bulk of those loci identified as variant by our custom arrays (less than five kb). Simply said: larger variants were called more often with greater success. A recent study by Sanders *et al* explored *de novo* CNV in autism utilizing two Illumina arrays to assess copy number. In initial attempts to determine their '*de novo* prediction thresholds', they found a 47% false call rate for variants identified with 20 or more probes or an average size of 60 kb or greater. A majority (82%) of the failures were false positive. The authors compared their data set where it overlapped with another study that had utilized NimbleGen's 2.1 million arrays for genome wide CNV detection (42 million probes) to validate their findings. Again, a minimum criterion of 20 probes or 60 kb on average was used, and 55/58 (95%) rare, de novo CNV were identified by both platforms. However, analysis of rare, de novo CNV *less than* 20 probes was not nearly as positive for the 31 CNV called between both studies. Only 4/31 (13%) CNV were identified by both platforms. These data suggest that our efforts to identify fine-scale structural variation, or variants less than even 25 kb in size, would have been plagued with multiple false calls no matter the platform or rigor used.

It appears that aCGH is a promising tool and has been advanced to the level of whole genome analysis with perhaps reliable performance limited to structural variants greater than 60 kb**.** However, for the purposes of screening a genome for variants smaller than this, alternative technologies such as those offered by next-generation sequencing (NGS) platforms may be the more suitable alternative. Technically, NGS can capture breakpoint

or junction sequence, however, mining for those junction sequences captured by the experiment is highly challenging. However, a few recent studies report successful application and identification of breakpoints ranging from large chromosomal breaks to smaller CNV. [28-30] For the purposes of furthering our own work, a NGS platform could be harnessed to identify the sequence breaks of structural changes. Re-evaluating the X chromosome of our sample population could identify insertion/deletion events (less than one kb), nucleotide substitution, inversion events, as well as the structural changes that we had hoped to identify comprehensively. Furthermore, NGS is capable of identifying sequence changes from one base to megabases encompassing the detection range of both BAC- and oligonucleotide-based arrays.

b. Consistency of sample phenotyping in autism

Despite the technical logistics that went into evaluating fine-scale structural variation on the X chromosome, our net findings left few if any compelling loci as being possibly involved with ASD. Our initial hypothesis leading to the screening of chromosome X specifically was based on compelling evidence for such X-linked loci. Further, such loci were hypothesized to confer a highly penetrant contribution to disease. It is possible that the samples used were themselves not appropriate for this type of experimental hypothesis and strategy.

Our sample selection strategy encompassed the entirety of the AGRE collection, and selection was made in a step-wise fashion to select for the most severely affected individuals (e.g., families containing only affected males, transmitted through females in

extended pedigrees, and, diagnosed with ASD by both ADI-R and ADOS). The majority of the pedigrees selected were nuclear in structure. Additionally, extensive prenatal and familial histories were collected. Chromosomal analysis and molecular testing for Fragile X Syndrome were conducted to exclude those pedigrees having individuals with an identifiable chromosomal or molecular alteration that would explain the observed ASD traits. The AGRE collection is primarily comprised of multiplex families and represented the largest portion of our sample population (100/164 samples assessed by the chromosome X CGH arrays with an additional 200 samples for genotype frequency studies); all samples were obtained from unrelated individuals. The multiplex nature of this cohort allowed us to select for pedigrees likely segregating a disease locus on the X chromosome.

For any genetic study, phenotyping of the affected individual(s) as well as their family members is a critical aspect of pedigree selection. As in our two sample populations, AGRE and SSC, we chose those affected individuals identified as autistic, excluding those with "Broad Spectrum" or "Not Quite Autism". We speculated that individuals with a more mild presentation such as Broad Spectrum may have an underlying genetic lesion that is a more subtle change compared with the deletions or duplications for which we were screening. For example, a mis-sense mutation that leaves a gene product intact but with diminished functionality might explain the more modest phenotype. We also excluded families in which notes suggesting autistic-like phenotypes were present in 'unaffected' members as some of these may represent families segregating for autosomal genes instead of X-linked ones.

However, we also observed some inconsistencies in the phenotyping of affected individuals. An example comes from the AGRE pedigree AU0983. In this family, the three boys are scored as having 'Autism'. One of the siblings (not the proband we used for our studies) was listed as 'not Spectrum or Autism' in the ADOS category. We understand the 'score' to be derived from the ADI-R test. The other two brothers are listed as having 'Autism' both by score and by ADOS. This is simply an example of an inconsistency that, while not frequent, suggests that phenotyping of individuals for autism is not nearly as straightforward a process one might suspect.

In effort to find a genetic basis for disease, the researcher must first depend on the phenotyping capabilities of clinicians and counselors in the collection of the diseased sample population. The goal is to identify a locus that has been disrupted in some form and shared among individuals with a common disorder. Depending on search strategy, that susceptibility locus is shared more often among cases than controls or segregates with disease within families. This search is based on the assumption that the underlying genetic lesion is responsible for the phenotype observed. However, if a patient's phenotype differs from clinician to clinician, then the underlying genetic contribution to that individual's particular traits might be different than that of another patient having the same diagnosis.

This scenario is not all that unlikely. Our study samples were derived from the AGRE and SSC collections of families affected with ASD. Affected individuals are evaluated by

what are currently considered the gold standards for ASD diagnosis, the Autism

Diagnostic Interview, Revised (ADI-R) and/or the Autism Diagnostic Observation

Schedule (ADOS). [31, 32] Cathy Lord developed both of these autism diagnostic

metrics. Earlier this year, Lord reported ASD diagnostic variability across 12 different

sites despite clinical training and correct implementation of ADI-R and ADOS.

Specifically, children that met both ADI-R and ADOS criteria for ASD did not share the

same final diagnostic conclusions. [33]

A second possibility for our failure to clearly identify autism susceptibility loci was that

our sample size was not sufficiently powered to identify a locus/loci that may contribute

to disease. Because we initially studied multiplex families, we presumed that we would

find inherited structural changes within one pedigree and that might be found with some

repeat occurrence in others. While such a CNV event could be rare and private to a single

family, the locus implicated might be disrupted in additional, different forms (other CNV

throughout the gene body, coding sequence changes, or aberrant methylation) in other

families affected with ASD. We could then use a normal population to compare genotype

frequencies (ASD versus control). For a CNV mediating autistic traits, we might find two

possible outcomes to CNV genotyping. First, we might find that a CNV occurs rarely or

only in our population with ASD. Or secondly, we might find enrichment or increased

frequency of occurrence of the given genotype in our ASD. For example, using our

sample of 300 males with ASD and 1,800 controls (1,500 NIMH controls as well as 300

fathers), assuming a minimum frequency of 2% in our control population and 80%

power, we would need to observe a frequency of 5.2% or more in our case population in

order to claim a statistically significant difference at a level of 0.05. The minimum

frequency observed for our ASD population stays about the same whether we include the

fathers as controls or not (5.3% minimum). However, a 5% frequency is quite high and

suggests a relatively common genetic change. While a locus may play a role in disease, it

does not necessarily have to share the same mechanism of mutation. Our genotyping is

based on a copy number change for a given locus, but a gene may be disrupted by

changes in coding sequence as well as by other modifiers (e.g., sequence based -

enhancers, repressors, post-synthesis alterations - methylation). But, by simply assaying

for copy number change, in our study we are requiring that more affected individuals

carry the same mutation in order to say with statistical certainty that a given locus is

involved with ASD.

An increase in our sample size would benefit our study design on a few different levels.

First, the power due to genetic heterogeneity introduced by phenotypic inconsistencies

might be minimized, as the greater the pool of individuals studied, the greater the chance

of selecting individuals who share the same susceptibility locus. Secondly, an increase of

the sample size of males with ASD would increase the power to detect a difference in

frequency of an observed structural change in our case/control comparisons. For

example, to detect a 1% increase (i.e., a 3% frequency in our ASD sample) as statistically

significant over a normal population frequency of 2% with 80% power, we would need

4,000 unrelated males with ASD as well as controls.

4. Alternative hypotheses that may explain our findings

A possible explanation for our lack of identified causal variants might lie in alternative explanations to the assumptions of the underlying genetic lesion(s). We hypothesized that a locus/loci resides on the X chromosome and defects in this locus/loci strongly contribute to ASD in individual carriers as well as the male bias in these disorders. While our hypothesis is based on empirically determined findings, it also presumes that a single locus largely contributes to ASD in the patient (monogenic), is highly penetrant, and is homogenous in its expressivity of the lesion(s). These assumptions are not unreasonable given that the X chromosome was selected to exist in a hemizygous state in our sample set (i.e., no Klinefelter's Syndrome or other sex chromosome anomalies were identified in our study population), and structural lesions on this chromosome would likely confer a more severe molecular and thus phenotypic response.

a. Our cohorts are comprised of a genetically heterogeneous population

We built our experimental design on the assumption that, in at least a subset of our cases, we could identify a highly penetrant, monogenic locus on the X chromosome that had been structurally altered and might therefore molecularly explain ASD. However, we also anticipated more than one locus was likely to play a role in ASD given how genetically heterogeneous ASD has already been demonstrated as being. With as much as 30% of all ASD bearing a known molecular etiology representing a minimum of 44 different genomic loci and 103 different genes, initially, we may have been able to predict or gauge how many loci we might reasonably expect to find on the X chromosome in a given sample size. [34-38]

b. Observed CNV may have reduced penetrance in a normal population

As the body of literature on structural changes in ASD grows, reduced penetrance of structural lesions has been used to explain observed deletion and/or duplication events that occur in both normal and affected populations. Under the model of a reduced penetrant genetic lesion, a structural change identified in an ASD cohort has an enriched frequency of occurrence as compared to the frequency of occurrence in a normal or phenotypically non-ASD control group. Additionally, this change can also be inherited from non-affected parents. One example of such a locus is 16p11 where recurrent microdeletions and microduplications have been identified and implicated as causal in ASD. Since its first identification in 2007, multiple studies of ASD cohorts have found structural changes at this locus to be significantly enriched in cases over controls.[38-42] However, Glessner *et al* found structural variants at this locus to be equally frequent in their ASDcases and controls. [43] While the findings of Glessner *et al* are valid, as in all studies, frequencies are based on the respective patient and control populations studied. The accumulating reports of structural changes at 16p11 enriched in ASD populations suggest a legitimate role for 16p11 gene(s) in ASD disorder.

If we were to relax our requirement for 'highly penetrant' genetic lesions, several variants identified in our AGRE and SSC studies would now be eligible for further examination. For example, we identified three duplications that overlapped the *synaptophysin* (*SYP*) gene, but we did not validate because they overlapped with other reported structural changes identified among those with different phenotypes. *SYP* encodes a membrane protein that localizes to small synaptic vesicles in brain and endocrine cells and has been

identified as the gene responsible in X-linked Mental Retardation (OMIM: 300802, MRX96) in four different pedigrees.[14] Currently, a 43kb and a 42kb inversion fully encompass *SYP* and as well as upstream neighboring genes and a 1.3 Mb deletion that includes *SYP* as well as *FTSJ1* (OMIM: 309549, MRX9 and/or MRX44) have been identified in normal populations. While an inversion event might change the genomic context in which a gene lies (*i.e.*, change expression pattern due to altered exposure to local enhancers and repressors), the deletion event might be considered more severe, especially in that it was identified in a supposedly normal male. No karyotypic information was made available for this individual, so this male could possibly be XXY. However the study utilized aCGH arrays with a female reference, and an increased X chromosome dosage should have been readily seen and presumably reported by the authors. Given that two X-linked Mental Retardation genes are encompassed in the deletion, it might not be unreasonable to either suspect this deletion call as a false positive or, at the least, a reduced penetrant lesion.

Our duplications are novel in that they directly involve the *SYP* gene. The three duplications overlap significantly and involve a substantial portion of the 3'UTR on one isoform and the last intron (intron 6) of the second isoform. (Figure 4.1) One could hypothesize that if these duplications validated, they may also introduce a subtle change to the function and expression of this gene product in our males with ASD distinct from that of a complete loss or complete inversion of this locus.

c. CNV identified in our cohorts may require additionally altered loci for a manifestation of autism

Finally, perhaps our assumption of a singe susceptibility locus on the X chromosome needs to be expanded to allow for the broader, multigenic model for ASD etiology.[44-51] Under an oligogenic model, a disruption of a single ASD susceptibility locus on the X chromosome may be required but additional genetic liabilities elsewhere in the genome are necessary for phenotypic expression of ASD. Within the context of a multiplex family, both parents would carry susceptibility alleles that result in ASD when they co-exist in a single individual. While a similar genetic set-up may exist in a simplex family, it may be more likely that one parent may carry such a susceptibility allele, and a secondmutation at another critical locus is required for manifestation of autism. Considering a multigenic model, those CNV that do not segregate with autism in our multiplex (or even simplex) families may still play a role in these disorders as the required '2nd hit' is elsewhere in the genome.

d. Partial inactivation of X chromosome loci can act as susceptibility loci to ASD in males.

Skuse et al suggest that there exists on the X chromosome an imprinted locus. When inherited from mothers, women with Turner Syndrome (45, X) display an impaired social and cognitive phenotype distinct from women with TS who inherit their father's chromosome X. Such a locus or loci would inherently cause males to be more susceptible to these social and cognitive phenotypes as males are hemizygous for the X chromosome.[52, 53] Because all males inherit the maternal X chromosome, this fact

159

may place males at an inherently higher risk than females for the ASD. Should a susceptibility locus of the inherited X be inactivated, this might result in an autistic phenotype. The mechanism of mutation is mediated by methylation-based gene silencing. Altered methylation regulation has already been described in a subset of females with autism (OMIM: 300496, autism susceptibility). [14, 54] The alterations and loss of the *methyl-CpG-Binding protein, MECP2* gene are know to cause Rett Syndrome, a regressive form of autism (OMIM: 312750, Rett Syndrome). [14]

5. Final Summary

In conclusion, we found no evidence to support our hypothesis that there are large structural variants on the X chromosome that are highly penetrant and have a causal role in ASD. However, we may very well have found susceptibility loci, but we are currently unable to discern a role for those loci due to difficulty validating CNV, as well as possibly being underpowered to identify susceptibility loci due to genetic heterogeneity and to the rarity of the variant. Additionally, we may not have sufficient molecular understanding of cellular processes to appreciate a role for a given locus and ASD, so a rare CNV found only once in our data set may indeed be causal in ASD. There may be human specific regulatory regions unidentified by conservation that are critical for 'normal' social and behavioral development (e.g., microRNAs, enhancers, repressor).

Despite what CNV we have and have not found, CNV remain to be further interrogated for validity and possibly functional involvement in ASD. Additionally, broadening our genotyping efforts to include all pedigrees from the AGRE (861 multiplex pedigrees) and

SSC cohorts would increase our power to detect an autistic locus that has reduced penetrance. Incorporating sequencing for coding changes of these loci would add to our power to identify a new X-linked autistic locus. While our hypothesis remains valid, it may be that search for a CNV based mechanism may not be the best mutational mechanism by which to capture new ASD loci. Identification of such loci may require a more subtle change such decreased or increased expression as those mediated by methylation changes or even subtle coding changes that do not abrogate but change gene product function.

6. References

1. *Prevalence of autism spectrum disorders - autism and developmental disabilities monitoring network, 14 sites, United States, 2008.* MMWR Surveill Summ, 2012. **61**(3): p. 1-19.
2. Boyle, C.A., et al., *Trends in the prevalence of developmental disabilities in US children, 1997-2008.* Pediatrics, 2011. **127**(6): p. 1034-42.
3. Kogan, M.D., et al., *Prevalence of parent-reported diagnosis of autism spectrum disorder among children in the US, 2007.* Pediatrics, 2009. **124**(5): p. 1395-403.
4. Yeargin-Allsopp, M., et al., *Prevalence of autism in a US metropolitan area.* Jama, 2003. **289**(1): p. 49-55.
5. Stone, J.L., et al., *Evidence for sex-specific risk alleles in autism spectrum disorder.* Am J Hum Genet, 2004. **75**(6): p. 1117-23.
6. Gauthier, J., et al., *Autism spectrum disorders associated with X chromosome markers in French-Canadian males.* Mol Psychiatry, 2006. **11**(2): p. 206-13.
7. Noor, A., et al., *Disruption at the PTCHD1 Locus on Xp22.11 in Autism spectrum disorder and intellectual disability.* Sci Transl Med, 2010. **2**(49): p. 49ra68.
8. Piton, A., et al., *Systematic resequencing of X-chromosome synaptic genes in autism spectrum disorder and schizophrenia.* Mol Psychiatry, 2011. **16**(8): p. 867-80.
9. Vincent, J.B., et al., *Genetic linkage analysis of the X chromosome in autism, with emphasis on the fragile X region.* Psychiatr Genet, 2005. **15**(2): p. 83-90.
10. Talebizadeh, Z., et al., *Brief report: non-random X chromosome inactivation in females with autism.* J Autism Dev Disord, 2005. **35**(5): p. 675-81.
11. Thomas, N.S., et al., *Xp deletions associated with autism in three females.* Hum Genet, 1999. **104**(1): p. 43-8.
12. Greenway, S.C., et al., *De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot.* Nat Genet, 2009. **41**(8): p. 931-5.
13. Malhotra, D., et al., *High frequencies of de novo CNVs in bipolar disorder and schizophrenia.* Neuron, 2011. **72**(6): p. 951-63.
14. Medicine, M.-N.I.o.G. *Online Mendelian Inheritance in Man.* 1987; Available from: omim.org/.
15. Gu, Y., et al., *A de novo deletion in FMR1 in a patient with developmental delay.* Hum Mol Genet, 1994. **3**(9): p. 1705-6.
16. Coffee, B., Ikeda, M, Budimirovic, DB, Hjelm, LN, Kaufmann, WE, and Warren, ST, *Mosaic FMR1 Deletion Causes Atypical Fragile X Syndrome and Can Lead to Molecular Misdiagnosis: A Report and Review.* American Journal of Medical Genetics. Part A, in press.
17. Freeman, J.L., et al., *Copy number variation: new insights in genome diversity.* Genome Res, 2006. **16**(8): p. 949-61.
18. Lockwood, W.W., et al., *Recent advances in array comparative genomic hybridization technologies and their applications in human genetics.* Eur J Hum Genet, 2006. **14**(2): p. 139-48.
19. Itsara, A., et al., *Population analysis of large copy number variants and hotspots of human genetic disease.* Am J Hum Genet, 2009. **84**(2): p. 148-61.

20.   McCarroll, S.A., et al., *Integrated detection and population-genetic analysis of SNPs and copy number variation.* Nat Genet, 2008. **40**(10): p. 1166-74.

21.   Zhang, Z.F., et al., *Detection of submicroscopic constitutional chromosome aberrations in clinical diagnostics: a validation of the practical performance of different array platforms.* Eur J Hum Genet, 2008. **16**(7): p. 786-92.

22.   Iafrate, A.J., et al. *Database of Genomic Variants.* 2004 Nov 02, 2010; Available from: projects.tcag.ca/variation/.

23.   Baross, A., et al., *Assessment of algorithms for high throughput detection of genomic copy number variation in oligonucleotide microarray data.* BMC Bioinformatics, 2007. **8**: p. 368.

24.   Locke, D.P., et al., *Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome.* Am J Hum Genet, 2006. **79**(2): p. 275-90.

25.   de Smith, A.J., et al., *Array CGH analysis of copy number variation identifies 1284 new genes variant in healthy white males: implications for association studies of complex diseases.* Hum Mol Genet, 2007. **16**(23): p. 2783-94.

26.   Kidd, J.M., et al., *Mapping and sequencing of structural variation from eight human genomes.* Nature, 2008. **453**(7191): p. 56-64.

27.   Conrad, D.F., et al., *Origins and functional impact of copy number variation in the human genome.* Nature, 2010. **464**(7289): p. 704-12.

28.   Sobreira, N.L., et al., *Characterization of complex chromosomal rearrangements by targeted capture and next-generation sequencing.* Genome Res, 2011. **21**(10): p. 1720-7.

29.   Huefner, N.D., et al., *Breadth by depth: expanding our understanding of the repair of transposon-induced DNA double strand breaks via deep-sequencing.* DNA Repair (Amst), 2011. **10**(10): p. 1023-33.

30.   Sun, R., et al., *Breakpointer: using local mapping artifacts to support sequence breakpoint discovery from single-end reads.* Bioinformatics, 2012.

31.   Lord, C., et al., *Autism diagnostic observation schedule: a standardized observation of communicative and social behavior.* J Autism Dev Disord, 1989. **19**(2): p. 185-212.

32.   Lord, C., M. Rutter, and A. Le Couteur, *Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders.* J Autism Dev Disord, 1994. **24**(5): p. 659-85.

33.   Lord, C., et al., *A multisite study of the clinical diagnosis of different autism spectrum disorders.* Arch Gen Psychiatry, 2012. **69**(3): p. 306-13.

34.   Muhle, R., S.V. Trentacoste, and I. Rapin, *The genetics of autism.* Pediatrics, 2004. **113**(5): p. e472-86.

35.   Freitag, C.M., *The genetics of autistic disorders and its clinical relevance: a review of the literature.* Mol Psychiatry, 2007. **12**(1): p. 2-22.

36.   Abrahams, B.S. and D.H. Geschwind, *Advances in autism genetics: on the threshold of a new neurobiology.* Nat Rev Genet, 2008. **9**(5): p. 341-55.

37.   Kumar, R.A. and S.L. Christian, *Genetics of autism spectrum disorders.* Curr Neurol Neurosci Rep, 2009. **9**(3): p. 188-97.

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

39. Kumar, R.A., et al., *Recurrent 16p11.2 microdeletions in autism.* Hum Mol Genet, 2007.

40. Weiss, L.A., et al., *Association between Microdeletion and Microduplication at 16p11.2 and Autism.* N Engl J Med, 2008.

41. Nord, A.S., et al., *Reduced transcript expression of genes affected by inherited and de novo CNVs in autism.* Eur J Hum Genet, 2011. **19**(6): p. 727-31.

42. Fernandez, B.A., et al., *Phenotypic spectrum associated with de novo and inherited deletions and duplications at 16p11.2 in individuals ascertained for diagnosis of autism spectrum disorder.* J Med Genet, 2010. **47**(3): p. 195-203.

43. Glessner, J.T., et al., *Autism genome-wide copy number variation reveals ubiquitin and neuronal genes.* Nature, 2009. **459**(7246): p. 569-73.

44. Risch, N., et al., *A genomic screen of autism: evidence for a multilocus etiology.* Am J Hum Genet, 1999. **65**(2): p. 493-507.

45. Geschwind, D.H. and P. Levitt, *Autism spectrum disorders: developmental disconnection syndromes.* Curr Opin Neurobiol, 2007. **17**(1): p. 103-11.

46. Happe, F., A. Ronald, and R. Plomin, *Time to give up on a single explanation for autism.* Nat Neurosci, 2006. **9**(10): p. 1218-20.

47. Goodman, R., *Infantile autism: a syndrome of multiple primary deficits?* J Autism Dev Disord, 1989. **19**(3): p. 409-24.

48. Bailey, A., W. Phillips, and M. Rutter, *Autism: towards an integration of clinical, genetic, neuropsychological, and neurobiological perspectives.* J Child Psychol Psychiatry, 1996. **37**(1): p. 89-126.

49. Veenstra-Vanderweele, J., S.L. Christian, and E.H. Cook, Jr., *Autism as a paradigmatic complex genetic disorder.* Annu Rev Genomics Hum Genet, 2004. **5**: p. 379-405.

50. Cook, E.H., Jr. and S.W. Scherer, *Copy-number variations associated with neuropsychiatric conditions.* Nature, 2008. **455**(7215): p. 919-23.

51. Fanciulli, M., E. Petretto, and T.J. Aitman, *Gene copy number variation and common human disease.* Clin Genet, 2010. **77**(3): p. 201-13.

52. Skuse, D.H., *Imprinting, the X-chromosome, and the male brain: explaining sex differences in the liability to autism.* Pediatr Res, 2000. **47**(1): p. 9-16.

53. Skuse, D.H., et al., *Evidence from Turner's syndrome of an imprinted X-linked locus affecting cognitive function.* Nature, 1997. **387**(6634): p. 705-8.

54. Amir, R.E. and H.Y. Zoghbi, *Rett syndrome: methyl-CpG-binding protein 2 mutations and phenotype-genotype correlations.* Am J Med Genet, 2000. **97**(2): p. 147-52.
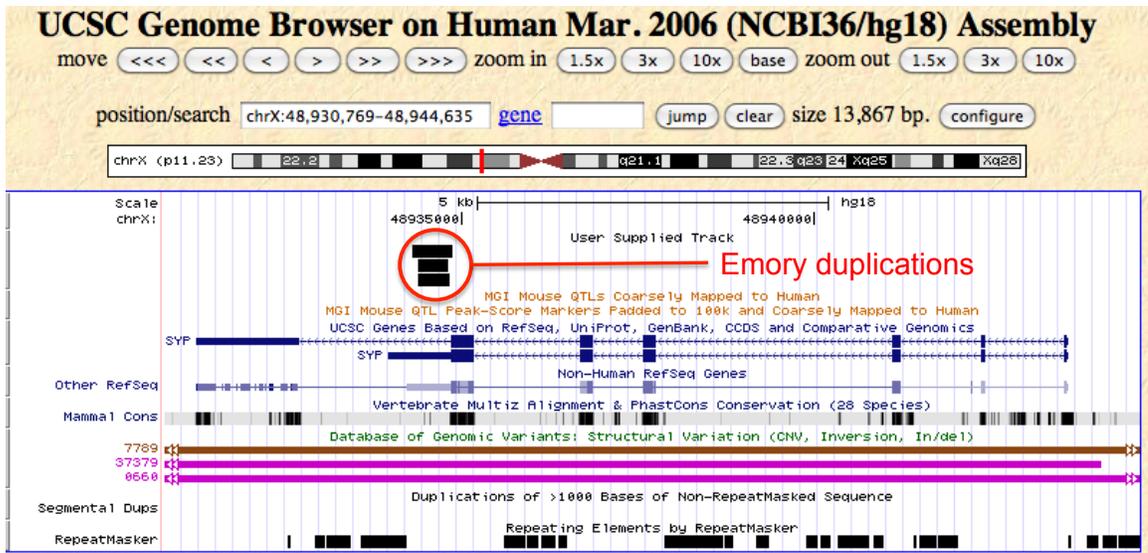
Figure 4.1 – Three duplication calls map to the *SYP* (X-linked Mental Retardation gene) gene.

Bars near bottom of image: Pink – inversion, brown – deletion

Chapter 5. Subjects, Materials, and Methods

1. Sample DNA

a. Autism Genetic Resource Exchange (AGRE)

The Autism Genetic Resource Exchange (AGRE, familyagre.org or research.agre.org) is

the product of the Autism Speaks scientific initiative, and houses an extensive collection

of biomaterials to be used for the study of autism spectrum disorder (ASD). At the time

we began our study, 949 families had been enrolled in their collection. AGRE consists

principally of multiplex families where affected individuals have been evaluated by one

or both of the current gold-standards for autistic diagnostic evaluation: the autism

Diagnostic Interview, Revised (ADI-R) and/or the Autism Diagnostic Observation

Schedule (ADOS) prior to enrollment.[1, 2] Extensive medical histories are collected

extending as far back as the conception and gestation of affected individuals. The racial

content of pedigrees we selected for CNV analysis is largely composed of families self-

identified as white, not Hispanic or Latino (60%), 20% as white Hispanic or Latino, and

the remaining 20% were listed as either Asian, African American, more than one race, or

unknown. We analyzed 300 pedigrees (300 probands, 300 fathers, and an assortment of

full families) from this collection.

DNA used in our studies was derived from immortalized cell lines.

b. Simons Simplex Collection (SSC)

Our second cohort of individuals with autism is derived from the Simons Simplex

Collection (SSC, sfari.org). The SSC houses biomaterials, phenotypic and clinical data

for families having only one child affected with ASD (simplex). Affected individuals are

evaluated by ADI-R and/or ADOS, and extensive family histories are taken. The ethnic

breakdown is similar to our first cohort in that a majority of samples are white, not

Hispanic (67%). The remaining twenty-one samples are a diverse representation of

African-American, Asian, more than one race, Native American, white Hispanic and

unknown ethnicities. We analyzed 64 samples from this collection.

DNA used in our studies was derived from immortalized cell lines.

c. National Institute of Mental Health (NIMH) control population

This cohort of Caucasian individuals was identified and collected by the National

Institute of Mental Health (NIMH) as part of a Human Genetics Initiative.

(nimhgenetics.org/available_data/controls/) The aim of this program was to create a

large biological resource of biomaterials and clinical data from individuals affected with

severe neuropsychiatric disorders (e.g., autism, schizophrenia, and bipolar are among the

disorders being targeted as part of this collection). In addition, a control set for this

program was collected for case and control studies. 3,828 individuals are in this group,

and they are derived from cord and venous blood. A normal phenotype was determined

from an "online, short self-report clinical assessment". We selected 1,500 males over the

age of 18 to be used for our studies.

We used DNA housed by this collection.

d. Array Comparative Genomic Hybridization (aCGH) Reference DNA: NA10851

We used the male CEPH sample NA10851 from the Human Genetic Cell Repository at

the Coriell Institute (www.coriell.org) for our sex-matched reference in our array

comparative genomic hybridization (aCGH) studies. We chose to use a sex-matched

reference as this strategy has been previously found to result in a better signal to noise

ratio, greater dynamic range of log2 values, and repeat sequences are more grossly

matched.[3] This male has been recommended for copy number studies in effort to enable

comparison of findings across studies. [4] Additionally, this DNA was used as a 'normal'

control for our PCR-based validation studies. It is often annotated as 'NA' in the PCR

protocols.

DNA from this individual is derived from a B-Lymphocyte cell line.

e. Validation control DNA

J1 and CM108 are two male DNAs that we used as additional controls for our validation

experiments. CM108 may be referred to as 'CM' in our PCR protocols. These DNAs are

housed in the Warren Cell Line collection, and DNA was derived from lymphoblasts.

2. AGRE Pedigree Selection

We selected families in a step-wise fashion in effort to identify those families likely

segregating an X-linked lesion. Each step or filter removes families because of a single

criterion.

Start: All pedigrees in collection (n=949)
Filter 1: Remove simplex families (n=861)
Filter 2: Remove families with at least one affected female (n=520)
Filter 3: Remove families with possible non-idiopathic diagnosis or chromosomal anomalies (n=493)
Filter 4: Remove affected individuals prematurely born + other notes (e.g., family reports autism/spectrum diagnosis of father/daughter) (n = 471)
Filter 5: Remove families with no parental DNA (n = 418)

Filter 6: Select families with ≥ two affecteds with first diagnosis by ADOS then ADI-R of 'Autism' or 'Spectrum' (n = 375)

Filter 7: Remove remaining families with clinically relevant notes; Remove if pedigree extends through male (i.e., brother-sister have affected children, 9 such pedigrees) (n = 359)

3. Database of Genomic Variants: CNV and indel analysis

Chapter 1. Figure 1A and Figure 1B.

Chapter 3. Figure 1.

We evaluated the Database of Genomic Variants (DGV, projects.tcag.ca/variation, build Hg18) by the number of references reporting in the DGV and the number of copy number variants (CNV, sequence having different copy number, greater than one kilobase (kb), benign or pathogenic) and indels (sequence having different copy number, 100 base pairs (bp) to one kb in size, benign or pathogenic) reported. In total, 42 different reports from the literature have contributed varying amounts of data to the DGV. All reports are assumed to be benign but validation and curation of reported variants is incomplete. There is overlap among sample populations, and various methodologies (e.g., BAC-based, SNP-based, aCGH, paired-end sequencing) were used to identify variants.

First, we categorized CNV and indels by chromosome and originating reference. Using this data, we were able to identify the percentage of references reporting CNV by chromosome ('green' line in Chapter 1, Figure 1A and Chapter 3, Figure 1). Second, we determined what percentage of all CNV and indel reported in DGV were found for each chromosome ('purple' line in Chapter 1, Figure 1A and 1B and Chapter 3, Figure 1).

4. Comparative Genomic Hybridization (CGH) Arrays

We utilized a two custom designed high-density array Comparative Genomic Hybridization microarrays (aCGH) manufactured by Roche NimbleGen. The first was a custom designed array that tiled 385,000 probes along the X chromosome with an average intermarker distance of 270 base pairs (bp) or 4 probes/kilobase (kb). Probes were photolythically synthesized on a glass slide. Probes were randomly distributed by position throughout the array area. This density array is no longer available for CNV detection.

The second custom CGH array we used tiled approximately two million probes (2,143,369 probes total, and 1,979,586 experimental probes) ranging from 50-75 bp in length along the X chromosome. (P/N 05223873001, nimblegen.com/products/cgh) 96.8 megabases (Mb) of unique sequence were targeted, and this resulted in a median spacing of one probe every 50 basepairs (bp). Prior to tiling-probe selection, the chromosomal sequence was first repeat masked, including masking of the PAR1 and PAR2 loci. This array is in fact comprised of three small arrays of photosynthesized probes built on a glass slide. In our custom micorarray, each 'subarray' represents one third of chromosome X in ascending order. Ie, subarry A01 contains probes from 2,709,520-47,369,192 bp, subarray A02 contains probes from 47,369,267-107,160,255 bp, and subarray A03 contains probes from 107,160,305-154,583,236 bp. While each subarray represents one third of the chromosome in sequential order, probes are randomized within the subarray. Ie, probes selected from the first third of the chromosome are physically randomized in (X,Y) position within the first subarray, the second third randomized within the second subarray, and the last third randomized within the third subarray.

5. aCGH Protocol and Scanning

a. NimbleGen's protocol for sample processing, array hybridization and wash,

and scanning

Please see Figure 5.1.

b. Optimized protocol for sample processing, array hybridization and wash

Please see Figure 5.2.

c. Array scanning

Use the Axon 4000B scanner at five micron. Scanner needs to be warmed up 15 minutes

prior to use. The scanner was calibrated once a month or more depending on overall

usage. The GenePixPro software is used to control the scanner and run scanner

diagnostics.

We followed the following steps to scan an array. 1) In the 'Image' tab, preview scan the

entire array area. 2) Once preview scan is complete, select the scan area. I scanned each

subarray individually. 3) After scan has begun, use the 'Histogram' tab to evaluate the

selected PMT settings. Make sure the y-axis is the log-normalized count. The 'Count

Ratio' (CR) measurement should be 1 +/- 0.15 to provide the greatest overlap of total

fluorescence, and the high intensity-tails targeted to the 1e-5 to 1e-4 range. Adjust PMT

settings for 635nm and 532nm lasers to reach these criteria. (635nm is the 'red' plotted

data and the numerator in the CR, 532nm is the 'green' plotted data and the denominator

in the CR) 4) Once scan is complete, save images as separate .tiff files, uncompressed.

Figure 5.3

6. CNV identification

   a. Create .pair reports in NimbleScan

Fluorescence intensity data must first be linked to probe information before CNV

identification can be conducted. We used the NimbleScan software (NS, v2.40) to create

.pair reports, and followed the manufacturer's instructions to create these first report files.

In brief, open NS, select the .tiff file to be gridded (532nm or 635nm). Then, select the

.ndf or design file for the subarray to be paired (A01: 071108_HG18_MI_CGH01.ndf ,

A02: 071108_HG18_MI_CGH02.ndf , A03: 071108_HG18_MI_CGH03.ndf). Select

auto brightness/contrast; select auto autoalign. Zoom in to the alignment oligo probes to

ensure gridding is accurate; otherwise, manually adjust gridding such that all grids are

centered over the alignment oligos. At minimum, I reviewed all four corners and the

central 'X' features before saving. Under the Analysis menu, the Reports menu allows

you to select 'Pair Report'. Reports can be generated singly or by batch. This report

option will generate two .pair files (one each for the 532nm and 635nm scans).

   b. Remove poorly behaving probes (5% of all experimental)

Using a custom perl script, poorly behaving probes were removed from .pair files prior to

segmentation analysis. The perl script used the list of 'poor' probes, searched each .pair

file for the 'bad' or 'poor' probe, and wrote a 2nd .pair file less the poor probes (and their

associated fluorescent data). The files containing the 'poor' probes are:

A01_bad_probes_6_5percent.txt, A02_bad_probes_5_4percent.txt,

A03_bad_probes_5_24percent.txt, chrX_HD_bad_probes.txt. These files list the probes

by Probe_ID for each of the subarrays and the 2.1M array in its entirety respectively.

c. Run NimbleScan analysis algorithm (segMNT)

We used the segMNT algorithm as the DNACopy algorithm reportedly generated a

tremendous amount of false calls. We followed the manufacturer's instructions for how

to analyze our .pair reports with the segMNT algorithm. In brief, under the Analysis

menu, the CGH menu allows you to select 'segMNT'. Choose the .pair files to be

analyzed, choose the destination folder for analyzed data, choose the .pos file, and hit

Run. The .pos files for the arrays with poor probes removed are:

071108_HG18_MI_CGH01.good6.5.pos, 071108_HG18_MI_CGH02.good5.4.pos,

071108_HG18_MI_CGH03.good5.24.pos

We used the default settings for analysis. They are as follows:

> Min segment difference (score units): 0.0
> Min segment length (number of probes): 2
> Acceptance percentile (0.4-0.999): 0.999
>
> Do not include non-unique probes in analysis.
> Do use 'Spatial Correction'.
> Do 'Normalize'.

d. Quality Control

Create MA Plots

Using the _segMNT.txt files generated by the segMNT analysis, we created MA

plots for each subarray in R (r-project.org). To automate this procedure, we

developed a custom perl script to create a 'source' document to be run in R. The

perl script essentially grabbed the file names of _segMNT.txt files to be analyzed, and the 'source' document was a single file that could be read in R to plot and save the MA plot files for each _segMNT.txt file.

In brief, the 'source' file read in each _segMNT.txt file with read.table. Four variables were created for each of the probes in the _segMNT.txt file. These variables were then used to make the MA plots: $\underline{M}$ = EXP_NORM – REF_NORM, $\underline{A}$ = 0.5 * (EXP_NORM + REF_NORM), $\underline{fit}$ = lm(M~A), and $\underline{fit.coef}$ = coef(fit). The variables 'M' versus 'A' were plotted and the linear regression line 'fit' was plotted in red as well. This allowed us to visualize and quantitatively assess overall data performance.

Create file containing mean and standard deviation of each subarray
Additionally, while each file was read into R, we generated a file that collected some basic statistics for the log2 values of the subarray. Using the RATIO_CORRECTED values, we generated a table listing the mean +/- one standard deviation (SD) as well as the mean +/- two SD. These values would be used to select segments or CNV called by the segMNT algorithm in the _segtable_segMNT.txt files.

7. Analysis of CNV identified by the NimbleGen and Optimized protocols

a. All CNV called by the NimbleGen and Optimized protocols

Chapter 2. Table 4 and Table 5.

174

Using the same number of 'good' probes for copy number analysis, CNV generated by 30 samples run on both protocols resulted in a significantly different call rate between the two protocols. We evaluated the number of CNV called per individual for each protocol. Using the Wilcoxon Signed Rank test, the p-value was less than 0.0011.

This significant difference was similarly reflected in our analysis of all deletions and duplications called by each protocol. By the Pearson's Chi-squared test with Yates' continuity correction, we observed a significant difference between the number of deletions (NimbleGen: 184, Tecan: 44) and duplications (NimbleGen: 118, Tecan: 130) with $p < 1.36e-13$.

b. *CNV characteristics generated by the NimbleGen and Optimized protocols*

Chapter 2. Table 5 and Figures 9 A-C.

We also compared several characteristics of the CNV called by either protocol. The same number of 'good' probes was used for copy number analysis as well as the same samples. We used the Welch Two Sample t-test with unequal variances to evaluate CNV size, the average number of probes/kb called for CNV by protocol, and the average GC content for all CNV, for deletions and for duplications.

Evaluation of CNV size for each protocol was not statistically different (all CNV: $p < 0.0537$, deletions: $p < 0.1627$, duplications: $p < 0.1627$). However, when we evaluated the probe coverage per variant (number of probes/CNV), we found that the diminished coverage observed in the Optimized protocol (all CNV: NimbleGen: 20.4, Optimized:

18.0; <u>deletions</u>: NimbleGen: 21.0, Optimized: 15.2; <u>duplications</u>: NimbleGen: 19.4, Optimized: 18.9) was statistically different for all categories (<u>all CNV</u>: p < 5.24E-10, <u>deletions</u>: p < 2.23E-12, <u>duplications</u>: p < 2.23E-12). And, GC content for these three categories was also found to be statistically significantly different between the two protocols (<u>all CNV</u>: p < 6.97E-21, <u>deletions</u>: p < 0.0310, <u>duplications</u>: p < 0.0310).

c. Analysis of false positive and true positive calls

Please see Table A.6.a and A.6.b for sample and locus data.

<u>False and True CNV calls</u>

<u>Chapter 2. Table 6 and Figure 10.</u>

For the same individuals and the same 'good' probes, we evaluated the false positive call rates for all CNV identified by the NimbleGen and Optimized protocols. The number of false positive (NimbleGen: 40, Optimized: 36) and true positive calls (NimbleGen: 22, Optimized: 26) was not statistically different between the two protocols by the Fisher's Exact Test for Count Data (p < 0.5804).

Characteristics of falsely and truly called CNV by the NimbleGen and Optimized protocols

<u>Chapter 2. Table 6.</u>

We also assessed the characteristics of all CNV falsely and truly called by size, number of probes/kb, and the average GC content. We found that the size of CNV falsely called was statistically significantly different between the two protocols by the Kolmogrov-Smirnov test (p < 0.0077). Not surprisingly, the truly called loci were not statistically different in size (p < 0.4282).

By the Kolmogrov-Smirnov test, the probe coverage for the true and false calls were statistically different with the number of probes/CNV for false calls $p < 8.01E-05$ and true calls $p < 0.0010$.

Chapter 2. Table 6 and Figure 11A and 11B.

Like CNV size, the GC content of true and false calls by each protocol was statistically different for false calls but not truly called CNV. Using the Kolmogrov-Smirnov test, the GC content for false calls was statistically different ($p < 7.45E-06$) but not for the truly called loci ($p < 0.4022$). We furthered our analysis by also assessing these call rates as compared to the X chromosome GC content. For this purpose, we binned the X chromosome by 1,000 bp bins and assessed the GC content for each of these bins as distributed across the chromosome. Figure 11 plots the cumulative percentage of the X chromosome 1,000 bp bins, the truly called loci, and the falsely called loci by the percent GC content.

8. Probe Analysis

a. Identification of 'poorly' behaving probes

We modeled our analysis of probe kinetics off of the methods reported by Mulle *et al*. [5] Specifically, we used the data from the 48 AGRE samples run on the 2.1M CGH arrays following the NimbleGen protocol. We used R to evaluate the log2 values for all probes. (r-project.org) For each subarray (714,457 probes), we determined the variance of the log2 value for each probe. We then sorted the probes by variance, and plotted the probes'

variance by index number. We chose a variance cutoff of 0.175 for all subarrays. Figure 5.4

b. Evaluation of 'poor' and 'good' probes

<u>Chapter 2. Table 2 and Figures 7 A-D.</u>

With the identification of the 'poor' probe set, we evaluated several properties of the 'poor' and 'good' probes. Specifically, we asked if the probes in the 'poor' probe set had qualities that were statistically different from the probes in the 'good' probe set. We evaluated probe length, purine content, GC content, and T_m (melting temperature) by the Welch t-test with unequal variance. All four parameters were statistically significant with the following pvalues: $p < 6.97E-49$, $p < 2.20E-16$, $p < 2.20E-16$, and $p < 2.20E-16$.

c. Evaluation of probe log2 values by processing protocol: NimbleGen versus Optimized

<u>Chapter 2. Table 3.</u>

We found that the optimized technical protocol produced average log2 values that were closer to the expected value of '0,' irrespective of subarray. However, when the log2 values for arrays tested by the NimbleGen or Optimized protocol (NimbleGen: 48 arrays, Optimized: 75-81 arrays) were evaluated, this difference was not significant different by the paired Wilcoxon Rank Sum test with Bonferroni correction.

d. Evaluation of probe coverage (probes/kb) by processing protocol: NimbleGen versus Optimized

178

Using the same number of 'good' probes for copy number analysis, the optimized

protocol returns a slightly diminished probe coverage per variant than that of the original

protocol. This decrease in coverage is significant by the Welch t-test with un-equal

variances (all CNV: $p < 5.24E\text{-}10$, deletions: $p < 2.23E\text{-}12$, duplications: $p < 2.23E\text{-}12$).

9. Select high-confidence segments or copy number variants

a. Select segments greater than one standard deviation from the mean

Using a custom perl script, we selected those segments listed in the

_segtable_segMNT.txt files generated by the segMNT algorithm. These segments were

more than one SD from the mean of each individual array. Selected segments were sent

to two output files. The first was simply another _segtable_segMNT.txt file containing

only those segments one SD from the mean, while the second file collected all segments,

irrespective of array origin, in one file. This became our working for CNV analysis.

b. Identify and merge multiple segments calling a single, variant locus

We found that occasionally, larger copy number variants (CNV) would be identified by

more than one 'segment' by the segMNT algorithm. We selected segments for merging

into a single call by the following criteria. 1) The 3' end of the 'upstream' segment is

within three kilobases (kb) of the 5' end of the next downstream segment. 2) The log2

values are shifted in the same direction. 3) The length of the new segment is determined

by the most upstream and downstream probes, the number of probes is the sum of all

probes within the original, smaller segments, and the new log2 value is a weighted

average of the number of probes and log2 values of the original segments. Specifically,

the number of probes for each original segment is multiplied by the log2 value for the

segment, the sum is taken of all these products and the sum divided by the sum of all

probes.

c. Select segments which have more than nine probes/kb

The average density of probes is 20 probes/kb. Given the dense nature of our custom

array, we believed that half the probes (10 probes/kb) would be sufficient for a call. With

this in mind, we purged the remaining segments in the total set of any segments called

with less than nine probes/kb. Figure 5.5

d. Remove samples with relatively high call rates

Finally, we determined the number of segments called/individual as well as the mean and

SD for the number of calls/individual for the set. For those individuals who had more

than two SD from the mean number of calls, we removed their calls from the data set

entirely. Figure 5.6

10. Validation of array identified CNV

a. PCR Confirmation

We used a PCR based strategy to validate aCGH identified structural variants. For

deletions, primers flanking the outsides of the breakpoints were designed for

amplification of the junction fragment. Primers were selected outside of the first probe

having a log2 value near '0'. Figure 5.7

One of two types of validations assays were developed to confirm duplications. All 'duplicated' sequence was assumed to be in tandem to the 'original' sequence. The first strategy was implemented for 'relatively' small duplication events (less than three kb) where primers were designed to flank the supposedly duplicated sequence (resulting in an approximate six kb sequence). For those duplications too large to be accommodated by this strategy, we used a PCR-based approach elegantly explained by Arlt *et al*.[6] Essentially two 'outward' facing primers are designed at the internal junction of duplicated sequence. If the duplication is real and in tandem, a junction fragment will be captured by these primers. Figure 5.8

For those duplications screened by the 'outward' facing primers and a junction fragment was not observed, an additional set of primers was designed to confirm that the 'duplication' primers actually worked. These additional primers sat 'outside' of the duplication boundary and would amplify a small region between the 'outside' primer and the internal 'duplication' primer. This assay also confirmed relative orientation of the duplicated sequences relative to flanking, single copy sequence. Figure 5.9

Because of the various sizes and sequence characteristics represented by all CNV, we employed several different Taq polymerases (Invitrogen Platinum Taq, Roche Expand Long taq, Roche High Fidelity Taq, and Takara Ex taq), varying $MgCl_2$ concentrations, various supplements (betaine (mono) hydrate – for GC rich regions, dimethyl sulfoxide (DMSO) – reduce secondary structure and for GC rich regions, and formamide – increase

specificity and for GC rich regions, tetramethylammonium chloride (TMAC) – reduce

non-specific priming), and various annealing temperatures to develop the validation

assays.[7-9]

b. Breakpoint sequencing

For those CNV junction fragments that validated, amplicons were gel eluted from sterile,

agarose gel and TA-cloned for bidirectional sequencing. We aimed to complete all steps

within one day to ensure the integrity of the 'A' overhang found at the ends of our

amplicon to increase efficiency of cloned product. For gel elution, we used either the

Promega (Wizard® SV Gel and PCR Clean-Up System, P/N A9280, promega.com) or

Qiagen (QIAquick® Gel Extraction Kit, P/N 28704, qiagen.com) based systems to

extract our PCR products. We followed the manufacturer's instructions to elute

amplicons. Both systems utilized a membrane to 'catch' the amplicon, and we used dH_2O

to elute.

TA-cloning was conducted using the pCR2.1-TOPO (amplicons up to three kb in size,

P/N K4500-01) and pCR-XL-TOPO (amplicons three to ten kb in size, P/N K4700-10)

systems from Invitrogen (invitrogen.com). We followed the manufacturer's instructions

to clone and transform TOP10 or DH5α-T1 chemically competent cells. In all cloning

reactions, we used the maximum volume of 4 μL of amplicon (no water), and for the 2.1-

TOPO clones, we used the optional one μL of salt solution to enhance incorporation of

our products. Transformed bacteria were grown on agar plates containing ampicillin (2.1-

TOPO clones) or kanamycin (XL-TOPO clones). Warm agar plates were spread with 40

µL X-gal (20 mg/mL) and 100 µL SOC half an hour prior to plating 10-50 µL of transformants.

Typically, five to ten colonies were picked for expansion in two mL of LB containing ampicillin (2.1-TOPO clones) or kanamycin (XL-TOPO clones). Plasmids were extracted using the Promega (PureYield™ Plasmid Miniprep System, P/N A1223) or Qiagen (Plasmid Mini Kit, P/N 12123) mini-prep systems. Plasmid extraction was conducted following the manufacturer's protocols.

Plasmids were then assayed for amplicon ligation by digestion with Hind III and Xho I enzymes. For each clone, we ran four conditions of no enzyme, only Hind III, only Xho I, and Hind III and Xho I. For a 1x reaction, we used one µL of BSA (1 mg/mL), one µL #2 Buffer (10x), 0.3 µL enzyme #1 (200 U/uL) (optional), 0.3 µL enzyme #2 (200 U/uL) (optional) and sterile, filtered water for a final volume of five uL. We added this five µL of enzyme mix to five µL of 200 ng of vector. Reaction was incubated at 37°C for minimum of one hour and were then run on an agarose gel to visualize cleavage products.

For those plasmids that had ligated our amplicon, first pass sequencing utilized the universal M13 Forward (-20) and M13 Reverse primers. These primers are positioned 100 bp and 292 bp respectively from the ligation site. Depending on length of sequence and size of amplicon product, new primers were designed to ensure bidirectional sequencing of the breakpoint junction. We obtained sequence data for 40 unique loci and bidirectional sequence for 28 of those loci.

11. Junction analysis

a. Literature breakpoint evaluation

We identified 681 (bidirectionally) sequenced breakpoints from 12 reports in the literature.[10-21] Populations and methods varied by study. For breakpoint/motif analysis, we removed two sets of variants from this data set. The first was based on variant size. In our preliminary analysis of the occurrence of motif frequency at the breakpoint, we evaluated various window sizes at the breaks. Because the window sizes were as large as 500 bp, we removed any indels that were less than one kb in size. This reduced the literature data set to 569 CNV. The second round of removals was based on our inability to match 'GC' content as well as 'N' content during random sequence generation. This reduced the final data set to 565 CNV.

b. Single nucleotide insertion or deletion at the breakpoint

For those CNV and indels that had homology and insertion data available, using Pearson's Chi-squared test, we asked if any one nucleotide was over- or under-represented at the breakpoint. For 57 CNV or indels showing a deletion of a single nucleotide, no single nucleotide occurred at a frequency statistically different from expected ($p < 0.727$). Similarly, for 21 CNV or indels showing an insertion of a single nucleotide, no single nucleotide occurred at a frequency statistically different from expected ($p < 0.6134$).

c. Development of random data set

For each literature CNV or 'Real' CNV that we used (n=565), we selected 1,000 random

positions on the Real CNV's chromosome and set those positions as the 'start' for the

'Random' CNV. These 1,000 Random CNV were then assigned the same length as the

Real CNV to determine the 'end' of the Random CNV. This process was repeated for all

565 Real CNV to generate 565,000 Random CNV representing 1,000 sets of Random

CNV having the same chromosomal and size distribution as the Real CNV data set.

Additionally, we ensured that the 'GC' and 'N' content of each Real CNV was matched

or closely matched by each of the Random CNV.

> *d. Evaluation of Real and Random CNV for the frequency of motif occurrence at*
>
> *the breakpoints*

First, we identified 39 motifs from the literature that had been previously identified as

enriched at the breakpoint junctions. [22-28] Second, we defined windows (defined by A,

B, C, and D in the figure) of varying sizes (25 bp, 50 bp, 100 bp, and 500 bp) for each

breakpoint, Real or Random. Figure 5.10 Third, for each window in each breakpoint

(Real or Random), we determined the frequency of occurrence for each of the 39 motifs

in both the forward and reverse strands. Because two motifs were greater than 26 bp in

size (Alu core element and Curved DNA signal), we eliminated evaluation of the 25 bp

window sizes.

Finally, the observed frequencies were then normalized by motif and window size.

Specifically, for each motif frequency, we divided the motif count by [2 * (Window Size

– Length of Motif + 1)]. The factor of two in the denominator is to account for searching both the forward and reverse strands.

For each of the motifs, we compared the normalized frequency for All windows (A+B+C+D), Outside windows (A+D) and Inside windows (B+C) in the Real breakpoints and Random breakpoints by the z-test.

12. Assess functional activity of *GRIA3* deletion: Luciferase Reporter Assay

a. Ligate region of interest into reporter backbone

We used the Promega luciferase reporter system (luciferase vectors: pGL3-Basic, pGL3-Promoter) to evaluate the region three kb upstream of the *GRIA3* gene. To generate pGL3-Promoter plasmids containing our two promoter sequence, we first PCR amplified our two regions of interest using primers with the Xho I and Hind III digestion sequences tagging the 5' end of the amplification primers. The forward and reverse primers had Xho I and Hind III tags, and the negative control primers (same sequence) had Hind III and Xho I tags (effectively, insert sequence 'backwards'). PCR amplicons and pGL3-Promoter plasmid were Hind III and Xho I digested. Digestion products were agarose gel eluted.

We then used the Takara DNA Ligation Kit v2.1 (Takara, takara-bio.com) and followed the manufacturer's instructions to ligate our amplicons to vector at two different ratios (0.01 pmol vector: 0.05 pmol amplicon, 0.03 pmol vector: 0.09 pmol amplicon). We made two sets of 0.01 pmol vector: 0.05 pmol amplicon; the second set did not receive

ligation agent and served as our negative control. After mixing the appropriate amounts

of vector (45.07 ng/uL) and amplicon, we added water to bring the final volume to 10 μL.

Then, 10 μL of Solution I was added to all combinations (with the exception of the

negative control – 10 μL of sterile, dH_2O was added). All mixes were incubated at 4°C

overnight. We ran 30 ng of each on an agarose gel to confirm ligation.

b. Transfect reporter plasmids into Neuro2A cells

The luciferase assay used three 6-well plates of Neuro2A cells, each plate had two rows

of three replicates. Plate one contained tranfection reagents and pGL3-Basic and pRL-TK

(Renilla Luciferase, internal control); plate two contained the negative control plasmids

(promoter sequence ligated in 'backwareds') and pRL-TK, and plate three had the

plasmid containing the promoter sequence ('forward' orientation) for the intact and

deleted allele as well as pRL-TK. Prior to transfection, PBS was incubated at 37°C for

five minutes, and 32 mL of Opti-MEM were incubated at room temperature for five

minutes.

Transfection was conducted following the Mirus protocol (TransIT-LT1, P/N MIR 2304,

mirusbio.com). The luciferase and Renilla vectors were mixed at a 5:1 molar ratio for all

plates and incubated at room temperature for 5-20 minutes. Concurrently, the TransIT-

LI1 transfection reagent was mixed with the room temperature Opti-MEM and incubated

at room temperature for 5-20minutes (1x/well: 250 uL Opti-MEM and five uL TransIT-

LIT1). TransIT-LT1 mixtures was dropped along the sides of the vector mixture tube, the

tube flicked and shaken vigorously (to create micelles), and incubated at room temperature for 5-20 minutes.

Prior to transfection, Neuro2A cells were rinsed twice with two mL of PBS, and 1,250 uL of Opti-MEM were then added to each well. 285 uL of vector/transfection mixture were added to each well (three wells for each vector type) in a drop-wise fashion. Plates were rocked to distribute mixture evenly and then incubated at 37C for 24-72 hours.

c. Assay luciferase activity

We evaluated lucifierase activity following Promega's protocol for the Dual Luciferase Reporter Assay System (P/N E1910). We briefly describe the protocol. For cells grown in 6-well plate, first wash cells with 1 mL PBS – rock to wash. Aspirate PBS with seripipette. Lyse cells using 250 μL Passive Lysis Buffer (PLB) (1x); rock or shake plate at room temperature for 15 minutes. Collect cells by transferring lysate to 1.5 mL tube. Spin at top speed at 4°C for 5 minutes to clear lysate. Transfer supernatant to new 1.5 mL tube. Can be stored at -20°C.

To assess luciferase activity from the lysed cells, first, prepare Luciferase Assay Reagent II (LAR II). Resuspend LAR II in 10 mL of Luciferase Assay Buffer II. Aliquot to 1.3 mL (enough for 12 assays) to 1.5 mL tubes for future assays. Store -80°C. For a 1x preparation, prepare Stop & Glo Reagent (SG) by mixing 2 μL S&G (50x) + 100 uL S&G Buffer (3.8 mL total) in a glass/polypropylene tube.

Turn on the luminometer (five minutes to warm up). Confirm program is for Single (not Dual) read; MODE: <STD> with a two-second pre-measurement delay; 10-second measurement. To assay activity, add 20uL cell lysate (one ug total protein) to luminometer tubes. Add 100 μL LARII reagent to sample. Aspirate up and down three times; do not vortex. Place tube in reader. Hit 'READ'. Do this for all samples. Then, add 100 μL S&G reagent. Aspirate up and down three times; do not vortex. (FLICK well.) Place tube in reader. Hit 'READ'.

d. Evaluate luciferase data

The first and second luminometer readings were for the fluorescence generated by the luciferase vectors (Firely; containing no insert (pGL3-Basic), insert in 'backwards' orientation (Negative Control, pGL3-Promoter), and insert in 'forward' orientation (the test, pGL3-Promoter)) and the Renilla reporter (pRL-TK) respectively. The Firely:Renilla ratio was determined for each replicate and the average of all three replicates was then compared. Using the t-test of equal variance, the deleted allele was found to be significantly increased over that of the intact allele ($p < 4.10E-05$).

13. References

1. Lord, C., et al., *Autism diagnostic observation schedule: a standardized observation of communicative and social behavior.* J Autism Dev Disord, 1989. **19**(2): p. 185-212.

2. Lord, C., M. Rutter, and A. Le Couteur, *Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders.* J Autism Dev Disord, 1994. **24**(5): p. 659-85.

3. Yatsenko, S.A., et al., *Microarray-based comparative genomic hybridization using sex-matched reference DNA provides greater sensitivity for detection of sex chromosome imbalances than array-comparative genomic hybridization with sex-mismatched reference DNA.* J Mol Diagn, 2009. **11**(3): p. 226-37.

4. Scherer, S.W., et al., *Challenges and standards in integrating surveys of structural variation.* Nat Genet, 2007. **39**(7 Suppl): p. S7-15.

5. Mulle, J.G., et al., *Empirical evaluation of oligonucleotide probe selection for DNA microarrays.* PLoS One, 2010. **5**(3): p. e9921.

6. Arlt, M.F., et al., *Replication stress induces genome-wide copy number changes in human cells that resemble polymorphic and pathogenic variants.* Am J Hum Genet, 2009. **84**(3): p. 339-50.

7. Lieb, B. *PCR Additives.* Available from: www.staff.uni-mainz.de/lieb/additiva.html.

8. Frackman, S., et al., *Betaine and DMSO: Enhancing Agents for PCR.* Promega Notes, 1998. **65**.

9. Sarkar, G., S. Kapelner, and S.S. Sommer, *Formamide can dramatically improve the specificity of PCR.* Nucleic Acids Res, 1990. **18**(24): p. 7465.

10. Conrad, D.F., et al., *Mutation spectrum revealed by breakpoint sequencing of human germline CNVs.* Nature Genetics, 2010. **42**(5): p. 385-U43.

11. de Smith, A.J., et al., *Array CGH analysis of copy number variation identifies 1284 new genes variant in healthy white males: implications for association studies of complex diseases.* Hum Mol Genet, 2007. **16**(23): p. 2783-94.

12. Nichol Edamura, K. and C.E. Pearson, *DNA methylation and replication: implications for the "deletion hotspot" region of FMR1.* Hum Genet, 2005. **118**(2): p. 301-4.

13. Goldmann, R., et al., *Genomic characterization of large rearrangements of the LDLR gene in Czech patients with familial hypercholesterolemia.* BMC Med Genet, 2010. **11**: p. 115.

14. Kim, P.M., et al., *Analysis of copy number variants and segmental duplications in the human genome: Evidence for a change in the process of formation in recent evolutionary history.* Genome Res, 2008. **18**(12): p. 1865-74.

15. Korbel, J.O., et al., *Systematic prediction and validation of breakpoints associated with copy-number variants in the human genome.* Proc Natl Acad Sci U S A, 2007. **104**(24): p. 10110-5.

16. Lam, H.Y., et al., *Nucleotide-resolution analysis of structural variants using BreakSeq and a breakpoint library.* Nat Biotechnol, 2010. **28**(1): p. 47-55.

17.	Nobile, C., et al., *Analysis of 22 deletion breakpoints in dystrophin intron 49.* Hum Genet, 2002. **110**(5): p. 418-21.

18.	Park, H., et al., *Discovery of common Asian copy number variants using integrated high-resolution array CGH and massively parallel DNA sequencing.* Nat Genet, 2010. **42**(5): p. 400-5.

19.	Vissers, L.E., et al., *Rare pathogenic microdeletions and tandem duplications are microhomology-mediated and stimulated by local genomic architecture.* Hum Mol Genet, 2009. **18**(19): p. 3579-93.

20.	Woodward, K.J., et al., *Heterogeneous duplications in patients with Pelizaeus-Merzbacher disease suggest a mechanism of coupled homologous and nonhomologous recombination.* Am J Hum Genet, 2005. **77**(6): p. 966-87.

21.	Zhang, F., C.M.B. Carvalho, and J.R. Lupski, *Complex human chromosomal and genomic rearrangements.* Trends in Genetics, 2009. **25**(7): p. 298-307.

22.	Abeysinghe, S.S., et al., *Translocation and gross deletion breakpoints in human inherited disease and cancer I: Nucleotide composition and recombination-associated motifs.* Hum Mutat, 2003. **22**(3): p. 229-44.

23.	Chuzhanova, N., et al., *Translocation and gross deletion breakpoints in human inherited disease and cancer II: Potential involvement of repetitive sequence elements in secondary structure formation between DNA ends.* Hum Mutat, 2003. **22**(3): p. 245-51.

24.	Dohoney, K.M. and J. Gelles, *Chi-sequence recognition and DNA translocation by single RecBCD helicase/nuclease molecules.* Nature, 2001. **409**(6818): p. 370-4.

25.	Kvikstad, E.M., F. Chiaromonte, and K.D. Makova, *Ride the wavelet: A multiscale analysis of genomic contexts flanking small insertions and deletions.* Genome Res, 2009. **19**(7): p. 1153-64.

26.	Singh, G.B., J.A. Kramer, and S.A. Krawetz, *Mathematical model to predict regions of chromatin attachment to the nuclear matrix.* Nucleic Acids Res, 1997. **25**(7): p. 1419-25.

27.	Chuzhanova, N., et al., *Gene conversion causing human inherited disease: evidence for involvement of non-B-DNA-forming sequences and recombination-promoting motifs in DNA breakage and repair.* Hum Mutat, 2009. **30**(8): p. 1189-98.

28.	Aoki, K., et al., *A novel gene, Translin, encodes a recombination hotspot binding protein associated with chromosomal translocations.* Nat Genet, 1995. **10**(2): p. 167-74.

Figure 5.1 – NimbleGen aCGH Protocol used.

Sample labeling and hybridization for NimbleGen array Comparative Genomic Hybridization

PROTOCOL FOR:
Processing and hybridization of NimbleGen aCGH by the manufacturer's instructions

Morna Ikeda and Stephen T. Warren
Department of Human Genetics, Emory University, Atlanta, GA USA

LEGEND

⇨ *ATTENTION*

* *HINT*

🖐 *REST*

REAGENTS AND CONSUMABLES
Deionized, filtered (0.22 um) sterile water
Tris-HCl (1M, pH 7.4, any vendor)
MgCl2 (1M, any vendor)
β-Mercaptoethanol (98%, any vendor)
Cy3/Cy5 9mer Wobble (Cy3: P/N N46-0001-50, Cy5: N46-0002-50, Trilink Biotechnologies, San Diego, CA, USA)
Ethanol (100%, any vendor)
Tris-HCl (1M, any vendor)
EDTA (0.5M, any vendor)
dNTP Set (100mM dNTP Set, P/N 10297-018, Invitrogen)
Klenow Fragment (3´→5´ exo–, 50 U/μL, P/N M0212M, New England Biolabs, Ipswich, MA, USA)
Sodium Chloride (5 M, any vendor)
Isopropanol (100%, any vendor)
Alexa Fluor Reactive Dye Decapacks (P/N 10297-018, Invitrogen)
Sodium Bicarbonate (any vendor)
DMSO (100%, any vendor)
Hybridization Kit (P/N 05583683001, Roche NimbleGen, Madison, WI, USA)
      2X Hybridization Buffer
      Hybridization Component A
      Alignment Oligo
Wash Buffer Kit (P/N 05584507001, Roche NimbleGen)
      Wash Buffer I, II, and III
      DTT (Use is optional. Follow manufacturer's instructions for amount to add.)
CP100 Pipette Tips (P/N F148414, Gilson Inc., Middleton, WI, USA, pipetman.com)

EQUIPMENT
Misonix sonicator with microtip (QSonica, Newton, CT, USA)
Thermocycler (any vendor)
Benchtop centrifuge (any vendor)
Vortex (any vendor)
Spectrophotometer (P/N ND-1000, NanoDrop)
Heat block (95°C, any vendor)
Speed-vac (Thermo Savant)
Precision Mixer Alignment Tool (PMAT, P/N 00832, Roche NimbleGen)
Positive Displacement Microman (10-100 uL, P/N M100, Gilson)
Microarray Dryer (Array-Go-Round, P/N 00898, Roche NimbleGen)
Axon 4000B (P/N 4000B, Molecular Devices)

Reduced Ozone environment (optional)

PROCEDURE

⇨ USE FILTERED STERILIZED WATER FOR ALL STEPS

RANDOMER PREPATION
1. Centrifuge lyophilized primer to collect at the bottom of tube.
2. Using fresh Randomer Buffer, add enough Buffer to rehydrate to 1 OD/42 μL.
3. Invert tube several times.
4. Centrifuge to collect liquid at bottom of tube.
5. Incubate at room temperature for 5-10 minutes protected from light.
6. Invert and centrifuge tube once again.
7. Aliquot 42 μL of primer to fresh, sterile thin-walled PCR strip tubes (0.2 mL).

✋

Primers can be stored at -20°C protected from light until ready for sample labeling.
Thaw frozen primers protected from light at room temperature.
Centrifuge tubes before adding sample.

DNA SONICATION & QC
8. In separate 1.5 mL tubes, bring 2 μg of genomic DNA to a final volume of 80 μL using dH$_2$O for both the test sample and a reference sample.
9. Flick tubes to mix contents, and centrifuge to collect liquid at bottom.
10. Incubate tubes at 55°C for 5-10 minutes to ensure DNA is in solution.
11. Flick tubes to mix contents and centrifuge to collect liquid at bottom again.
12. Store DNA on ice while preparing for sonication.
13. Using a Misonix sonicator, clean tip with 10% bleach followed by 80%-100% ethanol.
14. To sonicate DNA, place tip near bottom of tube and follow sonication protocol.

   Sonication Protocol:

Total Time	11 sec
on/off	0.5 sec
Amp	1.0

15. Return DNA to ice after sonication. Repeat tip cleaning with bleach and ethanol before sonicating next sample.
16. Run 5 μL of sonicated genomic DNA on 1% agarose gel. Confirm a smear between 200-2,000 bp.

✋ DNA can be stored at 4°C (less than one week) or -20°C (long term storage).

DNA AMPLIFICATION & LABELING
17. Add 40 μL sonicated DNA to 42 μL pre-aliquoted Cy3/Cy5 randomers. (Test DNA to Cy3, Reference DNA to Cy5)
18. Incubate mixture at 98°C for 5-10 min.
19. Ice water bath to snap-cool mixture while making master mix for at least 2 min.
20. Add 20 μL master mix to cooled mixture.
21. Aspirate up and down 10x using pipette. Centrifuge to collect liquid at bottom of tube.
22. Incubate at 37°C for 4 hours.

DNA PRECIPITATION
23. Add 10 μL EDTA (0.5M) to stop reaction.
24. Add 11.5 μL NaCL (5 M) to each tube.
25. Vortex briefly and centrifuge to collect liquid at bottom of tube.
26. Add 110 μL room temperature Isopropanol to 1.5mL tube.
27. Transfer stopped reaction (121.5 μL) to 1.5 mL tube containing Isopropanol.
28. Vortex briefly and centrifuge to collect liquid at bottom of tube.
29. Incubate at room temperature for 10 minutes shielded from light.

   *At this point, I pool the references together to minimize loss.

   REFERENCE: 3 references
   Add 330 μL Isopropanol to 1.5 mL tube.
   Transfer three stopped reactions to same tube.
   Vortex briefly and centrifuge to collect liquid at bottom of tube.
   Incubate at room temperature for 10 minutes shielded from light).

30. Centrifuge at max speed for 10 min.
31. Remove supernatant with pipette (232 μL).
32. Rinse with 500 μL of ice cold ethanol (80%) – be sure to dislodge pellet.
33. Centrifuge at max speed for 2 min.
34. Remove supernatant with pipette (500 μL).
35. Speedvac-low for 5 min to dry pellet.

✋ DNA pellet can be stored at -20°C protected from light and in sealed bag with desiccant (optional).

LABELED DNA QUANTIFICATION AND HYBRIDIZATION PREP
48. Rehydrate precipitated DNA in 20 μL dH$_2$O water.
49. Quantify labeled DNA on Nanodrop.

   *Use the Microarray Module to quantify Cy3 or Cy5 activity and DNA concentration.
   The linear range of this module is 740 ng/ μL.
   To ensure the measured DNA is within the limit of the reader, dilute DNA 1:4 and adjust calculations accordingly.
   Specific Activity should be between 10-20 pmol/μg.

50. Merge 34 μg of test, reference each.

51. Dry down.

Mixed DNA pellet can be stored at -20°C protected from light and in sealed bag with desiccant (optional).

52. Re-hydrate mixed DNA pellet in 12.3 μL H$_2$O.
53. Incubate at 55°C for 5-10 minutes. Gently tap bottom of tube to mix. Centrifuge.

HYBRIDIZATION PREPARATION
54. Make hybridization solution 31.7 μL/array).
55. Add 31.7 μL Hybridization Solution to sample.
56. Incubate Blank and Samples at 95°C for five minutes.
57. Hold at 42°C until ready.
58. Prepare slide and mixer (MAUI disposable lid) assembly using PMAT and place on MAUI bay (42°C) while preparing remaining slides.
59. Before adding sample, re-bray adhesive area to ensure complete annealing.
60. Using Pipetman, aspirate 41 μL and slowly dispense onto array.
61. Using Kim-wipe, wipe residual hybridization solution from each port before applying gel-stickers to enclose array chamber. Be careful not to wick solution from within.
62. Place sample and array assembly on MAUI. Once all samples loaded, close lid, and select hybridization program 'B'.
63. Hybridize 60 hours or more.

WASH SOLUTION
64. Make Wash Solution I, II, III.
65. Load jig into Wash I in upturned 1000p lid. Remove MAUI hybridization mixer while submerged. Agitate array 10-15 seconds in Wash I.
66. Move array to Wash I in staining dish while removing lids for additional arrays. Process up to eight arrays/batch.
67. Once all arrays are in Wash I, agitate vigorously up and down for two minutes. Tranfer to Wash II.
68. Agitate vigorously up and down for one minute. Tranfer to Wash III.
69. Agitate vigorously up and down for 30 secones. Tap array holder on paper towel to remove excess liquid. Load into balanced Array-Go-Round.
70. Dry for two minutes. Wipe residual liquid from edge of each slide. Store in dark dessicator until ready to scan.

71. **RECIPES**

Randomer Buffer (*make fresh every time*)

Component	1x (uL)
DI, sterile H2O	861
Tris-HCl (1M) (pH 7.4)	125
MgCl2 (1M)	12.5
B-MercaptoEtOH (98%)	1.75
TOTAL	1000

Rehydrate 9mers to one OD/42 μL
Add 1050 μL of Randomer Buffer to 25 ODs.
This gives a final concentration of one OD/42 μL.

TE (10x) (*store at room temperature*)

Components	1x (mL)
Tris-HCl (1M)	1.5
EDTA (.5M)	0.3
DI, sterile H2O	13.2
Total	15

dNTP Mix (50x) (*store at -20°C*)

Component	1x (uL)	Final Concentration
DI, sterile H2O	250	
TE (10x)	50	
dATP (100mM)	50	10 mM
dGTP (100mM)	50	10 mM
dTTP (100mM)	50	10 mM
dCTP (100mM)	50	10 mM
SUM	500	

Aliquot by 200 uL to minimize freeze-thaws.

Master Mix (for DNA amplification and labeling)

Component	1x (uL)
DI, sterile H2O	8
dNTPs 50x, (10mM)	10
Klenow (50 U/uL)	2
Total	20

Hybridization Solutions

Component	1x (uL)
Merged labl'd sample	12.3
Hybridization Buffer (2x)	22.0
Hybridization Component A	8.80
Alignment oligo	0.90
MM - total	31.7

Wash Buffers

Follow manufacturer's instructions to make 1x solutions. Use filtered (0.2um), sterile water to dilute. Two sets of 500 mL Wash Buffer I are needed (place one in upturned, 1000p lid, place second in Tissue-Tek dish). In brief, for 25 mL of Wash Buffer (10x), 225 mL of deionized, sterile water, and 25 uL of DTT (1 M).

Figure 5.2 – Optimized aCGH Protocol.

Modified sample labeling and hybridization for NimbleGen array Comparative Genomic Hybridization

PROTOCOL FOR:
Increased sample fluorescence labeling and stringency of hybridization using a Tecan HSPro 4800 for NimbleGen aCGH

Morna Ikeda and Stephen T. Warren
Department of Human Genetics, Emory University, Atlanta, GA USA

LEGEND

⇨*ATTENTION*

* *HINT*

✋ *REST*

REAGENTS
Deionized, filtered (0.22 um) sterile water
Tris-HCl (1M, pH 7.4, any vendor)
MgCl2 (1M, any vendor)
β-Mercaptoethanol (98%, any vendor)
Cy3/Cy5 9mer Wobble (Cy3: P/N N46-0001-50, Cy5: N46-0002-50, Trilink Biotechnologies, San Diego, CA, USA)
Ethanol (100%, any vendor)
Tris-HCl (1M, any vendor)
EDTA (0.5M, any vendor)
5-aminohexylacrylamido-dUTP (aha-dUTP, 50 mM in TE/50 μL, P/N A32761, Invitrogen, Carlsbad, CA, USA)
dNTP Set (100mM dNTP Set, P/N 10297-018, Invitrogen)
Klenow Fragment (3´→5´ exo–, 50 U/μL, P/N M0212M, New England Biolabs, Ipswich, MA, USA)
Sodium Chloride (5 M, any vendor)
Isopropanol (100%, any vendor)
Alexa Fluor Reactive Dye Decapacks (P/N 10297-018, Invitrogen)
Sodium Bicarbonate (any vendor)
DMSO (100%, any vendor)
Hybridization Kit (P/N 05583683001, Roche NimbleGen, Madison, WI, USA)
        2X Hybridization Buffer
        Hybridization Component A
        Alignment Oligo (I did not use the packaged oligo – I ordered my own.)

Cy3/Cy5 Alignment oligo (5' end - Cy3/Cy5 labeling,
TTCCTCTCGCTGTAATGACCTCTATGAATAATCCTATCAAACAACTCA, IDT, Coralville,
Iowa, USA)
Wash Buffer Kit (P/N 05584507001, Roche NimbleGen)
Wash Buffer I, II, and III
DTT (Use is optional. Follow manufacturer's instructions for amount to add.)

EQUIPMENT
Misonix sonicator with microtip (QSonica, Newton, CT, USA)
Thermocycler (any vendor)
Benchtop centrifuge (any vendor)
Vortex (any vendor)
Spectrophotometer (P/N ND-1000, NanoDrop)
Heat block (95°C) (any vendor)
Tecan HS4800 Pro (Tecan, Durham, NC, USA)
Speed-vac (Thermo Savant)
Axon 4000B (P/N 4000B, Molecular Devices)

Reduced Ozone environment (optional)

PROCEDURE

➡ USE FILTERED STERILIZED WATER FOR ALL STEPS

RANDOMER PREPATION
36. Centrifuge lyophilized primer to collect at the bottom of tube.
37. Using fresh Randomer Buffer, add enough Buffer to rehydrate to 1 OD/42 μL.
38. Invert tube several times.
39. Centrifuge to collect liquid at bottom of tube.
40. Incubate at room temperature for 5-10 minutes protected from light.
41. Invert and centrifuge tube once again.
42. Aliquot 42 μL of primer to fresh, sterile thin-walled PCR strip tubes (0.2 mL).

✋
Primers can be stored at -20°C protected from light until ready for sample labeling.
Thaw frozen primers protected from light at room temperature.
Centrifuge tubes before adding sample.

DNA SONICATION & QC
43. In separate 1.5 mL tubes, bring 2 μg of genomic DNA to a final volume of 80 μL using dH₂O for both the test sample and a reference sample.
44. Flick tubes to mix contents, and centrifuge to collect liquid at bottom.
45. Incubate tubes at 55°C for 5-10 minutes to ensure DNA is in solution.
46. Flick tubes to mix contents and centrifuge to collect liquid at bottom again.
47. Store DNA on ice while preparing for sonication.
48. Using a Misonix sonicator, clean tip with 10% bleach followed by 80%-100% ethanol.

49. To sonicate DNA, place tip near bottom of tube and follow sonication protocol.
    Sonication Protocol:

Total Time	11 sec
on/off	0.5 sec
Amp	1.0

50. Return DNA to ice after sonication. Repeat tip cleaning with bleach and ethanol before sonicating next sample.
51. Run 5 µL of sonicated genomic DNA on 1% agarose gel. Confirm a smear between 200-2,000 bp.

DNA can be stored at 4°C (less than one week) or -20°C (long term storage).

DNA AMPLIFICATION & LABELING
52. Add 40 µL sonicated DNA to 42 µL pre-aliquoted Cy3/Cy5 randomers. (Test DNA to Cy3, Reference DNA to Cy5)
53. Incubate mixture at 98°C for 5-10 min.
54. Ice water bath to snap-cool mixture while making master mix for at least 2 min.
55. Add 20 µL master mix to cooled mixture.
56. Aspirate up and down 10x using pipette. Centrifuge to collect liquid at bottom of tube.
57. Incubate at 37°C for 4 hours.
58. Incubate at 98°C for 5-10 min. Ice water bath to snap-cool for at least 2 min.
59. Add 1 µL Klenow, mix well by flicking bottom of tube. Centrifuge to collect liquid.
60. Incubate 37°C over night.

DNA PRECIPITATION
61. Add 10 µL EDTA (0.5M) to stop reaction.
62. Add 11.5 µL NaCL (5 M) to each tube.
63. Vortex briefly and centrifuge to collect liquid at bottom of tube.
64. Add 110 µL room temperature Isopropanol to 1.5mL tube.
65. Transfer stopped reaction (121.5 µL) to 1.5 mL tube containing Isopropanol.
66. Vortex briefly and centrifuge to collect liquid at bottom of tube.
67. Incubate at room temperature for 10 minutes shielded from light.

   * At this point, I pool the references together to minimize loss.
   REFERENCE: 3 references
   Add 330 µL Isopropanol to 1.5 mL tube.
   Transfer three stopped reactions to same tube.
   Vortex briefly and centrifuge to collect liquid at bottom of tube.
   Incubate at room temperature for 10 minutes shielded from light).

68. Centrifuge at max speed for 10 min.
69. Remove supernatant with pipette (232 µL).
70. Rinse with 500 µL of ice cold ethanol (80%) – be sure to dislodge pellet.
71. Centrifuge at max speed for 2 min.
72. Remove supernatant with pipette (500 µL).
73. Speedvac-low for 5 min to dry pellet.

✋ DNA pellet can be stored at -20°C protected from light and in sealed bag with desiccant (optional).

CONJUGATE REACTIVE DYES TO AMPLIFIED DNA

74. Rehydrate DNA in 10 μL dH$_2$O and warm in 52°C heat block for 5 min.
75. Add 6 μL of Labeling Buffer to rehydrated DNA. Mix gently and centrifuge to collect liquid at bottom of tube.
76. Add 4 μL of DMSO to reactive dye. Vortex for 10 sec and centrifuge to collect liquid at bottom of tube.
77. Add the prepared DNA (16 μL) to dye. Vortex to ensure well mixed and centrifuge to collect liquid at bottom of tube.
78. Incubate for 1 hour at room temperature in the dark.
79. Add 90 μL dH$_2$O and follow precipitation steps starting with addition of NaCl (Step 27).

✋ Once precipitated, DNA pellet can be stored at -20°C protected from light and in sealed bag with desiccant (optional).

LABELED DNA QUANTIFICATION AND HYBRIDIZATION PREP

48. Rehydrate precipitated DNA in 20 μL dH$_2$O water.
49. Quantify labeled DNA on Nanodrop.

✱ Use the Microarray Module to quantify Cy3 or Cy5 activity and DNA concentration.

> The linear range of this module is 740 ng/ μL.
> To ensure the measured DNA is within the limit of the reader, dilute DNA 1:4 and adjust calculations accordingly.
> Specific Activity should be > 20 pmol/μg, minimum 22 μg amplified.

72. Merge equal amounts of test, reference each.
73. Dry down.

✋ Mixed DNA pellet can be stored at -20°C protected from light and in sealed bag with desiccant (optional).

74. Re-hydrate mixed DNA pellet in 26 μL H$_2$O.
75. Incubate at 55°C for 5-10 minutes. Gently tap bottom of tube to mix. Centrifuge.

HYBRIDIZATION SOLUTION, WASH SOLUTION, AND MACHINE PREPARATION

76. Make Wash Solution I, II, III.
77. De-gas 1x Wash buffers:
    a. Allow Wash I (1x) to heat to 37°C for at least 30 minutes.
    b. Place Wash II (1x) and III (1x) under vacuum for at least 20 minutes.
78. Make blank hybridization solution (130 μL/array).
79. Make hybridization solution (64.0 μL/array).
80. Add 64 μL Hybridization Solution to sample.
81. Incubate Blank and Samples at 95°C for five minutes.
82. Hold at 52°C until ready.
83. Place correct lines to designated Wash Buffer.

84. Prime each line for 30 seconds.
85. Begin hybridization/wash program: Emory v10.

Randomer Buffer (*make fresh every time*)

Component	1x (uL)
DI, sterile H2O	861
Tris-HCl (1M) (pH 7.4)	125
MgCl2 (1M)	12.5
B-MercaptoEtOH (98%)	1.75
TOTAL	1000

Rehydrate 9mers to 1 OD/42 μL
Add 1050 μL of Randomer Buffer to 25 ODs.
This gives a final concentration of 1 OD/42 μL.

TE (10x) (*store at room temperature*)

Components	1x (mL)
Tris-HCl (1M)	1.5
EDTA (.5M)	0.3
DI, sterile H2O	13.2
Total	15

Aha-dNTP Mix (50x) (*store at -20°C*)

Component	1x (uL)	Final Concentration
DI, sterile H2O	91	
TE (10x)	21	
dATP (100mM)	21	10 mM
dGTP (100mM)	21	10 mM
dTTP (100mM)	7	3.33 mM
dCTP (100mM)	21	10 mM
dUPT (50 mM)	28	6.67 mM
SUM	210	

Master Mix (for DNA amplification and labeling)

Component	1x (uL)
DI, sterile H2O	9
aha-dNTPs (50x)	10
Klenow (50U/uL)	1
Total	20

Labeling Buffer (for reactive dye conjugation to aha-amplified sample)

Component	1x
DI, sterile H2O	1 mL
Sodium Bicarbonate	25 mg

Add 25 mg of Sodium Bicarbonate to 1 mL of deionized water. Vortex to get into solution. Aliquot 200μL to 1.5mL tubes, store at –20°C.

CPK6 Alignment Oligo Dilutions

Combine and dry down 50 uM and 500 nM dilutions for future use.

Hybridization Solutions

Blank Solution (to load onto arrays prior to sample loading)

Component	1x Tecan
H2O	37.73
2X Hybridization Buffer	65.00
Hybridization Component A	26.00
Alignment oligo-Cy3 (50 nM)	1.13
Alignment oligo-Cy5 (100 nM)	0.27
Total	130

Make BLANK hyb solutions (130 uL / array).

Hybridization Solution (to add to re-hydrated sample)

Component	1x (Tecan)
Merged labl'd sample	26.0
2X Hybridization Buffer	45
Hybridization Component A	18
Alignment oligo-Cy3 (50 nM)	0.78
Alignment oligo-Cy5 (100 nM)	0.18
Total	64.0
Hyb mix + sample - total	90.0

Wash Buffers

Follow manufacturer's instructions to make 1x solution. Use filtered (0.2um), sterile water to dilute.

For the Tecan HS4800Pro and Emory wash program v10:
Bottle 1: Wash Buffer 1
Bottle 2: Wash Buffer 2
Bottle 3: Wash Buffer 3

Bottle 5: dH$_2$0

Volumes estimates (11mL/minute) :
        Bottle 1: 580mL + (11mL/minute * 6.5minute * number of arrays)
        Bottle 2: 305mL + (11mL/minute * 0.5minute * number of arrays)
        Bottle 3: 305mL + (11mL/minute * 0.5minute * number of arrays)
        Bottle 5: 380mL + (11mL/minute * 1.0minute * number of arrays)
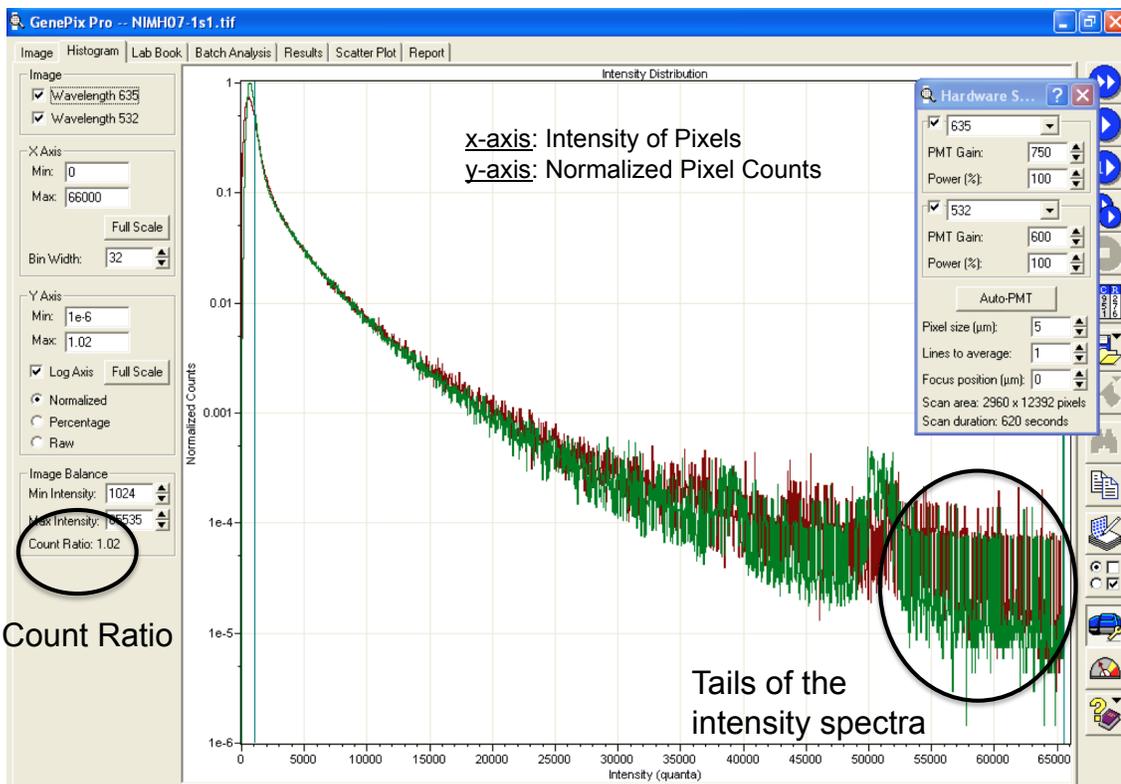
Figure 5.3 – Scanning parameters as informed by the Intensity Distribution Histogram

Log2 Variance of Probes: Subarray A01

Cutoff Variance: > 0.175
Probes Eliminated: 46,266 (6.48%)

Probe Variance
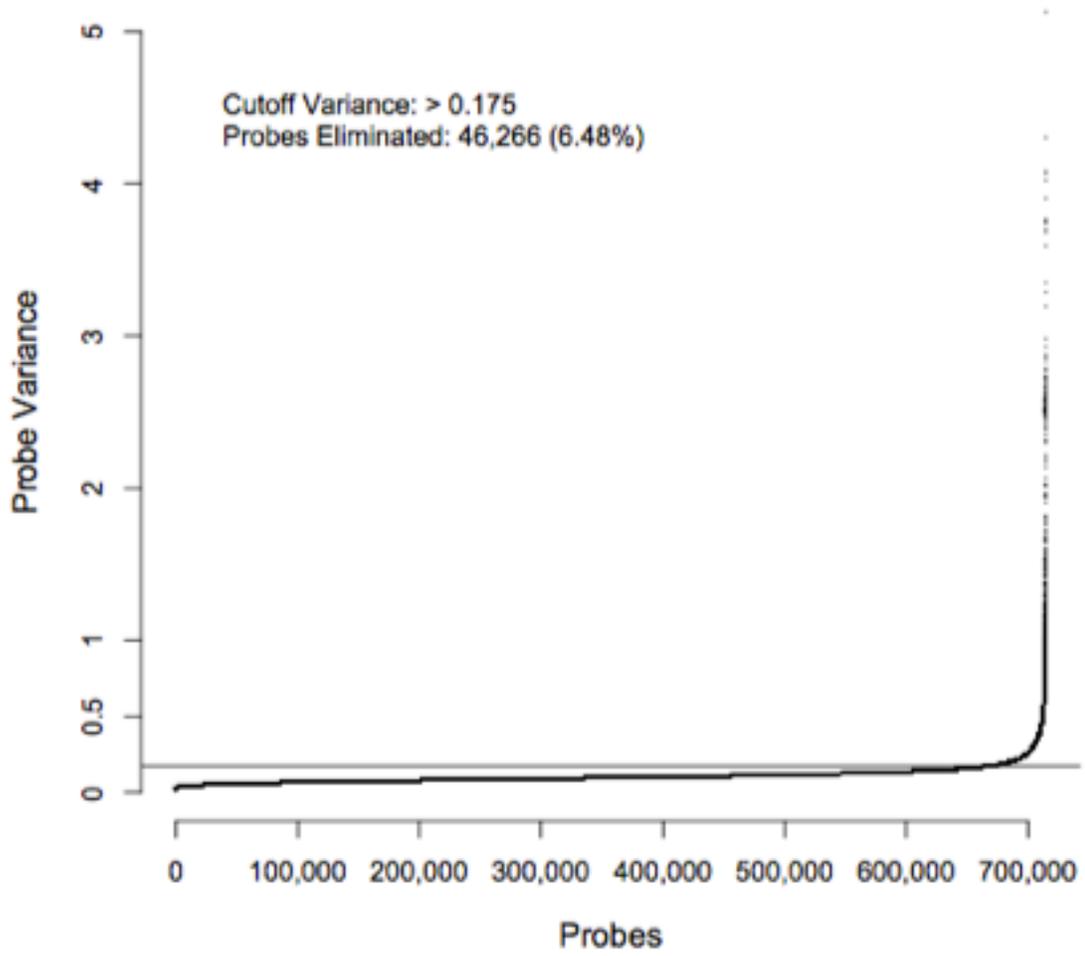
Probes

Figure 5.4 – Probe Variance Analysis
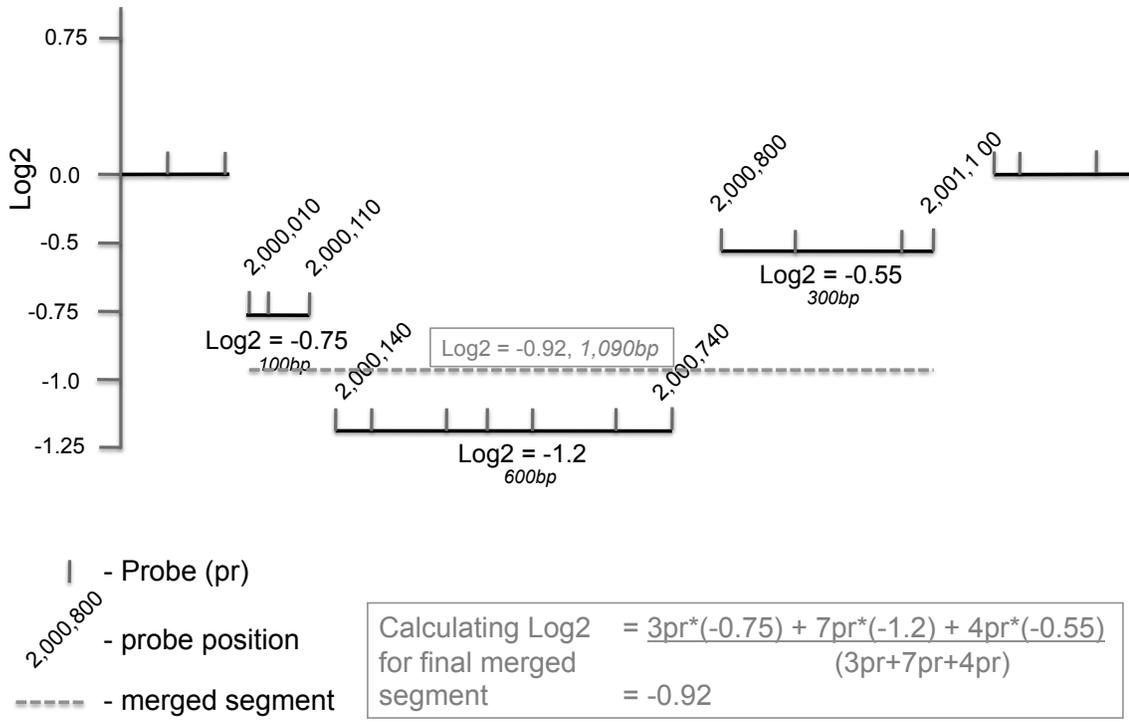
Figure 5.5 - Hypothetical example of merging multiple segments representing a single, deleted locus.
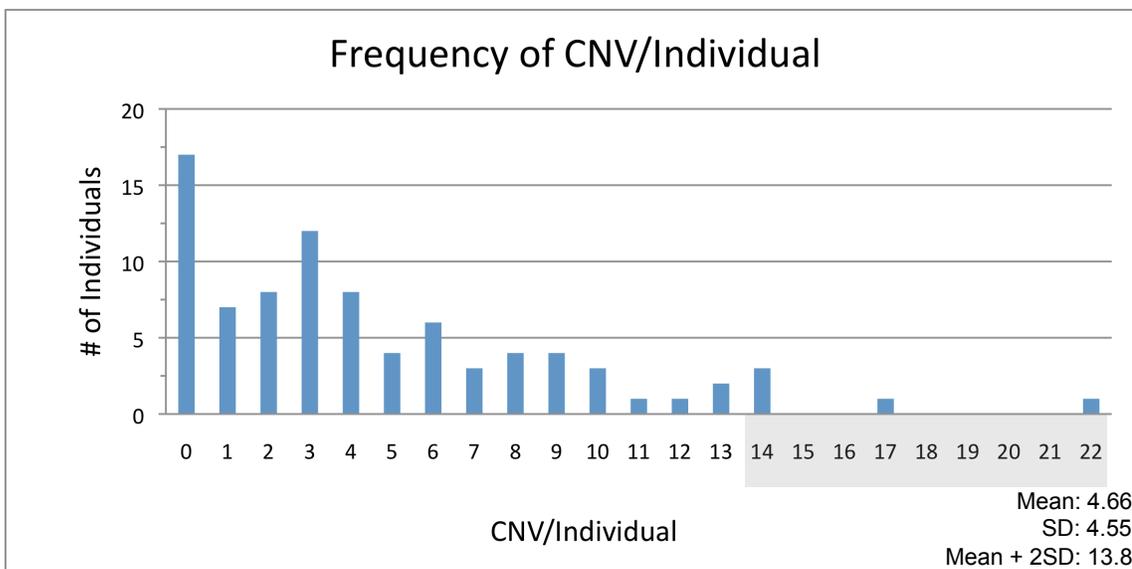
Figure 5.6 – Histogram of the number of CNV calls per individual.

Control: AGTCAACATA Ggtgggaata … aatgtaaaca gGTCAACTAA

29,861 bp

AU065404: AGTCAACATA GGTCAACTAA
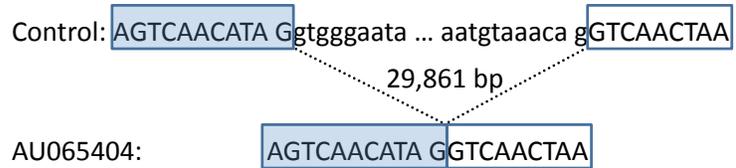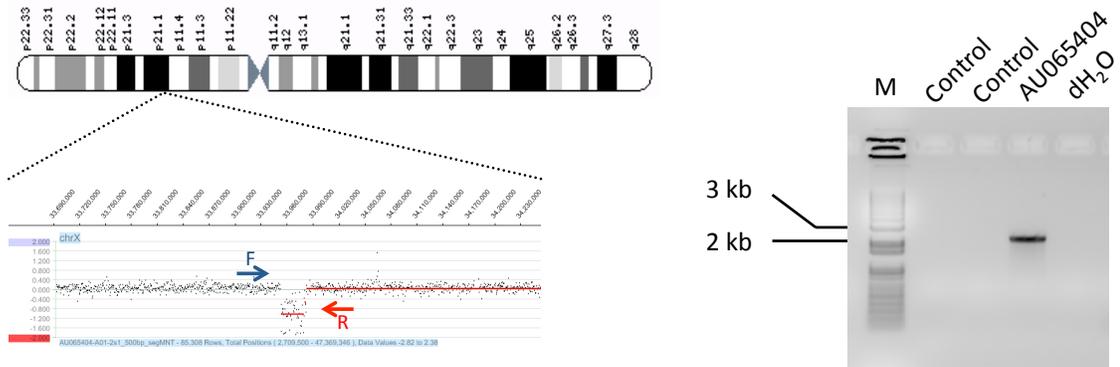
Figure 5.7 - Validation of an array identified deletion.
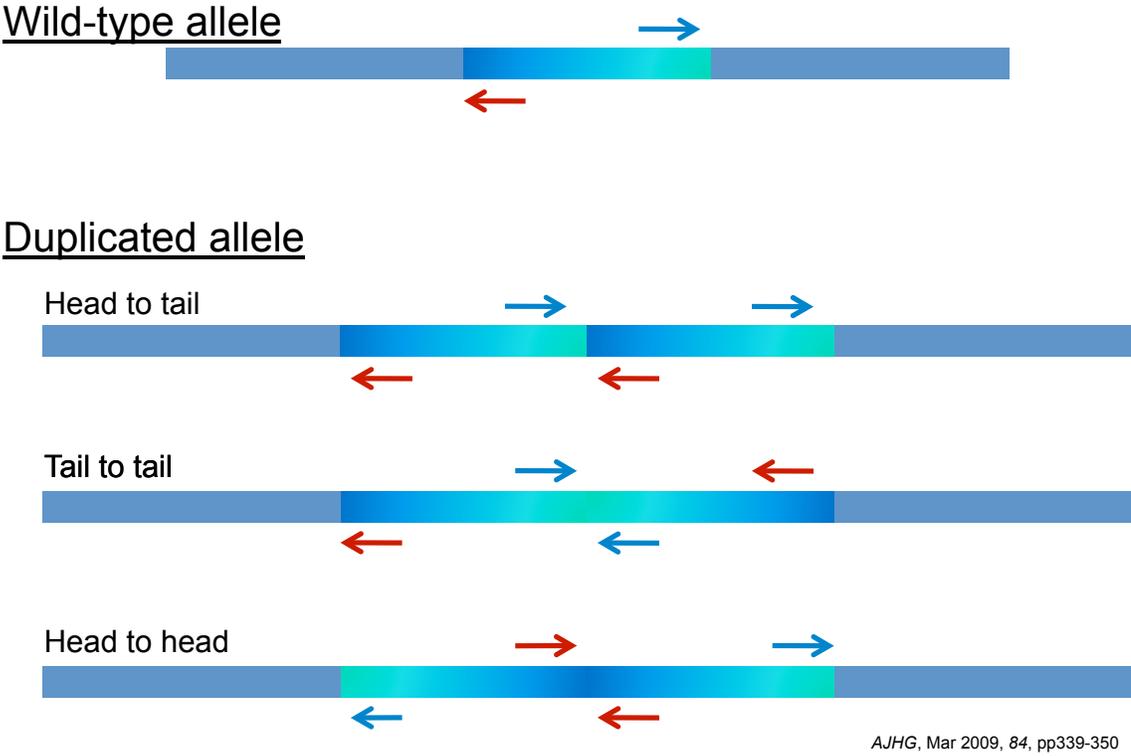
Figure 5.8 – Validation strategy for tandem duplications.
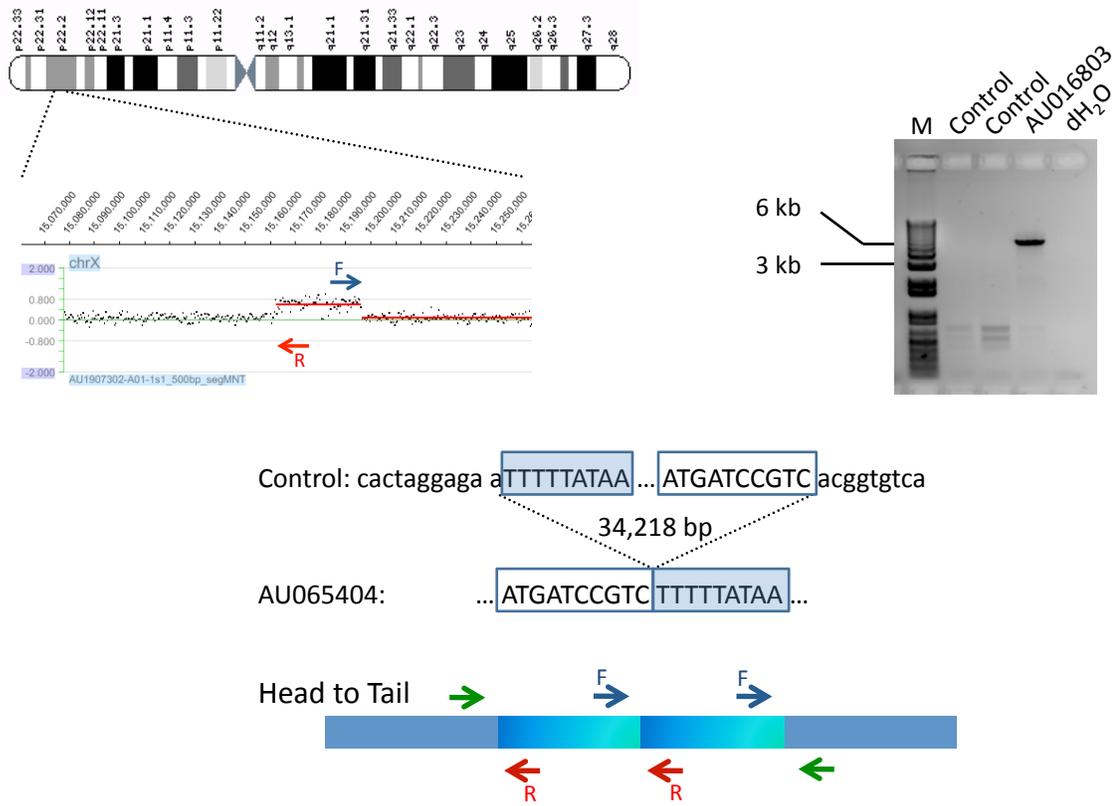
Figure 5.9 - Validation of an array identified duplication.
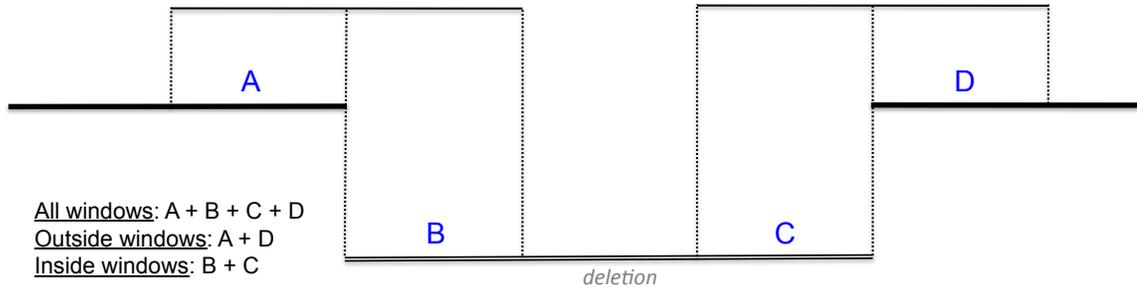
Figure 5.10 – Schematic for breakpoint junction analysis.

Appendix

1. Tables A.1 – A.4

Tables A.1-A.4 list the copy number variants (CNV) identified by the four combinations of array resolution (385,000 (385K) or 2,100,000 probes (2.1M)), array processing protocol (NimbleGen's or our optimized protocol), and sample population (Autism Genetic Resource (AGRE) males with autism, Simons Simplex Collection (SSC) males with autism, or National Institute of Mental Health (NIMH) control population).

CNV shaded in 'grey' indicate CNV or indel grouped together because 1) CNV or indel upstream and/or downstream breaks end within three kilobases of the next segment or 2) CNV or indel are called with complete overlap. 'Yellow' log2 values indicate a minority state (e.g., single deletion among several duplications) within a grouping.

Column headers for these tables are defined as:

ARRAY_ID   Name of original scan file (.tiff)
Sample   Sample ID
CHR   Chromosome
START   Start position (Hg18)
STOP   End position (Hg18)
SIZE   Length of segment in base pairs
Probes   Number of probes that called the segment
Probes/kb   Number of probes per kilobase
LOG2_RATIO   The average Log2 ratio of the Cy3 (test) to Cy5 (reference) intensity
GC   The percent GC content of the segment

2. Table A.5

Table A.5 list CNV and indel identified in the literature with breakpoint sequencing. The data are derived from the following 12 reports:

1   Conrad, D.F., et al., Origins and functional impact of copy number variation in the human genome. Nature, 2010. 464(7289): p. 704-12.

2   de Smith, A.J., et al., Array CGH analysis of copy number variation identifies 1284 new genes variant in healthy white males: implications for association studies of complex diseases. Hum Mol Genet, 2007. 16(23): p. 2783-94.

3   Nichol Edamura, K. and C.E. Pearson, DNA methylation and replication: implications for the "deletion hotspot" region of FMR1. Hum Genet, 2005. 118(2): p. 301-4.

4   Korbel, J.O., et al., Systematic prediction and validation of breakpoints associated with copy-number variants in the human genome. Proc Natl Acad Sci U S A, 2007. 104(24): p. 10110-5.

5   Goldmann, R., et al., Genomic characterization of large rearrangements of the LDLR gene in Czech patients with familial hypercholesterolemia. BMC Med Genet, 2010. 11: p. 115.

6   Kim, P.M., et al., Analysis of copy number variants and segmental duplications in the human genome: Evidence for a change in the process of formation in recent evolutionary history. Genome Res, 2008. 18(12): p. 1865-74.

7   Lam, H.Y., et al., Nucleotide-resolution analysis of structural variants using BreakSeq and a breakpoint library. Nat Biotechnol, 2010. 28(1): p. 47-55.

8   Nobile, C., et al., Analysis of 22 deletion breakpoints in dystrophin intron 49. Hum Genet, 2002. 110(5): p. 418-21.

9   Park, H., et al., Discovery of common Asian copy number variants using integrated high-resolution array CGH and massively parallel DNA sequencing. Nat Genet, 2010. 42(5): p. 400-5.

10   Vissers, L.E., et al., Rare pathogenic microdeletions and tandem duplications are microhomology-mediated and stimulated by local genomic architecture. Hum Mol Genet, 2009. 18(19): p. 3579-93.

11   Woodward, K.J., et al., Heterogeneous duplications in patients with Pelizaeus-Merzbacher disease suggest a mechanism of coupled homologous and nonhomologous recombination. Am J Hum Genet, 2005. 77(6): p. 966-87.

12   Zhang, F., et al., Mechanisms for nonrecurrent genomic rearrangements associated with CMT1A or HNPP: rare CNVs as a cause for missing heritability. Am J Hum Genet, 2010. 86(6): p. 892-903.

13   Emory - CNV and indel identified, validated, and breakpoint sequenced (AGRE and SSC cohorts)

Column headers for this table are defined as:

Ref	The first author of the reference from which the CNV/indel came.
Chr	Chromosome
Start	Start Position (Hg18)
Stop	Stop Position (Hg18)
state	Deletion or duplication or inversion
homology	Any homology noted at the breakpoints. '?' indicates that the original authors/I did not search for homologies

<u>hom. Length</u>   Length of the observed homology
<u>insertion</u>   Any insertion sequence found at the breakpoint
<u>insertion (remove spaces)</u>   Any insertion sequence found at the breakpoint - no spaces
                     in the sequence
<u>ins. Length</u>   Length of the observed insertion
<u>ID</u>   ID the authors may have assigned to the CNV/indel

3. Tables A.6.a-b

These two tables list validated calls by the NimbleGen and Optimized protocols. The

NimbleGen protocol was run on 50 AGRE samples, and the Optimized protocol was run

on 100 AGRE and 64 SSC samples. Both protocols were applied to the 2.1M CCGH

array.

The headers for the true (A.6.a) and falsely (A.6.b) called tables are defined as:

<u>ARRAY_ID</u>   Name of original scan file (.tiff)
<u>Sample</u>   Sample ID
<u>CHR</u>   Chromosome
<u>START</u>   Start position (Hg18)
<u>STOP</u>   End position (Hg18)
<u>SIZE (bp)</u> Length of segment in base pairs (bp)
<u>Probes</u>   Number of probes that called the segment
<u>Probes/kb</u> Number of probes per kilobase (kb)
<u>Mean_Log2</u> The average Log2 ratio of the Cy3 (test) to Cy5 (reference) intensity
<u>GC</u>   The percent GC content of the segment
<u>Primer Set</u> The primer set used to validate the locus
<u>CGH_Protocol</u> Which CGH protocol was used: NG (NimbleGen) or Opt (Optimized)

Table A.1 – CNV from four AGRE samples run on 385K arrays by the NimbleGen protocol.

ARRAY_ID	Sample	CHR	START	STOP	SIZE	Probes	Probes/kb	LOG2_RATIO	GC
AU028304-1s3	AU028304	chrX	34,063,446	34,065,771	2,325	8	3.4	0.55	34.0
AU004404-1s2	AU004404	chrX	67,040,957	67,046,757	5,800	16	2.8	-0.61	37.2
AU028304-1s3	AU028304	chrX	69,648,398	69,649,328	930	4	4.3	-0.88	45.8
AU004404-1s2	AU004404	chrX	78,808,938	78,811,924	2,986	10	3.3	0.69	34.5
AU028304-1s3	AU028304	chrX	78,809,398	78,811,444	2,046	7	3.4	0.63	36.4
AU015004-1s1	AU015004	chrX	78,809,398	78,811,193	1,795	6	3.3	0.87	38.3
AU028304-1s3	AU028304	chrX	115,051,644	115,066,806	15,162	30	2.0	0.37	38.2
AU004404-1s2	AU004404	chrX	115,051,999	115,057,702	5,703	18	3.2	0.54	36.1
AU002204-1s3	AU002204	chrX	115,052,833	115,057,262	4,429	14	3.2	0.41	34.6
AU015004-1s1	AU015004	chrX	115,053,308	115,056,847	3,539	12	3.4	0.64	33.7
AU004404-1s2	AU004404	chrX	126,425,995	126,430,002	4,007	11	2.7	0.62	31.3
AU028304-1s3	AU028304	chrX	126,425,995	126,428,962	2,967	8	2.7	0.64	32.1
AU004404-1s2	AU004404	chrX	147,317,361	147,318,588	1,227	5	4.1	-0.93	38.9
AU028304-1s3	AU028304	chrX	147,317,361	147,318,588	1,227	5	4.1	-0.85	38.9
AU004404-1s2	AU004404	chrX	148,451,913	148,462,535	10,622	30	2.8	-0.46	42.9
AU004404-1s2	AU004404	chrX	148,686,217	148,836,095	149,878	298	2.0	0.27	40.7

Table A.2 - CNV from 50 AGRE samples run on 2.1M arrays by the NimbleGen protocol.

AU0780301-1s1A01	AU0780301	chrX	34,063,531	34,065,771	2,240	33	14.7	0.43	34.1
AU016803-1s1A01	AU016803	chrX	34,063,631	34,065,771	2,140	31	14.5	0.46	34.4
AU002403-1s1A01	AU002403	chrX	34,063,839	34,065,771	1,932	28	14.5	0.67	35.3
AU009503-1s1A01	AU009503	chrX	34,063,839	34,065,771	1,932	28	14.5	0.69	35.3
AU029303-1s1A01	AU029303	chrX	34,063,839	34,065,846	2,007	29	14.4	0.78	36.1
AU083504-1s1A01	AU083504	chrX	34,063,839	34,065,846	2,007	29	14.4	0.57	36.1
AU0875302-1s2A01	AU0875302	chrX	34,063,839	34,065,771	1,932	28	14.5	0.57	35.3
AU012004-1s2A01	AU012004	chrX	36,652,992	36,653,157	165	5	30.3	-0.78	47.1
AU014505-1s1A01	AU014505	chrX	36,652,992	36,653,217	225	6	26.7	-1.64	46.4
AU016803-1s1A01	AU016803	chrX	36,652,992	36,653,217	225	6	26.7	-1.63	46.4
AU080803-1s1A01	AU080803	chrX	36,652,992	36,653,217	225	6	26.7	-1.63	46.4
AU083504-1s1A01	AU083504	chrX	36,652,992	36,653,217	225	6	26.7	-1.91	46.4
AU028903-1s1A01	AU028903	chrX	38,029,782	38,032,254	2,472	38	15.4	-0.26	48.5
AU034904-1s1A01	AU034904	chrX	38,208,962	38,210,787	1,825	39	21.4	-0.34	37.5
AU009503-1s1A01	AU009503	chrX	38,209,627	38,210,842	1,215	26	21.4	-0.50	37.9
AU055503-1s1A01	AU055503	chrX	38,547,627	38,550,411	2,784	59	21.2	0.31	63.0
AU034604-1s1A01	AU034604	chrX	39,525,943	39,549,870	23,927	385	16.1	0.47	45.9
AU009503-1s1A01	AU009503	chrX	39,628,002	39,628,334	332	8	24.1	-0.72	41.5
AU024003-1s1A01	AU024003	chrX	41,096,920	41,097,246	326	8	24.5	-0.69	42.4
AU074704-1s1A01	AU074704	chrX	42,163,664	42,164,864	1,200	26	21.7	-0.31	42.7
AU009503-1s1A01	AU009503	chrX	43,830,073	43,834,898	4,825	81	16.8	-0.32	43.5
AU018304-1s1A01	AU018304	chrX	43,844,730	43,845,335	605	14	23.1	-0.55	47.6
AU0852304-1s1A01	AU0852304	chrX	44,265,421	44,265,881	460	11	23.9	-0.52	36.7
AU024003-1s1A01	AU024003	chrX	45,703,426	45,704,121	695	16	23.0	0.44	53.7
AU034604-1s1A01	AU034604	chrX	46,190,999	46,193,830	2,831	53	18.7	0.31	54.1
AU034904-1s1A01	AU034904	chrX	46,580,348	46,581,804	1,456	25	17.2	0.44	56.1
AU024003-1s1A01	AU024003	chrX	46,888,299	46,888,847	548	9	16.4	0.75	52.9
AU034604-1s1A01	AU034604	chrX	46,888,299	46,888,812	513	8	15.6	0.98	53.2
AU030603-1s3A01	AU030603	chrX	46,888,484	46,888,812	328	7	21.3	1.11	54.4
AU062203-1s1A01	AU062203	chrX	46,888,484	46,888,847	363	8	22.0	0.72	53.8
AU083603-1s1A01	AU083603	chrX	46,888,484	46,888,847	363	8	22.0	0.60	53.8
AU032705-1s1A01	AU032705	chrX	46,888,529	46,888,847	318	7	22.0	0.67	54.9
AU065404-1s1A01	AU065404	chrX	46,888,599	46,888,812	213	5	23.5	1.11	60.2
AU038805-1s1A01	AU038805	chrX	46,963,626	46,965,349	1,723	37	21.5	0.29	53.0
AU043803-1s1A01	AU043803	chrX	46,964,064	46,965,589	1,525	33	21.6	-0.35	50.1
AU016803-1s1A01	AU016803	chrX	47,226,156	47,226,756	600	14	23.3	-0.49	46.0
AU058103-1s1A01	AU058103	chrX	47,226,241	47,226,756	515	12	23.3	-0.53	47.9
AU074704-1s1A01	AU074704	chrX	47,226,241	47,226,756	515	12	23.3	-0.56	47.9
AU018003-1s1A01	AU018003	chrX	47,226,341	47,226,756	415	10	24.1	-0.73	48.2
AU028903-1s1A01	AU028903	chrX	47,322,258	47,322,828	570	13	22.8	-0.58	51.3
AU004803-1s1A02	AU004803	chrX	47,377,313	47,379,754	2,441	39	16.0	-0.28	46.6
AU056603-1s1A02	AU056603	chrX	47,377,313	47,379,754	2,441	39	16.0	-0.40	46.6
AU043803-1s1A02	AU043803	chrX	47,378,054	47,379,930	1,876	39	20.8	-0.47	45.8
AU016803-1s1A02	AU016803	chrX	47,378,074	47,379,704	1,630	35	21.5	-0.35	46.1
AU028903-1s1A02	AU028903	chrX	47,468,996	47,469,816	820	18	22.0	-0.57	55.2
AU038805-1s1A02	AU038805	chrX	47,469,046	47,469,741	695	16	23.0	0.43	55.6
AU056003-1s1A02	AU056003	chrX	47,469,046	47,469,741	695	16	23.0	0.46	55.6
AU066103-1s1A02	AU066103	chrX	47,469,071	47,469,866	795	17	21.4	-0.53	55.2
AU008504-1s1A02	AU008504	chrX	47,814,113	47,818,002	3,889	56	14.4	0.28	49.2
AU018003-1s1A02	AU018003	chrX	47,829,938	47,830,368	430	10	23.3	-0.85	44.8
AU058103-1s1A02	AU058103	chrX	48,219,896	48,221,146	1,250	27	21.6	-0.40	47.2
AU066103-1s1A02	AU066103	chrX	48,219,966	48,221,126	1,160	25	21.6	-0.61	45.9
AU009503-1s1A02	AU009503	chrX	48,220,021	48,221,126	1,105	24	21.7	-0.50	45.5
AU028903-1s1A02	AU028903	chrX	48,220,021	48,221,221	1,200	26	21.7	-0.42	46.3
AU034904-1s1A02	AU034904	chrX	48,220,021	48,221,051	1,030	23	22.3	-0.35	45.0
AU038805-1s1A02	AU038805	chrX	48,220,021	48,221,126	1,105	24	21.7	0.30	45.5
AU043803-1s1A02	AU043803	chrX	48,220,021	48,221,126	1,105	24	21.7	-0.59	45.5
AU055503-1s1A02	AU055503	chrX	48,220,021	48,221,146	1,125	25	22.2	-0.41	45.6
AU034904-1s1A02	AU034904	chrX	48,271,876	48,272,356	480	11	22.9	-0.48	48.3
AU018003-1s1A02	AU018003	chrX	48,271,941	48,272,406	465	11	23.7	-0.67	49.4
AU018003-1s1A02	AU018003	chrX	48,276,384	48,278,319	1,935	40	20.7	-0.39	51.0
AU028903-1s1A02	AU028903	chrX	48,276,384	48,278,284	1,900	39	20.5	-0.45	51.1
AU043803-1s1A02	AU043803	chrX	48,276,384	48,278,379	1,995	41	20.6	-0.48	50.6
AU056003-1s1A02	AU056003	chrX	48,276,454	48,278,944	2,490	44	17.7	0.28	50.2
AU058503-1s1A02	AU058503	chrX	48,453,617	48,458,412	4,795	100	20.9	-0.36	44.6
AU034904-1s1A02	AU034904	chrX	48,453,752	48,458,562	4,810	100	20.8	-0.34	44.7
AU056803-1s1A02	AU056803	chrX	48,453,917	48,462,698	8,781	162	18.4	-0.23	45.1
AU043803-1s1A02	AU043803	chrX	48,454,787	48,458,412	3,625	76	21.0	-0.59	44.7
AU058103-1s1A02	AU058103	chrX	48,456,222	48,462,698	6,476	114	17.6	-0.30	44.8
AU028903-1s1A02	AU028903	chrX	48,530,120	48,531,142	1,022	22	21.5	-0.56	48.3
AU014803-1s1A02	AU014803	chrX	48,530,197	48,531,062	865	19	22.0	-0.33	47.9
AU018003-1s1A02	AU018003	chrX	48,530,197	48,531,062	865	19	22.0	-0.49	47.9
AU034904-1s1A02	AU034904	chrX	48,530,197	48,531,062	865	19	22.0	-0.44	47.9
AU043803-1s1A02	AU043803	chrX	48,530,197	48,530,987	790	18	22.8	-0.91	47.4
AU055503-1s1A02	AU055503	chrX	48,530,197	48,531,062	865	19	22.0	-0.62	47.9
AU056603-1s1A02	AU056603	chrX	48,530,197	48,531,087	890	20	22.5	-0.55	48.1
AU058103-1s1A02	AU058103	chrX	48,530,197	48,531,087	890	20	22.5	-0.55	48.1
AU058503-1s1A02	AU058503	chrX	48,530,197	48,531,062	865	19	22.0	-0.52	47.9
AU009503-1s1A02	AU009503	chrX	48,530,217	48,530,987	770	17	22.1	-0.61	47.1

AU021503-1s1A02	AU021503	chrX	48,530,217	48,531,062	845	18	21.3	-0.42	47.6
AU066103-1s1A02	AU066103	chrX	48,530,217	48,530,987	770	17	22.1	-0.97	47.1
AU0780301-1s1A02	AU0780301	chrX	48,530,217	48,531,087	870	19	21.8	-0.37	47.8
AU043803-1s1A02	AU043803	chrX	48,788,958	48,799,173	10,215	98	9.6	-0.39	50.3
AU018003-1s1A02	AU018003	chrX	48,794,839	48,795,977	1,138	19	16.7	-0.52	48.0
AU066103-1s1A02	AU066103	chrX	48,794,874	48,797,608	2,734	39	14.3	-0.52	46.7
AU028903-1s1A02	AU028903	chrX	48,798,558	48,799,268	710	16	22.5	-0.53	50.8
AU028903-1s1A02	AU028903	chrX	48,823,450	48,824,287	837	18	21.5	-0.54	53.9
AU008504-1s1A02	AU008504	chrX	48,908,480	48,909,147	667	15	22.5	0.49	56.3
AU028903-1s1A02	AU028903	chrX	48,941,474	48,943,347	1,873	40	21.4	-0.39	55.1
AU043803-1s1A02	AU043803	chrX	48,967,547	48,968,172	625	14	22.4	-0.71	51.7
AU038805-1s1A02	AU038805	chrX	48,995,154	48,996,304	1,150	25	21.7	0.39	53.1
AU028903-1s1A02	AU028903	chrX	48,995,264	48,996,414	1,150	25	21.7	-0.60	52.7
AU043803-1s1A02	AU043803	chrX	48,995,294	48,996,284	990	21	21.2	-0.79	52.5
AU055503-1s1A02	AU055503	chrX	48,995,294	48,996,304	1,010	22	21.8	-0.55	52.3
AU066103-1s1A02	AU066103	chrX	48,995,294	48,996,304	1,010	22	21.8	-0.62	52.3
AU058103-1s1A02	AU058103	chrX	48,995,364	48,996,304	940	21	22.3	-0.50	52.1
AU018003-1s1A02	AU018003	chrX	48,995,439	48,996,284	845	18	21.3	-0.47	52.5
AU080803-1s1A02	AU080803	chrX	49,012,390	49,013,090	700	16	22.9	0.59	59.2
AU0920301-1s1A02	AU0920301	chrX	49,012,490	49,013,180	690	16	23.2	0.39	59.2
AU066103-1s1A02	AU066103	chrX	49,720,668	49,721,053	385	9	23.4	-0.71	39.3
AU034904-1s1A02	AU034904	chrX	50,145,837	50,146,797	960	21	21.9	-0.64	44.7
AU066103-1s1A02	AU066103	chrX	50,145,837	50,146,797	960	21	21.9	-0.66	44.7
AU017504-1s1A02	AU017504	chrX	50,145,957	50,146,797	840	19	22.6	-0.60	44.4
AU021503-1s1A02	AU021503	chrX	50,145,957	50,146,797	840	19	22.6	-0.52	44.4
AU055503-1s1A02	AU055503	chrX	50,145,957	50,146,797	840	19	22.6	-0.57	44.4
AU018003-1s1A02	AU018003	chrX	50,146,172	50,146,797	625	14	22.4	-0.73	44.3
AU030603-1s3A02	AU030603	chrX	50,146,422	50,146,797	375	9	24.0	-0.82	41.8
AU008504-1s1A02	AU008504	chrX	50,146,517	50,146,797	280	7	25.0	-0.80	42.9
AU074704-1s1A02	AU074704	chrX	50,146,517	50,146,797	280	7	25.0	-0.85	42.9
AU083504-1s1A02	AU083504	chrX	50,146,517	50,146,797	280	7	25.0	-0.89	42.9
AU012004-1s2A02	AU012004	chrX	50,146,567	50,146,797	230	6	26.1	-0.85	43.2
AU034904-1s1A02	AU034904	chrX	50,572,385	50,574,411	2,026	42	20.7	0.33	60.1
AU018003-1s1A02	AU018003	chrX	51,627,974	51,629,858	1,884	28	14.9	-0.46	44.4
AU043803-1s1A02	AU043803	chrX	51,628,034	51,629,199	1,165	24	20.6	-0.59	41.8
AU034904-1s1A02	AU034904	chrX	51,628,124	51,629,129	1,005	22	21.9	-0.47	42.1
AU058503-1s1A02	AU058503	chrX	51,628,194	51,629,199	1,005	22	21.9	-0.51	42.0
AU009503-1s1A02	AU009503	chrX	51,628,384	51,629,104	720	16	22.2	-0.61	42.3
AU0780301-1s1A02	AU0780301	chrX	53,081,197	53,082,112	915	20	21.9	-0.35	43.3
AU009503-1s1A02	AU009503	chrX	53,097,412	53,108,214	10,802	158	14.6	-0.35	45.1
AU043803-1s1A02	AU043803	chrX	53,097,412	53,107,174	9,762	140	14.3	-0.38	45.6
AU0920301-1s1A02	AU0920301	chrX	53,099,892	53,102,717	2,825	53	18.8	0.36	48.1
AU008504-1s1A02	AU008504	chrX	53,100,432	53,102,672	2,240	47	21.0	0.42	48.1
AU032705-1s1A02	AU032705	chrX	53,100,432	53,102,212	1,780	38	21.3	0.39	48.7
AU038805-1s1A02	AU038805	chrX	53,100,452	53,102,317	1,865	39	20.9	0.45	48.6
AU014505-1s1A02	AU014505	chrX	53,100,517	53,102,427	1,910	40	20.9	0.39	48.3
AU028903-1s1A02	AU028903	chrX	53,100,517	53,102,747	2,230	47	21.1	-0.63	47.6
AU056003-1s1A02	AU056003	chrX	53,100,517	53,102,557	2,040	43	21.1	0.33	48.3
AU058503-1s1A02	AU058503	chrX	53,100,517	53,107,369	6,852	89	13.0	-0.29	46.2
AU021503-1s1A02	AU021503	chrX	53,126,892	53,127,752	860	17	19.8	-0.43	61.1
AU066103-1s1A02	AU066103	chrX	53,439,228	53,439,718	490	11	22.4	-0.69	45.2
AU034904-1s1A02	AU034904	chrX	53,604,041	53,604,391	350	8	22.9	-0.71	46.9
AU018003-1s1A02	AU018003	chrX	54,524,898	54,525,338	440	10	22.7	-0.56	47.2
AU056803-1s1A02	AU056803	chrX	54,790,338	54,791,943	1,605	33	20.6	-0.35	44.2
AU004803-1s1A02	AU004803	chrX	54,790,608	54,791,943	1,335	29	21.7	-0.30	43.8
AU018003-1s1A02	AU018003	chrX	54,790,608	54,791,943	1,335	29	21.7	-0.54	43.8
AU056603-1s1A02	AU056603	chrX	54,790,608	54,791,753	1,145	25	21.8	-0.47	43.7
AU0852304-1s1A02	AU0852304	chrX	54,790,608	54,792,093	1,485	32	21.5	-0.35	43.4
AU016803-1s1A02	AU016803	chrX	54,790,693	54,791,943	1,250	27	21.6	-0.44	43.4
AU021503-1s1A02	AU021503	chrX	54,790,693	54,792,008	1,315	28	21.3	-0.43	42.8
AU034904-1s1A02	AU034904	chrX	54,790,693	54,792,033	1,340	29	21.6	-0.46	42.9
AU055503-1s1A02	AU055503	chrX	54,790,693	54,792,008	1,315	28	21.3	-0.47	42.8
AU0780301-1s1A02	AU0780301	chrX	54,790,693	54,791,893	1,200	26	21.7	-0.39	43.8
AU009503-1s1A02	AU009503	chrX	54,790,763	54,791,893	1,130	25	22.1	-0.48	43.7
AU043803-1s1A02	AU043803	chrX	54,790,763	54,791,943	1,180	26	22.0	-0.51	43.3
AU066103-1s1A02	AU066103	chrX	54,790,763	54,791,868	1,105	24	21.7	-0.67	43.7
AU058103-1s1A02	AU058103	chrX	54,790,858	54,791,868	1,010	22	21.8	-0.49	44.6
AU083504-1s1A02	AU083504	chrX	54,964,123	54,965,178	1,055	23	21.8	0.41	57.1
AU080803-1s1A02	AU080803	chrX	54,964,213	54,965,519	1,306	28	21.4	0.43	55.6
AU066103-1s1A02	AU066103	chrX	55,043,872	55,044,637	765	17	22.2	-0.48	49.5
AU030603-1s3A02	AU030603	chrX	55,087,790	55,090,034	2,244	42	18.7	-1.14	40.3
AU066103-1s1A02	AU066103	chrX	55,790,524	55,790,929	405	9	22.2	-0.94	36.4
AU017504-1s1A02	AU017504	chrX	55,790,644	55,790,929	285	7	24.6	-0.90	35.7
AU029303-1s1A02	AU029303	chrX	55,790,644	55,790,929	285	7	24.6	-1.20	35.7
AU030603-1s3A02	AU030603	chrX	55,790,644	55,790,929	285	7	24.6	-1.16	35.7
AU032705-1s1A02	AU032705	chrX	55,790,644	55,790,929	285	7	24.6	-0.67	35.7
AU034604-1s1A02	AU034604	chrX	55,790,644	55,790,929	285	7	24.6	-1.06	35.7
AU034904-1s1A02	AU034904	chrX	55,790,644	55,790,929	285	7	24.6	-1.07	35.7
AU056603-1s1A02	AU056603	chrX	55,790,644	55,790,929	285	7	24.6	-0.96	35.7

AU065404-1s1A02	AU065404	chrX	55,790,644	55,790,929	285	7	24.6	-0.90	35.7
AU017504-1s1A02	AU017504	chrX	56,804,741	56,805,726	985	19	19.3	-0.52	46.9
AU024003-1s1A02	AU024003	chrX	56,804,741	56,805,631	890	17	19.1	-0.41	47.7
AU1038303-1s1A02	AU1038303	chrX	56,804,741	56,805,081	340	8	23.5	-0.87	48.6
AU008504-1s1A02	AU008504	chrX	56,804,861	56,805,231	370	9	24.3	-0.70	51.4
AU083603-1s1A02	AU083603	chrX	56,804,861	56,805,081	220	6	27.3	-0.99	51.3
AU0920301-1s1A02	AU0920301	chrX	57,617,045	57,634,875	17,830	262	14.7	0.40	43.0
AU050703-1s1A02	AU050703	chrX	57,763,081	57,768,570	5,489	70	12.8	0.79	35.1
AU083504-1s1A02	AU083504	chrX	58,256,031	58,256,531	500	11	22.0	-0.59	46.1
AU083504-1s1A02	AU083504	chrX	62,386,739	62,392,087	5,348	94	17.6	0.47	31.6
AU083504-1s1A02	AU083504	chrX	62,410,230	62,436,937	26,707	343	12.8	0.38	40.6
AU014505-1s1A02	AU014505	chrX	62,410,410	62,415,998	5,588	110	19.7	0.70	43.2
AU0920301-1s1A02	AU0920301	chrX	62,410,410	62,415,948	5,538	109	19.7	0.43	43.1
AU009503-1s1A02	AU009503	chrX	63,363,769	63,364,774	1,005	22	21.9	-0.41	42.0
AU066103-1s1A02	AU066103	chrX	64,685,989	64,688,884	2,895	53	18.3	0.32	61.9
AU034904-1s1A02	AU034904	chrX	65,163,446	65,164,201	755	17	22.5	-1.33	42.0
AU016803-1s1A02	AU016803	chrX	65,382,556	65,399,668	17,112	328	19.2	0.39	38.4
AU034904-1s1A02	AU034904	chrX	66,705,960	66,706,865	905	20	22.1	-0.52	41.1
AU009503-1s1A02	AU009503	chrX	66,706,065	66,706,770	705	16	22.7	-0.59	40.4
AU018003-1s1A02	AU018003	chrX	67,044,882	67,046,967	2,085	44	21.1	-0.60	39.0
AU030603-1s3A02	AU030603	chrX	67,045,802	67,046,892	1,090	24	22.0	-1.43	41.6
AU021503-1s1A02	AU021503	chrX	67,045,887	67,046,967	1,080	23	21.3	-0.85	41.2
AU066103-1s1A02	AU066103	chrX	67,879,957	67,880,347	390	8	20.5	-0.66	47.7
AU066103-1s1A02	AU066103	chrX	68,661,986	68,662,690	704	15	21.3	-0.58	44.5
AU009503-1s1A02	AU009503	chrX	68,662,061	68,662,690	629	14	22.3	-0.46	45.1
AU067803-1s2A02	AU067803	chrX	69,201,942	69,203,705	1,763	38	21.6	0.39	58.6
AU080803-1s1A02	AU080803	chrX	69,582,062	69,582,692	630	14	22.2	0.59	61.6
AU066103-1s1A02	AU066103	chrX	69,644,495	69,645,125	630	14	22.2	-0.55	47.4
AU030603-1s3A02	AU030603	chrX	69,648,398	69,650,576	2,178	30	13.8	-1.26	43.0
AU014505-1s1A02	AU014505	chrX	70,057,777	70,057,992	215	6	27.9	1.02	39.2
AU018304-1s1A02	AU018304	chrX	70,057,777	70,057,992	215	6	27.9	1.04	39.2
AU056003-1s1A02	AU056003	chrX	70,281,553	70,282,617	1,064	23	21.6	0.32	59.4
AU0920301-1s1A02	AU0920301	chrX	70,281,553	70,282,835	1,282	26	20.3	0.43	58.1
AU080803-1s1A02	AU080803	chrX	70,281,623	70,282,463	840	19	22.6	0.65	58.7
AU043803-1s1A02	AU043803	chrX	70,292,124	70,293,024	900	20	22.2	-0.54	48.3
AU0920301-1s1A02	AU0920301	chrX	70,292,254	70,293,429	1,175	25	21.3	==0.44==	50.0
AU058103-1s1A02	AU058103	chrX	70,292,319	70,293,024	705	16	22.7	-0.47	48.6
AU018003-1s1A02	AU018003	chrX	70,292,354	70,292,994	640	14	21.9	-0.51	49.2
AU028903-1s1A02	AU028903	chrX	70,292,424	70,293,479	1,055	23	21.8	-0.50	49.7
AU083504-1s1A02	AU083504	chrX	70,714,603	70,716,868	2,265	48	21.2	0.43	60.0
AU008403-1s1A02	AU008403	chrX	70,714,778	70,716,728	1,950	41	21.0	0.51	61.3
AU050703-1s1A02	AU050703	chrX	70,714,913	70,716,928	2,015	42	20.8	0.30	61.6
AU008504-1s1A02	AU008504	chrX	70,715,008	70,716,928	1,920	40	20.8	0.36	61.5
AU032705-1s1A02	AU032705	chrX	70,715,008	70,716,868	1,860	39	21.0	0.36	62.3
AU067803-1s2A02	AU067803	chrX	70,715,008	70,716,403	1,395	29	20.8	0.36	63.3
AU009904-1s1A02	AU009904	chrX	70,715,118	70,716,868	1,750	37	21.1	0.37	62.2
AU080803-1s1A02	AU080803	chrX	70,715,118	70,716,773	1,655	35	21.1	0.61	62.1
AU005304-1s1A02	AU005304	chrX	70,715,168	70,716,728	1,560	33	21.2	0.28	62.7
AU014505-1s1A02	AU014505	chrX	70,715,248	70,716,868	1,620	34	21.0	0.37	63.2
AU038805-1s1A02	AU038805	chrX	70,715,248	70,716,773	1,525	32	21.0	0.33	63.2
AU014803-1s1A02	AU014803	chrX	71,076,415	71,077,915	1,500	32	21.3	-0.29	45.5
AU034904-1s1A02	AU034904	chrX	71,076,415	71,077,675	1,260	27	21.4	-0.54	45.4
AU066103-1s1A02	AU066103	chrX	71,076,415	71,077,770	1,355	29	21.4	-0.69	45.4
AU1069302-1s1A02	AU1069302	chrX	71,076,415	71,077,525	1,110	24	21.6	-0.46	46.1
AU043803-1s1A02	AU043803	chrX	71,076,460	71,077,675	1,215	26	21.4	-0.67	45.5
AU009503-1s1A02	AU009503	chrX	71,076,500	71,077,710	1,210	26	21.5	-0.47	45.6
AU056603-1s1A02	AU056603	chrX	71,076,500	71,077,770	1,270	27	21.3	-0.53	45.7
AU0780301-1s1A02	AU0780301	chrX	71,076,500	71,077,605	1,105	24	21.7	-0.39	46.0
AU055503-1s1A02	AU055503	chrX	71,076,560	71,077,710	1,150	25	21.7	-0.49	46.2
AU018003-1s1A02	AU018003	chrX	71,076,620	71,077,770	1,150	25	21.7	-0.53	46.6
AU058503-1s1A02	AU058503	chrX	71,076,620	71,077,710	1,090	24	22.0	-0.56	46.5
AU058103-1s1A02	AU058103	chrX	71,076,650	71,077,770	1,120	24	21.4	-0.55	46.6
AU056603-1s1A02	AU056603	chrX	72,262,655	72,263,025	370	9	24.3	-0.82	38.3
AU1038303-1s1A02	AU1038303	chrX	73,082,436	73,084,643	2,207	29	13.1	-0.63	38.8
AU0920301-1s1A02	AU0920301	chrX	75,918,087	75,919,468	1,381	23	16.7	0.50	40.2
AU024003-1s1A02	AU024003	chrX	76,541,200	76,541,585	385	9	23.4	-0.68	41.2
AU002403-1s2A02	AU002403	chrX	77,008,320	77,010,300	1,980	29	14.6	0.76	42.7
AU018304-1s1A02	AU018304	chrX	77,008,320	77,010,225	1,905	28	14.7	0.86	43.1
AU021503-1s1A02	AU021503	chrX	77,008,320	77,008,540	220	6	27.3	1.13	37.4
AU058103-1s1A02	AU058103	chrX	77,008,320	77,008,540	220	6	27.3	0.94	37.4
AU1038303-1s1A02	AU1038303	chrX	77,008,320	77,010,330	2,010	30	14.9	0.61	42.4
AU083504-1s1A02	AU083504	chrX	77,010,060	77,010,225	165	5	30.3	1.75	40.8
AU034904-1s1A02	AU034904	chrX	77,908,755	77,911,012	2,257	23	10.2	-0.46	37.2
AU018304-1s1A02	AU018304	chrX	78,485,431	78,490,725	5,294	52	9.8	0.37	34.7
AU1038303-1s1A02	AU1038303	chrX	82,086,497	82,094,944	8,447	91	10.8	-0.53	39.0
AU066103-1s1A02	AU066103	chrX	85,039,398	85,040,403	1,005	22	21.9	-0.47	42.4
AU065404-1s1A02	AU065404	chrX	86,344,937	86,345,922	985	21	21.3	0.51	48.9
AU055503-1s1A02	AU055503	chrX	88,186,430	88,193,757	7,327	136	18.6	-0.35	33.7
AU004803-1s1A02	AU004803	chrX	92,682,938	92,688,057	5,119	78	15.2	-0.63	39.8

AU014505-1s1A02	AU014505	chrX	92,682,938	92,684,510	1,572	32	20.4	-1.09	35.7
AU014803-1s1A02	AU014803	chrX	92,682,938	92,683,038	100	3	30.0	-1.25	34.4
AU017504-1s1A02	AU017504	chrX	92,682,938	92,684,690	1,752	33	18.8	-1.04	36.7
AU029303-1s1A02	AU029303	chrX	92,682,938	92,687,794	4,856	77	15.9	-0.86	39.8
AU030603-1s3A02	AU030603	chrX	92,682,938	92,687,794	4,856	77	15.9	-0.82	39.8
AU055503-1s1A02	AU055503	chrX	92,682,938	92,684,690	1,752	33	18.8	-0.87	36.7
AU0852304-1s1A02	AU0852304	chrX	93,989,084	93,990,989	1,905	39	20.5	0.24	28.1
AU018304-1s1A02	AU018304	chrX	94,346,925	94,347,740	815	18	22.1	-0.38	48.5
AU030603-1s3A02	AU030603	chrX	94,937,476	94,938,101	625	14	22.4	1.99	37.8
AU034904-1s1A02	AU034904	chrX	94,937,476	94,938,101	625	14	22.4	1.95	37.8
AU066103-1s1A02	AU066103	chrX	94,937,476	94,938,026	550	13	23.6	1.37	37.4
AU018003-1s1A02	AU018003	chrX	94,937,506	94,938,026	520	12	23.1	0.85	37.8
AU024003-1s1A02	AU024003	chrX	94,937,506	94,938,101	595	13	21.8	1.58	38.2
AU050703-1s1A02	AU050703	chrX	94,937,506	94,938,026	520	12	23.1	0.93	37.8
AU055303-1s1A02	AU055303	chrX	94,937,506	94,938,026	520	12	23.1	1.22	37.8
AU058503-1s1A02	AU058503	chrX	94,937,506	94,938,026	520	12	23.1	1.44	37.8
AU1038303-1s1A02	AU1038303	chrX	94,937,506	94,938,101	595	13	21.8	1.22	38.2
AU066103-1s1A02	AU066103	chrX	94,940,516	94,942,651	2,135	40	18.7	-0.32	37.4
AU083504-1s1A02	AU083504	chrX	96,493,478	96,493,778	300	7	23.3	-0.74	41.3
AU0875302-1s1A02	AU0875302	chrX	96,985,461	96,986,463	1,002	22	22.0	-0.40	53.4
AU009503-1s1A02	AU009503	chrX	97,054,641	97,054,971	330	8	24.2	-1.44	37.1
AU066103-1s1A02	AU066103	chrX	99,821,471	99,822,616	1,145	24	21.0	-0.50	41.3
AU0852304-1s1A02	AU0852304	chrX	99,863,364	99,864,422	1,058	16	15.1	-0.41	42.1
AU080803-1s1A02	AU080803	chrX	100,359,562	100,361,746	2,184	39	17.9	0.31	44.4
AU024003-1s1A02	AU024003	chrX	100,868,724	100,869,338	614	13	21.2	-0.53	38.4
AU017504-1s1A02	AU017504	chrX	102,916,682	102,918,175	1,493	32	21.4	-0.43	47.4
AU021503-1s1A02	AU021503	chrX	103,062,566	103,110,269	47,703	566	11.9	0.53	40.9
AU021503-1s1A02	AU021503	chrX	103,147,977	103,192,016	44,039	624	14.2	0.71	43.0
AU1038303-1s1A02	AU1038303	chrX	103,295,654	103,296,615	961	15	15.6	-1.35	40.5
AU012004-1s2A02	AU012004	chrX	103,826,567	103,827,112	545	11	20.2	0.49	38.5
AU004803-1s1A02	AU004803	chrX	105,828,470	105,829,000	530	12	22.6	-0.59	44.3
AU066103-1s1A02	AU066103	chrX	106,045,447	106,047,742	2,295	42	18.3	-0.38	44.5
AU024003-1s1A02	AU024003	chrX	106,918,604	106,919,763	1,159	25	21.6	-0.37	41.0
AU018003-1s1A03	AU018003	chrX	107,366,406	107,367,451	1,045	22	21.1	-0.44	44.0
AU066103-1s1A03	AU066103	chrX	107,366,406	107,367,291	885	19	21.5	-0.74	44.0
AU1069302-1s1A03	AU1069302	chrX	107,366,406	107,367,451	1,045	22	21.1	-0.37	44.0
AU034904-1s1A03	AU034904	chrX	107,864,565	107,866,535	1,970	42	21.3	0.40	62.7
AU021503-1s1A03	AU021503	chrX	109,131,515	109,132,446	931	20	21.5	0.45	58.5
AU043803-1s1A03	AU043803	chrX	110,234,691	110,236,566	1,875	39	20.8	-0.44	40.1
AU056603-1s1A03	AU056603	chrX	111,172,139	111,172,529	390	9	23.1	-1.06	42.0
AU009904-1s1A03	AU009904	chrX	111,683,385	111,683,880	495	11	22.2	0.65	38.2
AU024003-1s1A03	AU024003	chrX	111,683,455	111,683,810	355	9	25.4	0.75	41.6
AU004803-1s1A03	AU004803	chrX	113,234,646	113,241,227	6,581	122	18.5	-0.83	37.8
AU066103-1s1A03	AU066103	chrX	114,211,054	114,212,014	960	21	21.9	-0.53	39.6
AU056603-1s1A03	AU056603	chrX	114,211,104	114,211,944	840	19	22.6	-0.59	39.7
AU024003-1s1A03	AU024003	chrX	114,211,134	114,212,199	1,065	23	21.6	-0.50	40.4
AU065404-1s1A03	AU065404	chrX	114,330,245	114,333,513	3,268	68	20.8	0.33	64.3
AU021503-1s1A03	AU021503	chrX	114,330,510	114,333,265	2,755	57	20.7	0.31	65.6
AU034604-1s1A03	AU034604	chrX	114,331,010	114,333,225	2,215	46	20.8	0.45	64.6
AU080803-1s1A03	AU080803	chrX	114,331,350	114,333,225	1,875	39	20.8	0.38	65.1
AU018003-1s1A03	AU018003	chrX	114,331,470	114,333,225	1,755	36	20.5	0.47	65.1
AU018304-1s1A03	AU018304	chrX	114,449,045	114,452,862	3,817	73	19.1	0.35	36.9
AU014505-1s1A03	AU014505	chrX	114,449,640	114,450,305	665	15	22.6	0.98	33.2
AU016803-1s1A03	AU016803	chrX	114,449,740	114,450,445	705	16	22.7	0.91	34.1
AU029303-1s1A03	AU029303	chrX	114,449,740	114,450,340	600	14	23.3	1.32	34.3
AU1038303-1s1A03	AU1038303	chrX	114,449,740	114,450,405	665	15	22.6	0.72	34.5
AU014803-1s1A03	AU014803	chrX	115,050,529	115,057,412	6,883	125	18.2	0.33	35.5
AU008403-1s1A03	AU008403	chrX	115,051,804	115,054,775	2,971	52	17.5	0.32	33.6
AU080803-1s1A03	AU080803	chrX	115,051,869	115,059,128	7,259	132	18.2	0.40	37.5
AU002403-1s2A03	AU002403	chrX	115,051,999	115,066,171	14,172	145	10.2	0.39	38.5
AU018304-1s1A03	AU018304	chrX	115,052,163	115,070,463	18,300	200	10.9	0.41	37.0
AU083504-1s1A03	AU083504	chrX	115,066,476	115,066,596	120	4	33.3	1.05	32.6
AU018304-1s1A03	AU018304	chrX	116,422,024	116,422,614	590	13	22.0	-0.80	36.1
AU055303-1s1A03	AU055303	chrX	116,422,024	116,422,614	590	13	22.0	-0.54	36.1
AU0875302-1s2A03	AU0875302	chrX	116,422,134	116,422,614	480	11	22.9	-0.68	35.0
AU066103-1s1A03	AU066103	chrX	117,574,923	117,575,298	375	9	24.0	-0.80	39.4
AU038805-1s1A03	AU038805	chrX	117,804,762	117,806,258	1,496	16	10.7	-0.36	40.4
AU065404-1s1A03	AU065404	chrX	117,845,440	117,846,430	990	22	22.2	0.34	54.8
AU0920301-1s1A03	AU0920301	chrX	118,254,638	118,255,609	971	20	20.6	0.50	58.8
AU014803-1s1A03	AU014803	chrX	120,665,480	120,666,390	910	19	20.9	-0.46	36.9
AU056003-1s1A03	AU056003	chrX	121,710,758	121,711,793	1,035	15	14.5	0.41	39.2
AU065404-1s1A03	AU065404	chrX	121,906,483	121,907,369	886	17	19.2	-0.41	37.7
AU083504-1s1A03	AU083504	chrX	122,143,410	122,144,175	765	17	22.2	-0.73	39.8
AU034904-1s1A03	AU034904	chrX	122,143,675	122,144,235	560	13	23.2	-0.95	40.5
AU066103-1s1A03	AU066103	chrX	122,191,374	122,192,284	910	20	22.0	-0.48	40.2
AU065404-1s1A03	AU065404	chrX	122,200,279	122,201,069	790	18	22.8	-1.03	41.8
AU032705-1s2A03	AU032705	chrX	122,200,614	122,201,069	455	11	24.2	-1.00	44.1
AU055303-1s1A03	AU055303	chrX	122,200,614	122,201,069	455	11	24.2	-1.09	44.1
AU050703-1s1A03	AU050703	chrX	122,778,733	122,780,189	1,456	23	15.8	-0.70	46.3

AU034904-1s1A03	AU034904	chrX	124,109,638	124,110,833	1,195	26	21.8	-0.36	38.5
AU032705-1s2A03	AU032705	chrX	124,787,689	124,788,214	525	12	22.9	-0.92	43.4
AU034904-1s1A03	AU034904	chrX	125,126,380	125,128,920	2,540	51	20.1	0.33	62.9
AU066103-1s1A03	AU066103	chrX	125,433,357	125,434,831	1,474	14	9.5	-1.14	43.9
AU043803-1s1A03	AU043803	chrX	125,433,392	125,434,831	1,439	13	9.0	-1.55	43.3
AU058103-1s1A03	AU058103	chrX	125,433,517	125,434,906	1,389	13	9.4	-1.04	43.3
AU034904-1s1A03	AU034904	chrX	126,420,614	126,430,002	9,388	137	14.6	0.30	33.2
AU029303-1s1A03	AU029303	chrX	126,425,055	126,430,052	4,997	83	16.6	0.55	31.4
AU083504-1s1A03	AU083504	chrX	126,425,325	126,430,002	4,677	77	16.5	0.41	31.5
AU066103-1s1A03	AU066103	chrX	128,291,066	128,291,662	596	13	21.8	-0.69	38.5
AU009904-1s1A03	AU009904	chrX	128,328,210	128,330,085	1,875	38	20.3	-0.64	40.5
AU083504-1s1A03	AU083504	chrX	128,328,210	128,330,085	1,875	38	20.3	-0.77	40.5
AU065404-1s1A03	AU065404	chrX	128,328,430	128,330,085	1,655	34	20.5	-0.72	40.2
AU030603-1s3A03	AU030603	chrX	128,328,535	128,330,085	1,550	33	21.3	-1.13	39.5
AU034904-1s1A03	AU034904	chrX	128,328,560	128,330,085	1,525	32	21.0	-0.93	39.7
AU034604-1s1A03	AU034604	chrX	128,892,826	128,894,009	1,183	25	21.1	0.57	64.7
AU066103-1s1A03	AU066103	chrX	129,306,360	129,306,695	335	8	23.9	-0.78	41.7
AU066103-1s1A03	AU066103	chrX	129,711,202	129,715,287	4,085	78	19.1	-0.45	41.5
AU018003-1s1A03	AU018003	chrX	129,711,252	129,714,877	3,625	68	18.8	-0.37	41.1
AU028903-1s1A03	AU028903	chrX	129,711,302	129,715,197	3,895	74	19.0	-0.32	41.4
AU056803-1s1A03	AU056803	chrX	129,711,402	129,715,172	3,770	71	18.8	-0.33	41.3
AU058103-1s1A03	AU058103	chrX	129,711,442	129,715,287	3,845	73	19.0	-0.35	41.3
AU0852304-1s1A03	AU0852304	chrX	129,713,932	129,715,267	1,335	27	20.2	-0.44	44.0
AU083603-1s1A03	AU083603	chrX	130,832,459	130,836,198	3,739	62	16.6	-0.49	40.2
AU020003-1s1A03	AU020003	chrX	130,832,744	130,835,768	3,024	47	15.5	-0.28	40.8
AU016803-1s1A03	AU016803	chrX	130,832,939	130,836,198	3,259	52	16.0	-0.83	41.3
AU002403-1s2A03	AU002403	chrX	130,833,279	130,836,198	2,919	51	17.5	-1.02	40.4
AU043803-1s1A03	AU043803	chrX	130,833,279	130,836,198	2,919	51	17.5	-1.09	40.4
AU067803-1s2A03	AU067803	chrX	130,833,279	130,836,198	2,919	51	17.5	-1.16	40.4
AU083504-1s1A03	AU083504	chrX	130,833,279	130,836,198	2,919	51	17.5	-1.05	40.4
AU1069302-1s1A03	AU1069302	chrX	130,833,279	130,836,198	2,919	51	17.5	-0.75	40.4
AU009904-1s1A03	AU009904	chrX	130,833,304	130,836,198	2,894	50	17.3	-0.91	40.6
AU080803-1s1A03	AU080803	chrX	130,833,304	130,836,198	2,894	50	17.3	-0.83	40.6
AU020003-1s1A03	AU020003	chrX	130,835,813	130,836,198	385	9	23.4	-1.07	42.8
AU056803-1s1A03	AU056803	chrX	131,572,555	131,573,010	455	11	24.2	-0.77	42.7
AU0875302-1s2A03	AU0875302	chrX	131,766,925	131,769,270	2,345	47	20.0	-0.79	41.7
AU018304-1s1A03	AU018304	chrX	131,917,783	131,919,816	2,033	43	21.2	0.40	65.1
AU034604-1s1A03	AU034604	chrX	133,133,179	133,135,279	2,100	44	21.0	0.33	59.5
AU056003-1s1A03	AU056003	chrX	133,346,142	133,350,722	4,580	83	18.1	-0.37	39.9
AU021503-1s1A03	AU021503	chrX	133,692,426	133,693,396	970	21	21.6	-0.86	43.7
AU008504-1s1A03	AU008504	chrX	134,119,690	134,159,690	40,000	568	14.2	0.41	40.2
AU034604-1s1A03	AU034604	chrX	135,056,972	135,058,348	1,376	30	21.8	0.53	71.3
AU050703-1s1A03	AU050703	chrX	135,057,257	135,058,152	895	20	22.3	0.42	70.8
AU080803-1s1A03	AU080803	chrX	135,057,642	135,058,248	606	14	23.1	0.67	65.7
AU066103-1s1A03	AU066103	chrX	135,359,543	135,360,598	1,055	23	21.8	-0.74	42.0
AU018003-1s1A03	AU018003	chrX	135,359,623	135,361,388	1,765	23	13.0	-0.40	39.5
AU034904-1s1A03	AU034904	chrX	135,767,106	135,767,811	705	10	14.2	-0.63	44.6
AU030603-1s3A03	AU030603	chrX	135,899,645	135,904,947	5,302	85	16.0	-1.02	36.3
AU028903-1s1A03	AU028903	chrX	135,941,581	135,942,831	1,250	27	21.6	-0.39	52.1
AU066103-1s1A03	AU066103	chrX	135,941,971	135,943,299	1,328	28	21.1	-0.42	50.3
AU066103-1s1A03	AU066103	chrX	135,943,359	135,944,189	830	18	21.7	0.38	69.3
AU018003-1s1A03	AU018003	chrX	137,810,150	137,810,705	555	12	21.6	-0.57	43.7
AU0852304-1s1A03	AU0852304	chrX	138,905,808	138,906,103	295	7	23.7	-0.53	49.3
AU034904-1s1A03	AU034904	chrX	139,000,488	139,002,311	1,823	39	21.4	0.33	62.3
AU083603-1s1A03	AU083603	chrX	139,410,814	139,411,996	1,182	26	22.0	0.34	54.1
AU028903-1s1A03	AU028903	chrX	139,623,380	139,624,140	760	16	21.1	-0.49	49.9
AU009503-1s1A03	AU009503	chrX	139,640,097	139,640,577	480	11	22.9	-0.69	43.6
AU009503-1s1A03	AU009503	chrX	139,814,885	139,816,340	1,455	31	21.3	-0.44	39.0
AU066103-1s1A03	AU066103	chrX	140,219,706	140,221,773	2,067	42	20.3	-0.35	40.7
AU1038303-1s1A03	AU1038303	chrX	142,179,913	142,181,007	1,094	23	21.0	-0.49	37.2
AU083603-1s1A03	AU083603	chrX	142,548,496	142,550,865	2,369	49	20.7	0.29	61.4
AU062203-1s1A03	AU062203	chrX	142,743,647	142,746,081	2,434	36	14.8	0.26	32.8
AU002403-1s2A03	AU002403	chrX	143,111,964	143,152,366	40,402	494	12.2	-0.53	35.4
AU002403-1s2A03	AU002403	chrX	143,163,522	143,175,922	12,400	211	17.0	-0.45	36.2
AU043803-1s1A03	AU043803	chrX	143,213,032	143,213,462	430	10	23.3	-0.68	37.0
AU009904-1s1A03	AU009904	chrX	143,436,349	143,441,886	5,537	106	19.1	-0.46	38.8
AU066103-1s1A03	AU066103	chrX	143,818,658	143,818,868	210	5	23.8	-1.23	37.7
AU055303-1s1A03	AU055303	chrX	143,962,926	143,964,923	1,997	43	21.5	-0.30	40.6
AU083504-1s1A03	AU083504	chrX	144,229,903	144,231,198	1,295	23	17.8	0.54	40.3
AU065404-1s1A03	AU065404	chrX	145,622,896	145,623,136	240	6	25.0	-0.78	41.8
AU056003-1s1A03	AU056003	chrX	145,658,097	145,658,347	250	6	24.0	-0.61	43.7
AU066103-1s1A03	AU066103	chrX	146,211,155	146,211,895	740	16	21.6	-0.54	40.7
AU056803-1s1A03	AU056803	chrX	147,316,191	147,318,913	2,722	56	20.6	-0.33	40.7
AU1069302-1s1A03	AU1069302	chrX	147,316,526	147,318,858	2,332	48	20.6	-0.51	39.9
AU073003-1s1A03	AU073003	chrX	147,316,601	147,318,858	2,257	47	20.8	-0.40	39.9
AU008504-1s1A03	AU008504	chrX	147,317,016	147,318,858	1,842	38	20.6	-0.63	39.1
AU002403-1s2A03	AU002403	chrX	147,317,161	147,318,913	1,752	36	20.5	-0.98	39.1
AU024003-1s1A03	AU024003	chrX	147,317,161	147,318,858	1,697	35	20.6	-0.83	39.3
AU005304-1s1A03	AU005304	chrX	147,317,206	147,318,858	1,652	34	20.6	-0.50	39.1

AU1038303-1s1A03	AU1038303	chrX	147,317,206	147,318,858	1,652	34	20.6	-0.69	39.1
AU080803-1s1A03	AU080803	chrX	147,317,251	147,318,858	1,607	33	20.5	-0.80	39.2
AU004803-1s1A03	AU004803	chrX	147,317,311	147,318,858	1,547	32	20.7	-0.92	39.2
AU014505-1s1A03	AU014505	chrX	147,317,311	147,318,858	1,547	32	20.7	-0.87	39.2
AU029303-1s1A03	AU029303	chrX	147,317,311	147,318,913	1,602	33	20.6	-1.19	38.9
AU030603-1s3A03	AU030603	chrX	147,317,361	147,318,858	1,497	31	20.7	-1.18	39.0
AU0920301-1s1A03	AU0920301	chrX	148,428,696	148,430,108	1,412	28	19.8	0.49	66.2
AU073003-1s1A03	AU073003	chrX	148,452,537	148,453,157	620	14	22.6	-1.91	55.2
AU034904-1s1A03	AU034904	chrX	148,543,469	148,543,924	455	9	19.8	-1.05	31.4
AU056003-1s1A03	AU056003	chrX	149,678,510	149,679,625	1,115	24	21.5	-0.58	55.4
AU029303-1s1A03	AU029303	chrX	149,678,535	149,679,480	945	20	21.2	-0.90	55.2
AU055303-1s1A03	AU055303	chrX	149,678,535	149,679,480	945	20	21.2	-0.67	55.2
AU1038303-1s1A03	AU1038303	chrX	149,678,535	149,679,545	1,010	21	20.8	-0.71	55.3
AU014505-1s1A03	AU014505	chrX	149,678,585	149,679,480	895	19	21.2	-0.71	55.7
AU018003-1s1A03	AU018003	chrX	149,766,445	149,767,145	700	14	20.0	-0.43	45.6
AU016803-1s1A03	AU016803	chrX	149,942,828	149,946,773	3,945	83	21.0	-0.31	37.3
AU066103-1s1A03	AU066103	chrX	149,942,828	149,946,773	3,945	83	21.0	-0.45	37.3
AU018003-1s1A03	AU018003	chrX	149,942,873	149,946,743	3,870	81	20.9	-0.36	37.3
AU034904-1s1A03	AU034904	chrX	149,942,873	149,946,958	4,085	86	21.1	-0.38	37.3
AU058103-1s1A03	AU058103	chrX	149,942,898	149,946,868	3,970	83	20.9	-0.34	37.3
AU0852304-1s1A03	AU0852304	chrX	149,942,968	149,946,833	3,865	81	21.0	-0.31	37.3
AU058503-1s1A03	AU058503	chrX	149,942,993	149,946,868	3,875	81	20.9	-0.30	37.3
AU024003-1s1A03	AU024003	chrX	150,044,196	150,046,376	2,180	20	9.2	0.98	38.9
AU056003-1s1A03	AU056003	chrX	150,044,246	150,046,534	2,288	21	9.2	0.56	38.5
AU014803-1s1A03	AU014803	chrX	150,045,071	150,046,514	1,443	18	12.5	0.80	38.9
AU017504-1s1A03	AU017504	chrX	150,045,071	150,046,514	1,443	18	12.5	0.85	38.9
AU029303-1s1A03	AU029303	chrX	150,045,071	150,046,514	1,443	18	12.5	0.98	38.9
AU0875302-1s2A03	AU0875302	chrX	150,045,071	150,046,281	1,210	15	12.4	0.88	40.4
AU043803-1s1A03	AU043803	chrX	150,045,266	150,046,281	1,015	14	13.8	0.99	39.6
AU056603-1s1A03	AU056603	chrX	150,045,266	150,046,376	1,110	16	14.4	0.87	39.3
AU012004-1s2A03	AU012004	chrX	150,045,576	150,046,281	705	13	18.4	0.58	39.2
AU018304-1s1A03	AU018304	chrX	150,045,576	150,046,351	775	14	18.1	1.02	39.1
AU020003-1s1A03	AU020003	chrX	150,045,576	150,046,534	958	17	17.7	0.73	37.1
AU050703-1s1A03	AU050703	chrX	150,045,576	150,046,514	938	16	17.1	0.87	37.2
AU073003-1s1A03	AU073003	chrX	150,045,576	150,046,281	705	13	18.4	0.70	39.2
AU1038303-1s1A03	AU1038303	chrX	150,045,576	150,046,534	958	17	17.7	0.77	37.1
AU016803-1s1A03	AU016803	chrX	150,045,626	150,046,376	750	14	18.7	0.99	38.7
AU034904-1s1A03	AU034904	chrX	150,608,758	150,610,143	1,385	29	20.9	-0.33	44.0
AU056603-1s1A03	AU056603	chrX	150,608,758	150,610,088	1,330	28	21.1	-0.43	43.7
AU009503-1s1A03	AU009503	chrX	150,608,963	150,610,618	1,655	36	21.8	-0.35	44.4
AU043803-1s1A03	AU043803	chrX	150,608,963	150,610,088	1,125	25	22.2	-0.39	44.5
AU058503-1s1A03	AU058503	chrX	150,608,963	150,610,143	1,180	26	22.0	-0.43	44.9
AU066103-1s1A03	AU066103	chrX	150,608,963	150,610,398	1,435	31	21.6	-0.46	44.7
AU018003-1s1A03	AU018003	chrX	150,815,800	150,816,480	680	15	22.1	-0.53	43.0
AU0920301-1s1A03	AU0920301	chrX	151,742,256	151,744,441	2,185	46	21.1	0.46	53.4
AU056003-1s1A03	AU056003	chrX	151,742,591	151,744,391	1,800	38	21.1	0.31	54.1
AU056603-1s1A03	AU056603	chrX	152,269,845	152,270,135	290	7	24.1	-1.03	44.1
AU062203-1s1A03	AU062203	chrX	152,269,845	152,270,135	290	7	24.1	-0.85	44.1
AU080803-1s1A03	AU080803	chrX	152,269,845	152,270,175	330	8	24.2	-0.91	44.4
AU014803-1s1A03	AU014803	chrX	152,269,895	152,270,175	280	7	25.0	-0.89	43.9
AU043803-1s1A03	AU043803	chrX	152,269,895	152,270,135	240	6	25.0	-1.26	43.4
AU055503-1s1A03	AU055503	chrX	152,269,895	152,270,175	280	7	25.0	-1.04	43.9
AU066103-1s1A03	AU066103	chrX	152,269,895	152,270,175	280	7	25.0	-1.08	43.9
AU0920301-1s1A03	AU0920301	chrX	152,301,264	152,303,214	1,950	35	17.9	0.49	51.8
AU080803-1s1A03	AU080803	chrX	152,465,584	152,466,834	1,250	27	21.6	0.60	61.4
AU055503-1s1A03	AU055503	chrX	152,649,628	152,650,821	1,193	23	19.3	-0.35	41.4
AU014803-1s1A03	AU014803	chrX	153,157,712	153,158,914	1,202	11	9.2	0.86	48.6
AU034904-1s1A03	AU034904	chrX	153,473,214	153,480,214	7,000	109	15.6	-0.60	48.2
AU004803-1s1A03	AU004803	chrX	153,540,502	153,541,207	705	16	22.7	0.52	54.9
AU083504-1s1A03	AU083504	chrX	154,047,451	154,056,710	9,259	152	16.4	-0.64	34.4
AU009503-1s1A03	AU009503	chrX	154,139,152	154,139,992	840	19	22.6	-1.25	41.1
AU014803-1s1A03	AU014803	chrX	154,139,152	154,139,952	800	18	22.5	-1.09	41.7
AU024003-1s1A03	AU024003	chrX	154,139,152	154,140,052	900	20	22.2	-1.06	40.8
AU067803-1s2A03	AU067803	chrX	154,139,152	154,140,052	900	20	22.2	-1.37	40.8
AU083504-1s1A03	AU083504	chrX	154,139,152	154,140,052	900	20	22.2	-1.34	40.8
AU0920301-1s1A03	AU0920301	chrX	154,139,152	154,140,052	900	20	22.2	-0.72	40.8
AU1038303-1s1A03	AU1038303	chrX	154,139,152	154,140,168	1,016	22	21.7	-0.93	40.6
AU004803-1s1A03	AU004803	chrX	154,139,332	154,139,952	620	14	22.6	-1.17	41.8
AU028903-1s1A03	AU028903	chrX	154,139,332	154,139,992	660	15	22.7	-0.60	41.0
AU058503-1s1A03	AU058503	chrX	154,428,778	154,429,538	760	17	22.4	-1.31	38.2
AU009503-1s1A03	AU009503	chrX	154,443,339	154,446,900	3,561	45	12.6	0.67	42.3
AU067803-1s2A03	AU067803	chrX	154,443,339	154,450,670	7,331	123	16.8	1.11	41.1
AU1038303-1s1A03	AU1038303	chrX	154,443,339	154,450,670	7,331	123	16.8	0.79	41.1
AU014803-1s1A03	AU014803	chrX	154,446,165	154,450,670	4,505	92	20.4	0.96	39.5
AU004803-1s1A03	AU004803	chrX	154,446,755	154,446,925	170	5	29.4	1.01	35.6
AU024003-1s1A03	AU024003	chrX	154,446,755	154,450,750	3,995	83	20.8	1.09	39.9
AU058103-1s1A03	AU058103	chrX	154,446,755	154,450,750	3,995	83	20.8	0.57	39.9
AU083504-1s1A03	AU083504	chrX	154,446,755	154,450,670	3,915	82	20.9	1.27	39.7

Table A.3 - CNV from 100 AGRE and 64 SSC samples run on 2.1M arrays by the Optimized protocol.

AU1953303-A01-1s1	AU1953303	chrX	23,694,128	23,695,658	1,530	31	20.3	0.44	61.05
AU0976303-1s1A01	AU0976303	chrX	23,952,568	23,954,076	1,508	26	17.2	-0.49	69.16
AU021203-1s1A01	AU021203	chrX	24,021,215	24,026,159	4,944	43	8.7	-0.72	41.67
AU056803-A01-2s1	AU056803	chrX	24,022,009	24,026,159	4,150	41	9.9	-0.74	38.96
SSC00460-A01-1s1	SSC00460	chrX	24,203,022	24,205,262	2,240	30	13.4	-1.12	44.1
SSC00264-A01-3s1	SSC00264	chrX	24,508,719	24,510,099	1,380	28	20.3	-1.46	45.3
AU1038303-A01-2s1	AU1038303	chrX	24,933,183	24,933,900	717	14	19.5	0.67	68.76
AU1898303-A01-1s2	AU1898303	chrX	26,273,052	26,276,195	3,143	54	17.2	-1.34	38.53
AU1534302-1s1A01	AU1534302	chrX	26,375,390	26,375,975	585	13	22.2	-0.97	43.08
AU1534302-1s1A01	AU1534302	chrX	26,390,763	26,391,753	990	21	21.2	-0.72	38.18
AU1054302-1s1A01	AU1054302	chrX	26,432,577	26,433,232	655	15	22.9	0.51	59.08
AU1056304-1s1A01	AU1056304	chrX	26,432,602	26,433,322	720	16	22.2	0.53	59.86
SSC00549-A01-1s1	SSC00549	chrX	26,720,715	26,731,591	10,876	195	17.9	0.58	32.9
AU080803-A01-2s1	AU080803	chrX	30,048,831	30,055,680	6,849	122	17.8	-0.71	35.99
AU1953303-A01-1s1	AU1953303	chrX	30,236,143	30,237,705	1,562	28	17.9	0.45	65.81
AU1069302-A01-2s1	AU1069302	chrX	30,236,575	30,237,705	1,130	22	19.5	0.48	67.17
AU1378304-2s1A01	AU1378304	chrX	30,257,653	30,258,043	390	8	20.5	-1.18	31.03
SSC00461-A01-2s1	SSC00461	chrX	30,716,521	30,734,701	18,180	322	17.7	0.74	52.1
AU065404-A01-2s1	AU065404	chrX	30,716,581	30,734,646	18,065	320	17.7	0.34	52.37
SSC00392-A01-2s1	SSC00392	chrX	30,716,696	30,734,701	18,005	320	17.8	0.45	52.1
A01AU1327305-1s3	AU1327305	chrX	30,716,811	30,734,701	17,890	318	17.8	-0.34	52.35
AU1211303-1s2A01	AU1211303	chrX	30,721,066	30,734,701	13,635	244	17.9	-0.62	52.12
AU032701-A01-3s1	AU032701	chrX	32,363,198	32,364,497	1,299	24	18.5	-0.59	37.18
AU1054302-1s1A01	AU1054302	chrX	32,897,210	32,898,212	1,002	19	19.0	-0.83	31.14
SSC00549-A01-1s1	SSC00549	chrX	32,897,851	32,898,716	865	10	11.6	-1.20	36.0
AU0875302-A01-2s1	AU0875302	chrX	33,749,081	33,756,552	7,471	69	9.2	-0.53	35.02
SSC00510-A01-2s1	SSC00510	chrX	33,952,881	33,980,077	27,196	451	16.6	-0.94	35.0
AU065404-A01-2s1	AU065404	chrX	33,952,911	33,982,646	29,735	497	16.7	-1.06	35.44
AU021203-1s1A01	AU021203	chrX	34,062,085	34,066,166	4,081	30	7.4	0.51	35.73
AU004803-A01-3s1	AU004803	chrX	34,063,446	34,065,846	2,400	21	8.8	0.73	35.00
AU0452303-A01-2s1	AU0452303	chrX	34,063,446	34,066,116	2,670	25	9.4	0.61	35.36
AU1038303-A01-2s1	AU1038303	chrX	34,063,446	34,066,661	3,215	30	9.3	0.55	33.56
AU1159302-2s1A01	AU1159303	chrX	34,063,446	34,065,466	2,020	20	9.9	0.66	33.51
AU065404-A01-2s1	AU065404	chrX	34,063,531	34,065,466	1,935	18	9.3	0.88	33.54
AU016803-A01-2s1	AU016803	chrX	34,063,576	34,065,846	2,270	18	7.9	0.65	35.24
AU0875302-A01-2s1	AU0875302	chrX	34,063,739	34,065,466	1,727	14	8.1	0.85	34.51
AU056803-A01-2s1	AU056803	chrX	34,064,509	34,065,916	1,407	13	9.2	1.03	37.03
AU1134303-A01-2s1	AU1134303	chrX	34,064,509	34,065,466	957	11	11.5	1.02	34.48
AU1947303-A01-1s1	AU1947303	chrX	34,349,231	34,351,198	1,967	19	9.7	-0.72	34.42
AU1953303-A01-1s1	AU1953303	chrX	34,349,231	34,351,378	2,147	20	9.3	-0.81	34.75
AU1267302-1s1A01	AU1267302	chrX	34,349,301	34,351,378	2,077	19	9.1	-0.56	34.81
AU1069302-A01-2s1	AU1069302	chrX	34,584,572	34,585,553	981	19	19.4	0.48	74.31
AU1069302-A01-2s1	AU1069302	chrX	34,870,749	34,871,874	1,125	18	16.0	0.43	60.80
SSC00518-A01-1s1	SSC00518	chrX	35,622,301	35,625,671	3,370	56	16.6	-0.49	38.8
AU1631303-1s1A01	AU1631303	chrX	36,516,858	36,540,028	23,170	295	12.7	-0.45	35.68
AU1953303-A01-1s1	AU1953303	chrX	36,885,407	36,886,107	700	11	15.7	0.62	72.57
AU1134303-A01-2s1	AU1134303	chrX	36,936,692	36,938,657	1,965	24	12.2	0.47	62.09
AU0852304-A01-2s1	AU0852304	chrX	36,936,957	36,938,657	1,700	20	11.8	0.69	63.00
SSC00452-A01-2s2	SSC00452	chrX	36,941,242	36,944,207	2,965	39	13.2	-0.72	45.2
AU1631303-1s1A01	AU1631303	chrX	36,995,555	37,004,231	8,676	116	13.4	-0.44	37.33
SSC00592-A01-3s1	SSC00592	chrX	37,285,506	37,288,191	2,685	53	19.7	0.44	57.7
AU1953303-A01-1s1	AU1953303	chrX	37,286,201	37,287,886	1,685	32	19.0	0.53	61.78
AU056803-A01-2s1	AU056803	chrX	38,547,557	38,549,428	1,871	39	20.8	0.59	65.15
AU1397313-A01-2s1	AU1397313	chrX	38,547,677	38,549,955	2,278	48	21.1	0.35	65.06
AU028903-A01-2s1	AU028903	chrX	38,547,707	38,550,411	2,704	55	20.3	0.46	63.76
AU1953303-A01-1s1	AU1953303	chrX	38,547,752	38,551,610	3,858	75	19.4	0.35	55.29
A01AU1799302-1s4	AU1799302	chrX	39,208,803	39,209,233	430	9	20.9	0.66	48.14
AU0976303-1s1A01	AU0976303	chrX	39,853,069	39,854,158	1,089	18	16.5	-0.54	73.74
AU1953303-A01-1s1	AU1953303	chrX	40,828,639	40,831,205	2,566	50	19.5	0.51	66.21
AU1947303-A01-1s1	AU1947303	chrX	40,828,679	40,830,870	2,191	43	19.6	0.42	69.97
AU065404-A01-2s1	AU065404	chrX	40,828,724	40,830,725	2,001	39	19.5	0.38	70.81
AU083504-A01-3s1	AU083504	chrX	40,828,724	40,830,965	2,241	44	19.6	0.52	69.79
AU1069302-A01-2s1	AU1069302	chrX	40,828,799	40,830,522	1,723	34	19.7	0.49	70.75
AU1397313-A01-2s1	AU1397313	chrX	40,828,839	40,830,870	2,031	40	19.7	0.45	70.06
AU1038303-A01-2s1	AU1038303	chrX	40,828,889	40,830,925	2,036	40	19.6	0.61	69.89
AU014803-A01-2s1	AU014803	chrX	40,828,919	40,830,407	1,488	29	19.5	0.53	69.49
AU1038303-A01-2s1	AU1038303	chrX	41,077,342	41,079,467	2,125	40	18.8	0.43	63.39
AU1267302-1s1A01	AU1267302	chrX	41,077,342	41,079,437	2,095	39	18.6	0.36	63.44
AU1346302-1s1A01	AU1346302	chrX	41,228,399	41,230,584	2,185	26	11.9	-0.87	44.35
AU016803-A01-2s1	AU016803	chrX	41,228,827	41,230,514	1,687	24	14.2	-1.20	42.44
SSC00137-A01-2s2	SSC00137	chrX	41,228,827	41,230,514	1,687	24	14.2	-0.90	42.1
SSC00379-A01-3s1	SSC00379	chrX	43,457,535	43,465,245	7,710	88	20.4	-0.35	49.9

SSC00035-3s1A01	SSC00035	chrX	43,829,768	43,831,743	1,975	37	18.7	-1.16	45.2
AU0852304-A01-2s1	AU0852304	chrX	44,264,996	44,266,216	1,220	20	16.4	-1.31	41.6
AU1397313-A01-2s1	AU1397313	chrX	44,617,076	44,618,261	1,185	24	20.3	0.46	69.4
SSC00461-A01-2s1	SSC00461	chrX	46,955,606	46,956,590	984	20	20.3	0.71	47.9
AU028903-A01-2s1	AU028903	chrX	46,962,238	46,964,449	2,211	44	19.9	0.48	62.7
AU1791303-1s2A01	AU1791303	chrX	47,226,241	47,226,731	490	11	22.4	-0.85	48.6
AU1211303-1s2A01	AU1211303	chrX	47,226,291	47,226,731	440	10	22.7	-0.97	49.8
AU0939304-1s1A02	AU0939304	chrX	47,394,069	47,395,282	1,213	26	21.4	0.44	61.3
AU056803-A02-2s1	AU056803	chrX	47,394,397	47,395,422	1,025	22	21.5	0.53	64.5
AU1069302-A02-2s1	AU1069302	chrX	47,394,397	47,395,552	1,155	25	21.6	0.54	63.6
AU1001202-1s1A02	AU1001202	chrX	47,758,088	47,878,128	120,040	1147	9.6	0.56	45.0
SSC00317-A02-2s1	SSC00317	chrX	47,765,076	47,870,767	105,691	1139	10.8	-1.26	44.6
AU1573302-1s2A02	AU1573302	chrX	47,778,068	47,887,478	109,410	987	9.0	-0.39	44.8
AU014803-A02-2s1	AU014803	chrX	48,208,660	48,211,205	2,545	54	21.2	0.39	58.9
AU1953303-A02-1s1	AU1953303	chrX	48,220,111	48,221,051	940	21	22.3	-0.42	44.5
AU1592301-1s2A02	AU1592301	chrX	48,276,454	48,278,154	1,700	33	19.4	-0.47	51.4
AU1622302-1s2A02	AU1622302	chrX	48,530,120	48,531,087	967	21	21.7	-0.52	48.6
AU1592301-1s2A02	AU1592301	chrX	48,530,197	48,531,062	865	19	22.0	-0.49	48.2
AU1953303-A02-1s1	AU1953303	chrX	48,530,197	48,531,087	890	20	22.5	-0.50	48.4
AU1424304-1s2A02	AU1424304	chrX	48,530,217	48,531,087	870	19	21.8	-0.54	48.3
AU1791303-1s2A02	AU1791303	chrX	48,530,217	48,531,142	925	20	21.6	-0.50	48.1
AU056003-A02-2s1	AU056003	chrX	48,533,289	48,534,370	1,081	17	15.7	0.74	53.7
AU056803-A02-2s1	AU056803	chrX	48,698,770	48,700,805	2,035	41	20.1	0.54	64.6
AU1953303-A02-1s1	AU1953303	chrX	48,714,198	48,714,773	575	13	22.6	-0.55	45.7
AU1592301-1s2A02	AU1592301	chrX	48,714,353	48,714,843	490	11	22.4	-0.75	44.7
AU0983302-1s1A02	AU0983302	chrX	48,738,809	48,739,434	625	13	20.8	-0.54	45.9
AU004803-A02-3s1	AU004803	chrX	48,934,299	48,934,869	570	12	21.1	0.78	53.3
AU1069302-A02-2s1	AU1069302	chrX	48,934,364	48,934,799	435	9	20.7	1.08	51.7
AU1134303-A02-2s1	AU1134303	chrX	48,934,364	48,934,824	460	10	21.7	0.84	52.4
AU1211303-1s2A02	AU1211303	chrX	48,967,547	48,967,942	395	9	22.8	-0.56	52.2
AU1791303-1s2A02	AU1791303	chrX	48,995,264	48,996,414	1,150	24	20.9	-0.48	53.2
AU1211303-1s2A02	AU1211303	chrX	48,995,294	48,996,414	1,120	23	20.5	-0.39	53.1
AU1346302-1s1A02	AU1346302	chrX	48,995,294	48,996,284	990	20	20.2	-0.50	52.8
AU1573302-1s2A02	AU1573302	chrX	48,995,294	48,996,304	1,010	21	20.8	-0.46	52.7
AU1953303-A02-1s1	AU1953303	chrX	48,995,294	48,996,359	1,065	22	20.7	-0.43	52.7
AU1424304-1s2A02	AU1424304	chrX	48,995,594	48,996,284	690	14	20.3	-0.60	52.3
AU065404-A02-2s1	AU065404	chrX	49,012,185	49,013,180	995	22	22.1	0.65	58.3
AU0939304-1s1A02	AU0939304	chrX	49,012,240	49,014,323	2,083	41	19.7	0.37	67.1
AU0983302-1s1A02	AU0983302	chrX	49,012,240	49,013,180	940	21	22.3	0.51	58.7
AU1338304-1s1A02	AU1338304	chrX	49,012,295	49,013,135	840	19	22.6	0.51	58.8
AU1056304-1s1A02	AU1056304	chrX	49,012,330	49,013,180	850	19	22.4	0.55	59.1
AU004803-A02-3s1	AU004803	chrX	49,012,390	49,013,180	790	18	22.8	0.63	59.4
AU055303-A02-2s1	AU055303	chrX	49,012,390	49,013,180	790	18	22.8	0.63	59.4
AU1001202-1s1A02	AU1001202	chrX	49,012,390	49,013,180	790	18	22.8	0.63	59.4
A02AU1143303-1s2	AU1143303	chrX	49,012,390	49,013,180	790	18	22.8	0.60	59.4
A02AU1327305-1s3	AU1327305	chrX	49,012,390	49,013,180	790	18	22.8	0.43	59.4
AU1397313-A02-2s2	AU1397313	chrX	49,012,390	49,013,615	1,225	25	20.4	0.58	66.3
AU1134303-A02-2s1	AU1134303	chrX	49,012,415	49,013,180	765	17	22.2	0.67	59.5
AU014803-A02-2s1	AU014803	chrX	49,012,490	49,014,388	1,898	37	19.5	0.50	68.4
AU018304-A02-2s1	AU018304	chrX	49,012,490	49,013,180	690	16	23.2	0.98	59.7
AU055503-A02-2s1	AU055503	chrX	49,012,490	49,013,180	690	16	23.2	0.75	59.7
AU083504-A02-3s1	AU083504	chrX	49,012,490	49,013,270	780	17	21.8	0.69	62.1
AU0875302-A02-2s1	AU0875302	chrX	49,012,490	49,013,180	690	16	23.2	0.87	59.7
AU0920301-A02-2s1	AU0920301	chrX	49,012,490	49,013,180	690	16	23.2	0.78	59.7
AU0780301-A02-2s1	AU0780301	chrX	49,012,510	49,013,180	670	15	22.4	0.77	59.4
AU083603-A02-2s1	AU083603	chrX	49,012,510	49,013,180	670	15	22.4	0.63	59.4
AU056003-A02-2s1	AU056003	chrX	49,012,575	49,013,180	605	14	23.1	0.77	60.2
AU063303-A02-1s1	AU063303	chrX	49,012,575	49,013,270	695	15	21.6	0.76	62.7
AU1054302-1s1A02	AU1054302	chrX	49,012,575	49,013,180	605	14	23.1	0.72	60.2
AU1344302-1s1A02	AU1344302	chrX	49,012,575	49,013,135	560	13	23.2	0.58	60.5
SSC00035-3s1A02	SSC00035	chrX	49,265,935	49,271,538	5,603	83	14.8	-0.45	44.0
AU1211303-1s2A02	AU1211303	chrX	49,529,973	49,531,486	1,513	29	19.2	0.41	67.5
AU1069302-A02-2s1	AU1069302	chrX	50,229,081	50,231,210	2,129	41	19.3	0.45	62.8
AU056803-A02-2s1	AU056803	chrX	50,229,416	50,231,140	1,724	34	19.7	0.46	61.8
SSC00379-A02-3s1	SSC00379	chrX	52,065,715	52,066,910	1,195	25	20.9	-1.04	37.8
AU056803-A02-2s1	AU056803	chrX	52,966,038	52,967,318	1,280	22	17.2	0.50	60.5
AU1134303-A02-2s1	AU1134303	chrX	52,966,813	52,967,078	265	7	26.4	0.80	63.8
SSC00098-A02-3s1	SSC00098	chrX	53,099,842	53,102,277	2,435	44	18.1	-0.55	48.5
AU008504-A02-2s1	AU008504	chrX	53,099,962	53,102,472	2,510	45	17.9	0.35	49.2
AU1346302-1s1A02	AU1346302	chrX	53,100,087	53,102,557	2,470	45	18.2	-0.50	49.5
A02AU1143303-1s2	AU1143303	chrX	53,100,432	53,102,672	2,240	46	20.5	-0.47	48.5
SSC00181-A02-2s1	SSC00181	chrX	53,100,432	53,102,317	1,885	39	20.7	-0.46	48.6
AU1054302-1s1A02	AU1054302	chrX	53,100,452	53,101,922	1,470	30	20.4	-0.39	49.7

AU1056304-1s1A02	AU1056304	chrX	53,100,452	53,102,672	2,220	45	20.3	-0.41	48.5
AU1324303-1s2A02	AU1324303	chrX	53,100,452	53,102,472	2,020	41	20.3	-0.36	49.2
AU1334303-1s1A02	AU1334303	chrX	53,100,452	53,102,277	1,825	37	20.3	-0.39	49.0
AU1953303-A02-1s1	AU1953303	chrX	53,100,452	53,102,622	2,170	44	20.3	-0.48	48.6
AU1073302-A02-2s1	AU1073302	chrX	53,100,517	53,102,317	1,800	37	20.6	-0.59	48.7
AU1267302-1s1A02	AU1267302	chrX	53,100,517	53,102,672	2,155	44	20.4	-0.44	48.2
A02AU1327305-1s3	AU1327305	chrX	53,100,517	53,102,717	2,200	45	20.5	-0.42	48.1
AU1344302-1s1A02	AU1344302	chrX	53,100,517	53,102,622	2,105	43	20.4	-0.46	48.3
AU1592301-1s2A02	AU1592301	chrX	53,100,517	53,103,956	3,439	57	16.6	-0.44	47.2
SSC00093-A02-1s1	SSC00093	chrX	53,100,517	53,102,472	1,955	40	20.5	-0.55	48.5
SSC00317-A02-2s1	SSC00317	chrX	53,100,517	53,102,882	2,365	49	20.7	-0.47	47.4
AU1414305-1s2A02	AU1414305	chrX	53,100,552	53,102,502	1,950	40	20.5	-0.44	48.9
AU004803-A02-3s1	AU004803	chrX	53,127,122	53,127,609	487	8	16.4	-0.99	68.0
AU1397313-A02-2s2	AU1397313	chrX	53,238,467	53,241,061	2,594	51	19.7	0.36	61.5
AU0939304-1s1A02	AU0939304	chrX	53,238,692	53,242,246	3,554	71	20.0	0.32	59.2
AU1344302-1s1A02	AU1344302	chrX	53,800,995	53,801,260	265	7	26.4	-1.35	53.2
AU1791303-1s2A02	AU1791303	chrX	54,573,144	54,573,631	487	8	16.4	0.75	62.6
AU056803-A02-2s1	AU056803	chrX	54,591,828	54,593,323	1,495	25	16.7	-0.45	48.4
AU1267302-1s1A02	AU1267302	chrX	54,863,979	54,868,109	4,130	52	12.6	0.32	54.6
AU021203-1s1A02	AU021203	chrX	54,867,594	54,868,109	515	11	21.4	0.57	59.4
A02AU1143303-1s2	AU1143303	chrX	55,069,357	55,070,762	1,405	30	21.4	-0.43	43.9
AU1134303-A02-2s1	AU1134303	chrX	56,826,211	56,827,896	1,685	35	20.8	0.37	55.8
AU1592301-1s2A02	AU1592301	chrX	56,845,230	56,846,164	934	18	19.3	0.63	54.4
AU0920301-A02-2s1	AU0920301	chrX	57,615,600	57,960,046	344,446	3789	11.0	0.47	38.6
AU050703-A02-2s1	AU050703	chrX	57,763,046	57,767,996	4,950	57	11.5	1.07	36.0
AU1324303-1s2A02	AU1324303	chrX	57,763,081	57,768,570	5,489	66	12.0	1.09	35.4
AU083504-A02-3s1	AU083504	chrX	58,256,056	58,256,496	440	8	18.2	-1.33	46.8
SSC00440-A02-2s1	SSC00440	chrX	62,386,739	62,422,538	35,799	443	12.4	0.67	38.5
AU0920301-A02-2s1	AU0920301	chrX	62,386,914	62,422,378	35,464	441	12.4	0.59	38.9
SSC00180-A02-2s1	SSC00180	chrX	62,386,914	62,422,378	35,464	441	12.4	0.58	38.5
SSC00460-A02-1s1	SSC00460	chrX	62,386,914	62,426,622	39,708	444	11.2	0.50	38.6
AU1038303-A02-2s1	AU1038303	chrX	62,494,361	62,520,001	25,640	294	11.5	-0.41	38.7
AU1134303-A02-2s1	AU1134303	chrX	62,921,207	62,922,007	800	18	22.5	0.51	62.4
AU1159302-2s1A02	AU1159303	chrX	64,002,046	64,013,621	11,575	201	17.4	-1.04	35.3
AU016803-A02-2s1	AU016803	chrX	65,382,441	65,415,702	33,261	530	15.9	0.51	38.9
AU014803-A02-2s1	AU014803	chrX	67,040,517	67,046,967	6,450	84	13.0	-0.48	38.7
AU1585301-1s2A02	AU1585301	chrX	67,040,517	67,047,042	6,525	85	13.0	-0.38	38.6
SSC00314-A02-2s1	SSC00314	chrX	67,040,517	67,046,967	6,450	84	13.0	-0.89	38.3
SSC00366-A02-2s1	SSC00366	chrX	67,040,517	67,047,042	6,525	85	13.0	-0.71	38.2
AU0983302-1s1A02	AU0983302	chrX	67,040,582	67,046,967	6,385	83	13.0	-1.13	38.6
AU018003-A02-2s1	AU018003	chrX	67,045,802	67,046,892	1,090	19	17.4	-0.74	41.8
AU0920301-A02-2s1	AU0920301	chrX	67,761,197	67,761,872	675	9	13.3	-0.75	28.9
AU0983302-1s1A02	AU0983302	chrX	67,879,047	67,880,967	1,920	26	13.5	-1.87	40.9
AU0983302-1s1A02	AU0983302	chrX	68,073,913	68,074,543	630	14	22.2	-0.42	39.5
AU1069302-A02-2s1	AU1069302	chrX	68,752,123	68,753,594	1,471	30	20.4	0.45	65.1
AU067803-A02-2s1	AU067803	chrX	69,201,902	69,205,364	3,462	72	20.8	0.48	57.9
AU0920301-A02-2s1	AU0920301	chrX	69,202,002	69,204,994	2,992	62	20.7	0.51	58.4
AU014803-A02-2s1	AU014803	chrX	69,202,022	69,205,124	3,102	64	20.6	0.39	58.6
AU1038303-A02-2s1	AU1038303	chrX	69,202,132	69,204,769	2,637	55	20.9	0.57	58.8
AU055503-A02-2s1	AU055503	chrX	69,588,687	69,592,509	3,822	75	19.6	0.39	59.0
SSC00505-A02-1s1	SSC00505	chrX	70,039,798	70,040,318	520	12	23.1	-1.47	35.2
AU020003-A02-2s1	AU020003	chrX	70,039,948	70,040,318	370	9	24.3	-1.18	37.0
AU0983302-1s1A02	AU0983302	chrX	70,039,948	70,040,318	370	9	24.3	-1.61	37.0
SSC00357-A02-2s1	SSC00357	chrX	70,039,948	70,040,368	420	10	23.8	-1.60	36.3
AU1378304-2s1A02	AU1378304	chrX	70,237,266	70,238,316	1,050	23	21.9	0.48	59.0
AU0939304-1s1A02	AU0939304	chrX	70,714,508	70,715,873	1,365	30	22.0	0.40	61.2
AU056803-A02-2s1	AU056803	chrX	70,714,913	70,716,073	1,160	25	21.6	0.56	64.2
AU1211303-1s2A02	AU1211303	chrX	70,714,913	70,715,873	960	21	21.9	0.42	64.3
AU067803-A02-2s1	AU067803	chrX	70,714,983	70,717,347	2,364	43	18.2	0.58	59.3
AU008504-A02-2s1	AU008504	chrX	70,715,008	70,716,773	1,765	37	21.0	0.41	62.4
AU1069302-A02-2s1	AU1069302	chrX	70,715,008	70,715,928	920	20	21.7	0.60	64.0
AU080803-A02-2s1	AU080803	chrX	70,715,118	70,716,773	1,655	35	21.1	0.47	62.4
AU0920301-A02-2s1	AU0920301	chrX	70,715,118	70,716,928	1,810	38	21.0	0.49	61.8
AU065404-A02-2s1	AU065404	chrX	70,715,168	70,716,838	1,670	35	21.0	0.55	62.9
AU083603-A02-2s1	AU083603	chrX	70,715,198	70,716,868	1,670	35	21.0	0.46	63.5
A02AU1327305-1s3	AU1327305	chrX	70,715,978	70,716,773	795	17	21.4	-0.45	60.3
AU1871302-1s1A02	AU1871302	chrX	70,740,201	70,740,796	595	6	10.1	0.90	50.8
AU1953303-A02-1s1	AU1953303	chrX	70,796,602	70,803,737	7,135	116	16.3	0.36	58.3
AU1346302-1s1A02	AU1346302	chrX	72,249,888	72,523,656	273,768	2927	10.7	0.39	38.2
AU1069302-A02-2s1	AU1069302	chrX	73,672,126	73,674,216	2,090	43	20.6	0.42	61.6
AU0939304-1s1A02	AU0939304	chrX	74,060,721	74,062,672	1,951	37	19.0	0.32	65.6
AU020003-A02-2s1	AU020003	chrX	75,282,723	75,283,263	540	12	22.2	-1.38	46.3
AU0920301-A02-2s1	AU0920301	chrX	75,917,772	75,919,918	2,146	35	16.3	0.64	41.0

AU0939304-1s1A02	AU0939304	chrX	77,404,660	77,410,841	6,181	74	12.0	-0.75	45.4
SSC00391-A02-3s1	SSC00391	chrX	77,406,085	77,410,671	4,586	56	12.2	-0.54	46.5
AU1862302-1s2A02	AU1862302	chrX	80,302,657	80,304,067	1,410	22	15.6	-0.61	39.6
SSC00035-3s1A02	SSC00035	chrX	80,303,147	80,304,212	1,065	20	18.8	-0.97	37.9
SSC00035-3s1A02	SSC00035	chrX	81,060,422	81,062,537	2,115	45	21.3	-0.93	35.7
AU1038303-A02-2s1	AU1038303	chrX	82,086,031	82,094,874	8,843	92	10.4	-1.04	39.0
AU0852304-A02-2s1	AU0852304	chrX	87,546,283	87,547,547	1,264	12	9.5	-0.95	36.0
AU055503-A02-2s1	AU055503	chrX	88,191,854	88,194,527	2,673	43	16.1	-1.26	36.3
SSC00452-A02-2s2	SSC00452	chrX	92,608,597	92,610,562	1,965	35	17.8	-1.07	48.9
AU083504-A02-3s1	AU083504	chrX	96,493,478	96,495,197	1,719	26	15.1	-0.73	37.2
AU1056304-1s1A02	AU1056304	chrX	96,493,478	96,495,197	1,719	26	15.1	-0.74	37.2
AU1001202-1s1A02	AU1001202	chrX	97,842,328	97,843,392	1,064	18	16.9	-1.06	34.5
SSC00093-A02-1s1	SSC00093	chrX	100,635,456	100,636,316	860	18	20.9	-0.83	53.2
AU004803-A02-3s1	AU004803	chrX	101,266,594	101,268,484	1,890	24	12.7	0.45	61.4
AU1953303-A02-1s1	AU1953303	chrX	101,853,198	101,854,673	1,475	26	17.6	0.50	67.1
AU021503-A02-2s1	AU021503	chrX	103,061,781	103,110,309	48,528	553	11.4	0.79	41.4
SSC00129-3s1A02	SSC00129	chrX	103,145,327	103,192,016	46,689	646	13.8	0.68	43.5
SSC00314-A02-2s1	SSC00314	chrX	103,145,327	103,192,016	46,689	646	13.8	0.65	43.5
AU021503-A02-2s1	AU021503	chrX	103,147,877	103,210,928	63,051	598	9.5	0.92	43.6
AU1622302-1s2A02	AU1622302	chrX	103,289,071	103,289,641	570	13	22.8	-0.92	42.3
AU1038303-A02-2s1	AU1038303	chrX	103,295,654	103,296,645	991	16	16.1	-1.40	40.9
SSC00317-A02-2s1	SSC00317	chrX	103,999,466	103,999,825	359	9	25.1	-1.45	36.5
AU0920301-A02-2s1	AU0920301	chrX	105,379,514	105,381,047	1,533	28	18.3	0.64	37.8
AU0895303-1s2A03	AU0895303	chrX	107,864,665	107,866,975	2,310	43	18.6	0.34	63.2
SSC00077-A03-3s1	SSC00077	chrX	108,183,989	108,184,388	399	6	15.0	-1.63	55.3
SSC00093-A03-1s2	SSC00093	chrX	108,183,989	108,184,388	399	6	15.0	-2.00	55.3
SSC00097-A03-1s1	SSC00097	chrX	108,183,989	108,184,388	399	6	15.0	-1.81	55.3
SSC00098-A03-3s1	SSC00098	chrX	108,183,989	108,184,388	399	6	15.0	-1.87	55.3
SSC00260-A03-2s1	SSC00260	chrX	108,183,989	108,184,388	399	6	15.0	-1.92	55.3
SSC00316-A03-2s1	SSC00316	chrX	108,183,989	108,184,388	399	6	15.0	-2.40	55.3
AU0852304-A03-2s1	AU0852304	chrX	109,131,565	109,134,047	2,482	48	19.3	0.37	64.8
AU014803-A03-2s1	AU014803	chrX	109,131,645	109,133,310	1,665	32	19.2	0.59	67.6
SSC00316-A03-2s1	SSC00316	chrX	110,234,756	110,236,566	1,810	38	21.0	-0.75	40.3
SSC00417-A03-1s1	SSC00417	chrX	111,501,112	111,569,590	68,478	1003	14.6	-0.61	37.4
SSC00264-A03-3s1	SSC00264	chrX	111,598,662	111,628,128	29,466	378	12.8	-0.91	36.0
SSC00332-A03-2s1	SSC00332	chrX	111,753,154	111,753,733	579	13	22.5	-0.90	28.2
AU0939304-1s1A03	AU0939304	chrX	111,970,302	111,971,664	1,362	28	20.6	0.39	62.6
AU1069302-A03-2s1	AU1069302	chrX	111,970,302	111,971,424	1,122	24	21.4	0.50	63.8
SSC00269-A03-2s1	SSC00269	chrX	112,047,857	112,052,442	4,585	68	14.8	-0.84	36.5
AU1344302-1s1A03	AU1344302	chrX	113,234,591	113,241,577	6,986	101	14.5	-0.86	38.4
AU004803-A03-3s1	AU004803	chrX	113,234,686	113,241,182	6,496	94	14.5	-0.40	38.1
AU1211303-1s2A03	AU1211303	chrX	114,330,335	114,333,305	2,970	56	18.9	0.35	65.2
AU1631303-1s1A03	AU1631303	chrX	114,330,725	114,333,350	2,625	50	19.0	0.45	65.8
AU018003-A03-2s2	AU018003	chrX	114,331,010	114,333,265	2,255	45	20.0	0.39	65.0
AU0895303-1s2A03	AU0895303	chrX	114,331,010	114,333,225	2,215	44	19.9	0.52	64.9
AU1069302-A03-2s1	AU1069302	chrX	114,331,010	114,333,350	2,340	47	20.1	0.54	65.0
AU065404-A03-2s1	AU065404	chrX	114,331,095	114,333,305	2,210	44	19.9	0.44	65.2
AU0852304-A03-2s1	AU0852304	chrX	114,331,095	114,333,225	2,130	42	19.7	0.59	65.1
AU1947303-A03-1s1	AU1947303	chrX	114,331,095	114,333,265	2,170	43	19.8	0.43	65.2
AU1953303-A03-1s1	AU1953303	chrX	114,331,095	114,333,140	2,045	41	20.0	0.57	65.2
SSC00296-A03-2s1	SSC00296	chrX	114,331,235	114,333,305	2,070	41	19.8	0.60	65.1
SSC00592-A03-3s1	SSC00592	chrX	114,331,350	114,333,265	1,915	38	19.8	0.62	65.1
AU1898303-1s1A03	AU1898303	chrX	114,331,375	114,333,140	1,765	35	19.8	0.51	65.4
AU0920301-A03-2s1	AU0920301	chrX	114,331,445	114,333,225	1,780	35	19.7	0.59	65.7
AU1875302-1s1A03	AU1875302	chrX	114,331,660	114,333,090	1,430	28	19.6	0.42	65.2
AU0452303-A03-2s1	AU0452303	chrX	114,331,805	114,333,090	1,285	25	19.5	0.70	65.4
AU0939304-1s1A03	AU0939304	chrX	114,332,135	114,333,350	1,215	24	19.8	0.45	65.5
AU1159302-2s1A03	AU1159303	chrX	114,448,415	114,450,696	2,281	28	12.3	0.55	33.7
SSC00426-A03-2s1	SSC00426	chrX	114,448,415	114,450,696	2,281	28	12.3	==-0.69==	33.4
AU021203-1s1A03	AU021203	chrX	114,448,780	114,450,696	1,916	26	13.6	0.45	30.6
AU1159302-2s1A03	AU1159303	chrX	115,635,001	115,639,653	4,652	81	17.4	-0.96	34.2
AU1069302-A03-2s1	AU1069302	chrX	115,823,448	115,931,741	108,293	1603	14.8	0.38	36.6
AU0852304-A03-2s1	AU0852304	chrX	116,588,255	116,590,742	2,487	44	17.7	-1.64	38.6
AU018304-A03-2s1	AU018304	chrX	117,133,965	117,135,095	1,130	24	21.2	0.47	65.0
AU0895303-1s2A03	AU0895303	chrX	117,991,249	117,994,889		65	17.9	0.31	63.9
AU028903-A03-2s1	AU028903	chrX	118,253,082	118,255,219	2,137	45	21.1	0.41	61.6
AU014803-A03-2s1	AU014803	chrX	118,253,848	118,255,144	1,296	27	20.8	0.51	68.8
AU0895303-1s2A03	AU0895303	chrX	118,253,873	118,254,949	1,076	22	20.4	0.50	70.7
AU021203-1s1A03	AU021203	chrX	118,940,776	118,948,472	7,696	81	10.5	0.59	41.9
AU1073302-A03-2s1	AU1073302	chrX	118,941,966	118,948,332	6,366	70	11.0	0.51	42.3
AU1378304-2s1A03	AU1378304	chrX	119,486,893	119,487,273	380	9	23.7	0.58	62.4
	AU074704	chrX	119,578,099	119,579,572	1,473	29	19.7	0.49	59.5
AU1001202-1s1A03	AU1001202	chrX	119,892,291	119,893,176	885	15	16.9	-0.97	59.9

AU1378304-2s1A03	AU1378304	chrX	120,416,082	120,416,943	861	8	9.3	-1.09	39.6
AU083504-A03-3s1	AU083504	chrX	122,143,675	122,144,235	560	11	19.6	-0.77	41.1
AU0983302-1s1A03	AU0983302	chrX	122,143,675	122,144,235	560	11	19.6	-1.47	41.1
AU1875302-1s1A03	AU1875302	chrX	122,778,683	122,780,424	1,741	19	10.9	-0.73	45.8
SSC00129-3s1A03	SSC00129	chrX	122,779,743	122,780,189	446	5	11.2	-1.65	44.1
SSC00346-A03-2s1	SSC00346	chrX	122,779,768	122,780,279	511	6	11.7	-1.20	43.4
AU0983302-1s1A03	AU0983302	chrX	122,815,537	122,819,376	3,839	59	15.4	0.72	44.9
AU083504-A03-3s1	AU083504	chrX	122,920,791	122,923,535	2,744	52	19.0	0.34	59.2
AU032701-A03-3s1	AU032701	chrX	124,787,664	124,788,414	750	16	21.3	-1.14	46.5
AU0895303-1s2A03	AU0895303	chrX	125,126,235	125,128,490	2,255	44	19.5	0.37	64.8
AU028903-A03-2s1	AU028903	chrX	125,126,325	125,128,380	2,055	40	19.5	0.41	65.8
AU1953303-A03-1s1	AU1953303	chrX	125,126,380	125,128,380	2,000	39	19.5	0.47	66.1
AU0895303-1s2A03	AU0895303	chrX	125,513,144	125,514,709	1,565	32	20.4	0.41	67.0
SSC00093-A03-1s2	SSC00093	chrX	125,514,204	125,515,019	815	18	22.1	-0.72	62.9
SSC00180-A03-2s1	SSC00180	chrX	125,687,154	125,687,444	290	7	24.1	-0.88	28.8
SSC00549-A03-1s1	SSC00549	chrX	126,158,051	126,158,891	840	19	22.6	-1.22	41.1
AU1334303-1s1A03	AU1334303	chrX	126,425,325	126,430,262	4,937	50	10.1	0.35	31.9
SSC00468-A03-2s1	SSC00468	chrX	128,327,945	128,329,860	1,915	25	13.1	-1.11	39.6
AU058103-A03-2s1	AU058103	chrX	128,328,110	128,329,810	1,700	21	12.4	-0.65	40.5
AU065404-A03-2s1	AU065404	chrX	128,328,110	128,329,710	1,600	20	12.5	-0.90	41.8
AU083504-A03-3s1	AU083504	chrX	128,328,110	128,329,860	1,750	22	12.6	-0.57	40.2
AU0920301-A03-2s1	AU0920301	chrX	128,328,110	128,329,625	1,515	18	11.9	-0.80	42.5
AU1054302-1s1A03	AU1054302	chrX	128,328,110	128,329,625	1,515	18	11.9	-0.99	42.5
AU1069302-A03-2s1	AU1069302	chrX	128,328,110	128,329,710	1,600	20	12.5	-0.91	41.8
AU1327305-1s2A03	AU1327305	chrX	128,328,110	128,329,625	1,515	18	11.9	-0.53	42.5
AU1338304-1s1A03	AU1338304	chrX	128,328,110	128,329,590	1,480	17	11.5	-0.66	42.6
SSC00331-A03-2s1	SSC00331	chrX	128,328,110	128,329,710	1,600	20	12.5	-0.76	41.2
SSC00391-A03-3s1	SSC00391	chrX	128,328,110	128,329,675	1,565	19	12.1	-1.05	41.5
SSC00460-A03-1s1	SSC00460	chrX	128,328,110	128,329,860	1,750	22	12.6	-0.80	39.7
SSC00549-A03-1s1	SSC00549	chrX	128,328,110	128,329,860	1,750	22	12.6	-1.27	39.7
SSC00035-3s1A03	SSC00035	chrX	128,328,160	128,329,625	1,465	17	11.6	-0.98	41.9
AU062203-A03-2s1	AU062203	chrX	128,328,360	128,329,625	1,265	15	11.9	-0.88	42.3
AU1211303-1s2A03	AU1211303	chrX	128,328,360	128,329,625	1,265	15	11.9	-0.91	42.3
SSC00518-A03-1s1	SSC00518	chrX	128,328,430	128,329,625	1,195	13	10.9	-1.54	42.1
AU0895303-1s2A03	AU0895303	chrX	128,483,910	128,485,657	1,747	35	20.0	0.33	62.6
AU1073302-A03-2s1	AU1073302	chrX	128,609,906	128,610,211	305	7	23.0	1.10	57.0
AU065404-A03-2s1	AU065404	chrX	128,893,016	128,894,561	1,545	29	18.8	0.57	71.8
AU083504-A03-3s1	AU083504	chrX	128,893,016	128,894,009	993	21	21.1	0.57	66.5
AU1056304-1s1A03	AU1056304	chrX	128,893,104	128,894,009	905	20	22.1	0.52	67.7
AU1143303-1s1A03	AU1143303	chrX	128,893,104	128,894,105	1,001	22	22.0	0.67	68.8
AU1378304-2s1A03	AU1378304	chrX	128,893,104	128,895,462	2,358	41	17.4	0.46	67.0
AU1159302-2s1A03	AU1159303	chrX	128,893,169	128,894,105	936	21	22.4	0.52	69.9
AU056803-A03-2s1	AU056803	chrX	128,941,017	128,946,320	5,303	93	17.5	0.40	67.4
AU014803-A03-2s1	AU014803	chrX	128,941,142	128,946,290	5,148	89	17.3	0.40	67.8
AU080803-A03-2s1	AU080803	chrX	128,942,154	128,945,257	3,103	59	19.0	0.39	65.1
AU018304-A03-2s1	AU018304	chrX	128,943,153	128,946,320	3,167	58	18.3	0.46	67.1
AU0976303-1s1A03	AU0976303	chrX	128,945,088	128,946,380	1,292	22	17.0	-0.49	72.7
SSC00391-A03-3s1	SSC00391	chrX	128,963,185	128,964,555	1,370	30	21.9	-0.41	52.8
AU056003-A03-2s1	AU056003	chrX	128,964,095	128,964,615	520	12	23.1	0.98	55.6
AU014803-A03-2s1	AU014803	chrX	129,070,418	129,073,992	3,574	65	18.2	0.40	65.2
SSC00093-A03-1s2	SSC00093	chrX	129,491,300	129,493,100	1,800	34	18.9	0.60	62.4
SSC00035-3s1A03	SSC00035	chrX	129,495,427	129,496,564	1,137	25	22.0	-1.05	49.4
SSC00518-A03-1s1	SSC00518	chrX	129,495,549	129,496,504	955	21	22.0	-1.10	49.4
AU1327305-1s2A03	AU1327305	chrX	129,711,402	129,712,522	1,120	24	21.4	-0.40	43.8
SSC00048-A03-1s1	SSC00048	chrX	129,711,442	129,715,197	3,755	68	18.1	0.50	41.3
SSC00366-A03-2s1	SSC00366	chrX	129,711,442	129,715,557	4,115	71	17.3	0.58	41.0
SSC00417-A03-1s1	SSC00417	chrX	130,029,230	130,030,433	1,203	25	20.8	0.74	55.4
AU0939304-1s1A03	AU0939304	chrX	130,640,034	130,641,821	1,787	16	9.0	0.55	48.6
AU028903-A03-2s1	AU028903	chrX	131,178,818	131,179,980	1,162	22	18.9	0.55	67.6
AU014803-A03-2s1	AU014803	chrX	131,178,888	131,180,176	1,288	24	18.6	0.50	69.2
AU0920301-A03-2s1	AU0920301	chrX	131,766,925	131,768,880	1,955	25	12.8	-0.89	42.2
AU1344302-1s1A03	AU1344302	chrX	131,766,925	131,769,270	2,345	31	13.2	-0.71	42.1
AU1327305-1s2A03	AU1327305	chrX	131,767,075	131,769,220	2,145	27	12.6	-0.57	41.9
SSC00129-3s1A03	SSC00129	chrX	131,767,075	131,769,220	2,145	27	12.6	-1.22	41.4
SSC00366-A03-2s1	SSC00366	chrX	131,767,075	131,769,220	2,145	27	12.6	-1.01	41.4
SSC00510-A03-2s1	SSC00510	chrX	131,767,075	131,769,220	2,145	27	12.6	-1.13	41.4
AU1378304-2s1A03	AU1378304	chrX	131,918,110	131,919,501	1,391	28	20.1	0.37	68.7
AU028903-A03-2s1	AU028903	chrX	133,133,179	133,135,589	2,410	46	19.1	0.35	59.2
AU1592301-1s2A03	AU1592301	chrX	133,346,272	133,350,722	4,450	73	16.4	-0.44	40.1
AU021503-A03-2s1	AU021503	chrX	133,692,526	133,693,326	800	17	21.3	-0.98	43.3
AU018304-A03-2s1	AU018304	chrX	133,757,887	133,758,510	623	14	22.5	0.73	60.5
AU1875302-1s1A03	AU1875302	chrX	133,758,086	133,758,445	359	9	25.1	0.70	57.7
AU008504-A03-2s1	AU008504	chrX	134,146,463	134,159,690	13,227	181	13.7	0.43	40.8

SSC00592-A03-3s1	SSC00592	chrX	134,393,837	134,400,858	7,021	118	16.8	0.40	63.4
AU0976303-1s1A03	AU0976303	chrX	134,394,595	134,400,410	5,815	96	16.5	-0.43	65.7
AU028903-A03-2s1	AU028903	chrX	134,482,494	134,483,554	1,060	23	21.7	0.54	68.6
SSC00093-A03-1s2	SSC00093	chrX	134,883,591	134,884,816	1,225	27	22.0	-0.77	62.5
SSC00097-A03-1s1	SSC00097	chrX	134,883,591	134,884,721	1,130	25	22.1	-0.88	64.3
SSC00077-A03-3s1	SSC00077	chrX	134,883,761	134,884,721	960	21	21.9	-0.91	65.7
AU0920301-A03-2s1	AU0920301	chrX	135,126,873	135,129,790	2,917	47	16.1	-0.97	43.6
SSC00129-3s1A03	SSC00129	chrX	135,126,908	135,129,790	2,882	46	16.0	-1.15	43.3
AU1592301-1s2A03	AU1592301	chrX	135,767,231	135,767,911	680	8	11.8	-0.78	46.0
SSC00417-A03-1s1	SSC00417	chrX	135,819,528	135,820,308	780	17	21.8	0.86	56.2
AU055303-A03-2s1	AU055303	chrX	135,819,673	135,820,148	475	11	23.2	0.68	57.1
AU028903-A03-2s1	AU028903	chrX	135,940,186	135,941,906	1,720	36	20.9	0.41	58.4
SSC00460-A03-1s1	SSC00460	chrX	135,941,711	135,943,359	1,648	34	20.6	0.35	50.5
AU1953303-A03-1s1	AU1953303	chrX	137,620,314	137,621,972	1,658	32	19.3	0.41	65.0
SSC00331-A03-2s1	SSC00331	chrX	139,323,442	139,328,934	5,492	87	15.8	-0.85	37.3
SSC00316-A03-2s1	SSC00316	chrX	139,623,340	139,638,648	15,308	219	14.3	-0.70	44.8
AU1378304-2s1A03	AU1378304	chrX	140,302,728	140,307,726	4,998	47	9.4	-0.61	34.6
AU1159302-2s1A03	AU1159303	chrX	140,304,158	140,306,721	2,563	29	11.3	-0.79	33.2
SSC00505-A03-1s1	SSC00505	chrX	140,519,717	140,592,330	72,613	1200	16.5	0.54	37.3
AU1334303-1s1A03	AU1334303	chrX	141,159,633	141,160,783	1,150	22	19.1	0.49	58.2
SSC00093-A03-1s2	SSC00093	chrX	142,550,345	142,551,155	810	17	21.0	-0.89	61.4
SSC00097-A03-1s1	SSC00097	chrX	142,550,405	142,551,155	750	16	21.3	-0.89	61.0
AU002403-A03-2s1	AU002403	chrX	143,165,878	143,166,968	178,000	23	21.1	-1.01	49.9
SSC00426-A03-2s1	SSC00426	chrX	143,289,593	143,290,073	480	11	22.9	-1.70	51.9
SSC00426-A03-2s1	SSC00426	chrX	143,429,179	143,445,368	16,189	253	15.6	0.46	35.2
AU1159302-2s1A03	AU1159303	chrX	143,436,414	143,445,418	9,004	132	14.7	-0.83	38.0
AU062203-A03-2s1	AU062203	chrX	143,436,794	143,445,418	8,624	125	14.5	-0.65	38.2
AU032701-A03-3s1	AU032701	chrX	144,229,938	144,231,343	1,405	21	14.9	0.51	39.6
AU0852304-A03-2s1	AU0852304	chrX	145,207,069	145,208,670	1,601	22	13.7	-0.85	44.8
AU1862302-1s1A03	AU1862302	chrX	145,928,240	145,938,424	10,184	141	13.8	0.67	37.7
SSC00366-A03-2s1	SSC00366	chrX	146,880,046	146,883,212	3,166	31	9.8	0.89	54.4
SSC00417-A03-1s1	SSC00417	chrX	146,880,076	146,883,237	3,161	31	9.8	0.61	54.3
AU1001202-1s1A03	AU1001202	chrX	148,424,029	148,430,302	6,273	69	11.0	0.38	57.9
SSC00417-A03-1s1	SSC00417	chrX	148,424,029	148,430,028	5,999	63	10.5	0.50	57.3
AU1143303-1s1A03	AU1143303	chrX	148,424,584	148,425,204	620	10	16.1	0.79	55.0
SSC00225-A03-2s1	SSC00225	chrX	148,451,773	148,464,425	12,652	149	11.8	-1.23	42.8
SSC00098-A03-3s1	SSC00098	chrX	148,452,512	148,462,190	9,678	129	13.3	-1.33	43.0
SSC00225-A03-2s1	SSC00225	chrX	148,683,832	148,701,373	17,541	273	15.6	0.59	46.6
AU058503-A03-2s1	AU058503	chrX	148,856,643	148,859,450	2,807	58	20.7	0.37	66.8
AU067803-A03-2s1	AU067803	chrX	148,857,223	148,863,113	5,890	66	11.2	0.44	59.8
AU014803-A03-2s1	AU014803	chrX	148,857,413	148,864,793	7,380	76	10.3	0.46	59.1
AU018304-A03-2s1	AU018304	chrX	148,857,493	148,859,665	2,172	44	20.3	0.60	67.4
SSC00417-A03-1s1	SSC00417	chrX	148,857,722	148,859,420	1,698	34	20.0	0.74	69.1
AU028903-A03-2s1	AU028903	chrX	148,992,605	148,992,915	310	8	25.8	-0.74	30.3
AU1338304-1s1A03	AU1338304	chrX	149,280,368	149,281,173	805	16	19.9	-0.53	67.1
AU028903-A03-2s1	AU028903	chrX	149,678,535	149,679,160	625	11	17.6	-1.15	55.5
AU055303-A03-2s1	AU055303	chrX	149,678,535	149,680,094	1,559	19	12.2	-0.52	53.9
AU0939304-1s1A03	AU0939304	chrX	149,678,535	149,679,160	625	11	17.6	-1.32	55.5
AU1875302-1s1A03	AU1875302	chrX	149,678,535	149,679,545	1,010	14	13.9	-1.07	55.7
SSC00366-A03-2s1	SSC00366	chrX	149,942,968	149,946,773	3,805	80	21.0	0.54	37.3
AU1378304-2s1A03	AU1378304	chrX	149,943,088	149,946,648	3,560	74	20.8	-0.35	37.8
AU1001202-1s1A03	AU1001202	chrX	150,457,842	150,462,001	4,159	49	11.8	-0.85	38.7
AU1159302-2s1A03	AU1159303	chrX	150,457,902	150,462,690	4,788	53	11.1	-1.04	39.0
AU058503-A03-2s1	AU058503	chrX	150,608,963	150,610,913	1,950	42	21.5	-0.45	45.3
AU1211303-1s2A03	AU1211303	chrX	150,614,032	150,614,924	892	15	16.8	0.47	75.4
AU1585301-1s2A03	AU1585301	chrX	150,627,527	150,629,067	1,540	27	17.5	-0.92	51.8
SSC00035-3s1A03	SSC00035	chrX	150,627,527	150,629,137	1,610	28	17.4	-1.90	50.9
AU1953303-A03-1s1	AU1953303	chrX	150,779,751	150,783,476	3,725	76	20.4	-0.74	43.0
SSC00316-A03-2s1	SSC00316	chrX	151,562,641	151,564,146	1,505	32	21.3	-0.94	41.7
SSC00316-A03-2s1	SSC00316	chrX	151,742,321	151,744,296	1,975	42	21.3	-0.76	53.3
SSC00417-A03-1s1	SSC00417	chrX	151,742,456	151,743,651	1,195	26	21.8	0.82	54.3
SSC00035-3s1A03	SSC00035	chrX	151,854,912	151,857,952	3,040	62	20.4	-1.24	53.0
AU1073302-A03-2s1	AU1073302	chrX	151,979,999	151,980,324	325	8	24.6	-0.64	38.5
AU0939304-1s1A03	AU0939304	chrX	152,120,424	152,192,189	71,765	807	11.2	0.46	43.4
AU1143303-1s1A03	AU1143303	chrX	152,238,877	152,240,102	1,225	17	13.9	0.59	60.0
SSC00093-A03-1s2	SSC00093	chrX	152,264,330	152,265,430	1,100	23	20.9	0.74	55.8
SSC00097-A03-1s1	SSC00097	chrX	152,264,330	152,265,335	1,005	21	20.9	0.75	55.8
SSC00316-A03-2s1	SSC00316	chrX	152,301,264	152,303,154	1,890	34	18.0	-0.75	51.7
AU1344302-1s1A03	AU1344302	chrX	152,363,112	152,364,047	935	17	18.2	-0.50	67.8
AU0895303-1s2A03	AU0895303	chrX	152,372,416	152,494,466	122,050	2291	18.8	0.26	59.6
AU014803-A03-2s1	AU014803	chrX	152,389,076	152,389,570	494	10	20.2	0.89	72.9
AU0895303-1s2A03	AU0895303	chrX	152,546,991	152,562,004	15,013	282	18.8	0.27	58.7

SSC00316-A03-2s1	SSC00316	chrX	152,560,776	152,561,914	1,138	24	21.1	-0.96	70.4
AU1622302-1s2A03	AU1622302	chrX	152,573,639	152,576,282	2,643	24	9.1	-0.65	47.6
SSC00093-A03-1s2	SSC00093	chrX	152,606,438	152,608,370	1,932	26	13.5	-0.56	75.9
AU0895303-1s2A03	AU0895303	chrX	152,642,042	152,724,158	82,116	1266	15.4	0.25	58.1
AU028903-A03-2s1	AU028903	chrX	152,690,489	152,730,588	40,099	726	18.1	0.33	61.6
SSC00077-A03-3s1	SSC00077	chrX	152,695,235	152,695,665	430	10	23.3	-0.98	74.6
SSC00093-A03-1s2	SSC00093	chrX	152,890,765	152,892,370	1,605	25	15.6	-0.59	62.7
SSC00549-A03-1s1	SSC00549	chrX	154,039,041	154,063,752	24,711	276	11.2	-1.30	36.7
AU0983302-1s1A03	AU0983302	chrX	154,044,880	154,057,130	12,250	179	14.6	-1.07	35.0
AU083504-A03-3s1	AU083504	chrX	154,047,946	154,057,440	9,494	147	15.5	-0.68	34.8
AU083504-A03-3s1	AU083504	chrX	154,067,596	154,071,868	4,272	60	14.0	-0.77	36.1
AU0983302-1s1A03	AU0983302	chrX	154,067,596	154,071,023	3,427	59	17.2	-1.12	35.3
AU1001202-1s1A03	AU1001202	chrX	154,278,014	154,583,237	305,223	2995	9.8	0.51	38.1
AU058103-A03-2s1	AU058103	chrX	154,431,718	154,434,720	3,002	33	11.0	-0.63	36.1
SSC00260-A03-2s1	SSC00260	chrX	154,582,223	154,583,237	1,014	12	11.8	-1.54	37.5

Table A.4 – CNV from 102 NIMH samples run on 2.1M arrays by NimbleGen protocol.

NIMH54-1s1	NIMH54	chrX	31,773,231	31,773,887	656	14	21.3	-0.48	40.5
NIMH53-A01-2s2	NIMH53	chrX	33,558,581	33,559,556	975	20	20.5	0.47	55.2
NIMH95-1s1	NIMH95	chrX	33,558,581	33,559,556	975	20	20.5	0.42	55.2
NIMH36-1s1	NIMH36	chrX	33,952,881	33,981,250	28,369	473	16.7	-0.79	34.8
NIMH40-1s1	NIMH40	chrX	33,952,911	33,964,182	11,271	183	16.2	-0.99	34.0
NIMH01-1s1	NIMH01	chrX	34,063,446	34,065,466	2,020	20	9.9	0.58	33.1
NIMH04-1s1	NIMH04	chrX	34,063,446	34,065,466	2,020	20	9.9	0.81	33.1
NIMH07-1s1	NIMH07	chrX	34,063,446	34,065,466	2,020	20	9.9	0.75	33.1
NIMH08-1s1	NIMH08	chrX	34,063,446	34,065,466	2,020	20	9.9	0.60	33.1
NIMH100-1s1	NIMH100	chrX	34,063,446	34,065,466	2,020	20	9.9	0.88	33.1
NIMH16-1s1	NIMH16	chrX	34,063,446	34,065,401	1,955	19	9.7	0.79	32.9
NIMH19-1s1	NIMH19	chrX	34,063,446	34,065,466	2,020	20	9.9	0.81	33.1
NIMH20-1s1	NIMH20	chrX	34,063,446	34,065,401	1,955	19	9.7	0.72	32.9
NIMH30-1s2	NIMH30	chrX	34,063,446	34,065,401	1,955	19	9.7	0.69	32.9
NIMH35-1s1	NIMH35	chrX	34,063,446	34,065,466	2,020	20	9.9	0.66	33.1
NIMH43-1s1	NIMH43	chrX	34,063,446	34,065,401	1,955	19	9.7	0.83	32.9
NIMH49-1s1	NIMH49	chrX	34,063,446	34,065,401	1,955	19	9.7	1.04	32.9
NIMH50-1s1	NIMH50	chrX	34,063,446	34,065,401	1,955	19	9.7	0.99	32.9
NIMH51-1s1	NIMH51	chrX	34,063,446	34,065,466	2,020	20	9.9	0.85	33.1
NIMH53-A01-2s2	NIMH53	chrX	34,063,446	34,065,401	1,955	19	9.7	0.86	32.9
NIMH55-1s1	NIMH55	chrX	34,063,446	34,065,274	1,828	18	9.8	0.74	32.5
NIMH56-1s1	NIMH56	chrX	34,063,446	34,065,466	2,020	20	9.9	0.74	33.1
NIMH57-1s1	NIMH57	chrX	34,063,446	34,065,401	1,955	19	9.7	0.92	32.9
NIMH63-1s1	NIMH63	chrX	34,063,446	34,065,401	1,955	19	9.7	0.90	32.9
NIMH66-1s1	NIMH66	chrX	34,063,446	34,065,401	1,955	19	9.7	0.69	32.9
NIMH67-1s1	NIMH67	chrX	34,063,446	34,065,466	2,020	20	9.9	0.82	33.1
NIMH69-2s1	NIMH69	chrX	34,063,446	34,065,401	1,955	19	9.7	0.77	32.9
NIMH75-1s1	NIMH75	chrX	34,063,446	34,065,401	1,955	19	9.7	1.03	32.9
NIMH76-1s1	NIMH76	chrX	34,063,446	34,065,401	1,955	19	9.7	1.05	32.9
NIMH80-1s1	NIMH80	chrX	34,063,446	34,065,466	2,020	20	9.9	0.81	33.1
NIMH81-1s1	NIMH81	chrX	34,063,446	34,066,116	2,670	25	9.4	0.61	35.0
NIMH82-1s1	NIMH82	chrX	34,063,446	34,065,401	1,955	19	9.7	0.67	32.9
NIMH84-1s1	NIMH84	chrX	34,063,446	34,066,116	2,670	25	9.4	0.72	35.0
NIMH88-1s1	NIMH88	chrX	34,063,446	34,065,401	1,955	19	9.7	1.02	32.9
NIMH90-1s1	NIMH90	chrX	34,063,446	34,065,401	1,955	19	9.7	0.86	32.9
NIMH93-1s1	NIMH93	chrX	34,063,446	34,065,466	2,020	20	9.9	0.90	33.1
NIMH94-1s1	NIMH94	chrX	34,063,446	34,065,466	2,020	20	9.9	0.73	33.1
NIMH96-1s1	NIMH96	chrX	34,063,446	34,065,466	2,020	20	9.9	0.81	33.1
NIMH99-A01-1s2	NIMH99	chrX	34,063,446	34,066,116	2,670	25	9.4	0.66	35.0
NIMH24-1s1	NIMH24	chrX	34,063,466	34,065,401	1,935	18	9.3	0.77	32.8
NIMH64-1s1	NIMH64	chrX	35,624,856	35,625,766	910	19	20.9	-1.11	37.8
NIMH69-2s1	NIMH69	chrX	35,689,759	35,690,664	905	17	18.8	0.38	50.6
NIMH68-1s1	NIMH68	chrX	35,921,767	36,190,195	268,428	3,995	14.9	0.66	34.9
NIMH32-1s1	NIMH32	chrX	36,361,872	36,363,155	1,283	22	17.1	-0.52	41.7
NIMH67-1s1	NIMH67	chrX	36,560,109	36,561,739	1,630	32	19.6	0.28	48.0
NIMH75-1s1	NIMH75	chrX	37,998,850	37,999,025	175	5	28.6	0.73	41.9
NIMH69-2s1	NIMH69	chrX	38,283,541	38,283,926	385	8	20.8	0.64	33.8
NIMH37-A01-2s2	NIMH37	chrX	38,283,561	38,283,926	365	7	19.2	0.81	34.6
NIMH75-1s1	NIMH75	chrX	38,305,103	38,306,959	1,856	37	19.9	0.22	61.2
NIMH84-1s1	NIMH84	chrX	39,494,577	39,502,493	7,916	128	16.2	0.63	49.6
NIMH17-1s1	NIMH17	chrX	39,810,721	39,864,432	53,711	972	18.1	0.34	53.7
NIMH19-1s1	NIMH19	chrX	39,823,805	39,853,234	29,429	560	19.0	0.36	57.5
NIMH21-1s1	NIMH21	chrX	39,833,606	39,848,822	15,216	291	19.1	0.28	57.8
NIMH32-1s1	NIMH32	chrX	39,834,985	39,835,866	881	19	21.6	-0.59	47.3
NIMH30-1s2	NIMH30	chrX	39,835,100	39,839,101	4,001	82	20.5	-0.25	55.0
NIMH28-1s1	NIMH28	chrX	39,848,822	39,854,158	5,336	96	18.0	0.63	68.5
NIMH44-1s1	NIMH44	chrX	39,913,004	39,913,864	860	17	19.8	-0.59	50.2
NIMH66-1s1	NIMH66	chrX	40,006,882	40,008,277	1,395	29	20.8	0.34	52.4
NIMH37-A01-2s2	NIMH37	chrX	40,324,376	40,325,651	1,275	20	15.7	0.37	60.1
NIMH96-1s1	NIMH96	chrX	40,828,639	40,831,205	2,566	50	19.5	0.31	65.8
NIMH37-A01-2s2	NIMH37	chrX	41,084,397	41,085,107	710	16	22.5	0.57	36.1
NIMH28-1s1	NIMH28	chrX	41,217,647	41,218,957	1,310	24	18.3	1.01	72.4
NIMH53-A01-2s2	NIMH53	chrX	43,398,261	43,398,971	710	15	21.1	-0.48	53.2
NIMH28-1s1	NIMH28	chrX	43,458,110	43,464,890	6,780	139	20.5	0.56	49.8
NIMH53-A01-2s2	NIMH53	chrX	43,499,243	43,501,120	1,877	39	20.8	0.27	48.3
NIMH74-A01-1s2	NIMH74	chrX	44,616,456	44,618,401	1,945	36	18.5	0.25	68.7
NIMH44-1s1	NIMH44	chrX	45,354,014	45,355,254	1,240	26	21.0	0.31	55.2
NIMH02-1s1	NIMH02	chrX	45,671,748	45,674,192	2,444	45	18.4	-1.02	33.7
NIMH28-1s1	NIMH28	chrX	46,317,849	46,319,871	2,022	40	19.8	0.76	69.6
NIMH37-A01-2s2	NIMH37	chrX	46,317,985	46,320,612	2,627	51	19.4	0.39	64.8

NIMH89-1s1	NIMH89	chrX	46,318,765	46,320,346	1,581	31	19.6	0.32	64.3
NIMH95-1s1	NIMH95	chrX	46,407,337	46,412,560	5,223	73	14.0	0.55	41.6
NIMH92-1s1	NIMH92	chrX	46,408,444	46,412,560	4,116	69	16.8	0.50	39.7
NIMH20-1s1	NIMH20	chrX	46,892,390	46,991,488	99,098	1,596	16.1	0.30	50.0
NIMH21-1s1	NIMH21	chrX	46,913,091	46,935,978	22,887	406	17.7	0.29	54.9
NIMH24-1s1	NIMH24	chrX	46,913,091	46,947,729	34,638	622	18.0	0.27	54.0
NIMH46-1s1	NIMH46	chrX	46,925,929	46,931,744	5,815	97	16.7	0.24	54.6
NIMH17-1s1	NIMH17	chrX	46,940,817	46,974,251	33,434	609	18.2	0.35	51.9
NIMH23-1s1	NIMH23	chrX	46,962,695	46,965,814	3,119	63	20.2	-0.33	56.5
NIMH44-1s1	NIMH44	chrX	46,963,535	46,965,744	2,209	47	21.3	-0.41	52.0
R085B11-A01-1s1	R085B11	chrX	46,963,756	46,965,524	1,768	38	21.5	-0.31	51.9
NIMH94-1s1	NIMH94	chrX	46,965,794	46,972,101	6,307	122	19.3	0.19	54.3
NIMH32-1s1	NIMH32	chrX	46,966,104	46,974,361	8,257	157	19.0	0.30	53.0
NIMH79-1s1	NIMH79	chrX	47,213,816	47,227,191	13,375	149	11.1	-0.25	41.5
NIMH39-1s1	NIMH39	chrX	47,214,228	47,227,451	13,223	151	11.4	-0.37	42.0
R085B11-A01-1s1	R085B11	chrX	47,225,263	47,226,731	1,468	21	14.3	-0.51	43.6
NIMH23-1s1	NIMH23	chrX	47,226,291	47,228,673	2,382	47	19.7	-0.33	52.4
NIMH19-1s1	NIMH19	chrX	47,361,748	47,373,433	11,685	230	19.7	0.34	52.9
NIMH01-1s1	NIMH01	chrX	47,364,422	47,371,993	7,571	150	19.8	0.23	52.4
NIMH33-1s1	NIMH33	chrX	47,369,703	47,376,220	6,517	120	18.4	0.30	54.3
NIMH31-1s2	NIMH31	chrX	47,465,711	47,468,996	3,285	67	20.4	0.32	55.5
NIMH24-1s1	NIMH24	chrX	47,465,856	47,471,201	5,345	103	19.3	0.25	55.1
NIMH05-2s1	NIMH05	chrX	48,202,382	48,208,560	6,178	97	15.7	0.40	51.7
NIMH92-1s1	NIMH92	chrX	48,201,697	48,208,745	7,048	113	16.0	0.27	52.2
NIMH17-1s1	NIMH17	chrX	48,201,887	48,267,211	65,324	1,050	16.1	0.28	51.1
R085B11-A02-1s1	R085B11	chrX	48,220,111	48,221,146	1,035	23	22.2	-0.50	45.3
NIMH05-2s1	NIMH05	chrX	48,251,623	48,257,519	5,896	117	19.8	0.37	54.3
NIMH23-1s1	NIMH23	chrX	48,282,983	48,284,033	1,050	23	21.9	-0.71	59.5
NIMH33-1s1	NIMH33	chrX	48,283,123	48,283,993	870	19	21.8	-0.72	57.3
NIMH24-1s1	NIMH24	chrX	48,283,188	48,283,888	700	16	22.9	-0.52	56.1
NIMH25-1s1	NIMH25	chrX	48,283,188	48,283,888	700	16	22.9	-0.59	56.1
NIMH38-1s1	NIMH38	chrX	48,283,188	48,284,033	845	19	22.5	-0.59	55.4
NIMH34-1s1	NIMH34	chrX	48,283,223	48,284,033	810	18	22.2	-0.61	54.4
R085B11-A02-1s1	R085B11	chrX	48,283,338	48,283,993	655	15	22.9	-0.57	52.1
NIMH08-1s1	NIMH08	chrX	48,341,265	48,345,123	3,858	71	18.4	0.38	57.7
NIMH40-1s1	NIMH40	chrX	48,342,305	48,345,558	3,253	59	18.1	0.37	57.2
NIMH24-1s1	NIMH24	chrX	48,426,969	48,452,477	25,508	471	18.5	0.23	52.0
NIMH08-1s1	NIMH08	chrX	48,435,103	48,449,215	14,112	256	18.1	0.20	51.4
NIMH23-1s1	NIMH23	chrX	48,439,518	48,441,098	1,580	29	18.4	-0.49	63.0
NIMH23-1s1	NIMH23	chrX	48,441,198	48,453,617	12,419	245	19.7	0.25	52.6
NIMH33-1s1	NIMH33	chrX	48,443,238	48,454,022	10,784	212	19.7	0.22	53.5
NIMH67-1s1	NIMH67	chrX	48,447,204	48,452,902	5,698	110	19.3	0.32	56.1
R085B10-1s1	R085B10	chrX	48,454,022	48,471,467	17,445	298	17.1	-0.34	43.3
NIMH28-1s1	NIMH28	chrX	48,456,967	48,464,854	7,887	137	17.4	0.34	43.4
NIMH17-1s1	NIMH17	chrX	48,528,440	49,005,879	477,439	6,344	13.3	0.27	50.4
R085B11-A02-1s1	R085B11	chrX	48,530,197	48,531,062	865	19	22.0	-0.58	47.9
NIMH13-1s1	NIMH13	chrX	48,557,175	48,571,137	13,962	234	16.8	0.29	55.8
NIMH10-1s1	NIMH10	chrX	48,557,395	48,569,507	12,112	199	16.4	0.36	53.9
NIMH96-1s1	NIMH96	chrX	48,566,205	48,571,232	5,027	97	19.3	0.23	59.8
NIMH25-1s1	NIMH25	chrX	48,575,220	48,578,431	3,211	54	16.8	-0.39	48.9
NIMH31-1s2	NIMH31	chrX	48,575,721	48,578,381	2,660	46	17.3	-0.36	43.9
NIMH38-1s1	NIMH38	chrX	48,575,721	48,578,431	2,710	47	17.3	-0.40	44.0
NIMH34-1s1	NIMH34	chrX	48,576,846	48,578,431	1,585	31	19.6	-0.40	45.5
R085B11-A02-1s1	R085B11	chrX	48,577,861	48,578,431	570	12	21.1	-0.57	50.2
NIMH08-1s1	NIMH08	chrX	48,631,753	48,668,438	36,685	613	16.7	0.21	52.9
NIMH04-1s1	NIMH04	chrX	48,632,601	48,653,495	20,894	341	16.3	0.18	52.9
NIMH44-1s1	NIMH44	chrX	48,698,440	48,699,240	800	18	22.5	-0.55	52.4
NIMH33-1s1	NIMH33	chrX	48,698,580	48,699,240	660	15	22.7	-0.74	53.4
NIMH25-1s1	NIMH25	chrX	48,698,635	48,699,240	605	14	23.1	-0.66	53.5
NIMH34-1s1	NIMH34	chrX	48,698,635	48,699,240	605	14	23.1	-0.66	53.5
NIMH32-1s1	NIMH32	chrX	48,698,660	48,699,240	580	13	22.4	-0.81	53.3
NIMH21-1s1	NIMH21	chrX	48,782,705	48,847,113	64,408	854	13.3	0.26	52.2
NIMH20-1s1	NIMH20	chrX	48,788,428	49,031,490	243,062	3,470	14.3	0.29	51.3
NIMH16-1s1	NIMH16	chrX	48,796,167	48,827,185	31,018	512	16.5	0.27	54.7
R085B11-A02-1s1	R085B11	chrX	48,797,828	48,798,448	620	13	21.0	0.39	73.8
NIMH30-1s2	NIMH30	chrX	48,798,448	48,799,503	1,055	22	20.9	-0.37	52.1
NIMH31-1s2	NIMH31	chrX	48,798,448	48,799,503	1,055	22	20.9	-0.33	52.1
NIMH32-1s1	NIMH32	chrX	48,798,448	48,799,243	795	16	20.1	-0.60	51.7
NIMH33-1s1	NIMH33	chrX	48,798,448	48,799,433	985	20	20.3	-0.56	52.1

NIMH34-1s1	NIMH34	chrX	48,798,448	48,799,433	985	20	20.3	-0.46	52.1
NIMH38-1s1	NIMH38	chrX	48,798,448	48,799,458	1,010	21	20.8	-0.43	52.3
NIMH40-1s1	NIMH40	chrX	48,798,448	48,799,368	920	19	20.7	-0.57	52.2
NIMH23-1s1	NIMH23	chrX	48,798,523	48,799,503	980	21	21.4	-0.62	51.8
NIMH25-1s1	NIMH25	chrX	48,798,523	48,799,503	980	21	21.4	-0.45	51.8
R085B11-A02-1s1	R085B11	chrX	48,798,523	48,799,268	745	16	21.5	-0.50	51.4
NIMH23-1s1	NIMH23	chrX	48,799,563	48,816,512	16,949	284	16.8	0.24	54.3
NIMH31-1s2	NIMH31	chrX	48,799,563	48,822,720	23,157	331	14.3	0.31	55.3
NIMH19-1s1	NIMH19	chrX	48,806,232	48,821,770	15,538	265	17.1	0.37	57.0
NIMH05-2s1	NIMH05	chrX	48,809,147	48,813,283	4,136	80	19.3	0.48	58.3
NIMH23-1s1	NIMH23	chrX	48,816,537	48,818,120	1,583	33	20.8	-0.46	61.5
NIMH23-1s1	NIMH23	chrX	48,818,195	48,823,690	5,495	99	18.0	0.29	56.0
NIMH19-1s1	NIMH19	chrX	48,898,879	48,937,479	38,600	595	15.4	0.30	54.0
NIMH05-2s1	NIMH05	chrX	48,898,899	49,018,033	119,134	1,979	16.6	0.25	52.9
NIMH21-1s1	NIMH21	chrX	48,899,204	48,936,714	37,510	575	15.3	0.30	54.1
NIMH04-1s1	NIMH04	chrX	48,899,359	48,935,279	35,920	549	15.3	0.19	54.4
NIMH16-1s1	NIMH16	chrX	48,899,404	49,005,129	105,725	1,719	16.3	0.22	53.3
NIMH23-1s1	NIMH23	chrX	48,906,630	48,907,975	1,345	26	19.3	-0.66	60.8
NIMH30-1s2	NIMH30	chrX	48,907,145	48,907,975	830	18	21.7	-0.56	57.2
NIMH44-1s1	NIMH44	chrX	48,907,145	48,907,975	830	18	21.7	-0.62	57.2
NIMH40-1s1	NIMH40	chrX	48,907,230	48,908,005	775	17	21.9	-0.72	55.3
NIMH25-1s1	NIMH25	chrX	48,907,350	48,908,070	720	16	22.2	-0.66	53.1
NIMH26-1s1	NIMH26	chrX	48,907,350	48,907,975	625	14	22.4	-0.61	52.8
NIMH31-1s2	NIMH31	chrX	48,907,350	48,907,755	405	10	24.7	-0.69	55.2
NIMH32-1s1	NIMH32	chrX	48,907,350	48,908,005	655	15	22.9	-0.67	52.6
NIMH33-1s1	NIMH33	chrX	48,907,350	48,907,975	625	14	22.4	-0.70	52.8
NIMH34-1s1	NIMH34	chrX	48,907,350	48,908,070	720	16	22.2	-0.66	53.1
NIMH38-1s1	NIMH38	chrX	48,907,350	48,908,070	720	16	22.2	-0.67	53.1
NIMH31-1s2	NIMH31	chrX	48,948,311	48,956,342	8,031	140	17.4	0.37	53.4
NIMH75-1s1	NIMH75	chrX	48,973,682	48,975,270	1,588	33	20.8	0.27	60.6
NIMH63-1s1	NIMH63	chrX	48,987,490	48,988,715	1,225	21	17.1	-0.26	50.2
NIMH92-1s1	NIMH92	chrX	48,988,570	48,999,259	10,689	192	18.0	0.24	56.3
NIMH63-1s1	NIMH63	chrX	48,989,090	49,005,464	16,374	309	18.9	0.23	56.4
NIMH48-1s1	NIMH48	chrX	48,989,869	49,005,184	15,315	293	19.1	0.23	57.3
NIMH50-1s1	NIMH50	chrX	48,989,869	48,996,304	6,435	130	20.2	0.37	59.5
R085B11-A02-1s1	R085B11	chrX	48,989,869	48,995,154	5,285	107	20.2	0.28	60.9
NIMH10-1s1	NIMH10	chrX	48,989,894	48,995,546	5,652	114	20.2	0.47	60.5
NIMH31-1s2	NIMH31	chrX	48,991,049	48,996,304	5,255	107	20.4	0.38	59.2
NIMH33-1s1	NIMH33	chrX	48,991,699	48,995,104	3,405	70	20.6	0.37	60.7
NIMH35-1s1	NIMH35	chrX	48,991,699	48,995,419	3,720	76	20.4	0.30	60.2
R085B11-A02-1s1	R085B11	chrX	48,995,219	48,996,359	1,140	24	21.1	-0.35	52.8
NIMH44-1s1	NIMH44	chrX	48,995,364	48,996,414	1,050	22	21.0	-0.28	52.6
NIMH92-1s1	NIMH92	chrX	48,999,329	48,999,754	425	9	21.2	-0.60	53.9
NIMH26-1s1	NIMH26	chrX	48,999,379	48,999,834	455	9	19.8	-0.58	53.6
NIMH32-1s1	NIMH32	chrX	48,999,379	48,999,834	455	9	19.8	-0.73	53.6
NIMH94-1s1	NIMH94	chrX	48,999,379	48,999,834	455	9	19.8	-0.65	53.6
NIMH95-1s1	NIMH95	chrX	48,999,379	48,999,754	375	8	21.3	-0.59	53.7
NIMH38-1s1	NIMH38	chrX	48,999,424	48,999,899	475	9	18.9	-0.62	52.2
NIMH53-2s1	NIMH53	chrX	48,999,424	48,999,754	330	7	21.2	-0.72	53.2
NIMH96-1s1	NIMH96	chrX	48,999,424	48,999,754	330	7	21.2	-0.82	53.2
NIMH38-1s1	NIMH38	chrX	48,999,924	49,005,224	5,300	105	19.8	0.24	56.4
NIMH32-1s1	NIMH32	chrX	49,012,855	49,013,180	325	8	24.6	-0.89	58.7
NIMH34-1s1	NIMH34	chrX	49,012,920	49,013,270	350	8	22.9	-0.82	63.9
NIMH38-1s1	NIMH38	chrX	49,012,920	49,013,135	215	6	27.9	-0.85	58.2
NIMH96-1s1	NIMH96	chrX	49,029,180	49,030,525	1,345	28	20.8	0.32	62.2
NIMH57-1s1	NIMH57	chrX	49,438,765	49,450,241	11,476	219	19.1	0.23	43.0
NIMH90-1s1	NIMH90	chrX	49,443,375	49,448,387	5,012	96	19.2	0.20	43.2
NIMH04-1s1	NIMH04	chrX	50,145,957	50,146,797	840	19	22.6	-0.59	44.4
NIMH21-1s1	NIMH21	chrX	50,228,557	50,232,495	3,938	78	19.8	0.32	56.5
NIMH67-1s1	NIMH67	chrX	51,069,986	51,070,966	980	22	22.4	0.38	55.4
NIMH60-1s1	NIMH60	chrX	51,166,273	51,167,998	1,725	37	21.4	0.35	64.1
NIMH28-1s1	NIMH28	chrX	51,627,939	51,629,289	1,350	27	20.0	0.96	41.5
R085B11-A02-1s1	R085B11	chrX	51,628,194	51,629,004	810	18	22.2	-0.60	42.2
NIMH05-2s1	NIMH05	chrX	51,654,500	51,656,155	1,655	35	21.1	0.36	52.2
NIMH06-1s1	NIMH06	chrX	52,966,578	52,967,078	500	12	24.0	0.77	63.7
NIMH07-1s1	NIMH07	chrX	53,099,962	53,102,357	2,395	43	18.0	-0.33	48.7
NIMH79-1s1	NIMH79	chrX	53,100,517	53,102,747	2,230	46	20.6	0.33	47.6
NIMH69-2s1	NIMH69	chrX	53,127,042	53,127,447	405	9	22.2	-0.76	62.7
NIMH04-1s1	NIMH04	chrX	53,236,172	53,242,761	6,589	130	19.7	0.19	53.7

NIMH26-1s1	NIMH26	chrX	53,238,248	53,243,646	5,398	107	19.8	0.30	57.2
NIMH35-1s1	NIMH35	chrX	53,238,268	53,242,691	4,423	88	19.9	0.31	57.7
NIMH33-1s1	NIMH33	chrX	53,367,506	53,368,361	855	19	22.2	-0.42	55.7
NIMH21-1s1	NIMH21	chrX	53,726,483	53,731,155	4,672	96	20.5	0.26	49.6
NIMH75-1s1	NIMH75	chrX	53,801,385	53,803,940	2,555	36	14.1	0.28	55.3
NIMH67-1s1	NIMH67	chrX	53,801,995	53,804,775	2,780	35	12.6	0.47	56.3
NIMH08-1s1	NIMH08	chrX	54,400,369	54,403,940	3,571	70	19.6	0.28	52.7
NIMH33-1s1	NIMH33	chrX	54,539,576	54,540,066	490	11	22.4	-0.69	49.5
NIMH20-1s1	NIMH20	chrX	54,572,044	54,576,106	4,062	74	18.2	0.41	51.7
NIMH31-1s2	NIMH31	chrX	54,790,608	54,792,468	1,860	39	21.0	-0.22	42.6
NIMH33-1s1	NIMH33	chrX	54,790,608	54,791,893	1,285	28	21.8	-0.38	44.2
NIMH44-1s1	NIMH44	chrX	54,790,608	54,792,193	1,585	33	20.8	-0.36	43.5
R085B11-A02-1s1	R085B11	chrX	54,790,693	54,791,893	1,200	26	21.7	-0.47	43.8
NIMH63-1s1	NIMH63	chrX	54,862,794	54,868,184	5,390	59	10.9	0.37	54.4
NIMH94-1s1	NIMH94	chrX	54,863,029	54,870,904	7,875	106	13.5	0.25	53.3
NIMH13-1s1	NIMH13	chrX	54,864,114	54,865,024	910	20	22.0	0.60	56.4
NIMH46-1s1	NIMH46	chrX	54,864,114	54,868,374	4,260	53	12.4	0.37	53.8
NIMH66-1s1	NIMH66	chrX	54,864,114	54,874,190	10,076	102	10.1	0.30	48.1
NIMH56-1s1	NIMH56	chrX	54,864,159	54,865,024	865	19	22.0	0.50	56.1
NIMH57-1s1	NIMH57	chrX	54,964,293	54,968,774	4,481	92	20.5	0.24	51.0
NIMH31-1s2	NIMH31	chrX	54,976,064	54,979,453	3,389	67	19.8	0.37	54.4
NIMH33-1s1	NIMH33	chrX	55,011,721	55,012,228	507	10	19.7	-1.23	29.5
NIMH33-1s1	NIMH33	chrX	56,605,975	56,607,060	1,085	21	19.4	-0.63	50.1
NIMH23-1s1	NIMH23	chrX	56,606,070	56,607,155	1,085	22	20.3	-0.62	53.3
NIMH30-1s2	NIMH30	chrX	56,606,070	56,607,060	990	20	20.2	-0.46	52.0
NIMH34-1s1	NIMH34	chrX	56,606,070	56,606,755	685	14	20.4	-0.63	43.3
NIMH40-1s1	NIMH40	chrX	56,606,070	56,607,155	1,085	22	20.3	-0.55	53.3
NIMH69-2s1	NIMH69	chrX	56,801,592	56,839,447	37,855	448	11.8	0.24	52.6
NIMH75-1s1	NIMH75	chrX	56,845,230	56,846,209	979	19	19.4	0.95	53.7
NIMH67-1s1	NIMH67	chrX	57,558,981	57,560,926	1,945	37	19.0	0.41	49.7
NIMH92-1s1	NIMH92	chrX	57,559,476	57,560,426	950	21	22.1	0.55	55.0
NIMH96-1s1	NIMH96	chrX	57,783,495	57,784,695	1,200	26	21.7	0.32	52.8
NIMH10-1s1	NIMH10	chrX	58,292,047	58,310,299	18,252	353	19.3	0.27	57.0
NIMH15-1s1	NIMH15	chrX	58,292,647	58,358,518	65,871	1,126	17.1	0.23	51.4
NIMH84-1s1	NIMH84	chrX	58,293,082	58,336,004	42,922	839	19.5	0.33	59.3
NIMH29-1s1	NIMH29	chrX	58,296,252	58,323,045	26,793	527	19.7	0.20	60.7
NIMH38-1s1	NIMH38	chrX	58,296,252	58,310,189	13,937	271	19.4	0.22	60.0
NIMH41-1s1	NIMH41	chrX	58,296,282	58,340,997	44,715	828	18.5	0.26	57.6
NIMH42-1s1	NIMH42	chrX	58,301,112	58,336,004	34,892	681	19.5	0.28	60.4
NIMH51-1s1	NIMH51	chrX	58,304,375	58,333,665	29,290	573	19.6	0.31	60.8
NIMH36-1s1	NIMH36	chrX	58,307,292	58,329,969	22,677	446	19.7	0.27	61.1
NIMH89-1s1	NIMH89	chrX	58,309,851	58,330,259	20,408	404	19.8	0.26	61.2
NIMH19-1s1	NIMH19	chrX	58,312,070	58,344,028	31,958	564	17.6	0.27	54.3
NIMH50-1s1	NIMH50	chrX	58,312,690	58,363,350	50,660	795	15.7	0.26	47.3
NIMH88-1s1	NIMH88	chrX	61,905,249	61,916,897	11,648	144	12.4	-0.60	36.6
NIMH17-1s1	NIMH17	chrX	62,386,739	62,413,236	26,497	272	10.3	0.45	36.9
NIMH53-2s1	NIMH53	chrX	62,386,739	62,415,948	29,209	322	11.0	0.61	37.5
NIMH56-1s1	NIMH56	chrX	62,386,739	62,415,998	29,259	323	11.0	0.66	37.5
NIMH59-1s1	NIMH59	chrX	62,386,739	62,422,378	35,639	442	12.4	0.78	38.6
NIMH38-1s1	NIMH38	chrX	63,169,441	63,171,216	1,775	36	20.3	0.28	52.1
NIMH67-1s1	NIMH67	chrX	63,169,541	63,171,216	1,675	34	20.3	0.37	52.3
NIMH53-2s1	NIMH53	chrX	63,568,900	63,571,275	2,375	48	20.2	0.33	54.9
NIMH94-1s1	NIMH94	chrX	63,568,900	63,571,350	2,450	49	20.0	0.27	54.9
NIMH95-1s1	NIMH95	chrX	63,568,900	63,571,350	2,450	49	20.0	0.26	54.9
NIMH89-1s1	NIMH89	chrX	63,568,975	63,570,395	1,420	31	21.8	0.34	56.1
NIMH44-1s1	NIMH44	chrX	64,389,740	64,390,835	1,095	24	21.9	0.47	56.8
NIMH30-1s2	NIMH30	chrX	64,803,458	64,804,533	1,075	21	19.5	-0.34	59.7
NIMH73-2s1	NIMH73	chrX	64,939,637	64,941,872	2,235	40	17.9	-1.38	32.4
NIMH75-1s1	NIMH75	chrX	65,190,638	65,191,913	1,275	27	21.2	0.34	55.0
NIMH26-1s1	NIMH26	chrX	65,634,335	65,635,710	1,375	30	21.8	0.35	52.3
NIMH67-1s1	NIMH67	chrX	66,413,571	66,414,941	1,370	28	20.4	0.41	52.3
NIMH57-1s1	NIMH57	chrX	66,574,116	66,645,439	71,323	768	10.8	0.47	35.9
NIMH21-1s1	NIMH21	chrX	66,682,160	66,685,094	2,934	58	19.8	0.25	51.1
NIMH34-1s1	NIMH34	chrX	66,682,595	66,684,030	1,435	28	19.5	-0.46	59.0
NIMH44-1s1	NIMH44	chrX	66,682,595	66,684,070	1,475	29	19.7	-0.46	58.6
NIMH35-1s1	NIMH35	chrX	66,682,765	66,684,070	1,305	26	19.9	-0.48	58.7
NIMH25-1s1	NIMH25	chrX	66,682,945	66,683,935	990	20	20.2	-0.52	58.9
NIMH26-1s1	NIMH26	chrX	66,682,945	66,684,070	1,125	22	19.6	-0.65	57.1
NIMH33-1s1	NIMH33	chrX	66,683,180	66,683,935	755	17	22.5	-0.71	54.1

NIMH69-2s1	NIMH69	chrX	66,893,410	66,894,459	1,049	16	15.3	0.48	39.3
NIMH100-1s1	NIMH100	chrX	67,040,582	67,046,892	6,310	82	13.0	-1.17	38.2
NIMH18-2s1	NIMH18	chrX	67,040,582	67,046,967	6,385	83	13.0	-1.00	38.2
NIMH50-1s1	NIMH50	chrX	67,040,582	67,046,967	6,385	83	13.0	-1.22	38.2
NIMH53-2s1	NIMH53	chrX	67,040,582	67,046,967	6,385	83	13.0	-1.04	38.2
NIMH75-1s1	NIMH75	chrX	67,040,582	67,046,967	6,385	83	13.0	-1.32	38.2
NIMH49-1s1	NIMH49	chrX	67,045,497	67,046,967	1,470	25	17.0	-1.83	41.0
NIMH79-1s1	NIMH79	chrX	67,046,867	67,046,967	100	3	30.0	-1.33	38.5
NIMH31-1s2	NIMH31	chrX	67,168,843	67,170,063	1,220	27	22.1	0.52	56.0
NIMH53-2s1	NIMH53	chrX	67,168,878	67,169,728	850	19	22.4	0.45	55.7
NIMH95-1s1	NIMH95	chrX	67,168,913	67,169,693	780	17	21.8	0.38	55.8
NIMH32-1s1	NIMH32	chrX	67,853,287	67,860,578	7,291	148	20.3	0.27	56.5
NIMH10-1s1	NIMH10	chrX	67,879,927	67,880,997	1,070	23	21.5	-0.35	46.8
NIMH44-1s1	NIMH44	chrX	67,879,927	67,880,467	540	12	22.2	-0.41	47.1
NIMH05-2s1	NIMH05	chrX	67,959,824	67,977,396	17,572	349	19.9	0.30	55.6
NIMH20-1s1	NIMH20	chrX	68,027,878	68,051,613	23,735	426	17.9	0.34	52.1
R085B11-A02-1s1	R085B11	chrX	68,073,913	68,074,853	940	19	20.2	-0.57	37.6
NIMH95-1s1	NIMH95	chrX	68,085,373	68,085,698	325	7	21.5	-0.56	46.8
NIMH20-1s1	NIMH20	chrX	68,750,193	68,753,924	3,731	74	19.8	0.30	52.1
NIMH10-1s1	NIMH10	chrX	69,333,412	69,334,482	1,070	23	21.5	-0.46	37.8
NIMH08-1s1	NIMH08	chrX	69,394,171	69,397,236	3,065	63	20.6	0.25	53.3
NIMH40-1s1	NIMH40	chrX	69,559,670	69,564,248	4,578	94	20.5	0.36	57.1
NIMH05-2s1	NIMH05	chrX	69,580,254	69,587,357	7,103	141	19.9	0.38	57.2
NIMH01-1s1	NIMH01	chrX	69,582,097	69,587,522	5,425	110	20.3	0.22	56.6
NIMH53-2s1	NIMH53	chrX	69,582,247	69,587,132	4,885	99	20.3	0.29	56.6
NIMH24-1s1	NIMH24	chrX	69,582,502	69,587,322	4,820	98	20.3	0.42	56.3
NIMH23-1s1	NIMH23	chrX	69,582,882	69,588,307	5,425	104	19.2	0.30	55.2
NIMH44-1s1	NIMH44	chrX	69,582,957	69,586,782	3,825	78	20.4	0.35	56.9
NIMH34-1s1	NIMH34	chrX	69,583,177	69,587,322	4,145	85	20.5	0.25	56.0
NIMH10-1s1	NIMH10	chrX	69,583,392	69,591,799	8,407	159	18.9	0.40	55.6
NIMH23-1s1	NIMH23	chrX	69,588,342	69,589,644	1,302	25	19.2	-0.41	61.0
NIMH26-1s1	NIMH26	chrX	69,771,733	69,773,308	1,575	34	21.6	0.36	58.1
NIMH96-1s1	NIMH96	chrX	69,771,793	69,773,378	1,585	34	21.5	0.26	58.7
NIMH57-1s1	NIMH57	chrX	70,064,599	70,065,219	620	13	21.0	0.57	51.2
NIMH40-1s1	NIMH40	chrX	70,232,943	70,240,470	7,527	152	20.2	0.24	52.4
NIMH56-1s1	NIMH56	chrX	70,237,391	70,241,239	3,848	78	20.3	0.28	53.1
NIMH01-1s1	NIMH01	chrX	70,240,695	70,241,094	399	8	20.1	0.67	49.3
NIMH57-1s1	NIMH57	chrX	70,240,695	70,241,094	399	8	20.1	0.73	49.3
NIMH93-1s1	NIMH93	chrX	70,268,884	70,269,809	925	19	20.5	-0.42	49.6
NIMH53-2s1	NIMH53	chrX	70,268,989	70,269,809	820	17	20.7	-0.41	48.4
NIMH66-1s1	NIMH66	chrX	70,269,054	70,269,809	755	16	21.2	-0.37	48.3
NIMH67-1s1	NIMH67	chrX	70,269,054	70,269,809	755	16	21.2	-0.38	48.3
NIMH95-1s1	NIMH95	chrX	70,269,054	70,269,809	755	16	21.2	-0.36	48.3
NIMH05-2s1	NIMH05	chrX	70,269,139	70,308,820	39,681	650	16.4	0.25	51.4
NIMH60-1s1	NIMH60	chrX	70,269,194	70,269,809	615	13	21.1	-0.44	47.3
NIMH06-1s1	NIMH06	chrX	70,270,335	70,292,614	22,279	387	17.4	0.26	53.0
NIMH24-1s1	NIMH24	chrX	70,270,933	70,308,748	37,815	625	16.5	0.22	51.5
NIMH08-1s1	NIMH08	chrX	70,281,273	70,282,333	1,060	21	19.8	0.60	61.4
NIMH67-1s1	NIMH67	chrX	70,281,428	70,292,354	10,926	208	19.0	0.28	55.1
NIMH07-1s1	NIMH07	chrX	70,283,385	70,291,824	8,439	158	18.7	0.23	55.3
NIMH34-1s1	NIMH34	chrX	70,283,780	70,291,824	8,044	150	18.6	0.22	55.3
NIMH06-1s1	NIMH06	chrX	70,351,393	70,395,754	44,361	644	14.5	0.21	50.4
NIMH47-1s1	NIMH47	chrX	70,380,938	70,392,200	11,262	226	20.1	0.20	54.7
NIMH35-1s1	NIMH35	chrX	70,556,721	70,558,233	1,512	24	15.9	-0.53	49.6
NIMH23-1s1	NIMH23	chrX	70,714,076	70,714,868	792	18	22.7	-1.05	49.3
NIMH34-1s1	NIMH34	chrX	70,714,076	70,714,868	792	18	22.7	-0.65	49.3
NIMH32-1s1	NIMH32	chrX	70,714,111	70,714,868	757	17	22.5	-1.05	49.8
NIMH33-1s1	NIMH33	chrX	70,714,111	70,714,868	757	17	22.5	-0.92	49.8
NIMH40-1s1	NIMH40	chrX	70,714,111	70,714,868	757	17	22.5	-1.02	49.8
NIMH24-1s1	NIMH24	chrX	70,751,857	70,755,120	3,263	66	20.2	0.37	59.4
NIMH01-1s1	NIMH01	chrX	70,753,802	70,754,720	918	20	21.8	0.37	59.2
NIMH05-2s1	NIMH05	chrX	70,753,927	70,755,005	1,078	23	21.3	0.54	58.8
NIMH05-2s1	NIMH05	chrX	70,795,042	70,809,426	14,384	234	16.3	0.31	56.0
NIMH30-1s2	NIMH30	chrX	70,796,237	70,807,097	10,860	176	16.2	0.41	58.4
NIMH45-1s1	NIMH45	chrX	70,796,237	70,804,362	8,125	136	16.7	0.30	57.8
NIMH32-1s1	NIMH32	chrX	70,796,287	70,802,657	6,370	101	15.9	0.29	57.1
NIMH61-1s1	NIMH61	chrX	70,796,287	70,803,537	7,250	118	16.3	0.30	57.8
NIMH90-1s1	NIMH90	chrX	70,796,467	70,816,402	19,935	269	13.5	0.19	54.4
NIMH60-1s1	NIMH60	chrX	70,796,602	70,803,697	7,095	115	16.2	0.31	58.0

NIMH76-1s1	NIMH76	chrX	70,796,602	70,803,537	6,935	112	16.1	0.33	57.9
NIMH21-1s1	NIMH21	chrX	70,796,687	70,817,797	21,110	290	13.7	0.28	54.4
NIMH23-1s1	NIMH23	chrX	71,002,886	71,012,160	9,274	167	18.0	0.32	56.1
NIMH44-1s1	NIMH44	chrX	71,002,886	71,010,435	7,549	132	17.5	0.34	57.8
NIMH47-1s1	NIMH47	chrX	71,002,886	71,010,575	7,689	135	17.6	0.30	57.7
NIMH38-1s1	NIMH38	chrX	71,002,956	71,010,500	7,544	132	17.5	0.26	57.8
NIMH31-1s2	NIMH31	chrX	71,002,991	71,010,550	7,559	132	17.5	0.39	57.7
NIMH29-1s1	NIMH29	chrX	71,003,031	71,010,550	7,519	131	17.4	0.30	57.8
NIMH75-1s1	NIMH75	chrX	71,003,106	71,010,575	7,469	131	17.5	0.25	57.7
NIMH76-1s1	NIMH76	chrX	71,003,251	71,010,550	7,299	127	17.4	0.33	57.9
NIMH96-1s1	NIMH96	chrX	71,003,251	71,010,390	7,139	124	17.4	0.28	58.0
NIMH07-1s1	NIMH07	chrX	71,076,155	71,077,815	1,660	34	20.5	-0.31	44.3
NIMH23-1s1	NIMH23	chrX	71,076,275	71,077,710	1,435	30	20.9	-0.49	43.7
NIMH25-1s1	NIMH25	chrX	71,076,415	71,077,605	1,190	25	21.0	-0.37	45.7
R085B11-A02-1s1	R085B11	chrX	71,076,415	71,077,845	1,430	30	21.0	-0.47	45.4
NIMH45-1s1	NIMH45	chrX	71,263,902	71,277,588	13,686	253	18.5	0.20	52.6
NIMH95-1s1	NIMH95	chrX	72,508,629	72,509,617	988	21	21.3	0.30	51.6
NIMH26-1s1	NIMH26	chrX	73,241,432	73,242,800	1,368	25	18.3	-0.52	48.8
NIMH35-1s1	NIMH35	chrX	73,241,532	73,242,770	1,238	22	17.8	-0.52	49.4
NIMH26-1s1	NIMH26	chrX	73,490,299	73,492,409	2,110	43	20.4	0.34	57.4
NIMH08-1s1	NIMH08	chrX	73,556,087	73,559,230	3,143	60	19.1	0.29	61.9
NIMH92-1s1	NIMH92	chrX	74,894,569	74,897,094	2,525	51	20.2	0.27	52.9
NIMH53-2s1	NIMH53	chrX	75,056,185	75,058,000	1,815	38	20.9	0.79	50.6
NIMH32-1s1	NIMH32	chrX	76,094,002	76,095,172	1,170	24	20.5	0.44	54.0
NIMH30-1s2	NIMH30	chrX	77,382,294	77,383,534	1,240	27	21.8	0.41	54.5
NIMH95-1s1	NIMH95	chrX	77,408,661	77,410,671	2,010	36	17.9	-1.40	34.7
NIMH96-1s1	NIMH96	chrX	77,418,594	77,419,609	1,015	22	21.7	0.38	47.4
NIMH02-1s1	NIMH02	chrX	77,670,756	77,674,677	3,921	76	19.4	-0.96	39.5
NIMH38-1s1	NIMH38	chrX	79,023,200	79,024,210	1,010	19	18.8	0.36	53.3
NIMH96-1s1	NIMH96	chrX	79,023,200	79,024,165	965	18	18.7	0.33	53.1
NIMH20-1s1	NIMH20	chrX	79,950,500	79,952,087	1,587	31	19.5	0.54	57.6
NIMH44-1s1	NIMH44	chrX	80,090,638	80,091,868	1,230	25	20.3	0.41	51.8
NIMH66-1s1	NIMH66	chrX	80,709,705	80,710,600	895	20	22.3	0.39	52.5
NIMH66-1s1	NIMH66	chrX	80,800,442	80,801,117	675	13	19.3	0.48	51.0
NIMH63-1s1	NIMH63	chrX	81,593,858	81,595,213	1,355	29	21.4	0.36	55.3
NIMH40-1s1	NIMH40	chrX	81,594,003	81,595,038	1,035	22	21.3	0.45	55.4
NIMH92-1s1	NIMH92	chrX	81,594,218	81,595,038	820	18	22.0	0.43	55.1
NIMH17-1s1	NIMH17	chrX	82,010,041	82,038,392	28,351	437	15.4	-0.93	35.6
NIMH18-2s1	NIMH18	chrX	82,086,347	82,094,874	8,527	86	10.1	-1.04	39.0
NIMH92-1s1	NIMH92	chrX	82,772,373	82,773,418	1,045	23	22.0	0.38	54.3
NIMH67-1s1	NIMH67	chrX	83,538,054	83,539,144	1,090	23	21.1	0.50	51.7
NIMH44-1s1	NIMH44	chrX	83,932,043	83,933,368	1,325	28	21.1	0.40	54.6
NIMH92-1s1	NIMH92	chrX	84,249,035	84,250,590	1,555	26	16.7	0.39	51.7
NIMH51-1s1	NIMH51	chrX	84,249,250	84,250,590	1,340	21	15.7	0.39	51.5
NIMH40-1s1	NIMH40	chrX	85,289,146	85,291,026	1,880	37	19.7	0.29	55.8
NIMH75-1s1	NIMH75	chrX	86,344,982	86,345,757	775	17	21.9	0.48	47.8
NIMH100-1s1	NIMH100	chrX	87,546,038	87,547,547	1,509	14	9.3	-0.68	35.5
NIMH55-1s1	NIMH55	chrX	87,546,038	87,547,547	1,509	14	9.3	-0.44	35.5
NIMH54-1s1	NIMH54	chrX	87,546,283	87,547,547	1,264	12	9.5	-0.68	35.6
NIMH100-1s1	NIMH100	chrX	88,187,301	88,194,167	6,866	108	15.7	-1.43	34.5
NIMH53-2s1	NIMH53	chrX	90,605,772	90,614,285	8,513	115	13.5	0.40	35.5
NIMH78-1s1	NIMH78	chrX	93,105,619	93,113,214	7,595	135	17.8	-1.00	36.9
NIMH48-1s1	NIMH48	chrX	93,851,036	93,859,017	7,981	90	11.3	-0.71	37.2
NIMH82-1s1	NIMH82	chrX	94,957,598	94,970,430	12,832	221	17.2	0.39	33.9
NIMH95-1s1	NIMH95	chrX	96,879,072	96,993,242	114,170	1,823	16.0	0.61	39.2
NIMH75-1s1	NIMH75	chrX	97,793,406	97,794,315	909	17	18.7	0.31	40.4
NIMH92-1s1	NIMH92	chrX	97,856,161	97,857,221	1,060	22	20.8	0.36	52.8
NIMH31-1s2	NIMH31	chrX	98,858,607	98,859,887	1,280	27	21.1	0.47	56.8
NIMH66-1s1	NIMH66	chrX	98,858,777	98,859,887	1,110	23	20.7	0.33	57.1
NIMH92-1s1	NIMH92	chrX	98,858,852	98,859,887	1,035	22	21.3	0.37	57.1
NIMH75-1s1	NIMH75	chrX	99,629,579	99,630,514	935	18	19.3	-0.33	42.6
NIMH33-1s1	NIMH33	chrX	100,112,514	100,118,106	5,592	54	9.7	-0.26	49.3
NIMH30-1s2	NIMH30	chrX	100,374,228	100,377,275	3,047	62	20.3	0.33	52.7
NIMH75-1s1	NIMH75	chrX	100,491,293	100,491,608	315	7	22.2	0.61	44.1
NIMH10-1s1	NIMH10	chrX	100,627,528	100,633,446	5,918	113	19.1	0.30	50.0
NIMH05-2s1	NIMH05	chrX	100,628,655	100,636,096	7,441	148	19.9	0.24	53.8
NIMH24-1s1	NIMH24	chrX	100,630,171	100,634,981	4,810	99	20.6	0.36	55.0
NIMH53-2s1	NIMH53	chrX	100,633,491	100,636,131	2,640	53	20.1	0.29	57.9
NIMH25-1s1	NIMH25	chrX	100,635,456	100,636,646	1,190	25	21.0	-0.50	48.7

NIMH23-1s1	NIMH23	chrX	100,635,486	100,636,131	645	13	20.2	-0.98	54.2
NIMH31-1s2	NIMH31	chrX	100,635,486	100,636,131	645	13	20.2	-0.63	54.2
NIMH33-1s1	NIMH33	chrX	100,635,486	100,636,131	645	13	20.2	-0.83	54.2
NIMH44-1s1	NIMH44	chrX	100,635,486	100,636,131	645	13	20.2	-0.70	54.2
NIMH30-1s2	NIMH30	chrX	100,635,556	100,636,316	760	16	21.1	-0.47	53.4
NIMH33-1s1	NIMH33	chrX	101,141,285	101,142,379	1,094	23	21.0	0.40	54.5
NIMH75-1s1	NIMH75	chrX	101,168,068	101,169,103	1,035	23	22.2	0.28	53.7
NIMH20-1s1	NIMH20	chrX	101,266,277	101,268,484	2,207	29	13.1	0.55	57.7
NIMH90-1s1	NIMH90	chrX	101,266,379	101,268,164	1,785	24	13.4	0.36	59.1
NIMH92-1s1	NIMH92	chrX	101,266,379	101,268,394	2,015	27	13.4	0.38	59.2
NIMH20-1s1	NIMH20	chrX	101,791,869	101,798,757	6,888	116	16.8	0.39	53.2
NIMH85-1s1	NIMH85	chrX	102,383,830	102,384,895	1,065	23	21.6	0.46	48.4
NIMH50-1s1	NIMH50	chrX	102,517,126	102,518,995	1,869	29	15.5	0.39	59.1
NIMH34-1s1	NIMH34	chrX	102,827,244	102,829,636	2,392	48	20.1	-0.30	45.7
NIMH31-1s2	NIMH31	chrX	102,827,379	102,828,205	826	18	21.8	-0.47	37.7
NIMH03-1s1	NIMH03	chrX	103,145,207	103,192,016	46,809	648	13.8	0.44	43.5
NIMH17-1s1	NIMH17	chrX	103,145,207	103,191,966	46,759	647	13.8	0.64	43.5
NIMH22-1s1	NIMH22	chrX	103,145,207	103,192,016	46,809	648	13.8	0.49	43.5
NIMH86-1s1	NIMH86	chrX	103,145,207	103,192,016	46,809	648	13.8	0.57	43.5
NIMH45-1s1	NIMH45	chrX	103,215,206	103,217,761	2,555	49	19.2	0.27	56.1
NIMH21-1s1	NIMH21	chrX	103,297,181	103,298,945	1,764	35	19.8	0.42	60.1
NIMH20-1s1	NIMH20	chrX	103,297,341	103,298,691	1,350	28	20.7	0.46	64.2
NIMH06-1s1	NIMH06	chrX	103,385,045	103,388,528	3,483	63	18.1	0.32	54.2
NIMH67-1s1	NIMH67	chrX	103,678,610	103,694,047	15,437	186	12.0	-1.68	37.8
NIMH92-1s1	NIMH92	chrX	103,772,179	103,773,544	1,365	29	21.2	0.34	53.6
NIMH89-1s1	NIMH89	chrX	104,304,142	104,304,767	625	13	20.8	0.45	44.6
NIMH93-1s1	NIMH93	chrX	104,304,172	104,304,612	440	9	20.5	0.55	44.0
NIMH31-1s2	NIMH31	chrX	105,452,848	105,453,913	1,065	23	21.6	0.51	58.8
NIMH30-1s2	NIMH30	chrX	105,452,893	105,454,303	1,410	30	21.3	0.35	55.9
NIMH32-1s1	NIMH32	chrX	105,452,958	105,454,008	1,050	23	21.9	0.47	59.4
NIMH44-1s1	NIMH44	chrX	105,452,958	105,454,208	1,250	27	21.6	0.44	57.5
NIMH69-2s1	NIMH69	chrX	105,660,793	105,663,095	2,302	48	20.9	0.22	47.8
NIMH40-1s1	NIMH40	chrX	105,696,230	105,697,495	1,265	26	20.6	0.40	55.5
NIMH40-1s1	NIMH40	chrX	106,029,797	106,034,197	4,400	90	20.5	0.25	51.1
NIMH05-2s1	NIMH05	chrX	106,954,805	106,957,654	2,849	53	18.6	0.33	57.1
NIMH19-1s1	NIMH19	chrX	106,955,075	106,957,634	2,559	47	18.4	0.39	57.9
NIMH26-1s1	NIMH26	chrX	106,956,814	106,957,489	675	15	22.2	-0.61	47.1
NIMH32-1s1	NIMH32	chrX	106,956,814	106,957,289	475	11	23.2	-0.77	47.8
NIMH34-1s1	NIMH34	chrX	106,956,814	106,957,394	580	13	22.4	-0.71	47.8
NIMH35-1s1	NIMH35	chrX	106,956,814	106,957,289	475	11	23.2	-0.61	47.8
NIMH21-1s1	NIMH21	chrX	107,862,375	107,867,535	5,160	100	19.4	0.40	56.6
NIMH63-1s1	NIMH63	chrX	108,023,544	108,024,794	1,250	27	21.6	0.36	54.7
NIMH32-1s1	NIMH32	chrX	108,859,431	108,864,197	4,766	86	18.0	-0.26	43.8
NIMH35-1s1	NIMH35	chrX	108,859,590	108,864,517	4,927	90	18.3	-0.25	43.7
NIMH30-1s2	NIMH30	chrX	108,863,302	108,863,917	615	13	21.1	-0.72	54.1
NIMH33-1s1	NIMH33	chrX	108,863,447	108,863,997	550	13	23.6	-0.65	47.0
NIMH73-1s1	NIMH73	chrX	109,097,243	109,110,882	13,639	224	16.4	-1.60	39.0
NIMH44-1s1	NIMH44	chrX	109,448,175	109,449,226	1,051	22	20.9	-0.60	43.5
NIMH08-1s1	NIMH08	chrX	109,924,373	109,927,083	2,710	56	20.7	0.27	52.0
NIMH20-1s1	NIMH20	chrX	110,225,483	110,228,547	3,064	62	20.2	0.35	52.4
NIMH28-1s1	NIMH28	chrX	110,235,171	110,236,751	1,580	34	21.5	0.82	40.4
NIMH29-2s1	NIMH29	chrX	110,235,171	110,236,786	1,615	35	21.7	0.43	40.0
NIMH44-1s1	NIMH44	chrX	111,171,799	111,172,864	1,065	17	16.0	-0.45	38.8
NIMH13-1s1	NIMH13	chrX	111,210,953	111,213,888	2,935	55	18.7	0.31	56.0
NIMH19-1s1	NIMH19	chrX	111,211,433	111,213,728	2,295	43	18.7	0.37	57.9
NIMH22-1s1	NIMH22	chrX	111,598,552	111,624,440	25,888	329	12.7	-0.85	36.4
NIMH55-1s1	NIMH55	chrX	111,753,119	111,753,733	614	14	22.8	-0.90	27.8
NIMH63-1s1	NIMH63	chrX	111,753,214	111,753,733	519	12	23.1	-1.38	28.2
NIMH54-1s1	NIMH54	chrX	111,753,259	111,753,733	474	11	23.2	-1.08	28.8
NIMH87-A03-2s2	NIMH87	chrX	111,753,259	111,753,733	474	11	23.2	-0.85	28.8
NIMH29-2s1	NIMH29	chrX	111,942,990	111,943,736	746	16	21.4	0.60	47.8
NIMH27-A03-2s2	NIMH27	chrX	111,943,160	111,943,761	601	14	23.3	0.54	46.8
NIMH33-1s1	NIMH33	chrX	113,165,532	113,166,117	585	12	20.5	-0.62	43.7
NIMH87-A03-2s2	NIMH87	chrX	113,165,642	113,166,012	370	8	21.6	-0.68	41.8
NIMH69-2s1	NIMH69	chrX	113,234,591	113,241,577	6,986	101	14.5	-1.79	38.1
NIMH69-2s1	NIMH69	chrX	113,396,218	113,396,843	625	12	19.2	-0.54	40.6
NIMH21-1s1	NIMH21	chrX	114,108,118	114,304,312	196,194	2,947	15.0	0.80	36.8
NIMH92-1s1	NIMH92	chrX	114,211,174	114,211,944	770	17	22.1	-0.54	39.7
NIMH25-1s1	NIMH25	chrX	114,211,334	114,211,844	510	12	23.5	-0.74	41.4

NIMH69-2s1	NIMH69	chrX	114,211,434	114,212,014	580	13	22.4	-0.96	39.2
NIMH80-1s1	NIMH80	chrX	114,211,434	114,212,014	580	13	22.4	-0.74	39.2
NIMH89-1s1	NIMH89	chrX	114,211,434	114,212,014	580	13	22.4	-0.81	39.2
NIMH100-1s1	NIMH100	chrX	114,211,569	114,211,944	375	9	24.0	-0.83	38.5
NIMH84-1s1	NIMH84	chrX	114,330,245	114,332,885	2,640	50	18.9	0.44	64.6
NIMH76-1s1	NIMH76	chrX	114,330,280	114,333,090	2,810	53	18.9	0.38	64.6
NIMH89-1s1	NIMH89	chrX	114,330,280	114,332,800	2,520	47	18.7	0.35	64.6
NIMH60-1s1	NIMH60	chrX	114,330,390	114,333,225	2,835	53	18.7	0.46	65.3
NIMH90-1s1	NIMH90	chrX	114,330,460	114,333,305	2,845	53	18.6	0.33	65.4
NIMH80-1s1	NIMH80	chrX	114,331,095	114,333,140	2,045	41	20.0	0.36	64.9
NIMH94-1s1	NIMH94	chrX	114,331,095	114,333,140	2,045	41	20.0	0.41	64.9
NIMH33-1s1	NIMH33	chrX	114,331,515	114,333,225	1,710	33	19.3	0.48	65.4
NIMH40-1s1	NIMH40	chrX	114,331,515	114,333,090	1,575	31	19.7	0.57	65.3
NIMH64-1s1	NIMH64	chrX	114,447,000	114,450,751	3,751	56	14.9	0.78	35.7
NIMH38-1s1	NIMH38	chrX	114,448,445	114,455,800	7,355	96	13.1	0.28	38.4
NIMH43-1s1	NIMH43	chrX	114,448,445	114,450,485	2,040	24	11.8	0.90	34.5
NIMH01-1s1	NIMH01	chrX	114,448,780	114,450,485	1,705	23	13.5	0.68	31.2
NIMH08-1s1	NIMH08	chrX	114,448,780	114,450,485	1,705	23	13.5	0.71	31.2
NIMH17-1s1	NIMH17	chrX	114,448,780	114,450,751	1,971	27	13.7	0.77	30.6
NIMH24-1s1	NIMH24	chrX	114,448,780	114,450,485	1,705	23	13.5	0.82	31.2
NIMH27-A03-2s2	NIMH27	chrX	114,448,780	114,450,485	1,705	23	13.5	0.82	31.2
NIMH29-2s1	NIMH29	chrX	114,448,780	114,450,485	1,705	23	13.5	0.93	31.2
NIMH32-1s1	NIMH32	chrX	114,448,780	114,450,485	1,705	23	13.5	0.70	31.2
NIMH34-1s1	NIMH34	chrX	114,448,780	114,450,751	1,971	27	13.7	0.64	30.6
NIMH44-1s1	NIMH44	chrX	114,448,780	114,450,485	1,705	23	13.5	1.01	31.2
NIMH48-1s1	NIMH48	chrX	114,448,780	114,450,485	1,705	23	13.5	1.20	31.2
NIMH50-1s1	NIMH50	chrX	114,448,780	114,450,751	1,971	27	13.7	0.95	30.6
NIMH51-1s1	NIMH51	chrX	114,448,780	114,450,836	2,056	28	13.6	0.78	30.6
NIMH53-2s1	NIMH53	chrX	114,448,780	114,450,485	1,705	23	13.5	1.02	31.2
NIMH54-1s1	NIMH54	chrX	114,448,780	114,450,751	1,971	27	13.7	0.81	30.6
NIMH55-1s1	NIMH55	chrX	114,448,780	114,450,485	1,705	23	13.5	0.84	31.2
NIMH57-1s1	NIMH57	chrX	114,448,780	114,450,485	1,705	23	13.5	0.98	31.2
NIMH60-1s1	NIMH60	chrX	114,448,780	114,450,485	1,705	23	13.5	0.90	31.2
NIMH63-1s1	NIMH63	chrX	114,448,780	114,450,751	1,971	27	13.7	1.06	30.6
NIMH67-1s1	NIMH67	chrX	114,448,780	114,450,485	1,705	23	13.5	0.86	31.2
NIMH69-2s1	NIMH69	chrX	114,448,780	114,450,485	1,705	23	13.5	1.20	31.2
NIMH75-1s1	NIMH75	chrX	114,448,780	114,450,445	1,665	22	13.2	1.04	31.1
NIMH76-1s1	NIMH76	chrX	114,448,780	114,450,485	1,705	23	13.5	1.02	31.2
NIMH80-1s1	NIMH80	chrX	114,448,780	114,450,485	1,705	23	13.5	1.11	31.2
NIMH86-1s2	NIMH86	chrX	114,448,780	114,450,485	1,705	23	13.5	0.73	31.2
NIMH87-A03-2s2	NIMH87	chrX	114,448,780	114,450,696	1,916	26	13.6	0.52	30.3
NIMH88-1s1	NIMH88	chrX	114,448,780	114,450,485	1,705	23	13.5	1.11	31.2
NIMH89-1s1	NIMH89	chrX	114,448,780	114,450,445	1,665	22	13.2	1.09	31.1
NIMH92-1s1	NIMH92	chrX	114,448,780	114,450,485	1,705	23	13.5	0.70	31.2
NIMH04-1s1	NIMH04	chrX	114,448,850	114,450,751	1,901	26	13.7	0.78	30.6
NIMH100-1s1	NIMH100	chrX	114,448,850	114,450,445	1,595	21	13.2	1.17	31.0
NIMH42-1s1	NIMH42	chrX	114,448,850	114,450,485	1,635	22	13.5	0.94	31.1
NIMH61-1s1	NIMH61	chrX	114,448,850	114,450,485	1,635	22	13.5	1.00	31.1
NIMH81-1s1	NIMH81	chrX	114,448,850	114,450,485	1,635	22	13.5	0.91	31.1
NIMH82-1s1	NIMH82	chrX	114,448,850	114,450,485	1,635	22	13.5	0.97	31.1
NIMH84-1s1	NIMH84	chrX	114,448,850	114,450,751	1,901	26	13.7	0.76	30.6
NIMH85-1s1	NIMH85	chrX	114,448,850	114,450,445	1,595	21	13.2	0.96	31.0
NIMH90-1s1	NIMH90	chrX	114,448,850	114,450,485	1,635	22	13.5	1.00	31.1
NIMH93-1s1	NIMH93	chrX	114,448,850	114,450,485	1,635	22	13.5	1.15	31.1
NIMH96-1s1	NIMH96	chrX	114,448,850	114,450,445	1,595	21	13.2	1.12	31.0
NIMH35-1s1	NIMH35	chrX	114,542,598	114,542,714	116	4	34.5	-1.10	29.7
NIMH31-1s1	NIMH31	chrX	114,925,429	114,937,770	12,341	248	20.1	0.24	51.8
NIMH39-1s1	NIMH39	chrX	114,925,594	114,937,685	12,091	243	20.1	0.25	51.8
NIMH49-1s1	NIMH49	chrX	114,928,184	114,937,715	9,531	191	20.0	0.33	52.2
NIMH27-A03-2s2	NIMH27	chrX	114,937,090	114,937,715	625	14	22.4	0.63	48.3
NIMH93-1s1	NIMH93	chrX	114,937,090	114,937,685	595	13	21.8	0.59	48.7
NIMH95-1s1	NIMH95	chrX	114,937,090	114,937,685	595	13	21.8	0.49	48.7
NIMH96-1s1	NIMH96	chrX	115,051,924	115,052,093	169	4	23.7	0.75	33.7
NIMH87-A03-2s2	NIMH87	chrX	115,103,362	115,104,537	1,175	25	21.3	0.31	53.1
NIMH53-2s1	NIMH53	chrX	115,623,901	115,624,841	940	21	22.3	0.39	53.1
NIMH96-1s1	NIMH96	chrX	116,022,711	116,023,328	617	11	17.8	-0.44	34.4
NIMH05-2s1	NIMH05	chrX	116,022,771	116,023,121	350	8	22.9	-0.65	35.6
NIMH40-1s1	NIMH40	chrX	116,087,691	116,089,407	1,716	29	16.9	0.45	54.7
NIMH44-1s1	NIMH44	chrX	116,266,983	116,267,663	680	13	19.1	-0.47	36.1
NIMH87-A03-2s2	NIMH87	chrX	116,322,669	116,323,804	1,135	24	21.1	0.33	54.7

NIMH56-1s1	NIMH56	chrX	116,675,035	116,675,610	575	12	20.9	0.57	39.6
NIMH53-2s1	NIMH53	chrX	117,185,118	117,185,538	420	9	21.4	-0.68	48.3
NIMH01-1s1	NIMH01	chrX	117,511,504	117,513,635	2,131	40	18.8	0.33	51.1
NIMH13-1s1	NIMH13	chrX	117,650,854	117,651,159	305	7	23.0	-1.15	47.3
NIMH20-1s1	NIMH20	chrX	117,841,298	117,847,010	5,712	108	18.9	0.41	52.2
NIMH87-A03-2s2	NIMH87	chrX	117,845,710	117,846,530	820	18	22.0	0.45	55.2
NIMH75-1s1	NIMH75	chrX	117,882,071	117,882,401	330	8	24.2	0.44	50.3
NIMH16-1s1	NIMH16	chrX	117,991,019	117,994,644	3,625	64	17.7	0.34	62.1
NIMH27-A03-2s2	NIMH27	chrX	117,993,789	117,994,019	230	5	21.7	0.90	46.3
NIMH23-1s1	NIMH23	chrX	118,241,004	118,242,158	1,154	21	18.2	-0.77	47.1
NIMH26-1s1	NIMH26	chrX	118,241,004	118,241,509	505	8	15.8	-1.23	58.5
NIMH30-1s2	NIMH30	chrX	118,241,004	118,241,509	505	8	15.8	-1.37	58.5
R085B11-A03-1s1	R085B11	chrX	118,255,144	118,255,534	390	9	23.1	-0.78	51.8
NIMH66-1s1	NIMH66	chrX	118,290,504	118,291,409	905	20	22.1	0.36	58.9
NIMH63-1s1	NIMH63	chrX	118,290,584	118,291,369	785	17	21.7	0.49	59.1
NIMH96-1s1	NIMH96	chrX	118,290,749	118,291,604	855	17	19.9	0.40	63.9
NIMH26-1s1	NIMH26	chrX	118,416,189	118,417,149	960	20	20.8	-0.62	44.6
NIMH32-1s1	NIMH32	chrX	118,416,189	118,417,149	960	20	20.8	-0.64	44.6
NIMH44-1s1	NIMH44	chrX	118,416,189	118,417,149	960	20	20.8	-0.63	44.6
NIMH30-1s2	NIMH30	chrX	118,416,264	118,417,149	885	19	21.5	-0.68	45.3
NIMH38-1s1	NIMH38	chrX	118,416,264	118,417,149	885	19	21.5	-0.54	45.3
NIMH40-1s1	NIMH40	chrX	118,416,264	118,417,149	885	19	21.5	-0.60	45.3
NIMH25-1s1	NIMH25	chrX	118,416,359	118,417,149	790	17	21.5	-0.66	46.4
NIMH31-1s1	NIMH31	chrX	118,416,359	118,417,149	790	17	21.5	-0.61	46.4
NIMH35-1s1	NIMH35	chrX	118,416,359	118,417,149	790	17	21.5	-0.61	46.4
NIMH08-1s1	NIMH08	chrX	118,485,037	118,488,352	3,315	57	17.2	0.34	54.7
NIMH54-1s1	NIMH54	chrX	118,537,640	118,543,287	5,647	70	12.4	0.58	46.3
NIMH57-1s1	NIMH57	chrX	118,537,640	118,542,233	4,593	63	13.7	0.55	46.2
NIMH91-2s1	NIMH91	chrX	118,537,640	118,543,287	5,647	70	12.4	0.69	46.3
NIMH55-1s1	NIMH55	chrX	118,537,765	118,543,287	5,522	68	12.3	0.64	46.6
NIMH30-1s2	NIMH30	chrX	118,709,537	118,711,112	1,575	30	19.0	-0.38	56.0
NIMH32-1s1	NIMH32	chrX	118,709,922	118,711,157	1,235	24	19.4	-0.35	61.5
R085B11-A03-1s1	R085B11	chrX	118,710,563	118,713,743	3,180	63	19.8	0.31	63.9
NIMH32-1s1	NIMH32	chrX	118,711,182	118,713,683	2,501	50	20.0	0.35	62.2
R085B11-A03-1s1	R085B11	chrX	119,326,556	119,329,544	2,988	59	19.7	0.26	57.2
NIMH75-1s1	NIMH75	chrX	120,001,137	120,001,522	385	9	23.4	0.59	37.5
NIMH27-A03-2s2	NIMH27	chrX	120,047,910	120,049,601	1,691	35	20.7	0.26	44.2
NIMH87-A03-2s2	NIMH87	chrX	120,259,772	120,260,452	680	13	19.1	0.47	38.0
NIMH57-1s1	NIMH57	chrX	120,476,163	120,477,746	1,583	34	21.5	0.34	41.5
NIMH29-2s1	NIMH29	chrX	120,476,188	120,477,406	1,218	26	21.3	0.39	43.1
NIMH27-A03-2s2	NIMH27	chrX	120,476,406	120,477,601	1,195	26	21.8	0.49	42.8
NIMH21-1s1	NIMH21	chrX	120,814,894	120,816,694	1,800	37	20.6	0.35	55.0
NIMH53-2s1	NIMH53	chrX	121,132,578	121,142,362	9,784	162	16.6	-1.16	33.0
NIMH67-1s1	NIMH67	chrX	121,400,522	121,403,442	2,920	57	19.5	0.35	52.3
NIMH31-1s1	NIMH31	chrX	121,400,547	121,403,442	2,895	56	19.3	0.26	52.4
NIMH45-1s1	NIMH45	chrX	121,400,547	121,403,502	2,955	57	19.3	0.26	52.1
NIMH76-1s1	NIMH76	chrX	121,400,597	121,403,367	2,770	53	19.1	0.33	52.3
NIMH66-1s1	NIMH66	chrX	121,400,842	121,401,582	740	14	18.9	0.39	51.6
NIMH96-1s1	NIMH96	chrX	121,400,842	121,403,582	2,740	53	19.3	0.31	51.9
NIMH01-1s1	NIMH01	chrX	121,400,947	121,401,881	934	16	17.1	0.36	46.3
NIMH92-1s1	NIMH92	chrX	121,401,277	121,403,502	2,225	42	18.9	0.30	51.6
NIMH69-2s1	NIMH69	chrX	121,906,343	121,907,519	1,176	23	19.6	-0.35	38.8
NIMH53-2s1	NIMH53	chrX	121,906,403	121,907,369	966	19	19.7	-0.43	38.3
NIMH54-1s1	NIMH54	chrX	121,906,483	121,907,149	666	12	18.0	-0.50	37.2
NIMH40-1s1	NIMH40	chrX	121,906,548	121,907,434	886	17	19.2	-0.52	37.9
NIMH21-1s1	NIMH21	chrX	122,521,915	122,525,059	3,144	48	15.3	0.41	50.8
NIMH17-1s1	NIMH17	chrX	122,523,174	122,525,249	2,075	42	20.2	0.46	57.2
NIMH20-1s1	NIMH20	chrX	122,523,174	122,525,955	2,781	44	15.8	0.49	52.0
NIMH54-1s1	NIMH54	chrX	122,778,683	122,780,279	1,596	16	10.0	-0.66	45.9
NIMH67-1s1	NIMH67	chrX	122,778,733	122,780,254	1,521	14	9.2	-0.83	46.2
NIMH90-1s1	NIMH90	chrX	122,778,733	122,780,279	1,546	15	9.7	-0.66	45.9
NIMH75-1s1	NIMH75	chrX	124,281,534	124,284,444	2,910	46	15.8	0.21	57.5
NIMH96-1s1	NIMH96	chrX	124,582,612	124,584,007	1,395	30	21.5	0.30	53.2
NIMH19-1s1	NIMH19	chrX	124,819,774	124,821,879	2,105	41	19.5	-1.11	34.9
NIMH20-1s1	NIMH20	chrX	125,126,155	125,128,960	2,805	56	20.0	0.43	62.0
NIMH87-A03-2s2	NIMH87	chrX	125,494,962	125,495,342	380	8	21.1	-0.63	37.2
NIMH95-1s1	NIMH95	chrX	125,513,234	125,514,094	860	18	20.9	0.42	63.9
NIMH25-1s1	NIMH25	chrX	125,514,154	125,514,924	770	17	22.1	-0.71	65.2
NIMH35-1s1	NIMH35	chrX	125,514,254	125,514,924	670	15	22.4	-0.72	64.2

NIMH40-1s1	NIMH40	chrX	125,514,299	125,515,044	745	17	22.8	-0.75	60.5
NIMH10-1s1	NIMH10	chrX	126,420,769	126,429,852	9,083	100	11.0	0.29	33.3
NIMH08-1s1	NIMH08	chrX	126,425,325	126,429,852	4,527	46	10.2	0.66	31.4
NIMH17-1s1	NIMH17	chrX	126,425,325	126,429,852	4,527	46	10.2	0.64	31.4
NIMH21-1s1	NIMH21	chrX	126,425,325	126,429,852	4,527	46	10.2	0.72	31.4
NIMH01-1s1	NIMH01	chrX	126,425,570	126,429,852	4,282	45	10.5	0.58	31.6
NIMH05-2s1	NIMH05	chrX	126,425,570	126,429,852	4,282	45	10.5	0.79	31.6
NIMH43-1s1	NIMH43	chrX	126,425,570	126,429,852	4,282	45	10.5	0.81	31.6
NIMH48-1s1	NIMH48	chrX	126,425,570	126,429,852	4,282	45	10.5	1.02	31.6
NIMH50-1s1	NIMH50	chrX	126,425,570	126,429,852	4,282	45	10.5	0.94	31.6
NIMH61-1s1	NIMH61	chrX	126,425,570	126,429,852	4,282	45	10.5	0.89	31.6
NIMH66-1s1	NIMH66	chrX	126,425,570	126,429,852	4,282	45	10.5	0.75	31.6
NIMH76-1s1	NIMH76	chrX	126,425,570	126,429,852	4,282	45	10.5	0.92	31.6
NIMH92-1s1	NIMH92	chrX	126,425,570	126,429,852	4,282	45	10.5	0.75	31.6
NIMH94-1s1	NIMH94	chrX	126,425,570	126,429,852	4,282	45	10.5	0.94	31.6
NIMH56-1s1	NIMH56	chrX	126,426,100	126,429,852	3,752	44	11.7	0.97	31.1
NIMH27-A03-2s2	NIMH27	chrX	126,426,125	126,429,852	3,727	43	11.5	0.80	31.2
NIMH96-1s1	NIMH96	chrX	126,426,375	126,429,852	3,477	42	12.1	0.95	31.0
NIMH54-1s1	NIMH54	chrX	126,426,630	126,429,852	3,222	41	12.7	0.78	30.5
NIMH53-2s1	NIMH53	chrX	126,426,740	126,429,852	3,112	40	12.9	0.94	30.6
NIMH75-1s1	NIMH75	chrX	126,426,740	126,429,852	3,112	40	12.9	0.96	30.6
NIMH86-1s2	NIMH86	chrX	126,426,740	126,429,852	3,112	40	12.9	0.80	30.6
NIMH69-2s1	NIMH69	chrX	126,426,830	126,429,852	3,022	39	12.9	1.17	30.3
NIMH95-1s1	NIMH95	chrX	126,426,830	126,429,852	3,022	39	12.9	0.89	30.3
NIMH04-1s1	NIMH04	chrX	126,426,896	126,429,852	2,956	38	12.9	0.84	30.2
NIMH100-1s1	NIMH100	chrX	126,426,896	126,429,852	2,956	38	12.9	1.00	30.2
NIMH57-1s1	NIMH57	chrX	126,426,896	126,429,852	2,956	38	12.9	0.98	30.2
NIMH24-1s1	NIMH24	chrX	126,427,236	126,429,852	2,616	37	14.1	0.81	30.3
NIMH29-2s1	NIMH29	chrX	126,427,236	126,429,852	2,616	37	14.1	0.88	30.3
NIMH44-1s1	NIMH44	chrX	126,427,236	126,429,852	2,616	37	14.1	1.03	30.3
NIMH46-1s1	NIMH46	chrX	126,427,236	126,429,852	2,616	37	14.1	0.84	30.3
NIMH80-1s1	NIMH80	chrX	126,427,236	126,429,852	2,616	37	14.1	1.14	30.3
NIMH82-1s1	NIMH82	chrX	126,427,236	126,429,852	2,616	37	14.1	0.83	30.3
NIMH93-1s1	NIMH93	chrX	126,427,236	126,429,852	2,616	37	14.1	1.02	30.3
NIMH67-1s1	NIMH67	chrX	127,479,117	127,480,277	1,160	23	19.8	0.38	53.0
R085B11-A03-1s1	R085B11	chrX	127,583,976	127,585,656	1,680	34	20.2	0.30	57.1
NIMH92-1s1	NIMH92	chrX	127,618,762	127,619,987	1,225	26	21.2	0.36	49.5
NIMH31-1s1	NIMH31	chrX	127,742,195	127,743,640	1,445	31	21.5	0.35	55.1
R085B11-A03-1s1	R085B11	chrX	127,853,268	127,855,568	2,300	48	20.9	-0.32	40.0
R085B11-A03-1s1	R085B11	chrX	127,959,718	127,960,733	1,015	17	16.7	-0.50	37.8
NIMH44-1s1	NIMH44	chrX	127,959,988	127,960,818	830	17	20.5	-0.55	39.5
R085B11-A03-1s1	R085B11	chrX	128,065,929	128,066,638	709	16	22.6	-0.42	38.5
NIMH92-1s1	NIMH92	chrX	128,148,260	128,155,077	6,817	124	18.2	-1.57	39.6
NIMH06-1s1	NIMH06	chrX	128,148,360	128,155,102	6,742	123	18.2	-1.22	39.5
NIMH05-2s1	NIMH05	chrX	128,328,110	128,329,810	1,700	21	12.4	-1.09	40.2
NIMH50-1s1	NIMH50	chrX	128,328,110	128,329,860	1,750	22	12.6	-1.21	39.7
NIMH54-1s1	NIMH54	chrX	128,328,110	128,329,810	1,700	21	12.4	-1.08	40.2
NIMH55-1s1	NIMH55	chrX	128,328,110	128,329,810	1,700	21	12.4	-0.91	40.2
NIMH60-1s1	NIMH60	chrX	128,328,110	128,329,860	1,750	22	12.6	-0.91	39.7
NIMH64-1s1	NIMH64	chrX	128,328,110	128,329,810	1,700	21	12.4	-1.25	40.2
NIMH66-1s1	NIMH66	chrX	128,328,110	128,329,810	1,700	21	12.4	-0.98	40.2
NIMH67-1s1	NIMH67	chrX	128,328,110	128,329,860	1,750	22	12.6	-1.28	39.7
NIMH90-1s1	NIMH90	chrX	128,328,110	128,329,710	1,600	20	12.5	-1.10	41.2
NIMH88-1s1	NIMH88	chrX	128,328,405	128,329,860	1,455	18	12.4	-1.49	39.3
NIMH05-2s1	NIMH05	chrX	128,608,000	128,615,641	7,641	152	19.9	0.35	55.5
NIMH05-2s1	NIMH05	chrX	128,753,457	128,757,338	3,881	77	19.8	0.40	55.8
NIMH10-1s1	NIMH10	chrX	128,938,256	128,944,781	6,525	118	18.1	0.26	57.9
NIMH20-1s1	NIMH20	chrX	128,938,435	128,944,992	6,557	117	17.8	0.45	58.1
NIMH17-1s1	NIMH17	chrX	128,940,617	128,945,425	4,808	87	18.1	0.44	64.1
NIMH92-1s1	NIMH92	chrX	128,940,617	128,945,227	4,610	83	18.0	0.26	63.8
NIMH96-1s1	NIMH96	chrX	128,941,017	128,945,108	4,091	74	18.1	0.33	65.4
NIMH38-1s1	NIMH38	chrX	128,974,008	128,978,163	4,155	86	20.7	0.29	58.8
NIMH23-1s1	NIMH23	chrX	129,071,518	129,073,247	1,729	29	16.8	-0.29	72.2
NIMH26-1s1	NIMH26	chrX	129,134,946	129,135,581	635	14	22.0	-0.37	42.6
NIMH44-1s1	NIMH44	chrX	129,135,191	129,135,711	520	12	23.1	-0.48	39.2
NIMH66-1s1	NIMH66	chrX	129,490,335	129,494,135	3,800	67	17.6	0.22	57.0
NIMH07-1s1	NIMH07	chrX	129,711,077	129,715,267	4,190	77	18.4	-0.27	41.5
NIMH29-2s1	NIMH29	chrX	129,711,302	129,715,197	3,895	71	18.2	0.33	41.4
NIMH27-A03-2s2	NIMH27	chrX	129,711,347	129,715,677	4,330	75	17.3	0.39	41.0
NIMH44-1s1	NIMH44	chrX	129,711,347	129,715,287	3,940	72	18.3	-0.29	41.4

243

NIMH57-1s1	NIMH57	chrX	129,711,402	129,715,047	3,645	67	18.4	0.27	41.3
NIMH28-1s1	NIMH28	chrX	129,711,442	129,715,197	3,755	68	18.1	0.87	41.3
NIMH23-1s1	NIMH23	chrX	129,713,742	129,715,267	1,525	30	19.7	-0.53	44.0
R085B11-A03-1s1	R085B11	chrX	129,771,750	129,773,040	1,290	26	20.2	0.31	58.2
NIMH30-1s2	NIMH30	chrX	130,139,847	130,140,407	560	12	21.4	-0.65	48.5
NIMH60-1s1	NIMH60	chrX	130,239,122	130,240,995	1,873	35	18.7	0.31	53.9
NIMH29-2s1	NIMH29	chrX	130,239,507	130,243,140	3,633	68	18.7	0.26	50.9
NIMH30-1s2	NIMH30	chrX	130,473,430	130,476,783	3,353	46	13.7	0.28	48.6
NIMH76-1s1	NIMH76	chrX	130,640,914	130,641,821	907	12	13.2	-1.81	45.3
NIMH46-1s1	NIMH46	chrX	130,724,611	130,725,466	855	13	15.2	0.42	51.4
NIMH24-1s1	NIMH24	chrX	130,985,386	130,986,591	1,205	25	20.7	-1.10	48.0
R085B10-1s2	R085B10	chrX	131,757,779	131,796,875	39,096	649	16.6	-0.49	40.8
NIMH38-1s1	NIMH38	chrX	131,766,925	131,769,220	2,295	30	13.1	-0.93	41.7
NIMH90-1s1	NIMH90	chrX	131,766,925	131,769,220	2,295	30	13.1	-1.25	41.7
NIMH05-2s1	NIMH05	chrX	131,766,995	131,769,220	2,225	29	13.0	-1.27	41.4
NIMH38-1s1	NIMH38	chrX	132,514,136	132,515,072	936	20	21.4	-0.40	38.9
R085B11-A03-1s1	R085B11	chrX	132,514,136	132,514,761	625	14	22.4	-0.61	39.8
NIMH35-1s1	NIMH35	chrX	132,840,193	132,841,109	916	18	19.7	-1.86	38.7
NIMH26-1s1	NIMH26	chrX	133,421,463	133,423,144	1,681	31	18.4	-0.58	58.3
NIMH23-1s1	NIMH23	chrX	133,421,538	133,423,144	1,606	30	18.7	-0.67	58.7
NIMH30-1s2	NIMH30	chrX	133,421,538	133,423,144	1,606	30	18.7	-0.61	58.7
NIMH35-1s1	NIMH35	chrX	133,422,003	133,423,144	1,141	24	21.0	-0.55	54.4
NIMH38-1s1	NIMH38	chrX	133,422,164	133,423,144	980	22	22.4	-0.66	50.3
NIMH44-1s1	NIMH44	chrX	133,422,164	133,423,529	1,365	29	21.2	-0.57	48.1
NIMH01-1s1	NIMH01	chrX	133,904,983	133,906,470	1,487	24	16.1	0.54	42.2
NIMH23-1s1	NIMH23	chrX	134,055,097	134,059,687	4,590	75	16.3	0.32	57.9
NIMH95-1s1	NIMH95	chrX	134,055,317	134,058,122	2,805	51	18.2	0.25	57.0
NIMH21-1s1	NIMH21	chrX	134,057,817	134,060,683	2,866	42	14.7	0.44	64.5
NIMH50-1s1	NIMH50	chrX	134,119,690	134,160,090	40,400	552	13.7	0.62	40.3
NIMH32-1s1	NIMH32	chrX	134,305,630	134,306,235	605	14	23.1	-0.50	52.6
NIMH89-1s1	NIMH89	chrX	134,382,794	134,388,717	5,923	98	16.5	0.27	60.1
NIMH87-A03-2s2	NIMH87	chrX	134,384,099	134,388,897	4,798	81	16.9	0.27	58.5
NIMH34-1s1	NIMH34	chrX	134,882,626	134,885,031	2,405	50	20.8	-0.54	56.4
NIMH33-1s1	NIMH33	chrX	134,883,591	134,885,031	1,440	31	21.5	-0.79	59.9
NIMH24-1s1	NIMH24	chrX	134,883,641	134,884,781	1,140	25	21.9	-0.65	63.7
NIMH25-1s1	NIMH25	chrX	134,883,641	134,885,031	1,390	30	21.6	-0.80	60.1
NIMH23-1s1	NIMH23	chrX	134,883,686	134,884,816	1,130	25	22.1	-1.06	63.4
NIMH30-1s2	NIMH30	chrX	134,883,686	134,884,966	1,280	28	21.9	-0.95	61.6
NIMH42-1s1	NIMH42	chrX	134,883,686	134,884,966	1,280	28	21.9	-0.76	61.6
NIMH43-1s1	NIMH43	chrX	134,883,686	134,885,031	1,345	29	21.6	-0.64	60.4
NIMH44-1s1	NIMH44	chrX	134,883,731	134,884,966	1,235	27	21.9	-0.83	61.7
NIMH40-1s1	NIMH40	chrX	134,883,761	134,884,966	1,205	26	21.6	-0.82	61.5
NIMH31-1s1	NIMH31	chrX	134,884,431	134,884,966	535	12	22.4	-0.98	45.8
NIMH46-1s1	NIMH46	chrX	134,987,293	134,990,779	3,486	54	15.5	0.56	42.9
NIMH45-1s1	NIMH45	chrX	135,129,255	135,129,790	535	12	22.4	-1.53	43.0
NIMH29-2s1	NIMH29	chrX	135,382,010	135,382,385	375	8	21.3	0.81	49.2
NIMH45-1s1	NIMH45	chrX	135,939,306	135,941,481	2,175	45	20.7	0.30	51.7
NIMH17-1s1	NIMH17	chrX	135,939,756	135,944,189	4,433	88	19.9	0.37	56.1
NIMH20-1s1	NIMH20	chrX	135,939,856	135,944,189	4,333	86	19.8	0.39	56.4
NIMH67-1s1	NIMH67	chrX	135,939,906	135,941,751	1,845	39	21.1	0.33	56.2
NIMH35-1s1	NIMH35	chrX	135,941,811	135,944,899	3,088	61	19.8	-0.33	54.1
NIMH40-1s1	NIMH40	chrX	135,941,811	135,945,219	3,408	67	19.7	-0.34	53.3
NIMH44-1s1	NIMH44	chrX	135,941,811	135,945,219	3,408	67	19.7	-0.46	53.3
NIMH23-1s1	NIMH23	chrX	135,941,971	135,945,219	3,248	64	19.7	-0.45	53.3
R085B11-A03-1s1	R085B11	chrX	135,941,971	135,943,419	1,448	30	20.7	-0.46	49.6
NIMH34-1s1	NIMH34	chrX	135,941,996	135,943,419	1,423	29	20.4	-0.55	49.7
NIMH25-1s1	NIMH25	chrX	135,942,056	135,943,419	1,363	28	20.5	-0.57	50.0
NIMH30-1s2	NIMH30	chrX	135,942,056	135,945,084	3,028	60	19.8	-0.46	53.7
NIMH32-1s1	NIMH32	chrX	135,942,096	135,943,419	1,323	27	20.4	-0.57	49.9
NIMH33-1s1	NIMH33	chrX	135,942,096	135,943,419	1,323	27	20.4	-0.56	49.9
R085B11-A03-1s1	R085B11	chrX	135,943,464	135,944,189	725	13	17.9	0.36	72.1
R085B11-A03-1s1	R085B11	chrX	135,944,214	135,945,469	1,255	26	20.7	-0.32	45.9
NIMH47-1s1	NIMH47	chrX	136,336,931	136,339,448	2,517	48	19.1	0.38	61.3
NIMH60-1s1	NIMH60	chrX	136,337,136	136,339,239	2,103	43	20.4	0.35	59.5
NIMH25-1s1	NIMH25	chrX	136,351,701	136,352,176	475	11	23.2	-0.94	51.1
NIMH26-1s1	NIMH26	chrX	136,351,736	136,352,176	440	10	22.7	-1.02	48.4
NIMH30-1s2	NIMH30	chrX	136,351,736	136,352,176	440	10	22.7	-0.95	48.4
NIMH31-1s1	NIMH31	chrX	136,351,796	136,352,711	915	18	19.7	-0.59	44.8
NIMH34-1s1	NIMH34	chrX	136,351,796	136,352,176	380	9	23.7	-1.02	43.9

NIMH35-1s1	NIMH35	chrX	136,351,796	136,352,176	380	9	23.7	-1.00	43.9
NIMH38-1s1	NIMH38	chrX	136,351,796	136,352,176	380	9	23.7	-0.90	43.9
NIMH40-1s1	NIMH40	chrX	136,351,796	136,352,176	380	9	23.7	-1.03	43.9
NIMH44-1s1	NIMH44	chrX	136,351,796	136,352,176	380	9	23.7	-1.03	43.9
NIMH32-1s1	NIMH32	chrX	136,351,846	136,352,176	330	8	24.2	-1.13	40.8
NIMH20-1s1	NIMH20	chrX	136,455,746	136,479,863	24,117	401	16.6	0.28	50.7
NIMH17-1s1	NIMH17	chrX	136,472,889	136,479,398	6,509	128	19.7	0.28	57.4
NIMH30-1s2	NIMH30	chrX	136,787,471	136,788,931	1,460	30	20.5	-0.78	37.3
NIMH32-1s1	NIMH32	chrX	136,787,726	136,788,686	960	20	20.8	-0.80	37.3
NIMH23-1s1	NIMH23	chrX	136,787,771	136,788,611	840	17	20.2	-1.09	37.1
NIMH75-1s1	NIMH75	chrX	136,787,771	136,788,686	915	19	20.8	-0.85	37.4
NIMH67-1s1	NIMH67	chrX	136,787,811	136,788,686	875	18	20.6	-0.52	37.7
NIMH67-1s1	NIMH67	chrX	136,958,642	136,959,487	845	18	21.3	0.53	55.8
NIMH20-1s1	NIMH20	chrX	138,840,885	138,843,292	2,407	45	18.7	0.48	66.0
NIMH90-1s1	NIMH90	chrX	139,323,202	139,328,934	5,732	92	16.1	-1.47	36.8
NIMH57-1s1	NIMH57	chrX	139,635,320	139,639,633	4,313	79	18.3	0.32	44.1
NIMH66-1s1	NIMH66	chrX	139,875,780	139,876,215	435	10	23.0	-0.63	42.6
R085B11-A03-1s1	R085B11	chrX	140,073,060	140,075,337	2,277	43	18.9	-0.32	39.3
NIMH80-1s1	NIMH80	chrX	140,304,966	140,306,721	1,755	26	14.8	-0.78	31.3
NIMH40-1s1	NIMH40	chrX	140,324,648	140,325,403	755	15	19.9	0.55	57.6
NIMH92-1s1	NIMH92	chrX	140,867,426	140,868,641	1,215	23	18.9	0.34	55.6
NIMH67-1s1	NIMH67	chrX	141,088,874	141,091,339	2,465	44	17.8	0.33	55.3
NIMH96-1s1	NIMH96	chrX	141,159,238	141,160,823	1,585	30	18.9	0.32	56.4
NIMH40-1s1	NIMH40	chrX	141,159,358	141,160,928	1,570	30	19.1	0.43	55.4
NIMH31-1s1	NIMH31	chrX	141,159,383	141,160,783	1,400	26	18.6	0.44	57.3
NIMH38-1s1	NIMH38	chrX	141,159,383	141,160,658	1,275	23	18.0	0.47	57.3
NIMH30-1s2	NIMH30	chrX	141,159,598	141,160,753	1,155	22	19.0	0.43	57.6
NIMH51-1s1	NIMH51	chrX	141,159,743	141,160,968	1,225	24	19.6	0.45	55.6
NIMH87-A03-2s2	NIMH87	chrX	141,159,743	141,160,823	1,080	21	19.4	0.42	58.4
NIMH30-1s2	NIMH30	chrX	141,242,203	141,244,103	1,900	35	18.4	0.26	52.8
NIMH100-1s1	NIMH100	chrX	141,511,509	141,512,182	673	14	20.8	-0.53	36.2
NIMH25-1s1	NIMH25	chrX	141,896,498	141,897,473	975	20	20.5	-1.47	39.8
NIMH87-A03-2s2	NIMH87	chrX	142,179,513	142,181,007	1,494	28	18.7	-0.49	36.9
NIMH56-1s1	NIMH56	chrX	142,179,778	142,180,887	1,109	21	18.9	-0.50	37.6
NIMH89-1s1	NIMH89	chrX	142,179,853	142,180,952	1,099	20	18.2	-0.34	37.4
NIMH94-1s1	NIMH94	chrX	142,179,853	142,180,952	1,099	20	18.2	-0.55	37.4
NIMH32-1s1	NIMH32	chrX	142,549,885	142,551,290	1,405	28	19.9	-0.51	60.9
NIMH25-1s1	NIMH25	chrX	142,550,345	142,551,155	810	17	21.0	-0.72	61.4
NIMH33-1s1	NIMH33	chrX	142,550,345	142,551,155	810	17	21.0	-0.78	61.4
NIMH23-1s1	NIMH23	chrX	142,550,405	142,551,155	750	16	21.3	-0.93	61.0
NIMH26-1s1	NIMH26	chrX	142,550,405	142,551,290	885	18	20.3	-0.64	58.9
NIMH30-1s2	NIMH30	chrX	142,550,405	142,551,155	750	16	21.3	-0.71	61.0
NIMH38-1s1	NIMH38	chrX	142,550,435	142,551,260	825	16	19.4	-0.55	59.4
NIMH31-1s1	NIMH31	chrX	142,550,580	142,551,155	575	12	20.9	-0.85	57.4
NIMH33-1s1	NIMH33	chrX	142,667,231	142,669,186	1,955	40	20.5	-0.41	46.7
NIMH34-1s1	NIMH34	chrX	142,667,271	142,669,116	1,845	37	20.1	-0.41	47.3
NIMH23-1s1	NIMH23	chrX	142,667,371	142,669,116	1,745	35	20.1	-0.53	48.0
NIMH02-1s1	NIMH02	chrX	143,436,349	143,441,766	5,417	98	18.1	-0.67	38.9
NIMH27-A03-2s2	NIMH27	chrX	143,436,349	143,441,886	5,537	100	18.1	-1.09	38.8
NIMH38-1s1	NIMH38	chrX	143,436,349	143,445,804	9,455	140	14.8	-0.70	37.3
NIMH44-1s1	NIMH44	chrX	143,436,349	143,440,936	4,587	85	18.5	-0.82	39.2
NIMH65-1s1	NIMH65	chrX	143,436,349	143,441,826	5,477	99	18.1	-1.04	38.9
NIMH77-1s1	NIMH77	chrX	143,436,349	143,445,899	9,550	141	14.8	-0.95	37.2
NIMH90-1s1	NIMH90	chrX	143,517,134	143,518,082	948	17	17.9	0.37	47.9
NIMH60-1s1	NIMH60	chrX	143,517,362	143,518,022	660	13	19.7	0.51	48.0
NIMH19-1s1	NIMH19	chrX	144,229,763	144,232,169	2,406	34	14.1	0.80	36.1
NIMH56-1s1	NIMH56	chrX	144,229,938	144,232,169	2,231	31	13.9	0.74	34.8
NIMH45-1s1	NIMH45	chrX	144,519,774	144,521,607	1,833	33	18.0	0.34	60.7
NIMH40-1s1	NIMH40	chrX	144,705,601	144,706,821	1,220	25	20.5	-0.38	46.7
NIMH96-1s1	NIMH96	chrX	145,365,689	145,366,856	1,167	19	16.3	0.42	52.0
NIMH80-1s1	NIMH80	chrX	146,119,632	146,120,626	994	19	19.1	0.41	39.6
NIMH93-1s1	NIMH93	chrX	146,119,632	146,120,626	994	19	19.1	0.42	39.6
NIMH75-1s1	NIMH75	chrX	146,119,862	146,120,626	764	14	18.3	0.41	40.1
NIMH44-1s1	NIMH44	chrX	146,869,511	146,871,161	1,650	27	16.4	0.40	53.0
NIMH61-1s1	NIMH61	chrX	147,738,157	147,738,794	637	12	18.8	-0.90	42.3
NIMH27-A03-2s2	NIMH27	chrX	148,152,917	148,153,923	1,006	22	21.9	0.66	42.7
NIMH79-1s1	NIMH79	chrX	148,153,077	148,153,682	605	14	23.1	0.58	41.9
NIMH38-1s1	NIMH38	chrX	148,518,636	148,522,185	3,549	66	18.6	-0.26	51.8
NIMH25-1s1	NIMH25	chrX	148,520,217	148,521,945	1,728	32	18.5	-0.41	61.3

NIMH30-1s2	NIMH30	chrX	148,520,312	148,522,185	1,873	34	18.2	-0.38	59.3
NIMH53-2s1	NIMH53	chrX	148,540,884	148,542,429	1,545	32	20.7	0.32	58.5
NIMH19-1s1	NIMH19	chrX	148,543,514	148,543,959	445	6	13.5	-1.21	31.0
NIMH31-1s1	NIMH31	chrX	148,646,679	148,653,039	6,360	81	12.7	0.31	56.9
NIMH05-2s1	NIMH05	chrX	148,649,444	148,652,699	3,255	47	14.4	0.46	57.5
NIMH96-1s1	NIMH96	chrX	148,649,544	148,653,299	3,755	57	15.2	0.29	56.5
NIMH53-2s1	NIMH53	chrX	148,650,624	148,652,969	2,345	38	16.2	0.32	59.0
NIMH18-2s1	NIMH18	chrX	148,686,097	148,852,282	166,185	1,578	9.5	-1.10	41.4
NIMH67-1s1	NIMH67	chrX	148,693,656	148,698,741	5,085	101	19.9	0.32	55.6
NIMH67-1s1	NIMH67	chrX	148,747,339	148,748,844	1,505	31	20.6	0.44	54.8
NIMH31-1s1	NIMH31	chrX	148,747,539	148,748,844	1,305	28	21.5	0.52	55.5
NIMH32-1s1	NIMH32	chrX	148,747,739	148,748,789	1,050	23	21.9	0.53	55.2
NIMH93-1s1	NIMH93	chrX	148,857,878	148,859,520	1,642	33	20.1	-0.28	68.2
NIMH10-1s1	NIMH10	chrX	149,276,358	149,285,385	9,027	173	19.2	0.22	59.1
NIMH05-2s1	NIMH05	chrX	149,281,173	149,285,941	4,768	88	18.5	0.32	62.3
NIMH17-1s1	NIMH17	chrX	149,281,238	149,285,605	4,367	81	18.5	0.40	63.4
NIMH30-1s2	NIMH30	chrX	149,421,904	149,428,571	6,667	128	19.2	0.22	53.6
NIMH38-1s1	NIMH38	chrX	149,433,561	149,435,156	1,595	33	20.7	0.41	55.5
NIMH05-2s1	NIMH05	chrX	149,678,535	149,679,445	910	13	14.3	-1.63	54.8
NIMH100-1s1	NIMH100	chrX	149,678,535	149,679,445	910	13	14.3	-1.72	54.8
NIMH25-1s1	NIMH25	chrX	149,678,535	149,679,160	625	11	17.6	-1.31	55.2
NIMH33-1s1	NIMH33	chrX	149,678,535	149,679,160	625	11	17.6	-1.69	55.2
NIMH35-1s1	NIMH35	chrX	149,678,535	149,679,160	625	11	17.6	-1.18	55.2
NIMH88-1s1	NIMH88	chrX	149,678,535	149,679,625	1,090	16	14.7	-1.56	55.3
NIMH93-1s1	NIMH93	chrX	149,678,535	149,679,325	790	12	15.2	-1.70	53.8
NIMH69-2s1	NIMH69	chrX	149,678,585	149,679,325	740	11	14.9	-1.60	54.4
NIMH80-1s1	NIMH80	chrX	149,678,585	149,679,445	860	12	14.0	-1.77	55.3
NIMH89-1s1	NIMH89	chrX	149,678,585	149,679,160	575	10	17.4	-1.75	56.0
NIMH19-1s1	NIMH19	chrX	149,678,940	149,693,844	14,904	280	18.8	0.26	53.7
NIMH27-A03-2s2	NIMH27	chrX	149,772,205	149,773,310	1,105	24	21.7	0.43	47.6
NIMH21-1s1	NIMH21	chrX	149,817,013	149,818,719	1,706	34	19.9	0.48	66.7
R085B11-A03-1s1	R085B11	chrX	149,942,898	149,946,958	4,060	85	20.9	-0.41	37.3
NIMH01-1s1	NIMH01	chrX	150,033,568	150,035,033	1,465	31	21.2	0.43	48.6
NIMH38-1s1	NIMH38	chrX	150,089,543	150,090,288	745	15	20.1	-0.46	43.3
NIMH81-1s1	NIMH81	chrX	150,147,769	150,155,706	7,937	148	18.6	-1.65	42.2
NIMH57-1s1	NIMH57	chrX	150,166,784	150,167,314	530	11	20.8	0.69	44.6
NIMH27-A03-2s2	NIMH27	chrX	150,187,985	150,190,496	2,511	51	20.3	0.31	44.3
NIMH07-1s1	NIMH07	chrX	150,458,117	150,459,512	1,395	18	12.9	-1.63	43.9
NIMH69-2s1	NIMH69	chrX	150,598,190	150,598,655	465	11	23.7	0.55	50.6
NIMH82-1s1	NIMH82	chrX	150,598,270	150,598,655	385	9	23.4	0.62	49.2
NIMH57-1s1	NIMH57	chrX	150,598,330	150,598,655	325	8	24.6	0.77	48.8
NIMH19-1s1	NIMH19	chrX	150,612,380	150,623,387	11,007	196	17.8	0.25	53.2
NIMH23-1s1	NIMH23	chrX	150,614,684	150,615,759	1,075	24	22.3	-0.53	51.5
NIMH25-1s1	NIMH25	chrX	150,614,684	150,615,974	1,290	28	21.7	-0.48	49.8
NIMH26-1s1	NIMH26	chrX	150,614,684	150,616,680	1,996	33	16.5	-0.40	48.6
NIMH34-1s1	NIMH34	chrX	150,614,684	150,615,614	930	21	22.6	-0.58	54.1
NIMH35-1s1	NIMH35	chrX	150,614,684	150,615,669	985	22	22.3	-0.43	53.9
NIMH38-1s1	NIMH38	chrX	150,623,822	150,624,692	870	18	20.7	-0.38	41.9
NIMH53-2s1	NIMH53	chrX	150,831,076	150,843,793	12,717	137	10.8	0.34	59.8
NIMH67-1s1	NIMH67	chrX	150,831,076	150,843,873	12,797	139	10.9	0.37	59.7
NIMH84-1s1	NIMH84	chrX	150,831,076	150,843,498	12,422	132	10.6	0.40	59.9
NIMH40-1s1	NIMH40	chrX	150,831,146	150,845,088	13,942	161	11.5	0.34	58.2
NIMH89-1s1	NIMH89	chrX	150,831,146	150,843,473	12,327	130	10.5	0.29	59.9
NIMH31-1s1	NIMH31	chrX	150,831,221	150,843,908	12,687	138	10.9	0.37	59.8
NIMH12-1s1	NIMH12	chrX	150,831,296	150,843,873	12,577	136	10.8	0.39	59.8
NIMH38-1s1	NIMH38	chrX	150,831,296	150,845,533	14,237	168	11.8	0.25	57.8
NIMH81-1s1	NIMH81	chrX	150,831,296	150,843,123	11,827	121	10.2	0.48	60.2
NIMH43-1s1	NIMH43	chrX	150,831,316	150,843,873	12,557	135	10.8	0.34	59.8
NIMH17-1s1	NIMH17	chrX	150,831,386	150,843,473	12,087	126	10.4	0.39	59.9
NIMH08-1s1	NIMH08	chrX	150,892,265	150,896,392	4,127	77	18.7	0.32	55.1
NIMH63-1s1	NIMH63	chrX	151,033,250	151,040,081	6,831	129	18.9	0.31	56.7
NIMH87-A03-2s2	NIMH87	chrX	151,033,620	151,039,981	6,361	124	19.5	0.27	58.2
NIMH94-1s1	NIMH94	chrX	151,034,600	151,040,231	5,631	111	19.7	0.26	57.9
NIMH66-1s1	NIMH66	chrX	151,034,960	151,039,481	4,521	90	19.9	0.27	58.8
NIMH84-1s1	NIMH84	chrX	151,050,291	151,058,378	8,087	149	18.4	0.21	50.6
NIMH24-1s1	NIMH24	chrX	151,053,858	151,058,268	4,410	89	20.2	0.30	57.0
NIMH13-1s1	NIMH13	chrX	151,054,058	151,058,223	4,165	84	20.2	0.33	57.3
NIMH19-1s1	NIMH19	chrX	151,054,203	151,058,418	4,215	85	20.2	0.34	57.4
NIMH60-1s1	NIMH60	chrX	151,055,933	151,058,268	2,335	47	20.1	0.38	59.2

NIMH05-2s1	NIMH05	chrX	151,057,189	151,058,143	954	21	22.0	0.53	62.6
NIMH44-1s1	NIMH44	chrX	151,370,560	151,371,784	1,224	26	21.2	-0.45	40.5
NIMH27-A03-2s2	NIMH27	chrX	151,562,411	151,564,426	2,015	42	20.8	0.32	42.5
NIMH23-1s1	NIMH23	chrX	151,571,158	151,572,213	1,055	23	21.8	0.53	56.1
NIMH81-1s1	NIMH81	chrX	151,650,913	151,654,360	3,447	53	15.4	0.43	58.1
NIMH100-1s1	NIMH100	chrX	151,651,103	151,653,980	2,877	46	16.0	0.24	58.6
NIMH19-1s1	NIMH19	chrX	151,651,128	151,653,980	2,852	45	15.8	0.43	58.7
NIMH05-2s1	NIMH05	chrX	151,651,538	151,653,980	2,442	38	15.6	0.45	58.5
NIMH89-1s1	NIMH89	chrX	151,652,083	151,654,615	2,532	34	13.4	0.33	58.0
NIMH49-1s1	NIMH49	chrX	151,789,191	151,854,852	65,661	1,208	18.4	0.22	53.1
NIMH10-1s1	NIMH10	chrX	151,790,486	151,857,907	67,421	1,244	18.5	0.20	53.1
NIMH76-1s1	NIMH76	chrX	151,810,012	151,849,216	39,204	729	18.6	0.21	56.1
NIMH16-1s1	NIMH16	chrX	151,813,088	151,829,025	15,937	303	19.0	0.30	57.1
NIMH23-1s1	NIMH23	chrX	151,821,209	151,857,952	36,743	671	18.3	0.29	55.1
NIMH40-1s1	NIMH40	chrX	151,829,400	151,857,907	28,507	525	18.4	0.29	54.6
NIMH32-1s1	NIMH32	chrX	151,830,490	151,840,416	9,926	193	19.4	0.30	57.7
R085B11-A03-1s1	R085B11	chrX	151,831,735	151,838,976	7,241	143	19.7	0.29	58.5
NIMH87-A03-2s2	NIMH87	chrX	151,846,351	151,849,096	2,745	57	20.8	0.35	58.7
NIMH05-2s1	NIMH05	chrX	151,868,084	151,868,494	410	8	19.5	-0.57	51.7
NIMH06-1s1	NIMH06	chrX	151,902,891	151,913,415	10,524	123	11.7	0.30	57.3
NIMH69-2s1	NIMH69	chrX	151,946,060	151,952,575	6,515	114	17.5	0.23	59.1
NIMH100-1s1	NIMH100	chrX	151,949,465	151,952,325	2,860	53	18.5	0.22	60.5
NIMH60-1s1	NIMH60	chrX	151,974,280	151,979,944	5,664	104	18.4	0.32	60.7
NIMH76-1s1	NIMH76	chrX	152,129,668	152,140,580	10,912	207	19.0	0.20	52.5
NIMH17-1s1	NIMH17	chrX	152,136,403	152,140,580	4,177	83	19.9	0.46	59.7
NIMH82-1s1	NIMH82	chrX	152,262,390	152,265,335	2,945	57	19.4	0.26	52.9
NIMH53-2s1	NIMH53	chrX	152,262,600	152,265,525	2,925	57	19.5	0.32	53.1
NIMH95-1s1	NIMH95	chrX	152,263,015	152,265,560	2,545	50	19.6	0.38	53.0
NIMH80-1s1	NIMH80	chrX	152,263,130	152,265,270	2,140	42	19.6	0.40	52.5
NIMH56-1s1	NIMH56	chrX	152,264,255	152,265,270	1,015	21	20.7	0.40	55.0
NIMH92-1s1	NIMH92	chrX	152,269,585	152,270,080	495	9	18.2	-0.54	45.1
NIMH96-1s1	NIMH96	chrX	152,269,730	152,270,175	445	7	15.7	-0.72	43.6
NIMH26-1s1	NIMH26	chrX	152,271,504	152,278,265	6,761	87	12.9	0.24	50.7
NIMH14-1s1	NIMH14	chrX	152,291,280	153,544,603	1,253,323	16,299	13.0	0.20	53.6
NIMH09-1s1	NIMH09	chrX	152,293,929	153,553,301	1,259,372	16,292	12.9	0.23	53.6
NIMH15-1s1	NIMH15	chrX	152,293,929	152,942,034	648,105	10,427	16.1	0.29	55.4
NIMH05-2s1	NIMH05	chrX	152,302,579	153,371,776	1,069,197	10,190	9.5	0.31	54.0
NIMH22-1s1	NIMH22	chrX	152,305,682	152,899,701	594,019	9,678	16.3	0.27	55.7
NIMH70-2s1	NIMH70	chrX	152,305,857	152,893,525	587,668	9,650	16.4	0.17	55.7
NIMH12-1s1	NIMH12	chrX	152,305,917	152,948,952	643,035	10,362	16.1	0.27	55.4
NIMH04-1s1	NIMH04	chrX	152,306,027	152,969,114	663,087	10,683	16.1	0.25	55.2
NIMH19-1s1	NIMH19	chrX	152,310,872	152,982,600	671,728	10,846	16.1	0.25	55.1
NIMH17-1s1	NIMH17	chrX	152,358,146	152,664,254	306,108	5,139	16.8	0.30	56.4
NIMH16-1s1	NIMH16	chrX	152,379,378	152,742,303	362,925	5,983	16.5	0.31	56.6
NIMH10-1s1	NIMH10	chrX	152,384,882	152,807,186	422,304	6,894	16.3	0.26	56.1
NIMH20-1s1	NIMH20	chrX	152,541,535	153,018,526	476,991	7,238	15.2	0.33	53.9
NIMH21-1s1	NIMH21	chrX	152,541,770	153,014,725	472,955	7,167	15.2	0.30	53.9
NIMH84-1s1	NIMH84	chrX	152,551,110	152,941,984	390,874	5,935	15.2	0.21	55.3
NIMH26-1s1	NIMH26	chrX	152,306,027	152,312,552	6,525	74	11.3	0.32	57.2
NIMH24-1s1	NIMH24	chrX	152,306,027	152,317,901	11,874	167	14.1	0.29	57.3
NIMH57-1s1	NIMH57	chrX	152,355,657	152,357,672	2,015	27	13.4	-0.32	45.0
NIMH60-1s1	NIMH60	chrX	152,305,757	152,496,312	190,555	3,413	17.9	0.27	57.0
NIMH20-1s1	NIMH20	chrX	152,320,275	152,531,936	211,661	3,790	17.9	0.39	56.2
NIMH26-1s1	NIMH26	chrX	152,358,806	152,502,824	144,018	2,683	18.6	0.25	58.4
NIMH08-1s1	NIMH08	chrX	152,358,926	152,478,463	119,537	2,216	18.5	0.29	58.9
NIMH45-1s1	NIMH45	chrX	152,359,306	152,510,297	150,991	2,774	18.4	0.25	58.1
NIMH67-1s1	NIMH67	chrX	152,363,927	152,522,269	158,342	2,902	18.3	0.26	58.0
NIMH23-1s1	NIMH23	chrX	152,364,082	152,561,864	197,782	3,562	18.0	0.27	57.0
NIMH39-1s1	NIMH39	chrX	152,366,252	152,500,629	134,377	2,515	18.7	0.28	58.6
NIMH50-1s1	NIMH50	chrX	152,369,801	152,486,224	116,423	2,177	18.7	0.26	59.0
NIMH21-1s1	NIMH21	chrX	152,372,679	152,533,266	160,587	2,937	18.3	0.38	57.7
NIMH53-2s1	NIMH53	chrX	152,373,314	152,428,828	55,514	1,039	18.7	0.21	60.4
NIMH61-1s1	NIMH61	chrX	152,373,524	152,463,286	89,762	1,675	18.7	0.32	59.9
NIMH32-1s1	NIMH32	chrX	152,376,007	152,467,389	91,382	1,701	18.6	0.28	59.9
NIMH84-1s1	NIMH84	chrX	152,376,252	152,496,312	120,060	2,255	18.8	0.27	59.2
NIMH48-1s1	NIMH48	chrX	152,379,378	152,500,774	121,396	2,291	18.9	0.29	59.3
NIMH81-1s1	NIMH81	chrX	152,380,470	152,505,239	124,769	2,331	18.7	0.30	59.0
NIMH76-1s1	NIMH76	chrX	152,380,945	152,488,524	107,579	2,019	18.8	0.27	59.8
NIMH24-1s1	NIMH24	chrX	152,381,015	152,512,517	131,502	2,441	18.6	0.33	58.8

NIMH49-1s1	NIMH49	chrX	152,382,595	152,487,429	104,834	1,968	18.8	0.34	59.8
NIMH01-1s1	NIMH01	chrX	152,384,707	152,488,750	104,043	1,948	18.7	0.24	59.9
NIMH90-1s1	NIMH90	chrX	152,386,817	152,435,018	48,201	881	18.3	0.24	61.5
NIMH89-1s1	NIMH89	chrX	152,388,662	152,468,759	80,097	1,492	18.6	0.27	60.8
NIMH40-1s1	NIMH40	chrX	152,389,076	152,505,239	116,163	2,163	18.6	0.31	59.1
NIMH33-1s1	NIMH33	chrX	152,389,525	152,518,694	129,169	2,376	18.4	0.26	58.7
NIMH31-1s1	NIMH31	chrX	152,390,050	152,420,552	30,502	565	18.5	0.30	61.8
NIMH06-1s1	NIMH06	chrX	152,390,120	152,500,544	110,424	2,077	18.8	0.28	59.5
NIMH96-1s1	NIMH96	chrX	152,391,499	152,495,352	103,853	1,951	18.8	0.25	59.8
NIMH95-1s1	NIMH95	chrX	152,391,739	152,467,389	75,650	1,406	18.6	0.23	60.8
NIMH30-1s2	NIMH30	chrX	152,394,634	152,518,184	123,550	2,270	18.4	0.22	58.6
NIMH47-1s1	NIMH47	chrX	152,408,136	152,477,348	69,212	1,291	18.7	0.32	59.8
NIMH63-1s1	NIMH63	chrX	152,409,671	152,500,774	91,103	1,712	18.8	0.27	58.9
NIMH69-2s1	NIMH69	chrX	152,415,276	152,452,150	36,874	672	18.2	0.23	59.1
NIMH33-1s1	NIMH33	chrX	152,389,076	152,389,476	400	8	20.0	-0.51	72.4
NIMH30-1s2	NIMH30	chrX	152,518,279	152,518,739	460	11	23.9	-0.47	48.0
NIMH66-1s1	NIMH66	chrX	152,531,936	152,532,326	390	9	23.1	-0.69	48.2
NIMH94-1s1	NIMH94	chrX	152,541,600	152,560,721	19,121	346	18.1	0.24	58.3
NIMH95-1s1	NIMH95	chrX	152,541,600	152,560,721	19,121	346	18.1	0.22	58.3
NIMH31-1s1	NIMH31	chrX	152,541,600	152,564,158	22,558	409	18.1	0.22	59.8
NIMH24-1s1	NIMH24	chrX	152,541,655	152,560,721	19,066	345	18.1	0.32	58.2
NIMH26-1s1	NIMH26	chrX	152,541,770	152,560,776	19,006	343	18.0	0.27	58.3
NIMH90-1s1	NIMH90	chrX	152,541,770	152,560,776	19,006	343	18.0	0.24	58.3
NIMH60-1s1	NIMH60	chrX	152,541,840	152,572,226	30,386	547	18.0	0.30	59.3
NIMH38-1s1	NIMH38	chrX	152,557,416	152,560,721	3,305	69	20.9	0.31	60.4
NIMH33-1s1	NIMH33	chrX	152,560,776	152,561,819	1,043	22	21.1	-0.42	70.9
NIMH26-1s1	NIMH26	chrX	152,560,851	152,561,479	628	14	22.3	-0.51	74.5
NIMH08-1s1	NIMH08	chrX	152,572,081	152,572,894	813	14	17.2	-0.35	33.2
NIMH44-1s1	NIMH44	chrX	152,572,081	152,572,999	918	16	17.4	-0.47	35.0
NIMH47-1s1	NIMH47	chrX	152,585,682	152,952,822	367,140	5,581	15.2	0.26	55.4
NIMH24-1s1	NIMH24	chrX	152,602,949	152,934,894	331,945	4,996	15.1	0.26	55.4
NIMH39-1s1	NIMH39	chrX	152,603,044	152,735,519	132,475	1,992	15.0	0.24	57.5
NIMH67-1s1	NIMH67	chrX	152,605,699	152,742,303	136,604	2,035	14.9	0.27	57.1
NIMH56-1s1	NIMH56	chrX	152,585,515	152,591,432	5,917	114	19.3	0.23	61.3
NIMH90-1s1	NIMH90	chrX	152,585,515	152,591,897	6,382	123	19.3	0.26	61.5
NIMH33-1s1	NIMH33	chrX	152,585,682	152,591,432	5,750	112	19.5	0.36	62.0
NIMH34-1s1	NIMH34	chrX	152,592,813	152,593,488	675	15	22.2	-0.49	58.3
NIMH40-1s1	NIMH40	chrX	152,602,949	152,627,475	24,526	333	13.6	0.41	61.4
NIMH96-1s1	NIMH96	chrX	152,603,044	152,613,915	10,871	159	14.6	0.28	65.9
NIMH95-1s1	NIMH95	chrX	152,603,434	152,629,535	26,101	355	13.6	0.31	61.1
NIMH54-1s1	NIMH54	chrX	152,604,004	152,616,167	12,163	164	13.5	0.21	64.2
NIMH26-1s1	NIMH26	chrX	152,608,390	152,633,215	24,825	324	13.1	0.37	57.8
NIMH32-1s1	NIMH32	chrX	152,616,167	152,633,215	17,048	223	13.1	0.31	56.2
NIMH96-1s1	NIMH96	chrX	152,648,483	152,650,876	2,393	32	13.4	-0.23	44.9
NIMH94-1s1	NIMH94	chrX	152,648,880	152,650,761	1,881	28	14.9	-0.21	43.5
NIMH53-2s1	NIMH53	chrX	152,649,628	152,650,761	1,133	19	16.8	-0.30	40.8
NIMH31-1s1	NIMH31	chrX	152,650,716	152,731,368	80,652	1,252	15.5	0.22	57.5
NIMH96-1s1	NIMH96	chrX	152,650,901	152,723,248	72,347	1,106	15.3	0.23	57.5
NIMH45-1s1	NIMH45	chrX	152,650,971	152,737,217	86,246	1,333	15.5	0.25	57.2
NIMH63-1s1	NIMH63	chrX	152,651,021	152,664,159	13,138	183	13.9	0.28	58.8
NIMH60-1s1	NIMH60	chrX	152,651,161	152,731,743	80,582	1,252	15.5	0.31	57.5
NIMH43-1s1	NIMH43	chrX	152,652,826	152,731,188	78,362	1,211	15.5	0.28	57.5
NIMH40-1s1	NIMH40	chrX	152,654,809	152,663,549	8,740	118	13.5	0.46	61.3
R085B11-A03-1s1	R085B11	chrX	152,655,094	152,664,254	9,160	114	12.4	0.33	61.2
NIMH32-1s1	NIMH32	chrX	152,656,484	152,664,792	8,308	88	10.6	0.38	60.3
NIMH50-1s1	NIMH50	chrX	152,664,887	152,724,158	59,271	936	15.8	0.28	57.4
NIMH63-1s1	NIMH63	chrX	152,676,387	152,750,355	73,968	1,250	16.9	0.27	57.7
NIMH66-1s1	NIMH66	chrX	152,680,498	152,720,628	40,130	738	18.4	0.22	58.2
NIMH76-1s1	NIMH76	chrX	152,680,733	152,735,519	54,786	984	18.0	0.28	58.5
NIMH17-1s1	NIMH17	chrX	152,680,733	152,937,261	256,528	4,026	15.7	0.31	59.3
R085B11-A03-1s1	R085B11	chrX	152,685,659	152,731,333	45,674	830	18.2	0.29	55.4
NIMH51-1s1	NIMH51	chrX	152,690,369	152,727,618	37,249	684	18.4	0.25	56.6
NIMH49-1s1	NIMH49	chrX	152,772,248	152,923,505	151,257	2,352	15.5	0.30	55.5
NIMH69-2s1	NIMH69	chrX	152,773,478	152,807,296	33,818	596	17.6	0.29	57.7
NIMH63-1s1	NIMH63	chrX	152,774,348	152,805,016	30,668	546	17.8	0.29	58.0
NIMH76-1s1	NIMH76	chrX	152,774,918	152,899,606	124,688	2,032	16.3	0.24	57.4
NIMH40-1s1	NIMH40	chrX	152,775,018	152,800,959	25,941	457	17.6	0.40	57.4
NIMH45-1s1	NIMH45	chrX	152,775,088	152,926,334	151,246	2,337	15.5	0.25	57.4
NIMH48-1s1	NIMH48	chrX	152,775,648	152,803,199	27,551	490	17.8	0.32	57.8
NIMH43-1s1	NIMH43	chrX	152,778,483	152,793,080	14,597	273	18.7	0.38	59.1

248

NIMH31-1s1	NIMH31	chrX	152,779,722	152,893,525	113,803	1,944	17.1	0.24	55.5
NIMH60-1s1	NIMH60	chrX	152,779,832	152,807,186	27,354	503	18.4	0.38	59.7
NIMH61-1s1	NIMH61	chrX	152,780,012	152,892,600	112,588	1,921	17.1	0.30	56.7
NIMH81-1s1	NIMH81	chrX	152,780,976	152,793,310	12,334	242	19.6	0.36	60.3
NIMH39-1s1	NIMH39	chrX	152,780,996	152,807,131	26,135	480	18.4	0.34	59.8
NIMH94-1s1	NIMH94	chrX	152,783,267	152,804,384	21,117	389	18.4	0.31	59.1
NIMH16-1s1	NIMH16	chrX	152,784,002	152,926,134	142,132	2,187	15.4	0.32	55.8
NIMH23-1s1	NIMH23	chrX	152,784,287	152,809,202	24,915	435	17.5	0.38	58.4
NIMH95-1s1	NIMH95	chrX	152,786,155	152,800,002	13,847	254	18.3	0.28	58.8
NIMH01-1s1	NIMH01	chrX	152,787,180	152,807,046	19,866	359	18.1	0.27	59.9
NIMH53-2s1	NIMH53	chrX	152,789,320	152,800,142	10,822	197	18.2	0.24	58.0
NIMH92-1s1	NIMH92	chrX	152,793,215	152,811,737	18,522	299	16.1	0.23	56.6
NIMH67-1s1	NIMH67	chrX	152,793,290	152,884,425	91,135	1,518	16.7	0.27	55.0
NIMH75-1s1	NIMH75	chrX	152,793,385	152,804,454	11,069	195	17.6	0.25	58.5
NIMH96-1s1	NIMH96	chrX	152,793,460	152,811,642	18,182	292	16.1	0.25	56.6
NIMH87-A03-2s2	NIMH87	chrX	152,794,560	152,807,321	12,761	222	17.4	0.25	59.4
NIMH01-1s1	NIMH01	chrX	152,820,455	152,889,681	69,226	1,204	17.4	0.24	58.9
NIMH96-1s1	NIMH96	chrX	152,820,495	152,829,786	9,291	179	19.3	0.33	63.5
NIMH10-1s1	NIMH10	chrX	152,820,570	152,951,167	130,597	2,024	15.5	0.28	62.3
NIMH60-1s1	NIMH60	chrX	152,820,570	152,886,154	65,584	1,136	17.3	0.33	61.1
NIMH07-1s1	NIMH07	chrX	152,820,570	152,890,330	69,760	1,213	17.4	0.26	61.6
NIMH40-1s1	NIMH40	chrX	152,820,650	152,886,199	65,549	1,136	17.3	0.30	62.0
NIMH63-1s1	NIMH63	chrX	152,820,650	152,830,174	9,524	182	19.1	0.35	63.0
NIMH48-1s1	NIMH48	chrX	152,820,810	152,829,761	8,951	173	19.3	0.39	63.6
NIMH50-1s1	NIMH50	chrX	152,820,885	152,829,786	8,901	172	19.3	0.40	63.7
NIMH69-2s1	NIMH69	chrX	152,821,820	152,890,200	68,380	1,187	17.4	0.25	58.3
NIMH06-1s1	NIMH06	chrX	152,828,921	152,959,283	130,362	1,979	15.2	0.27	58.4
NIMH30-1s2	NIMH30	chrX	152,848,103	152,889,925	41,822	737	17.6	0.26	58.4
NIMH26-1s1	NIMH26	chrX	152,848,128	152,892,035	43,907	772	17.6	0.26	58.4
NIMH90-1s1	NIMH90	chrX	152,864,867	152,898,157	33,290	568	17.1	0.24	58.4
NIMH23-1s1	NIMH23	chrX	152,866,352	152,890,445	24,093	467	19.4	0.38	55.7
NIMH30-1s2	NIMH30	chrX	152,889,950	152,892,060	2,110	37	17.5	-0.41	65.5
NIMH23-1s1	NIMH23	chrX	152,890,475	152,891,960	1,485	27	18.2	-0.59	67.3
NIMH33-1s1	NIMH33	chrX	153,014,407	153,017,466	3,059	58	19.0	-0.29	57.1
NIMH44-1s1	NIMH44	chrX	153,016,546	153,017,436	890	19	21.3	-0.53	50.6
NIMH53-2s1	NIMH53	chrX	152,826,496	152,826,811	315	8	25.4	0.77	59.7
NIMH33-1s1	NIMH33	chrX	152,848,413	152,851,298	2,885	58	20.1	0.41	59.3
NIMH69-2s1	NIMH69	chrX	153,210,565	153,365,011	154,446	2,150	13.9	0.23	57.6
NIMH19-1s1	NIMH19	chrX	153,211,532	153,537,538	326,006	3,471	10.6	0.29	54.8
NIMH21-1s1	NIMH21	chrX	153,213,307	153,417,824	204,517	2,807	13.7	0.37	56.6
NIMH22-1s1	NIMH22	chrX	153,213,307	153,429,930	216,623	2,973	13.7	0.30	56.5
NIMH08-1s1	NIMH08	chrX	153,213,447	153,354,185	140,738	1,947	13.8	0.29	58.2
NIMH20-1s1	NIMH20	chrX	153,213,572	153,537,638	324,066	3,439	10.6	0.40	54.9
NIMH43-1s1	NIMH43	chrX	153,213,597	153,354,435	140,838	1,949	13.8	0.27	58.3
NIMH12-1s1	NIMH12	chrX	153,213,597	153,369,728	156,131	2,185	14.0	0.34	58.0
NIMH50-1s1	NIMH50	chrX	153,213,597	153,416,769	203,172	2,781	13.7	0.26	56.6
NIMH51-1s1	NIMH51	chrX	153,213,597	153,257,756	44,159	607	13.7	0.24	60.5
NIMH95-1s1	NIMH95	chrX	153,213,597	153,266,406	52,809	763	14.4	0.24	60.1
NIMH61-1s1	NIMH61	chrX	153,213,642	153,364,341	150,699	2,081	13.8	0.31	57.9
NIMH70-2s1	NIMH70	chrX	153,213,642	153,417,684	204,042	2,797	13.7	0.23	56.6
NIMH47-1s1	NIMH47	chrX	153,213,642	153,427,651	214,009	2,922	13.7	0.29	56.4
NIMH49-1s1	NIMH49	chrX	153,213,642	153,429,785	216,143	2,963	13.7	0.31	56.5
NIMH76-1s1	NIMH76	chrX	153,213,697	153,394,781	181,084	2,511	13.9	0.23	57.1
NIMH60-1s1	NIMH60	chrX	153,213,697	153,426,460	212,763	2,903	13.6	0.31	56.4
NIMH26-1s1	NIMH26	chrX	153,213,697	153,259,031	45,334	632	13.9	0.31	60.3
NIMH89-1s1	NIMH89	chrX	153,213,697	153,266,361	52,664	760	14.4	0.27	60.2
NIMH10-1s1	NIMH10	chrX	153,215,322	153,477,589	262,267	3,094	11.8	0.28	55.7
NIMH06-1s1	NIMH06	chrX	153,215,817	153,429,835	214,018	2,919	13.6	0.27	56.5
NIMH67-1s1	NIMH67	chrX	153,215,847	153,364,636	148,789	2,041	13.7	0.27	57.9
NIMH63-1s1	NIMH63	chrX	153,216,327	153,424,866	208,539	2,835	13.6	0.26	56.5
NIMH13-1s1	NIMH13	chrX	153,230,615	153,417,579	186,964	2,682	14.3	0.23	56.5
NIMH24-1s1	NIMH24	chrX	153,231,380	153,481,115	249,735	3,050	12.2	0.27	55.3
NIMH39-1s1	NIMH39	chrX	153,234,970	153,427,286	192,316	2,715	14.1	0.24	56.1
NIMH84-1s1	NIMH84	chrX	153,236,245	153,432,835	196,590	2,766	14.1	0.22	56.0
NIMH23-1s1	NIMH23	chrX	153,236,245	153,259,076	22,831	408	17.9	0.28	61.2
NIMH17-1s1	NIMH17	chrX	153,237,660	153,493,933	256,273	3,093	12.1	0.32	54.7
NIMH46-1s1	NIMH46	chrX	153,237,730	153,247,253	9,523	170	17.9	0.32	60.3
NIMH04-1s1	NIMH04	chrX	153,247,063	153,429,565	182,502	2,535	13.9	0.27	55.9
NIMH16-1s1	NIMH16	chrX	153,250,328	153,417,774	167,446	2,313	13.8	0.33	55.9
NIMH23-1s1	NIMH23	chrX	153,259,447	153,261,601	2,154	42	19.5	-0.24	69.3

249

NIMH26-1s1	NIMH26	chrX	153,311,958	153,324,825	12,867	245	19.0	0.30	58.8
NIMH30-1s2	NIMH30	chrX	153,312,068	153,317,608	5,540	113	20.4	0.35	57.4
NIMH53-2s1	NIMH53	chrX	153,312,758	153,317,588	4,830	98	20.3	0.27	57.9
NIMH53-2s1	NIMH53	chrX	153,317,608	153,318,088	480	10	20.8	-0.34	51.0
NIMH30-1s2	NIMH30	chrX	153,317,723	153,318,774	1,051	18	17.1	-0.33	69.7
NIMH92-1s1	NIMH92	chrX	153,407,872	153,417,404	9,532	145	15.2	0.26	57.5
NIMH45-1s1	NIMH45	chrX	153,430,316	153,431,955	1,639	20	12.2	-0.23	42.3
NIMH57-1s1	NIMH57	chrX	154,044,835	154,056,820	11,985	177	14.8	-1.32	34.5
NIMH77-1s1	NIMH77	chrX	154,044,835	154,057,130	12,295	180	14.6	-1.30	34.6
NIMH09-1s1	NIMH09	chrX	154,051,356	154,071,868	20,512	234	11.4	-0.59	36.6
NIMH77-1s1	NIMH77	chrX	154,067,596	154,071,023	3,427	59	17.2	-1.27	35.0
NIMH83-1s1	NIMH83	chrX	154,208,939	154,219,124	10,185	144	14.1	0.51	35.5
NIMH46-1s1	NIMH46	chrX	154,425,197	154,431,258	6,061	91	15.0	-1.19	36.6
NIMH57-1s1	NIMH57	chrX	154,573,654	154,574,880	1,226	16	13.1	-0.59	30.6
NIMH90-1s1	NIMH90	chrX	154,573,654	154,574,880	1,226	16	13.1	-0.69	30.6

Table A.5 - Sequenced Breakpoints Identified in the Literature

Author	Chr	Start	End	Type	Seq1	Seq2	ID
Conrad	3	14,153,187	14,153,863	del	gcccag		CNVR1301.1
Conrad	3	15,820,177	15,821,743	del	ATG		CNVR1306.1
Korbel	3	22,067,362	22,072,811	del	AAAAATTCAATGTG		NA18505_del 498
Conrad	3	25,005,891	25,007,254	del		GGTGTAC	CNVR1326.1
Park	3	26,425,973	26,427,303	del	?		5_AK8
Conrad	3	26,425,974	26,427,303	del	t		CNVR1329.1
Conrad	3	28,162,038	28,172,790	del		CTGA	CNVR1334.1
Vissers	3	29176135(62)	30008810(37)	del	GTGGCTCACGCCTGTAATCCCAGCACT		del 27
Korbel	3	32,077,056	32,082,888	del	G		NA15510_del 281
Korbel	3	32,077,056	32,082,888	del	G		NA18505_del 518
Conrad	3	32,077,056	32,082,889	del	CC		CNVR1341.1
Kim	3	47,465,673	47,468,445	del	?		
Conrad	3	47,465,673	47,468,447	del	gc		CNVR1375.1
Korbel	3	56,582,769	56,596,625	del			NA15510_del 294
Kim	3	62,639,438	62,670,706	del	?		
Korbel	3	68,822,372	68,830,549	del			NA18505_del 489
Korbel	3	68,822,372	68,830,549	del			NA15510_del 283
Korbel	3	74,230,279	74,237,488	del	CTCT		NA18505_del 500
Park	3	78,862,108	78,862,409	del	?		6_NA18968
Lam	3	81,008,005	81,009,297	del	?		Wheeler
Conrad	3	84,187,632	84,190,074	del			CNVR1447.1
Korbel	3	89,592,654	89,598,697	del	AAAAAAGAGAGACAG		NA15510_del 296
Conrad	3	100,690,114	100,699,349	dup	AC		CNVR1468
Conrad	3	101,111,472	101,112,321	del		t	CNVR1469.1
Conrad	3	105,760,890	105,761,737	del	*A*		CNVR1480.1
Conrad	3	105,760,890	105,761,737	del	a		CNVR1480.1
Conrad	3	108,520,647	108,523,025	del	gat		CNVR1485.1
Korbel	3	109,202,507	109,223,395	del			NA15510_inv 525
Kim	3	121,644,332	121,647,642	del	?		
Conrad	3	131,614,714	131,615,594	del	t		CNVR1540.1
Park	3	133,190,943	133,196,075	del	?		7_AK18
Conrad	3	142,027,098	142,029,737	del	ccact		CNVR1562.1
Conrad	3	147,867,880	147,877,553	del	ca		CNVR1576.1
Conrad	3	150,751,166	150,752,841	del	aggga		CNVR1583.1
Park	3	153,247,620	153,248,061	del	?		8_NA18542
Conrad	3	157,574,864	157,576,385	del		TTAAA	CNVR1591.1
Conrad	3	163,994,828	164,109,029	del	c		CNVR1608.1
Vissers	3	165,316,928	165,701,172	del			del 2
Park	3	167,470,179	167,470,516	del	?		9_NA18968
Conrad	3	180,032,827	180,033,568	del	GCC		CNVR1648.2
deSmith	3	181137034(6)	181137500(2)	del	TC		
Kim	3	188,063,727	188,068,042	del	?		
Park	3	191,220,038	191,223,219	del	?		10_AK12
Conrad	3	191,220,042	191,223,222	del	ctca		CNVR1669.1
Conrad	3	192,547,379	192,554,361	del		TATCCAGA	CNVR1675.1
Korbel	3	192,547,379	192,554,362	del			NA15510_del 293
Korbel	3	193,420,025	193,471,341	del	TCA		NA15510_del 284
Conrad	3	194,358,026	194,368,102	del	TCT		CNVR1685.1
Conrad	3	198,418,970	198,423,652	del		CAGACT	CNVR1716.1
Korbel	4	33,805	39,019	del	T		NA18505_del 569
Korbel	4	3,583,174	3,588,419	del			NA15510_del 335
Korbel	4	3,883,363	3,964,476	del			NA15510_del 330
Conrad	4	6,733,297	6,736,765	del	GAA		CNVR1791.1
Korbel	4	9,820,365	9,843,671	del	T		NA18505_del 541
Korbel	4	9,820,365	9,843,671	del	T		NA15510_del 311
Korbel	4	16,553,336	16,559,843	del			NA18505_del 551
Conrad	4	18,666,569	18,667,067	del		ACAGTTAATTCAGTTACAGTTAA-TACAGTTAA	CNVR1838.1
Korbel	4	20,769,760	20,776,557	del	TTTAGTATAATTTTCT		NA15510_del 345
Korbel	4	20,769,760	20,776,557	del	TTTAGTATAATTTTCT		NA18505_del 548
Korbel	4	20,978,066	20,986,741	del			NA18505_del 552
Conrad	4	28,030,631	28,031,142	del	T		CNVR1859.1
Park	4	30,623,047	30,624,073	del	?		11_AK18
Korbel	4	39,911,421	39,913,454	del			NA15510_inv A
Kim	4	42,457,435	42,464,300	del	?		
Park	4	43,446,564	43,446,887	del	?		12_NA18949
Conrad	4	45,897,554	45,899,785	del	TC		CNVR1904.1
Kim	4	58,180,961	58,185,488	del	?		
Conrad	4	58,506,644	58,513,191	del	CC		CNVR1928.1
Conrad	4	61,621,786	61,624,840	del	AA		CNVR1937.1
Korbel	4	70,502,047	70,507,392	del	CCA		NA18505_del 559
Korbel	4	70,502,047	70,507,392	del	CCA		NA15510_del 347
Korbel	4	79,488,148	79,494,224	del	AAAGCTTCAAGATA		NA18505_del 543
Korbel	4	79,488,148	79,494,224	del	AAAGCTTCAAGATA		NA15510_del 327
Kim	4	79,488,158	79,494,220	del	?		
Korbel	4	81,107,078	81,113,124	del	AAGAATAGGGCGTCCG		NA15510_del 325
Conrad	4	87,195,401	87,198,978	del	CTA		CNVR1991.1
Conrad	4	92,150,579	92,154,859	del			CNVR2000.1
Conrad	4	98,573,314	98,578,237	del	g		CNVR2013.1
deSmith	4	98573314(5)	98578237(8)	del	G		
Kim	4	106,926,782	106,936,575	del	?		
Conrad	4	107,276,027	107,282,811	del		CA	CNVR2030.1
Conrad	4	107,823,749	107,828,706	del	AATA		CNVR2032.1
Korbel	4	108,347,262	108,351,169	del	AAAACATTAGGATTCTCTTT		NA18505_del 588
Kim	4	108,347,263	108,351,179	del	?		
Korbel	4	116,148,168	116,151,323	del	GTGA		NA15510_del 344
Conrad	4	117,587,246	117,587,942	del	CTC		CNVR2055.1
Conrad	4	135,652,452	135,655,151	del		GTACAT	CNVR2080.1
Kim	4	142,450,233	142,452,513	del	?		
Conrad	4	162,098,484	162,104,398	del	AATTT		CNVR2137.1
Conrad	4	163,613,455	163,615,872	del	GA		CNVR2141.1
Kim	4	165,024,355	165,039,560	del	?		
Conrad	4	165,422,493	165,425,669	del	aaat		CNVR2147.1
Park	4	165,422,493	165,425,670	del	?		13_NA18942
Conrad	4	167,223,766	167,224,369	del		TTAGCAC	CNVR2152.1
Korbel	4	168,129,302	168,260,424	del	TT		NA18505_del 555
Korbel	4	168,345,738	168,351,840	del			NA18505_del 589
Conrad	4	169,045,106	169,229,628	del	cttc		CNVR2161.1
Conrad	4	172,610,983	172,616,000	del		CTT	CNVR2168.1
Conrad	4	173,225,214	173,229,506	del		C	CNVR2171.1
Conrad	4	173,661,607	173,670,095	del		A	CNVR2172.1
Korbel	4	173,661,607	173,670,096	del			NA18505_del 576
Korbel	4	173,661,607	173,670,096	del			NA15510_del 334

Conrad	4	175,861,744	175,863,548	del	TGG		CNVR2174.1
Korbel	4	178,704,201	178,725,490	del			NA18505_del 546
Conrad	4	182,293,555	182,294,125	del	GT		CNVR2196.1
Conrad	4	185,395,045	185,406,164	del	AAAG		CNVR2209.1
Korbel	4	190,630,172	190,634,918	del	TTTT		NA18505_del 574
Korbel	4	190,822,807	190,850,110	del			NA18505_inv 931
Korbel	4	190,822,807	190,850,110	del			NA15510_inv 527
Kim	5	71,386	76,029	del	?		
Conrad	5	2,926,819	2,927,967	del			CNVR2318.1
Kim	5	10,579,961	10,585,291	del	?		
Conrad	5	16,177,808	16,178,491	del		aatagaa	CNVR2371.1
Conrad	5	19,411,028	19,412,156	del		atgg	CNVR2388.1
Conrad	5	21,485,956	21,488,250	del	TGGTCTCTGCTTCTTATTCTTCATGTGCTC		CNVR2395.1
Conrad	5	26,832,458	26,837,656	del	TTT		CNVR2407.1
Conrad	5	28,526,758	28,531,365	del	TGT		CNVR2412.1
Korbel	5	40,004,665	40,012,423	del			NA18505_del 622
Korbel	5	46,306,413	46,311,594	del	TTCT		NA18505_del 632
Korbel	5	46,306,413	46,311,594	del	TTCT		NA15510_del 363
Kim	5	49,471,345	49,476,325	del	?		
Korbel	5	57,359,234	57,369,537	del	GA		NA15510_del 373
Korbel	5	57,359,234	57,369,537	del	GA		NA18505_del 640
Kim	5	57,715,747	57,721,855	del	?		
Korbel	5	57,715,755	57,721,865	del	AAAAAACAC		NA18505_del 605
Conrad	5	63,734,102	63,737,115	del	GGC		CNVR2483.1
deSmith	5	65479417(40)	65479953(75)	del	ATTGTATAGTGCTATCATTATGT		
deSmith	5	78,145,556	78,147,626	del			
Vissers	5	78514996(8)	79167040(2)	del	GC		del 11
Conrad	5	86,281,760	86,282,899	del		tat	CNVR2522.1
Park	5	97,427,318	97,428,518	del	?		14_NA18997
Conrad	5	97,961,018	97,967,734	del	ATGTAGGAAAAT		CNVR2533.1
Korbel	5	99,541,738	99,548,258	del	A		NA18505_del 625
Conrad	5	106,352,650	106,354,342	del		T	CNVR2554.1
Korbel	5	108,622,427	108,629,161	del	CCATTCTTATTTCTT		NA18505_del 608
Conrad	5	111,967,481	111,972,239	del		ctaactaataatataactaataatataac-taatataataatataactaatataa	CNVR2566.1
Conrad	5	119,408,045	119,410,579	del		GAG	CNVR2587.1
Conrad	5	121,160,505	121,163,855	del	gaca		CNVR2590.1
Park	5	127,363,899	127,364,827	del	?		15_AK6
Conrad	5	127,435,338	127,438,815	del			CNVR2600.1
Conrad	5	133,162,590	133,164,387	del		ggatggaggaagag	CNVR2610.1
Conrad	5	140,202,321	140,219,115	del		cca	CNVR2622.1
Conrad	5	150,157,856	150,161,793	del	**A**		CNVR2646.1
Conrad	5	150,183,353	150,203,456	del		AAATAGA	CNVR2647.1
Korbel	5	151,436,615	151,442,650	del	CAAAATTACATGGTGGA		NA18505_del 606
Korbel	5	151,436,623	151,442,641	del	CAAAATTACATGGTGGA		NA15510_del 362
Conrad	5	155,244,367	155,245,159	del	TAATCT		CNVR2658.1
Conrad	5	162,794,351	162,795,869	del		aacaagatgattcac	CNVR2669.1
Park	5	162,794,351	162,795,870	del	?		16_AK6
Conrad	5	166,335,344	166,336,212	del	ACTG		CNVR2676.1
Park	5	170,062,496	170,063,968	del	?		17_NA18542
Kim	5	177,754,281	177,756,656	del	?		
Conrad	5	180,501,840	180,502,585	del	A		CNVR2723.1
Conrad	6	16,992,947	16,994,075	del	a		CNVR2789.1
Park	6	22,158,817	22,162,220	del	?		18_AK6
Conrad	6	23,851,238	23,853,924	del			CNVR2811.1
deSmith	6	24433341(6)	24435791(6)	del	TCTCC		
Conrad	6	31,464,139	31,561,087	del	tttctt		CNVR2842.1
Conrad	6	32,667,951	32,668,369	del	A		CNVR2845.3
Kim	6	34,045,807	34,050,676	del	?		
deSmith	6	34425086(9)	34427582(5)	del	AGA		
Conrad	6	35,734,365	35,737,724	del	cc		CNVR2859.1
Korbel	6	49,038,905	49,046,493	del	?		NA18505_del 675
Conrad	6	51,307,435	51,308,085	del	TTTC		CNVR2899.1
Conrad	6	51,844,070	51,844,767	del	ATT		CNVR2901.1
Conrad	6	55,933,916	55,954,671	del		TACT	CNVR2912.1
Korbel	6	57,405,149	57,409,206	del			NA15510_del 394
Korbel	6	57,531,030	57,537,083	del	TAAAAGACTATATACATC		NA15510_del 382
Conrad	6	67,097,521	67,099,910	del		AATTAATACATGCCAGATT	CNVR2946.2
Conrad	6	69,195,308	69,196,113	del		C	CNVR2951.1
Korbel	6	74,648,523	74,659,593	del	AA		NA18505_del 645
Conrad	6	77,073,432	77,085,882	del	GG		CNVR2971.1
Korbel	6	85,374,868	85,380,946	del	ACATTTCT		NA18505_del 660
Korbel	6	85,374,868	85,380,946	del	ACATTTCT		NA15510_del 381
Conrad	6	89,978,472	89,978,896	del			CNVR2996.1
Conrad	6	95,250,045	95,251,059	del	TA		CNVR3004.1
Conrad	6	96,494,322	96,496,142	del		AA	CNVR3006.1
Conrad	6	100,141,223	100,142,043	del		ATTGC	CNVR3009.1
Conrad	6	100,813,532	100,814,184	del	CAA		CNVR3010.1
Conrad	6	103,844,156	103,869,582	del		tgcca	CNVR3020.1
Korbel	6	107,277,574	107,277,728	del			NA15510_inv 558
Korbel	6	130,890,026	130,893,988	del			NA18505_inv 972
Korbel	6	133,383,515	133,389,581	del	AGGTGTGATGTTTT		NA18505_del 670
Korbel	6	133,383,515	133,389,581	del	AGGTGTGATGTTTT		NA15510_del 389
Conrad	6	134,310,673	134,311,422	del	AGAG		CNVR3074.1
Conrad	6	141,590,317	141,591,727	del	TTTC		CNVR3083.1
Conrad	6	154,000,064	154,003,221	del	ATTT		CNVR3112.1
Vissers	6	155746035(65)	157307963(93)	del	TGTTAGCCAGGATGGTCTTGATCTCCTGAC		del 28
deSmith	6	162,645,085	162,645,903	del		TTTT	
Conrad	6	164,463,437	164,467,181	del	C		CNVR3150.1
Korbel	6	164,463,437	164,467,182	del	G		NA15510_del 384
Kim	6	165,644,659	165,652,123	del	?		
Korbel	6	168,836,086	168,836,577	del			NA18505_inv 973
Korbel	6	168,836,086	168,836,577	del			NA15510_inv 556
Conrad	7	1,822,569	1,830,102	del			CNVR3242.1
Korbel	7	6,866,324	6,889,694	del			NA18505_del 731
Conrad	7	13,244,637	13,247,240	del	c		CNVR3295.1
Conrad	7	16,138,298	16,140,597	del			CNVR3304.1
Conrad	7	20,715,311	20,720,104	del	GC		CNVR3311.1
Conrad	7	22,401,328	22,403,288	del	CAC		CNVR3314.1
Conrad	7	52,930,697	52,932,410	del	ACAG		CNVR3389.2
Conrad	7	54,563,980	54,564,927	del			CNVR3394.1
Conrad	7	69,834,135	69,840,971	del			CNVR3443.1
deSmith	7	82856583(4)	82857509(10)	del	A		

Conrad	7	89,648,345	89,650,537	del	gc		CNVR3485.1
Korbel	7	90,869,019	90,880,521	del	G		NA15510_del 405
Korbel	7	96,313,826	96,319,935	del	GCAACTGGAACTTTC		NA18505_del 722
Korbel	7	96,313,826	96,319,935	del	GCAACTGGAACTTTC		NA15510_del 401
Conrad	7	98,066,481	98,068,689	del	ttg		CNVR3503.1
Korbel	7	106,847,671	106,850,014	del			NA18505_inv 936
Korbel	7	106,847,671	106,850,014	del			NA15510_inv 559
Korbel	7	113,203,397	113,209,444	del	CATAATGGCATTTTT		NA18505_del 694
Korbel	7	113,203,397	113,209,444	del	CATAATGGCATTTTT		NA15510_del 408
Kim	7	113,203,412	113,209,444	del	?		
Korbel	7	113,439,881	113,446,502	del	TTGT		NA18505_del 714
Conrad	7	115,718,878	115,728,486	del	ag		CNVR3539.1
Conrad	7	125,833,125	125,838,686	del	TAC		CNVR3559.1
Conrad	7	127,002,204	127,005,218	del		ACAC	CNVR3563.1
Park	7	131,923,553	131,924,090	del	?		19_NA18526
Conrad	7	133,435,543	133,448,872	del	TGC		CNVR3573.1
Vissers	7	145,936,946	146,244,399	del		A	del 3
Conrad	7	147,703,799	147,707,262	del	TGGATC		CNVR3609.1
Korbel	7	151,620,490	151,704,326	del	TTTG		NA18505_del 696
Korbel	7	151,620,490	151,704,326	del	TTTG		NA15510_del 397
Conrad	7	158,810,205	158,815,487	del		aa	CNVR3683.1
Kim	8	584,397	589,415	del	?		
Korbel	8	584,453	589,415	del			NA18505_del 762
Kim	8	2,116,965	2,122,377	del	?		
Conrad	8	5,582,924	5,593,187	del		ac	CNVR3760.1
Conrad	8	11,282,971	11,284,620	del	aag		CNVR3781.1
Korbel	8	13,658,494	13,696,112	del	AA		NA18505_del 759
Korbel	8	16,245,681	16,251,905	del			NA18505_del 763
Vissers	8	24447284(5)	31302615(6)	del	GGTG		del 5
Conrad	8	25,028,351	25,046,862	del	ac		CNVR3830.1
Korbel	8	25,122,595	25,126,570	del			NA18505_del 743
Kim	8	25,122,602	25,126,570	del	?		
Conrad	8	25,122,602	25,126,576	del	GGCTCAG		CNVR3831.1
Conrad	8	39,351,231	39,506,385	del		T	CNVR3859.1
Conrad	8	40,004,842	40,009,998	del		tgaccagc	CNVR3860.1
Korbel	8	40,893,763	40,898,989	del			NA15510_del 421
Korbel	8	42,309,595	42,313,559	del			NA18505_del 754
Conrad	8	49,255,939	49,256,532	del	G		CNVR3878.1
Conrad	8	51,193,644	51,200,882	del		GTGTTTCCTAAGTGCTTA	CNVR3882.1
Conrad	8	51,387,538	51,390,870	del	CTA		CNVR3884.1
Vissers	8	60249086(147)	62195805(66)	del	GGGGTCAGGGACCCACTTGAGGAGGCAGT-CTGCCCGTTCTCAGATCTCCAGCTGCGTGCTG		del 29
Vissers	8	61798469(71)	61889695(7)	del	TT		del 12
Park	8	62,197,914	62,198,447	del	?		20_NA18552
Korbel	8	73,950,326	73,956,385	del	TGCAAATCTT		NA18505_del 768
Kim	8	73,950,329	73,956,378	del	?		
Conrad	8	75,525,426	75,529,531	del		C	CNVR3935.1
Conrad	8	82,207,573	82,209,183	del		ttacgtgtac	CNVR3946.2
Conrad	8	85,423,525	85,431,728	del	CAAC		CNVR3952.1
Korbel	8	120,223,723	120,230,397	del	AT		NA18505_del 770
Conrad	8	120,223,725	120,230,432	del	ta		CNVR4031.1
Korbel	8	126,664,303	126,670,317	del	TGTGAGTG		NA15510_del 434
Korbel	8	126,664,303	126,670,317	del	AAACACTCACA		NA18505_del 748
Conrad	8	131,919,931	131,921,920	del		A	CNVR4052.1
Korbel	8	135,152,106	135,158,209	del	CATTCTTCAACATTTT		NA18505_del 773
Korbel	8	135,152,106	135,158,209	del	CATTCTTCAACATTTT		NA15510_del 425
Conrad	8	137,229,379	137,233,082	del	GC		CNVR4067.1
Korbel	8	144,771,577	144,785,836	del	TTT		NA15510_del 427
Conrad	9	4,366,421	4,367,644	del	tat		CNVR4153.2
Conrad	9	8,630,851	8,631,626	del	AAGA		CNVR4174.1
Conrad	9	9,506,969	9,507,964	del		GTTTTTCTGTA	CNVR4177.1
Conrad	9	10,394,564	10,395,094	del	CCA		CNVR4181.1
Vissers	9	14,196,884	16,342,939	del	GT		del 1
Conrad	9	23,352,801	23,367,685	del		GA	CNVR4213.1
Conrad	9	31,281,356	31,282,673	del	C		CNVR4234.1
Kim	9	70,927,942	70,933,175	del	?		
Korbel	9	70,927,942	70,933,177	del	GT		NA18505_del 788
Korbel	9	70,927,942	70,933,177	del	GT		NA15510_del 439
Kim	9	73,446,481	73,449,953	del	?		
Korbel	9	80,627,751	80,660,989	del	T		NA18505_del 807
Kim	9	84,854,269	84,860,328	del	?		
Korbel	9	85,698,412	87,640,324	del			NA18505_inv 980
Conrad	9	100,348,869	100,351,490	del	tctca		CNVR4414.1
Kim	9	112,516,996	112,519,927	del	?		
Vissers	9	117097067(8)	117798935(6)	del	A		del 6
Conrad	9	117,235,735	117,237,188	del	T		CNVR4452.1
Conrad	9	129,221,443	129,225,901	del	gatc		CNVR4479.1
Conrad	9	137,333,959	137,336,077	del	ag		CNVR4527.3
Korbel	9	137,353,885	137,357,426	del	TGA		NA15510_del 451
Conrad	9	137,353,887	137,357,424	del	tga		CNVR4527.1
Conrad	10	4,280,063	4,281,683	del	C		CNVR4594.1
Kim	10	4,427,701	4,431,391	del	?		
Conrad	10	4,698,520	4,700,525	del		ATAG	CNVR4596.1
Korbel	10	5,277,305	5,283,355	del	AAAAAAATAGTGTAAAGT dup		NA15510_del 64
Korbel	10	5,627,107	5,677,112	del			NA18505_del 119
Korbel	10	5,627,107	5,677,112	del			NA15510_del 67
Kim	10	5,627,110	5,677,111	del	?		
Korbel	10	6,451,584	6,457,651	del	CAGCAAATCATTTTCCT		NA18505_del 99
Conrad	10	6,694,413	6,704,196	del	AC		CNVR4606.1
Conrad	10	7,117,046	7,118,307	del		A	CNVR4609.1
Lam	10	7,117,046	7,118,307	del	?		Wheeler
Conrad	10	17,350,717	17,355,515	del		TATACTATGTGTAT	CNVR4634.1
Park	10	20,036,712	20,038,183	del	?		21_AK14
Conrad	10	58,879,793	58,881,859	del		gttaaagatcaatc	CNVR4760.1
Conrad	10	61,031,970	61,035,014	del	AATCA		CNVR4767.1
Conrad	10	65,179,256	65,181,407	del	AC		CNVR4773.1
Conrad	10	66,068,780	66,070,754	del	GATA		CNVR4776.1
Korbel	10	66,976,938	66,985,301	del	AC		NA15510_del 82
Park	10	66,976,938	66,985,301	del	?		22_AK6
Conrad	10	66,976,940	66,985,302	del	GT		CNVR4779.1
Korbel	10	77,925,582	77,931,030	del	TTCAGT		NA15510_del 72
Kim	10	84,117,799	84,120,345	del	?		
Conrad	10	93,623,318	93,624,545	del	cac		CNVR4855.1

Korbel	10	96,857,182	96,864,933	del			NA18505_del 106
Korbel	10	102,342,367	102,354,417	del	TCAA		NA18505_del 120
Conrad	10	107,940,672	107,941,586	del	T		CNVR4884.1
Park	10	107,940,672	107,941,586	del	?		23_AK10
Conrad	10	107,985,984	107,987,724	del			CNVR4885.1
Conrad	10	108,020,308	108,022,533	del		A	CNVR4886.1
Kim	10	114,102,173	114,106,649	del	?		
Conrad	10	114,102,173	114,106,650	del	GG		CNVR4893.1
Kim	10	128,578,838	128,582,206	del	?		
Park	10	130,726,861	130,727,265	del	?		24_NA18537
Conrad	10	132,799,052	132,802,769	del	TTCA		CNVR4949.1
Korbel	11	4,924,739	4,933,353	del			NA18505_del 158
Korbel	11	4,924,739	4,933,353	del			NA15510_del 103
Conrad	11	5,478,208	5,479,749	del	gt		CNVR5044.1
Conrad	11	5,716,663	5,718,941	del		T	CNVR5048.1
Korbel	11	5,741,150	5,765,860	del	AC		NA15510_del 96
Conrad	11	5,741,151	5,765,860	del	AC		CNVR5049.1
Conrad	11	7,091,294	7,093,075	del	ccag acTTTAAGAC		CNVR5056.1
Conrad	11	11,336,197	11,337,606	del	gcc		CNVR5072.1
Conrad	11	22,427,462	22,429,349	del	CAAATA TG		CNVR5103.1
Conrad	11	24,399,728	24,408,863	del		AGCAGA	CNVR5109.1
Conrad	11	29,924,149	29,925,042	del	t		CNVR5125.1
Conrad	11	45,386,425	45,388,192	del	GCC		CNVR5158.1
Conrad	11	47,014,165	47,020,283	del		AAAGTGGGATAGTGGAA	CNVR5162.1
Conrad	11	48,557,434	48,560,860	del	TC		CNVR5165.1
Conrad	11	58,388,350	58,393,131	del	GC		CNVR5189.1
Korbel	11	58,388,350	58,393,132	del	C		NA15510_del 98
Conrad	11	59,984,740	59,985,962	del	t		CNVR5192.1
Conrad	11	69,660,203	69,662,138	del	CTC		CNVR5218.1
Conrad	11	81,534,058	81,540,993	del		CAGTTACAAATATGTCTGTTTCT	CNVR5250.1
Conrad	11	85,981,878	85,984,206	del			CNVR5255.1
Korbel	11	92,791,149	92,800,594	del			NA18505_del 149
Conrad	11	95,641,573	95,643,043	del	C		CNVR5278.1
Korbel	11	101,071,002	101,079,503	del			NA15510_del 90
Conrad	11	103,772,961	103,778,439	del	ATA		CNVR5294.1
Korbel	11	103,772,961	103,778,440	del	ATA		NA18505_del 159
Korbel	11	104,798,682	104,804,116	del			NA18505_del 153
Conrad	11	128,187,926	128,188,620	del		C	CNVR5351.1
Conrad	11	134,107,192	134,112,875	del	gtgt		CNVR5372.1
Conrad	11	134,238,110	134,239,324	del	GT		CNVR5375.1
Conrad	12	247,761	255,094	del	agca		CNVR5382.1
Conrad	12	5,092,083	5,093,713	del			CNVR5408.1
Conrad	12	6,111,699	6,118,340	del	ag		CNVR5412.1
Kim	12	11,075,858	11,142,017	del	?		
Conrad	12	11,917,682	11,918,420	del			CNVR5436.1
Kim	12	15,909,933	15,912,931	del	?		
deSmith	12	20,859,912	20,859,936	del	AATA	TAG	
Conrad	12	22,310,208	22,315,497	del		CCA	CNVR5456.1
Korbel	12	22,310,232	22,315,498	del			NA15510_del 111
Conrad	12	23,830,789	23,831,356	del	t		CNVR5458.1
Conrad	12	33,606,390	33,608,237	del	AACAA		CNVR5492.1
Conrad	12	36,294,988	36,304,145	del			CNVR5501.2
Kim	12	38,587,965	38,602,082	del	?		
Park	12	49,259,982	49,261,778	del	?		25_AK4
Kim	12	55,618,220	55,663,208	del	?		
Korbel	12	57,008,350	57,016,840	del			NA18505_del 173
Korbel	12	57,008,350	57,016,840	del			NA15510_del 108
Conrad	12	58,808,112	58,811,308	del	TGTCTA		CNVR5559.1
Conrad	12	66,213,619	66,214,389	del	AGA		CNVR5577.1
Korbel	12	68,881,119	68,883,851	del	GAAGTGTCATACTTTTTT		NA15510_del 112
Conrad	12	69,158,533	69,164,484	del		C	CNVR5582.1
Korbel	12	79,379,862	79,380,365	del			NA15510_inv 548
Conrad	12	83,117,245	83,120,148	del	gcctca		CNVR5607.1
Conrad	12	85,128,811	85,133,374	dup	AG		CNVR5595
Kim	12	94,757,723	94,760,459	del	?		
Conrad	12	98,318,079	98,326,901	del	G		CNVR5639.1
Conrad	12	100,626,541	100,631,221	del	ACAG		CNVR5644.1
Korbel	12	100,626,541	100,631,222	del	CTGT		NA15510_del 109
Kim	12	128,624,266	128,628,228	del	?		
Vissers	13	31154817(20)	37648937(40)	del	TGC		del 14
Conrad	13	33,033,730	33,042,821	del			CNVR5837.1
Kim	13	33,033,730	33,042,822	del	?		
Conrad	13	37,955,319	37,958,205	del	GGAA		CNVR5850.1
Park	13	38,832,183	38,833,482	del	?		26_AK18
Conrad	13	49,411,121	49,412,891	del	gt		CNVR5870.1
Kim	13	56,650,541	56,686,865	del	?		
Conrad	13	69,633,730	69,673,451	del	tc		CNVR5918.1
Conrad	13	71,375,261	71,378,578	del	c		CNVR5920.1
Kim	13	71,705,623	71,710,360	del	?		
Korbel	13	71,741,038	71,744,926	del	T		NA15510_del 122
Korbel	13	80,703,808	80,712,934	del	Wimpy polyA		NA18505_del 206
Conrad	13	82,063,937	82,070,208	del	TG		CNVR5946.1
Conrad	13	88,219,968	88,220,938	del		cattattagcagc	CNVR5960.1
Conrad	13	89,660,856	89,662,812	del	AAAT		CNVR5962.1
Lam	13	103,695,322	103,695,323	del	?		Levy
Park	13	108,159,746	108,160,439	del	?		27_NA18942
Park	14	21,951,506	21,952,100	del	?		28_AK10
Conrad	14	21,951,507	21,952,100	del	G		CNVR6084.1
Kim	14	34,184,839	34,192,011	del	?		
Park	14	38,074,269	38,074,779	del	?		29_NA18542
Conrad	14	39,679,566	39,687,423	del		caggctcctttgtaaataa	CNVR6133.1
Korbel	14	40,883,961	40,929,395	del	AACT		NA18505_del 219
Conrad	14	42,057,278	42,062,684	del	CT		CNVR6140.1
Conrad	14	53,780,099	53,783,438	del	TGA		CNVR6158.1
Conrad	14	69,086,749	69,092,243	del	gtgt		CNVR6178.1
deSmith	14	72402705(7)	72403561(3)	del	GT		
deSmith	14	72615517(24)	72616685(92)	del	TTTTTTT		
Kim	14	73,076,457	73,108,631	del	?		
Conrad	14	79,176,043	79,184,853	del	tg		CNVR6203.1
Conrad	14	80,947,966	80,950,367	del		CAGAGTTAAGATAAGCA	CNVR6209.1
Conrad	14	81,568,863	81,573,083	del		AACATAAATC	CNVR6211.1
Kim	14	81,568,863	81,573,084	del	?		
Park	14	81,568,863	81,573,084	del	?		30_AK18

Korbel	14	84,366,861	84,371,909	del	TTAGAAAC		NA15510_del 126
Park	14	84,366,861	84,371,909	del	?		31_NA18973
Conrad	14	84,366,869	84,371,916	del	GTTTCTAA		CNVR6220.1
Korbel	14	105,282,153	105,397,046	del	AGC		NA18505_del 225
Kim	14	105,282,154	105,397,044	del	?		
Korbel	14	105,311,003	105,398,170	del			NA15510_del 133
Korbel	15	18,841,481	18,849,632	del			NA18505_del 241
Korbel	15	18,841,481	18,849,632	del			NA15510_del 135
Conrad	15	21,606,470	21,612,738	del		ac	CNVR6306.1
Kim	15	22,009,161	22,111,478	del	?		
Korbel	15	25,589,686	25,602,096	del			NA18505_del 233
Conrad	15	33,947,426	33,949,511	del	ttat		CNVR6355.1
Park	15	37,531,682	37,532,152	del	?		32_NA18592
Conrad	15	37,531,683	37,532,239	del	**A**		CNVR6357.1
Korbel	15	42,941,408	42,947,386	del			NA18505_del 247
Park	15	44,647,999	44,648,461	del	?		33_AK10
Conrad	15	60,493,387	60,495,087	del	ATT		CNVR6418.1
Kim	15	68,808,907	68,814,563	del	?		
Conrad	15	69,495,663	69,500,744	del		acc	CNVR6438.1
Conrad	15	69,668,624	69,670,021	del		T	CNVR6439.1
Conrad	15	82,332,438	82,334,852	del		tttc	CNVR6484.1
deSmith	15	83858012(6)	83860206(10)	del	TTTT		
Conrad	15	97,474,127	97,474,924	del	ctg		CNVR6543.1
Park	15	99,159,012	99,159,896	del	?		34_AK10
Conrad	15	99,159,013	99,159,896	del	TGC		CNVR6552.1
Conrad	16	13,201,968	13,203,997	del	ct		CNVR6642.1
Korbel	16	14,407,231	14,412,704	del			NA18505_del 277
Korbel	16	16,841,867	16,847,919	del	AAGCTTAATTCTTTT		NA18505_del 265
Conrad	16	22,955,276	22,957,032	del		GATTCT	CNVR6670.1
deSmith	16	22,955,277	22,957,032	del		GATTCT	
Korbel	16	25,247,611	25,250,630	del	GG		NA15510_del 160
Kim	16	29,167,046	86,811,700	del	?		
Korbel	16	44,955,510	44,985,518	del			NA18505_inv 922
Conrad	16	55,923,812	55,924,619	del		T	CNVR6737.1
deSmith	16	56282299(301)	56285908(10)	del	AAT		
Conrad	16	61,101,835	61,108,169	del		GT	CNVR6752.1
Conrad	16	75,096,634	75,101,526	del		C	CNVR6782.1
Lam	16	75,096,634	75,101,526	del	?		Wheeler
deSmith	16	76115170(4)	76115184(8)	del	GGGG		
Conrad	16	76,929,139	76,942,399	del			CNVR6791.2
Kim	16	76,929,139	76,942,400	del	?		
Vissers	16	84275151(5)	86275753(7)	del	AGCC		del 17
Vissers	16	84374208(26)	85277007(25)	del	GAGACCAGCCTGGCCAAC		del 25
Vissers	16	84402571(9)	85435712(20)	del	TGAGCCAC		del 20
Vissers	16	85157840(3)	85288901(4)	del	GCC		del 15
Conrad	16	86,385,864	86,388,966	del	cca		CNVR6850.1
Vissers	16	87676976(96)	88037131(51)	del	CCAAAGTGCTGGGATTACAG		del 26
deSmith	16	88,089,521	88,095,227	del			
Conrad	17	193,767	197,056	del		aaatggttattaatt	CNVR6908.1
Conrad	17	5,536,431	5,538,222	del	TCC		CNVR6961.1
Conrad	17	11,190,428	11,200,650	del	tctgc		CNVR6988.1
Conrad	17	11,352,392	11,353,062	del		acacaggtccataaagaaagaa	CNVR6990.1
Zhang	17	14215232(59)	15509173(200)	del	AGCCTCCCAAAGTGCTGGGATTACAGG		C1292 and C2405
Zhang	17	14,553,352	15,089,591	del			A26
Zhang	17	14,890,749	15,328,218	del		TAAAATTATCTTTTAGTCATTAA	SP951
Zhang	17	15057049(59)	15468624(34)	del	GTTTCACCAT		SP54C
Zhang	17	15079030(1)	15096344(5)	del	T	CAT	SPR2
Zhang	17	15096987(9)	15110285(7)	del	TC		A29
Zhang	17	15,101,735	15,106,908	del	A		A23
Zhang	17	15118189(94)	15311614(9)	del	TCTCT		SPR1
Zhang	17	15143662(3)	15329785(6)	del	A		SP3672 and SP3840
Conrad	17	15,730,280	15,734,511	del		TAGTT	CNVR7004.1
Korbel	17	15,730,283	15,734,503	del	T		NA15510_del 181
Conrad	17	22,560,680	22,564,522	del	aaattcacatggca		CNVR7043.1
Conrad	17	23,804,514	23,807,620	del	ct		CNVR7047.1
Park	17	27,130,737	27,131,657	del	?		35_NA18582
Korbel	17	49,513,662	49,523,001	del			NA15510_del 185
Conrad	17	53,042,810	53,044,915	del	ct		CNVR7144.1
Korbel	17	63,284,492	63,722,780	del			NA18505_del 312
Korbel	17	65,966,692	65,972,772	del	CACAAAATCTT		NA15510_del 179
Conrad	17	71,873,832	71,876,903	del	GGA		CNVR7201.1
Conrad	18	5,314,645	5,316,244	del	CA		CNVR7241.1
Korbel	18	14,541,533	14,559,237	del	TGGAAC		NA18505_del 320
Kim	18	14,542,177	14,558,726	del	?		
Conrad	18	22,825,618	22,826,339	del			CNVR7283.1
Conrad	18	28,142,161	28,143,489	del	gaa		CNVR7292.3
Conrad	18	28,749,691	28,755,259	del	tggag		CNVR7293.1
Korbel	18	32,507,071	32,513,858	del			NA18505_del 322
Park	18	33,560,058	33,560,631	del	?		36_AK10
Conrad	18	36,513,891	36,520,748	del	T		CNVR7307.1
Conrad	18	40,230,705	40,236,072	del		G	CNVR7312.1
Vissers	18	44917833(6)	45160155(8)	del	TAA		del 16
Kim	18	45,948,971	45,952,385	del	?		
Park	18	45,948,975	45,952,385	del	?		37_AK4
Korbel	18	46,124,317	46,130,364	del			NA18505_del 323
Park	18	48,716,563	48,717,029	del	?		38_NA18564
Conrad	18	49,390,403	49,391,772	del	CT		CNVR7336.1
Conrad	18	49,459,967	49,464,573	del		GAATAGGTGGCTATCTTCTAGGTGGC-TAAACACCT	CNVR7337.1
Korbel	18	50,206,961	50,210,958	del	AGCCA		NA18505_del 325
Conrad	18	53,097,735	53,099,715	del			CNVR7344.1
Park	18	53,097,735	53,099,716	del	?		39_NA18564
Korbel	18	56,067,879	56,076,109	del			NA18505_del 324
Conrad	18	61,874,818	61,883,348	del	agattgc		CNVR7364.1
Conrad	18	61,917,853	61,920,186	del	ttgg		CNVR7365.1
Korbel	18	63,109,994	63,118,242	del	G		NA18505_del 328
Park	18	72,476,184	72,476,990	del	?		40_NA18552
Conrad	18	73,395,986	73,397,148	del	t		CNVR7401.1
Conrad	18	75,158,734	75,159,972	del			CNVR7428.1
Goldman	19	11,059,401	11,072,587	del	CTCCTGCCTCAGCCTCCCGAGTAGCTGG-GACTACAGGCACC		Fig1A
Goldman	19	11065247(58)	11079476(87)	dup	TTTAGTAGAGA		Fig1B
Goldman	19	11076446(9)	1108456(6)9	dup	AGA		Fig1D

Goldman	19	11078170(4)	11085805(9)	del	AAAA		Fig2A
Goldman	19	110847(53)85	11098862(95)	del	CCTCAGCCTCCCAAAGTGCTGGGATTACAGGT		Fig2C
Goldman	19	11093997(4002)	11094300(5)	del	CTCCT		Fig2B
Conrad	19	14,907,392	14,910,477	del	t		CNVR7551.1
deSmith	19	35979321(08)	35981593(606)	del	TGCACTCCAGCCT		
Conrad	19	39,419,441	39,422,452	del		tctagcatgtctactcagcatgcag	CNVR7620.1
Korbel	19	46,047,867	46,079,914	del			NA15510_del 202
Conrad	19	51,314,576	51,320,153	del			CNVR7673.1
Park	19	59,548,033	59,548,601	del	?		41_NA18999
Conrad	19	61,558,270	61,561,735	del	a		CNVR7740.1
Conrad	20	1,337,143	1,338,817	del	ACA		CNVR7762.1
Korbel	20	4,393,321	4,397,531	del			NA18505_del 429
Kim	20	7,044,793	7,050,847	del	?		
Korbel	20	14,719,512	14,887,609	del	GT		NA15510_del 265
Conrad	20	14,719,514	14,887,611	del	ca		CNVR7793.1
Conrad	20	15,249,734	15,251,771	del	TTT		CNVR7794.1
Kim	20	28,122,727	28,149,711	del	?		
Conrad	20	36,488,476	36,489,226	del	cc		CNVR7842.1
Kim	20	42,760,727	42,762,938	del	?		
Conrad	20	61,195,144	61,196,043	del	c		CNVR7927.1
Korbel	21	10,118,922	10,130,309	del			NA18505_del 465
Korbel	21	10,118,922	10,130,309	del			NA15510_del 267
Conrad	21	15,510,251	15,513,325	del	at		CNVR7956.2
Kim	21	19,758,801	19,765,198	del	?		
Korbel	21	20,722,534	20,767,095	del	A		NA18505_del 463
Korbel	21	26,942,867	26,943,597	del			NA18505_inv 928
Park	21	28,634,908	28,635,998	del	?		42_AK8
Conrad	21	42,223,191	42,226,560	del	GAGGAA		CNVR8021.1
Conrad	21	43,647,128	43,662,451	dup	TCCCACCCA		CNVR8000
Conrad	21	43,794,796	43,797,732	del	GAGA		CNVR8030.1
Conrad	21	43,983,814	43,985,302	del	CAGG		CNVR8032.1
Conrad	21	44,443,627	44,445,149	del	TGAAC		CNVR8037.1
Kim	22	27,963,089	27,965,391	del	?		
deSmith	22	32085555(72)	32090063(80)	del	TTTTTTTTTTTTTGAGA		
Conrad	22	32,110,447	32,113,338	del		AGAT	CNVR8140.1
Conrad	22	37,624,054	37,628,634	del	taaa		CNVR8163.1
Korbel	22	47,355,685	47,366,304	del			NA18505_del 486
Vissers	X	18297583(94)	18454991(5002)	del	TTTTGTATTTT		del 21
Vissers	X	18360769(71)	18498469(71)	del	TC		del 13
Vissers	X	18367759(74)	18442628(43)	del	TCCCAGCTACTCGGG		del 23
Edamura	X	146,801,147	146,801,336	del	?		Mila
Edamura	X	146,801,175	146,809,912	del	?		Quan
Edamura	X	146,801,176	146,801,333	del	?		deGraaff
Edamura	X	146,801,185	146,801,399	del	?		Petek
Edamura	X	146,801,187	146,801,332	del	?		deGraaff
Edamura	X	146,801,187	146,801,752	del	?		deGraaff
Edamura	X	146,801,207	146,801,487	del	?		deVries
Edamura	X	146,801,208	146,801,486	del	?		deGraaff
Edamura	X	146,801,213	146,801,322	del	?		Snow
Edamura	X	146,801,229	146,801,370	del	?		Mannermaa
Edamura	X	146,801,255	146,801,338	del	?		deGraaff
Edamura	X	146801090/92	146801565/67	del	?		Schmucker
Edamura	X	146801150/54	146801475/79	del	?		Grasso
Edamura	X	146801176+7	146,801,333	del	?		Grasso
Edamura	X	146801193/95	146801339/43	del	?		Gronskov
Edamura	X	146801195/97	146801395/97	del	?		Fan
Emory	X	24,021,195	24,026,627	del	Alu element		AU120.4
Emory	X	30,257,687	30,258,519	del	AG	A	AU065.4
Emory	X	33,952,914	33,982,773	del	AGGT		AU030.3
Emory	X	38,271,449	38,272,275	del	AAAAT	CT	AU087.5
Emory	X	44,264,761	44,266,313	del	Alu element, ~90 bp		AU096.5
Emory	X	58,256,036	58,256,488	del	AGGCATTCTAATGATAGAGACACctgtggtga		AU111.8
Emory	X	64,002,062	64,014,239	del	ACACT		AU122.1
Emory	X	65,381,485	65,415,746	dup	C, pyrimidine track		AU133.7
Emory	X	80,303,149	80,304,088	del	ATA		AU030.1
Emory	X	82,085,921	82,094,944	del	GCCTCCC, CAGAG		AU080.3
Emory	X	87,545,357	87,547,591	del	ATTA, AT		AU097.8
Emory	X	96,493,413	96,495,398	del	AG		AU113.2
Emory	X	97,842,279	97,843,597	del	ATT		AU116.6
Emory	X	103,289,074	103,289,688	del	ATTGCCCT		AU131.8
Emory	X	103,295,675	103,296,655	del	CATT		AU033.95
Emory	X	111,753,107	111,753,796	del	AA, poly T		AU103.3
Emory	X	113,234,594	113,241,627	del	GGC		AU145.2
Emory	X	116,588,176	116,591,267	del	GCT	CTTC	AU149.7
Emory	X	120,416,100	120,416,999	del		AATCAA	AU038.3
Emory	X	122,143,696	122,144,257	del	AG, GT		AU064.0B
Emory	X	131,767,115	131,769,226	del	GCC		AU082.1
Emory	X	133,692,157	133,693,471	del	AA		AU143.4
Emory	X	143,436,372	143,445,447	del	ATATCC		AU103.29
Emory	X	145,207,078	145,208,683	del	ATTT		AU058.3
Emory	X	149,678,591	149,679,508	del	TCTG		AU024.0
Emory	X	150,457,729	150,462,994	del	~220 nt		AU044.3
Emory	X	30,048,859(61)	30,056,096(98)	del	TT		AU150.5
Kim	X	92,682,955	92,688,161	del	?		
Korbel	X	35,537,365	35,544,320	del	TT		NA18505_del 829
Korbel	X	48,902,127	48,906,009	del			NA15510_inv 564
Korbel	X	154,570,791	154,574,920	del			NA18505_del 833
Nobile	X	31,750,900	31,832,080	del	CT		Junct 2
Nobile	X	31,753,361	31,810,985	del	CA	GCACATATCTCAGCACACATATCAGCACA	Junct 3
Nobile	X	31,640,226(9)	31762772(5)	del	GCG		Junct 5
Nobile	X	31,752,508(10)	31,823,801(4)	del	AGC		Junct 1
Woodward	X	102,082,641	103,009,563	dup	?		P026
Woodward	X	102,259,865	106,837,041	dup	?		P110
Woodward	X	102,362,654	103,127,653	dup	?		PMD9
Woodward	X	102,435,939	102,945,236	dup	?		P116
Woodward	X	102,490,121	103,444,951	dup	?		P114
Woodward	X	102,491,424	103,296,881	dup	?		PMD24
Woodward	X	102,537,145	103,011,041	dup	?		P134
Woodward	X	102,602,533	102,994,508	dup	?		P224
Woodward	X	102,682,198	104,356,130	dup	?		P176/PMD7
Woodward	X	102,816,004	102,992,894	dup	?		P015
Woodward	X	102,830,045	103,110,073	dup	?		P255
Woodward	X	102,831,298	102,942,004	dup	?		P348

Table A.6.a – Truly called CNV by the NimbleGen and Optimized protocols.

ARRAY_ID	Sample	CHR	START	STOP	SIZE(bp)	Probes	Probes/kb	Mean_Log2	GC	Primer	Platform
AU0780301.1s1A01	AU0780301	chrX	8,746,518	8,746,898	380	8	21.1	-1.68	43.6	AU008.7	NG
AU1038303.1s1A02	AU1038303	chrX	103,295,654	103,296,615	961	15	15.6	-1.35	40.5	AU103.3	NG
AU021503.1s1A03	AU021503	chrX	133,692,426	133,693,396	970	21	21.6	-0.86	43.7	AU133.7	NG
AU004803.1s1A03	AU004803	chrX	113,234,646	113,241,227	6,581	122	18.5	-0.83	37.8	AU113.2	NG
AU016803.1s1A01	AU016803	chrX	19,376,031	19,378,453	2,422	47	19.4	-0.82	40.6	AU019.38	NG
AU080803.1s1A01	AU080803	chrX	30,054,252	30,055,680	1,428	30	21.0	-0.81	37.6	AU030.1	NG
AU0875302.1s2A03	AU0875302	chrX	131,766,925	131,769,270	2,345	47	20.0	-0.79	41.7	AU131.8	NG
AU083504.1s1A02	AU083504	chrX	96,493,478	96,493,778	300	7	23.3	-0.74	41.3	AU096.5	NG
AU083504.1s1A03	AU083504	chrX	122,143,410	122,144,175	765	17	22.2	-0.73	39.8	AU122.1	NG
AU018003.1s1A01	AU018003	chrX	19,375,946	19,378,453	2,507	49	19.5	-0.72	40.7	AU019.38	NG
AU1038303.1s1A03	AU1038303	chrX	149,678,535	149,679,545	1,010	21	20.8	-0.71	55.3	AU149.7	NG
AU050703.1s1A03	AU050703	chrX	122,778,733	122,780,189	1,456	23	15.8	-0.70	46.3	AU122.78	NG
AU055303.1s1A03	AU055303	chrX	149,678,535	149,679,480	945	20	21.2	-0.67	55.2	AU149.7	NG
AU083504.1s1A03	AU083504	chrX	154,047,451	154,056,710	9,259	152	16.4	-0.64	34.4	AU154.0A	NG
AU083504.1s1A02	AU083504	chrX	58,256,031	58,256,531	500	11	22.0	-0.59	46.1	AU058.3	NG
AU056003.1s1A03	AU056003	chrX	149,678,510	149,679,625	1,115	24	21.5	-0.58	55.4	AU149.7	NG
AU1038303.1s1A02	AU1038303	chrX	82,086,497	82,094,944	8,447	91	10.8	-0.53	39.0	AU082.1	NG
AU0852304.1s1A01	AU0852304	chrX	44,265,421	44,265,881	460	11	23.9	-0.52	36.7	AU044.3	NG
AU065404.1s1A01	AU065404	chrX	33,980,127	33,980,127	27,121	482	17.8	-0.52	35.0	AU033.95	NG
AU056003.1s1A01	AU056003	chrX	8,746,203	8,749,728	3,525	63	17.9	-0.26	37.9	AU008.7	NG
AU016803.1s1A02	AU016803	chrX	65,382,556	65,399,668	17,112	328	19.2	0.39	38.4	AU065.4	NG
AU0920301.1s1A02	AU0920301	chrX	57,617,045	57,634,875	17,830	262	14.7	0.40	43.0	AU057.6	NG
AU0852304-A03-2s1	AU0852304	chrX	116,588,255	116,590,742	2,487	44	17.7	-1.64	38.6	AU116.6	Opt
AU1038303-A02-2s1	AU1038303	chrX	103,295,654	103,296,645	991	16	16.1	-1.40	40.9	AU103.3	Opt
AU083504-A02-3s1	AU083504	chrX	58,256,056	58,256,496	440	8	18.2	-1.33	46.8	AU058.3	Opt
AU0852304-A01-2s1	AU0852304	chrX	44,264,996	44,266,216	1,220	20	16.4	-1.31	41.6	AU044.3	Opt
AU020003-A02-2s1	AU020003	chrX	70,039,948	70,040,318	370	9	24.3	-1.18	37.0	AU070.0	Opt
AU028903-A03-2s1	AU028903	chrX	149,678,535	149,679,160	625	11	17.6	-1.15	55.5	AU149.7	Opt
AU065404-A01-2s1	AU065404	chrX	33,952,911	33,982,646	29,735	497	16.7	-1.06	35.4	AU033.95	Opt
AU1038303-A02-2s1	AU1038303	chrX	82,086,031	82,094,874	8,843	92	10.4	-1.04	39.0	AU082.1	Opt
AU021503-A03-2s1	AU021503	chrX	133,692,526	133,693,326	800	17	21.3	-0.98	43.3	AU133.7	Opt
AU0920301-A03-2s1	AU0920301	chrX	135,126,873	135,129,790	2,917	47	16.1	-0.97	43.6	AU135.1	Opt
AU0852304-A02-2s1	AU0852304	chrX	87,546,283	87,547,547	1,264	12	9.5	-0.95	36.0	AU087.5	Opt
AU0920301-A03-2s1	AU0920301	chrX	131,766,925	131,768,880	1,955	25	12.8	-0.89	42.2	AU131.8	Opt
AU0852304-A03-2s1	AU0852304	chrX	145,207,069	145,208,670	1,601	22	13.7	-0.85	44.8	AU145.2	Opt
AU083504-A03-3s1	AU083504	chrX	122,143,675	122,144,235	560	11	19.6	-0.77	41.1	AU122.1	Opt
AU056803-A01-2s1	AU056803	chrX	24,022,009	24,026,159	4,150	41	9.9	-0.74	39.0	AU024.0	Opt
AU016803-A01-2s1	AU016803	chrX	19,375,516	19,379,418	3,902	41	10.5	-0.73	42.3	AU019.38	Opt
AU004803-A01-3s1	AU004803	chrX	8,746,458	8,748,798	2,340	27	11.5	-0.73	35.3	AU008.7	Opt
AU083504-A02-3s1	AU083504	chrX	96,493,478	96,495,197	1,719	26	15.1	-0.73	37.2	AU096.5	Opt
AU080803-A01-2s1	AU080803	chrX	30,048,831	30,055,680	6,849	122	17.8	-0.71	36.0	AU030.1	Opt
AU083504-A03-3s1	AU083504	chrX	154,047,946	154,057,440	9,494	147	15.5	-0.68	34.8	AU154.0A	Opt
AU062203-A03-2s1	AU062203	chrX	143,436,794	143,445,418	8,624	125	14.5	-0.65	38.2	AU143.4	Opt
AU055303-A03-2s1	AU055303	chrX	149,678,535	149,680,094	1,559	19	12.2	-0.52	53.9	AU149.7	Opt
AU004803-A03-3s1	AU004803	chrX	113,234,686	113,241,182	6,496	94	14.5	-0.40	38.1	AU113.2	Opt
AU0920301-A02-2s1	AU0920301	chrX	57,615,600	57,960,046	344,446	3789	11.0	0.47	38.6	AU057.6	Opt
AU016803-A02-2s1	AU016803	chrX	65,382,441	65,415,702	33,261	530	15.9	0.51	38.9	AU065.4	Opt
AU0920301-A02-2s1	AU0920301	chrX	105,379,514	105,381,047	1,533	28	18.3	0.64	37.8	AU105.4	Opt

Table A.6.b – Falsely called CNV by the NimbleGen and Optimized protocols.

ARRAY_ID	Sample	CHR	START	STOP	SIZE (bp)	Probes	Probes/kb	Mean_Log2	GC	Primer Set	CGH_Protocol
AU021503.1s1A02	AU021503	chrX	67,045,887	67,046,967	1,080	23	21.3	-0.85	41.2	AU067.0	NG
AU028903.1s1A02	AU028903	chrX	53,100,517	53,102,747	2,230	47	21.1	-0.63	47.6	AU053.10	NG
AU055503.1s1A02	AU055503	chrX	48,530,197	48,531,062	865	19	22.0	-0.62	47.9	AU048.5/AU048.533	NG
AU028903.1s1A02	AU028903	chrX	48,995,264	48,996,414	1,150	25	21.7	-0.60	52.7	AU048.995	NG
AU018003.1s1A02	AU018003	chrX	67,044,882	67,046,967	2,085	44	21.1	-0.60	39.0	AU067.0	NG
AU028903.1s1A02	AU028903	chrX	48,530,120	48,531,142	1,022	22	21.5	-0.56	48.3	AU048.5/AU048.533	NG
AU055503.1s1A02	AU055503	chrX	48,995,294	48,996,304	1,010	22	21.8	-0.55	52.3	AU048.995	NG
AU058103.1s1A02	AU058103	chrX	48,530,197	48,531,087	890	20	22.5	-0.55	48.1	AU048.5/AU048.533	NG
AU058503.1s1A02	AU058503	chrX	48,530,197	48,531,062	865	19	22.0	-0.52	47.9	AU048.5/AU048.533	NG
AU058103.1s1A02	AU058103	chrX	48,995,364	48,996,304	940	21	22.3	-0.50	52.1	AU048.995	NG
AU018003.1s1A02	AU018003	chrX	48,530,197	48,531,062	865	19	22.0	-0.49	47.9	AU048.5/AU048.533	NG
AU028903.1s1A03	AU028903	chrX	139,623,380	139,624,140	760	16	21.1	-0.49	49.9	AU139.6A	NG
AU018003.1s1A02	AU018003	chrX	48,995,439	48,996,284	845	18	21.3	-0.47	52.5	AU048.995	NG
AU002403.1s2A03	AU002403	chrX	143,163,522	143,175,922	12,400	211	17.0	-0.45	36.2	AU143.2	NG
AU028903.1s1A02	AU028903	chrX	48,276,384	48,278,284	1,900	39	20.5	-0.45	51.1	AU048.28	NG
AU0852304.1s1A03	AU0852304	chrX	129,713,932	129,715,267	1,335	27	20.2	-0.44	44.0	AU129.7	NG
AU021503.1s1A02	AU021503	chrX	53,126,892	53,127,752	860	17	19.8	-0.43	61.1	AU053.127	NG
AU021503.1s1A02	AU021503	chrX	48,530,217	48,531,062	845	18	21.3	-0.42	47.6	AU048.5/AU048.533	NG
AU028903.1s1A02	AU028903	chrX	48,220,021	48,221,221	1,200	26	21.7	-0.42	46.3	FTSJ1	NG
AU055503.1s1A02	AU055503	chrX	48,220,021	48,221,146	1,125	25	22.2	-0.41	45.6	FTSJ1	NG
AU058103.1s1A02	AU058103	chrX	48,219,896	48,221,146	1,250	27	21.6	-0.40	47.2	FTSJ1	NG
AU028903.1s1A03	AU028903	chrX	135,941,581	135,942,831	1,250	27	21.6	-0.39	52.1	AU135.94	NG
AU018003.1s1A03	AU018003	chrX	129,711,252	129,714,877	3,625	68	18.8	-0.37	41.1	AU129.7	NG
AU056003.1s1A03	AU056003	chrX	133,346,142	133,350,722	4,580	83	18.1	-0.37	39.9	AU133.3	NG
AU0780301.1s1A02	AU0780301	chrX	48,530,217	48,531,087	870	19	21.8	-0.37	47.8	AU048.5/AU048.533	NG
AU058103.1s1A03	AU058103	chrX	129,711,442	129,715,287	3,845	73	19.0	-0.35	41.3	AU129.7	NG
AU056803.1s1A03	AU056803	chrX	129,711,402	129,715,172	3,770	71	18.8	-0.33	41.3	AU129.7	NG
AU014803.1s1A02	AU014803	chrX	48,530,197	48,531,062	865	19	22.0	-0.33	47.9	AU048.5/AU048.533	NG
AU028903.1s1A03	AU028903	chrX	129,711,302	129,715,197	3,895	74	19.0	-0.32	41.4	AU129.7	NG
AU058503.1s1A03	AU058503	chrX	53,100,517	53,107,369	6,852	89	13.0	-0.29	46.2	AU053.10	NG
AU056003.1s1A02	AU056003	chrX	48,276,454	48,278,944	2,490	44	17.7	0.28	50.2	AU048.28	NG
AU021503.1s1A03	AU021503	chrX	114,330,510	114,333,265	2,755	57	20.7	0.31	65.6	AU114.3	NG
AU065404.1s1A03	AU065404	chrX	114,330,245	114,333,513	3,268	68	20.8	0.33	64.3	AU114.3	NG
AU056003.1s1A02	AU056003	chrX	53,100,517	53,102,557	2,040	43	21.1	0.33	48.3	AU053.10	NG
AU058103.1s1A01	AU058103	chrX	20,194,030	20,194,994	964	20	20.7	0.34	72.6	AU020.2	NG
AU0920301.1s1A02	AU0920301	chrX	53,099,892	53,102,717	2,825	53	18.8	0.36	48.1	AU053.10	NG
AU080803.1s1A03	AU080803	chrX	114,331,350	114,333,225	1,875	39	20.8	0.38	65.1	AU114.3	NG
AU067803.1s2A02	AU067803	chrX	69,201,942	69,203,705	1,763	38	21.6	0.39	58.6	AU069.2	NG
AU008504.1s1A02	AU008504	chrX	53,100,432	53,102,672	2,240	47	21.0	0.42	48.1	AU053.10	NG
AU018003.1s1A03	AU018003	chrX	114,331,470	114,333,225	1,755	36	20.5	0.47	65.1	AU114.3	NG
AU002403-A03-2s1	AU002403	chrX	143,165,878	143,166,968	178,000	23	21.1	-1.01	49.9	AU143.2	Opt
AU004803-A02-3s1	AU004803	chrX	53,127,122	53,127,609	487	8	16.4	-0.99	68.0	AU053.127	Opt
AU0920301-A02-2s1	AU0920301	chrX	67,761,197	67,761,872	675	9	13.3	-0.75	28.9	AU067.76	Opt
AU028903-A03-2s1	AU028903	chrX	148,992,605	148,992,915	310	8	25.8	-0.74	30.3	AU148.99	Opt
AU018003-A02-2s1	AU018003	chrX	67,045,802	67,046,892	1,090	19	17.4	-0.74	41.8	AU067.0	Opt
AU014803-A02-2s1	AU014803	chrX	67,040,517	67,046,967	6,450	84	13.0	-0.48	38.7	AU067.0	Opt
AU083504-A03-3s1	AU083504	chrX	122,920,791	122,923,535	2,744	52	19.0	0.34	59.2	AU122.9	Opt
AU008504-A02-2s1	AU008504	chrX	53,099,962	53,102,472	2,510	45	17.9	0.35	49.2	AU053.10	Opt
AU083603-A01-2s1	AU083603	chrX	19,300,606	19,302,311	1,705	34	19.9	0.37	61.6	AU019.3	Opt
AU014803-A02-2s1	AU014803	chrX	48,208,660	48,211,205	2,545	54	21.2	0.39	58.9	AU048.21	Opt
AU014803-A02-2s1	AU014803	chrX	69,202,022	69,205,124	3,102	64	20.6	0.39	58.6	AU069.2	Opt
AU055503-A02-2s1	AU055503	chrX	69,588,687	69,592,509	3,822	75	19.6	0.39	59.0	AU069.6	Opt
AU028903-A03-2s1	AU028903	chrX	135,940,186	135,941,906	1,720	36	20.9	0.41	58.4	AU135.94	Opt
AU1069302-A02-2s1	AU1069302	chrX	73,672,126	73,674,216	2,090	43	20.6	0.42	61.6	AU073.7	Opt
AU1038303-A01-2s1	AU1038303	chrX	41,077,342	41,079,467	2,125	40	18.8	0.43	63.4	AU041.1	Opt
AU1069302-A01-2s1	AU1069302	chrX	34,870,749	34,871,874	1,125	18	16.0	0.43	60.8	AU034.9	Opt
AU065404-A03-2s1	AU065404	chrX	114,331,095	114,333,305	2,210	44	19.9	0.44	65.2	AU114.3	Opt
AU018304-A01-2s1	AU018304	chrX	20,193,739	20,196,889	3,150	64	20.3	0.45	65.0	AU020.2	Opt
AU018304-A03-2s1	AU018304	chrX	117,133,965	117,135,095	1,130	24	21.2	0.47	65.0	AU117.1	Opt
AU067803-A02-2s1	AU067803	chrX	69,201,902	69,205,364	3,462	72	20.8	0.48	57.9	AU069.2	Opt
AU056803-A01-2s1	AU056803	chrX	20,193,659	20,197,128	3,469	71	20.5	0.48	63.0	AU020.2	Opt
AU018003-A01-2s1	AU018003	chrX	20,194,318	20,194,994	676	13	19.2	0.48	73.4	AU020.2	Opt
AU056803-A02-2s1	AU056803	chrX	52,966,038	52,967,318	1,280	22	17.2	0.50	60.5	AU052.97	Opt
AU1069302-A03-2s1	AU1069302	chrX	111,970,302	111,971,424	1,122	24	21.4	0.50	63.8	AU111.97	Opt
AU0920301-A02-2s1	AU0920301	chrX	69,202,002	69,204,994	2,992	62	20.7	0.51	58.4	AU069.2	Opt
AU1069302-A01-2s1	AU1069302	chrX	20,193,474	20,194,599	1,125	25	22.2	0.53	61.2	AU020.2	Opt
AU1069302-A03-2s1	AU1069302	chrX	114,331,010	114,333,350	2,340	47	20.1	0.54	65.0	AU114.3	Opt
AU1038303-A02-2s1	AU1038303	chrX	69,202,132	69,204,769	2,637	55	20.9	0.57	58.8	AU069.2	Opt
AU0920301-A03-2s1	AU0920301	chrX	114,331,445	114,333,225	1,780	35	19.7	0.59	65.7	AU114.3	Opt
AU0852304-A03-2s1	AU0852304	chrX	114,331,095	114,333,225	2,130	42	19.7	0.59	65.1	AU114.3	Opt
AU055303-A03-2s1	AU055303	chrX	135,819,673	135,820,148	475	11	23.2	0.68	57.1	AU135.82	Opt
AU0852304-A01-2s1	AU0852304	chrX	36,936,957	36,938,657	1,700	20	11.8	0.69	63.0	AU036.93	Opt
AU018304-A03-2s1	AU018304	chrX	133,757,887	133,758,510	623	14	22.5	0.73	60.5	AU133.8	Opt
AU056003-A02-2s1	AU056003	chrX	48,533,289	48,534,370	1,081	17	15.7	0.74	53.7	AU048.533	Opt
AU056003-A03-2s1	AU056003	chrX	128,964,095	128,964,615	520	12	23.1	0.98	55.6	AU128.96	Opt
AU018003-A03-2s2	AU018003	chrX	114,331,010	114,333,265	2,255	45	20.0	0.39	65.0	AU114.3	Opt