Recurrent 16p11.2 microdeletions in autism

Ravinesh A. Kumar1, Samer KaraMohamed1, Jyotsna Sudi1, Donald F. Conrad1, Camille Brune5, Judith A. Badner4, T. Conrad Gilliam1, Norma J. Nowak6, Edwin H. Cook Jr5, William B. Dobyns1,2,3 and Susan L. Christian1,*

1Department of Human Genetics, 2Department of Neurology, 3Department of Pediatrics and 4Department of Psychiatry, University of Chicago, Chicago, IL 60637, USA, 5Department of Psychiatry, University of Illinois at Chicago, Chicago, IL 60612, USA and 6Department of Cancer Genetics, Roswell Park Cancer Institute, Buffalo, NY 14236, USA

Received November 21, 2007; Revised and Accepted December 19, 2007

Autism is a childhood neurodevelopmental disorder with a strong genetic component, yet the identification of autism susceptibility loci remains elusive. We investigated 180 autism probands and 372 control subjects by array comparative genomic hybridization (aCGH) using a 19K whole-genome tiling path bacterial artificial chromosome microarray to identify submicroscopic chromosomal rearrangements specific to autism. We discovered a recurrent 16p11.2 microdeletion in two probands with autism and none in controls. The deletion spans ~500-kb and is flanked by ~147-kb segmental duplications (SDs) that are >99% identical, a common characteristic of genomic disorders. We assessed the frequency of this new autism genomic disorder by screening an additional 532 probands and 465 controls by quantitative PCR and identified two more patients but no controls with the microdeletion, indicating a combined frequency of 0.6% (4/712 autism versus 0/837 controls; Fisher exact test \( P = 0.044 \)). We confirmed all 16p11.2 deletions using fluorescence in situ hybridization, microsatellite analyses and aCGH, and mapped the approximate deletion breakpoints to the edges of the flanking SDs using a custom-designed high-density oligonucleotide microarray. Bioinformatic analysis localized 12 of the 25 genes within the microdeletion to nodes in one interaction network. We performed phenotype analyses and found no striking features that distinguish patients with the 16p11.2 microdeletion as a distinct autism subtype. Our work reports the first frequency, breakpoint, bioinformatic and phenotypic analyses of a de novo 16p11.2 microdeletion that represents one of the most common recurrent genomic disorders associated with autism to date.

INTRODUCTION

Autism (OMIM 209850) is a childhood neurodevelopmental disorder apparent by 3 years of age and characterized by qualitative impairments in reciprocal social interaction, deficits in verbal communication, restricted interests and repetitive behaviors. Autism comprises the severe end of the autism spectrum disorders (ASD), which also include Asperger syndrome, pervasive developmental disorder not otherwise specified and Rett syndrome. Prevalence rates for autism and ASD are 0.2 and 0.6%, respectively. Firm evidence of a genetic basis for autism has been demonstrated by twin studies that show 60–91% concordance rates in monozygotic twins (1,2), making autism the most heritable of all complex neuropsychiatric disorders. The search for autism susceptibility loci has involved several approaches including genome-wide linkage analysis, and association and mutation studies of candidate genes (3,4). However, the molecular basis of autism remains largely unknown.

A complementary approach to identify genetic susceptibility variants for autism involves searching for chromosomal abnormalities in patients. Microscopically visible chromosomal rearrangements such as deletions and duplications have been identified in 1–3% of autism subjects (5). The most

© 2007 The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
frequent are maternal 15q11-13 duplications; other known recurrent rearrangements include duplications of 17p12, and deletions of 7q11.23, 17p11.2 and 22q11.2 (3). Recent high resolution microarray-based studies in autism have identified a number of novel submicroscopic copy number variants (CNVs) including both deletions and duplications (6,7), although all are rare and few are recurrent. We hypothesize that more remain to be discovered.

We undertook a genome-wide investigation of autism-specific CNVs in 180 autism probands and 372 control subjects using a 19K whole-genome tiling path bacterial artificial chromosome (BAC) microarray and identified a 16p11.2 microdeletion in two patients but no controls. We screened an additional 532 autism probands and 465 controls by quantitative PCR (qPCR), and identified two additional probands deleted for 16p11.2. Here we provide phenotype data for patients with microdeletions at this locus, present detailed analyses and quantitative real-time PCR (qPCR) confirmed the new microdeletions by hybridization (FISH). Microsatellite analyses and quantitative real-time PCR (qPCR) confirmed paternity and showed that the deletions were de novo and maternally derived (Fig. 1C–F). Chromosome analysis was reported as normal for HI2977 and not listed for HI2467. We demonstrated that the deletion in HI2977 was consistent with our previously estimated telomeric and centromeric breakpoint boundaries at ~29.6 and ~30.2 Mb, respectively, which are in close proximity to the flanking SDs, thereby supporting our hypothesis of SD-mediated recombination.

To better map the deletion breakpoints, we performed replicate hybridization experiments in subject HI2467 and combined these results using segmentation analysis (Fig. 3C). Although complicated by the very high homology between the SDs, these results demonstrate a sharp breakpoint at the centromeric or internal end of the telomeric SD (arrow 1). However, the centromeric SD has two apparent breakpoints located near the internal (arrow 2) and external (arrow 3) ends of the repeat with an intermediate log2 ratio between them. We interpret the intermediate log2 ratio to reflect reduction of sequence within the SD from the normal four copies to three copies, as many of the probes from the SD regions have very high homology and cannot be reliably distinguished. This analysis unambiguously maps the breakpoints of the 16p11.2 microdeletion to the edges of flanking SDs, consistent with our previously estimated telomeric and centromeric breakpoint boundaries.

The microdeletion contains ~25 genes (Fig. 3A), with another four located within the SDs. To study the potential contribution of individual or sets of genes in the microdeletion to autism, we investigated whether these genes interact biologically using the Ingenuity Pathway Analysis (IPA) tool, which uses known functional relationships and interactions between gene products reported in the literature. Twelve of the 25 genes were mapped to a single genetic network (Fig. 4) that included genes involved with cell-to-cell signaling and interaction. The pathways for 3 of the 25 genes (D Coul2A, MAPK3 and ALDOA) included post-synaptic density genes that have been hypothesized to underlie autism (8).

We analyzed the behavioral profile on all patients with 16p11.2 microdeletions based on the phenotype data currently

RESULTS

We investigated 180 unrelated autism subjects and 372 control subjects by array comparative genomic hybridization (aCGH) using a whole-genome tiling path microarray comprised of ~19 000 BACs. In this initial study, we identified a recurrent ~500-kb microdeletion in two unrelated patients with autism (HI0646 and HI0624) that was not detected in 372 control subjects, indicating a frequency of 1.1% (2/180 autism versus 0/372 controls; Fisher exact test $P = 0.11$) (Fig. 1A and B). The microdeletion has not been reported as a CNV in the autosomal region of any disease gene. We confirmed both deletions by fluorescence in situ hybridization (FISH). Microsatellite analyses and quantitative real-time PCR (qPCR) confirmed paternity and showed that the deletions were de novo and maternally derived (Fig. 1C–F). Chromosome analysis was reported as normal for HI0624 and not listed for HI0646 on the Autism Genetics Resource Exchange (AGRE) website (http://www.agre.org/). A summary of our molecular studies is presented in Table 1. We did not identify the 16p11.2 microdeletion in an affected sibling of HI0646, or in one affected and one unaffected sibling of HI0624.

We assessed the frequency of the 16p11.2 microdeletion in autism by screening an additional 532 probands and 465 controls using qPCR. We identified two additional probands (HI2467 and HI2997) and no controls deleted for 16p11.2, giving a frequency of 0.4% in this cohort (2/532 autism versus 0/465 controls; Fisher exact test $P = 0.50$). The frequency of the 16p11.2 microdeletion combining both cohorts is 0.6% (4/712 autism versus 0/837 controls; Fisher exact test $P = 0.044$). We confirmed the new microdeletions by FISH, microsatellite analysis and aCGH (Table 1), also confirming paternity (data not shown). Chromosome analysis was reported as normal for HI2977 and not listed for HI2467. We demonstrated that the deletion in HI2977 was de novo and paternally derived; his affected sibling did not carry the microdeletion. The last family proved to be more complex; the deletion in HI2467 was paternally inherited and apparently de novo as the deletion was not found in either parent (or in her unaffected sister). However, her affected brother also had the microdeletion, demonstrating post-zygotic mosaicism for the 16p11.2 deletion and complete association of the microdeletion with autism in this family. The pedigree structures of all four families are shown in Figure 2. We also identified the reciprocal duplication product of the 16p11.2 microdeletion in one subject with autism (HI0128), her unaffected mother, and two controls (Table 1).

The genomic structure and gene content of the 16p11.2 microdeletion were interrogated using the UCSC Browser (http://genome.ucsc.edu; hg 18). Segmental duplications (SDs) of ~147 kb and >99% identity were found to flank each end of the microdeletion (Fig. 3A), a common characteristic of genomic disorders. We tested the hypothesis that the 16p11.2 microdeletion is mediated by recombination between the SDs by using a custom-designed chromosome 16p11.2 oligonucleotide array. Although the array data produced low signal-to-noise ratio due to the highly repetitive genomic structure of our target region, we could approximate the breakpoints in all five patients (Fig. 3B). We estimated the telomeric and centromeric breakpoint boundaries at ~29.6 and ~30.2 Mb, respectively, which are in close proximity to the flanking SDs, thereby supporting our hypothesis of SD-mediated recombination.

To better map the deletion breakpoints, we performed replicate hybridization experiments in subject HI2467 and combined these results using segmentation analysis (Fig. 3C). Although complicated by the very high homology between the SDs, these results demonstrate a sharp breakpoint at the centromeric or internal end of the telomeric SD (arrow 1). However, the centromeric SD has two apparent breakpoints located near the internal (arrow 2) and external (arrow 3) ends of the repeat with an intermediate log2 ratio between them. We interpret the intermediate log2 ratio to reflect reduction of sequence within the SD from the normal four copies to three copies, as many of the probes from the SD regions have very high homology and cannot be reliably distinguished. This analysis unambiguously maps the breakpoints of the 16p11.2 microdeletion to the edges of flanking SDs, consistent with our previously estimated telomeric and centromeric breakpoint boundaries.

The microdeletion contains ~25 genes (Fig. 3A), with another four located within the SDs. To study the potential contribution of individual or sets of genes in the microdeletion to autism, we investigated whether these genes interact biologically using the Ingenuity Pathway Analysis (IPA) tool, which uses known functional relationships and interactions between gene products reported in the literature. Twelve of the 25 genes were mapped to a single genetic network (Fig. 4) that included genes involved with cell-to-cell signaling and interaction. The pathways for 3 of the 25 genes (D Coul2A, MAPK3 and ALDOA) included post-synaptic density genes that have been hypothesized to underlie autism (8). We analyzed the behavioral profile on all patients with 16p11.2 microdeletions based on the phenotype data currently

Downloaded from https://academic.oup.com/hmg/article-abstract/17/4/628/2356715 by guest on 18 January 2019
available in the AGRE database (Table 2). We found no striking features that distinguish subjects with the 16p11.2 microdeletion as a distinct subtype of autism, although a majority had a trend toward behavioral difficulties involving aggression and overactivity. While all patients in the AGRE database have had physical examinations, data regarding facial appearance were entered for only one sibling pair. The boy with the 16p11.2 microdeletion (HI0646) had minor facial dysmorphism consisting of downslanting palpebral fissures, prominent ears and broad nasal root. However, his brother with autism who did not have the microdeletion also had minor facial dysmorphism, which suggests that the minor facial dysmorphism may be familial rather than related to the microdeletion.

We also reviewed available phenotype data for subjects with duplications of 16p11.2. One girl (HI0128) with autism and duplication 16p11.2 had severe compulsions and rituals consistent with obsessive-compulsive disorder. However, she also had a 1.0-Mb deletion of 7q31.2 and 890-kb duplication of 16q22.1, making interpretation of her phenotype difficult.

Figure 1. Identification of a new recurrent de novo 16p11.2 microdeletion in autism. (A and B) BAC aCGH demonstrates the presence of a deletion at 16p11.2 in two patients (black arrows highlight the vertical line of dots that indicate the presence of a deletion). The aCGH results show the log2 ratio of the reference versus patient DNA on the vertical axis. Each individual BAC is represented as a single blue dot. Known CNVs are represented as green dots. The horizontal axis shows the position of each BAC along the chromosome. (C and D) Microsatellite analysis demonstrates that the deletions are de novo and maternally inherited. (E and F) FISH results confirm the microdeletion at 16p11.2. BAC RP11-1107E4 (green arrow) and BAC RP11-114A15 (red arrow) show two copies of the green probe and a single copy of the red probe indicating a deletion of RP11-114A15. A similar result was obtained using RP11-74E23 (red) and RP11-455F5 (green) (data not shown).
only the 16q22.1 duplication was de novo. Her mother (HI0126), who carries both the 7q31.2 deletion and 16p11.2 duplication, has depressive and anxiety symptoms (onset 13–19 years for both) without formal diagnosis of depression or anxiety disorder, as well as undiagnosed learning disabilities and other unspecified behavioral problems. The two control subjects (04C34861 and 04C36902) with duplication 16p11.2 noted several minor behavioral abnormalities such as compulsions, generalized anxiety, specific phobias and panic attacks, but no mental health diagnoses. While both had compulsions—defined as an inability to resist certain activities in the presence of obsessions—the frequency was nominally or not significant in comparison to other controls.

**DISCUSSION**

The present study reports the recurrent nature and first frequency assessment of a novel and typically de novo 16p11.2 microdeletion that is significantly associated with autism: 4/712 autism versus 0/837 controls; Fisher exact test \( P = 0.044 \). We are aware of no other studies of autism that have

![Figure 2. Pedigree structure of families with 16p11.2 microdeletions. Squares indicate males and circles indicate females. Individuals affected with autism are presented as black symbols while white symbols are unaffected. The family ID is presented at the top of each pedigree. Individual IDs are presented beneath each symbol. Individuals with the 16p11.2 microdeletion are indicated.](https://academic.oup.com/hmg/article-abstract/17/4/628/2356715)

![Table 1. Summary of molecular studies in patients with 16p11.2 rearrangements](https://academic.oup.com/hmg/article-abstract/17/4/628/2356715)
reported a single microdeletion or duplication variant with $P < 0.10$. The frequency of the 16p11.2 microdeletion reported in this study is likely an underestimate of the actual occurrence in autism as we tested primarily multiplex families, which would be expected to have a higher frequency of inherited and lower frequency of de novo abnormalities compared with simplex families (6). Our study suggests that the 16p11.2 microdeletion is one of the most common recurrent genomic disorders associated with autism.

While our study is the first to report the recurrent nature of 16p11.2 microdeletions in autism, it is not the first report of this deletion. The 16p11.2 microdeletion was recently identified in a female with Asperger syndrome from a simplex family, without further details regarding the phenotype (6). A similar deletion was reported in monozygotic twins presenting with seizure disorder, mild mental retardation and aortic valve anomalies (9). Both twins and a third brother without the microdeletion had additional congenital anomalies, pointing to another genetic disorder segregating in the family that may account for the aortic valve anomalies. Neither twin had formal testing for autism, although one had more severe expressive and receptive language deficits than expected from his overall level of functioning, as is characteristic of individuals with ASD (Supplementary Materials, Data S1 for details).

Finally, one patient with a large ~8.7-Mb deletion of 16p11.2–p12.2 that overlaps the region deleted in our patients was recently reported (10). This deletion encompasses ~104 genes, so that a more severe phenotype would be expected. Accordingly, she was reported to have developmental abnormalities most consistent with moderate mental retardation, facial dysmorphism including Robin sequence, multiple hand anomalies and unexplained syncope; however, no mention of autism was included. The additional abnormalities described in this girl are not likely due to deletion of genes in our 16p11.2 microdeletion, as only one proband in our study showed minor and non-specific craniofacial dysmorphism. However, follow-up studies to obtain additional phenotype

---

**Figure 3.** Breakpoint characterization of a recurrent 16p11.2 microdeletion in autism. (A) Schematic of the 16p11.2 microdeletion. The location of the microdeletion is indicated on a diagram of Chromosome 16 followed by a horizontal black line representing an enlargement of the region. Large blue arrows represent the ~147-kb SDs. The genes located within this region, not drawn to scale, are listed above the black line. BACs labeled for FISH are represented by green (overlap SDs) and red (within microdeletion) lines. Known CNVs are represented by blue lines. (B) The approximate breakpoints (dashed lines) are reported for all five patients using a high density custom chromosome 16 oligonucleotide microarray. The short normal flanking regions outside the dashed lines contain probes that are located mostly above the single horizontal red line, whereas the region between the dashed lines contain a large number of probes below the horizontal line (albeit noisy data), indicating copy number loss that is consistent with the FISH analyses. The array results show the log$_2$ ratio of the reference versus patient DNA on the vertical axis. Each individual probe is represented as a single black dot. The horizontal axis shows the position of each probe along the chromosome. Vertical dashed lines represent the estimated telomeric and centromeric breakpoint boundaries. (C) The log$_2$ ratios from two replicate experiments on HI2467 were averaged to create a lower noise data set. The red line shows a 1000-probe moving average of these data, while the 4 underlying black lines depict the mean probe intensity of (from left to right): the telomeric segmental duplication (SD), the unique deleted region, the centromeric SD, and copy-normal sequence. A sharp breakpoint at the centromeric or internal end of the telomeric SD (arrow 1) is indicated, as well as two apparent centromeric breakpoints located near the internal (arrow 2) and external (arrow 3) ends of the repeat.
data on these and other patients with the microdeletion are indicated.

Our data raise several important questions regarding the 16p11.2 microdeletion. First, could the deletion be non-pathogenic, with the observed difference in frequency between the autism and control cohorts simply a chance finding? We think this is unlikely considering all of the data available. The microdeletion was found in none of 837 control patients in our study, in none of the control subjects listed in the Database of Genomic Variants (although the platforms varied), and in none of 360 patients with mental retardation (see below). It results in loss of at least 25 genes, a sufficiently large number that a phenotype would be expected, particularly considering that the list includes several good candidate genes (also see below). Finally, our $P$-value was significant ($P = 0.044$) when considering our entire cohort.

Next, is the 16p11.2 microdeletion specific for ASD, or can it cause other developmental disorders as well? The deletion was found in none of 360 patients with mental retardation studied using a whole-genome tiling BAC array (11,12). Several other large CNV studies of mental retardation have been done with no deletions of 16p11.2 found, although the platforms used included few or no probes in this region (13–15). While further studies are needed, these studies already suggest that the 16p11.2 microdeletion is not associated with mental retardation, or alternatively causes mental retardation less often than autism.

Finally, does the 16p11.2 microdeletion cause ASD by itself or only serve as a major risk factor in combination with other susceptibility variants? Only future studies will answer this question, as only a few genes (or non-genetic contributing factors) causing or contributing to autism have so far been identified. Perhaps the microdeletion alone is sufficient to cause mild ASD such as Asperger syndrome (6), while additional susceptibility variants are needed to cause severe autism. Our novel observation of de novo CNVs in patients with autism but not in several of their affected sibs suggests that other ASD risk factors may also be segregating in these families. If so, then the 16p11.2 microdeletion may represent a susceptibility variant contributing to the disorder in combination with other genetic variants shared or unshared by the affected siblings.

**Figure 4.** Functional network analysis of genes on 16p11.2. Genes are represented as nodes and the biological relationship between two nodes is represented as an edge (line). All edges are supported by at least one reference from the literature, from a textbook or from canonical information stored in the Ingenuity Pathways Knowledge Base. Nodes are displayed using various shapes that represent the functional class of the gene product. Shaded nodes represent genes residing within the microdeletion.
Considering all of these observations together, we hypothesize that 16p11.2 microdeletions are a risk factor for ASD generally, and may be causal for mild ASD in some families. Further, our study raises interesting questions that have not been discussed in previous reports of CNV in autism, such as the role of genetic heterogeneity within families. Had the current study been performed primarily in simplex families, the discordant genotypes between affected siblings would not have been detected. The present work emphasizes the importance of studying multiplex as well as simplex families for the identification of both inherited and spontaneous mutations.

For a few deletion syndromes such as 22q11.2, both the deletion and reciprocal duplication have been associated with autism (16–18) (Christian et al., submitted). It is not yet clear whether this will be true for the 16p11.2 locus as well. The reciprocal duplication product of the 16p11.2 microdeletion has been reported in a mother and daughter with mild mental retardation (19) and in two more unrelated patients with autistic behavior (20). However, both studies used lower resolution techniques such as karyotyping and FISH so that the duplicated region may be larger than the /C24 500-kb segment highlighted in the current study. Our work and that of others suggest that analysis of 16p11.2 duplications should be explored further in autism as well as obsessive-compulsive disorder.

Several of the genes that reside within the 16p11.2 microdeletion represent promising candidates for autism based on known expression and functional data. **MAPK3** is expressed in human fetal and adult brains; **Mapk3**−/− mice display abnormal avoidance behavior, hyperactivity, reduced long-term potentiation and immune system abnormalities (21). **MAZ** is expressed in several tissues with the highest expression in brain found in motor and midfrontal cortex. **MAZ** directly regulates genes involved in GABA signaling (22), neuronal differentiation (23) and the serotonin signaling pathway.
The same phenomenon may be found for other autism loci. Heterogeneity in these families, and leads us to hypothesize thatlings without the microdeletion supports a role for genetic het-
cognitive, language and social impairments. We hypothesize
evidence for the involvement of 16p11.2 in other
recent 16p11.2 microdeletion in autism and ASD, and strength-
et (32).

effects of maternal immune activation on fetal brain develop-
and has recently been identified as a key mediator of the
response to disease injury in the nervous system and regulation
to play a role in brain development, learning and memory,
response to disease injury in the nervous system and regulation
of the balance between neurogenesis and gliogenesis (30,31),
and has recently been identified as a key mediator of the
effects of maternal immune activation on fetal brain develop-
ment (32).

In summary, our work provides the first evidence of a recur-
current 16p11.2 microdeletion in autism and ASD, and strength-
en evidence for the involvement of 16p11.2 in other
cognitive, language and social impairments. We hypothesize
that perturbations in the network of genes identified in this
study, especially those within the microdeletion region,
underlie the autism phenotype. Our discovery of affected sib-
lings without the microdeletion supports a role for genetic het-
erogeneity in these families, and leads us to hypothesize that
the same phenomenon may be found for other autism loci and potential for other common complex disorders.

MATERIALS AND METHODS

Subjects

The Institutional Review Board at the University of Chicago
approved this study. The original 180 AGRE samples were
simplex in 19 and multiplex in 161 families. For the additional
230 AGRE samples, 11 were simplex and 219 multiplex
families, while the 302 National Institutes of Mental Health
(NIMH) Genetics Initiative autism samples included 57
simplex and 245 multiplex families. Therefore, the total
cohort of 712 autism probands consisted of 87 simplex and
625 multiplex families. DNA from 837 control subjects was
selected from the NIMH Genetics Initiative Control sample
set, and were used for both the aCGH (n = 372) and qPCR
(n = 465) experiments. All control subjects were characterized
for Axis I disorders. Although individuals were characterized
for Axis I disorders, individuals meeting criteria for diagnosis
of an Axis I disorder were not excluded from the sample.
Autistic traits were not assessed in this sample.

Phenotypic evaluation

Detailed data of all subjects with autism, their parents and
some unaffected siblings are available from the AGRE
website, and were reviewed, compiled and analyzed by the
authors (C.B. and E.H.C.). These data included detailed
medical histories and physical exams, and results from the
Autism Diagnostic Interview—Revised (33,34), the Autism
Diagnostic Observation Schedule (35) and other cognitive
testing. Data represented as missing in Table 2 was unavail-
able in the AGRE database.

BAC CGH array

The minimal tiling RPCI BAC array contains ~19 000 BAC
clones that were chosen by virtue of their STS content, paired
BAC end-sequence and association with heritable dis-
orders and cancer. The backbone of the array consists of
4603 BAC clones that were directly mapped to specific,
single chromosomal positions by FISH. Each clone was
printed in duplicate on amino-silanated glass slides type A+
(Schott Nexterion typeA+) using a MicroGrid II TAS
arrayer (Agena Discoveries). The BAC DNA printing solu-
tions were prepared from sequence connected RPCI-11
BACs by ligation-mediated PCR as described previously
(36). The BAC DNA spots are ~80 µm in diameter with a
150 µm center-to-center spacing creating an array of
~38 000 elements. The printed slides were dried overnight
and UV-crosslinked (350 mJ) in a Stratalinker 2400 (Strata-
gene) immediately before hybridization. A complete list of
the BAC clones printed on the 19K array can be found on
the RPCI microarray facility website. Reference and test
sample genomic DNA (1 µg each) was fluorescently labeled
using the BioArray CGH Labeling System (Enzo Life
Sciences). For this study, each control DNA was used as an
independent test sample against a reference DNA pool of 20
sex-mismatched individuals, and each AGRE patient was
used as the test sample against the same sex-mismatched refer-
dence DNA pools. The DNA was denatured in the presence of
random primers at 99°C for 10 min in a thermocycler, quickly
cooled to 4°C and transferred onto ice. The samples were
labeled with a dNTP-cyanine 3 mix (or dNTP-cyanine 5)
and Klenow followed by incubation overnight at 37°C. Unin-
corporated nucleotides were removed using a QIAquick PCR
purification column (Qiagen) and the labeled probe was
eluted with 25 µl washes. Prior to hybridization, the test
and reference probes were resuspended in 110 µl SlideHyb
Buffer #3 (Ambion) containing 5 µl of 20 µg/µl Cot-1 DNA
and 5 µl of 100 µg/µl Yeast tRNA (Invitrogen), heated to
95°C for 5 min and placed onto ice. Hybridization to the 19K
BAC microarrays was performed for 16 h at 55°C using a
GeneTAC hybridization station (Genomic Solutions, Inc.) as
described previously (37). After hybridization, the slides
were automatically washed in the GeneTAC station with redu-
cing concentrations of SSC and SDS. The slides were scanned
using a GenePix 4200AL Scanner (Molecular Devices) to gen-
erate high-resolution (5 µm) images for both Cy3 (test) and
Cy5 (control) channels. Image analysis was performed using
the ImaGene (version 6.1.0) software from BioDiscovery,
Inc. The log2 test/control ratios were normalized using a
sub-grid Loess correction. Mapping data for each BAC was
acquired from the March 2006 version of the human genome
and added to the resulting log2 test/control values. BACs over-
lapping SDs or published CNVs were flagged (38–41). To
select abnormal loci for further analysis, the aCGH data
were analyzed using two methods. (i) All clones outside of

(24). DOC2A is expressed predominantly in the brain, loca-
lizes to synaptic vesicles, and is hypothesized to regulate
synaptic activity though Ca-dependent mechanisms (25), con-
sistent with the proposed role of Ca2+ signaling in autism (26).
Doc2a−/− mice display alterations in synaptic transmission
and long-term potentiation as well as learning and behavioral
deficits that include abnormal passive avoidance behavior
(27). SEZ6L2 is expressed at high levels in human brain;
Sez6l2−/− mice display behavior and neurological abnor-
malities that include impaired motor coordination and abnormal
excitatory postsynaptic currents (28). HIRIP3 is a widely
expressed gene that interacts directly with HIRA, a major can-
didate for the DiGeorge syndrome and related developmental
disorders (29). IL6, which is not located within the deletion,
was found as a node in the identified network. IL6 is known
to play a role in brain development, learning and memory.

Human Molecular Genetics, 2008, Vol. 17, No. 4

635

Downloaded from https://academic.oup.com/hmg/article-abstract/17/4/628/2356715 by guest on 18 January 2019
five SD were identified using software developed by Shuang Liu. The autosomes and each sex chromosome were analyzed separately. All pseudautosomal BACs were included with the autosome analyses. (ii) Manual curation was also performed to identify loci with multiple abnormal clones that were detected as a single clone using a threshold of five SD. All loci with two contiguous abnormal BACs were then screened against the control data. Only CNVs not present in the controls were studied further.

Oligonucleotide CGH arrays and breakpoint analyses

We characterized the breakpoints using a NimbleGen Systems (Madison, WI) custom oligonucleotide Chr. 16 array (HG18; NCBI build 36) covering 29 318 000–30 304 400 Mb designed under the following specifications: isothermal probes (50–74 bp) with target Tm of 76; synthesis cycle limit of 148 cycles; filtered at average 15-mer frequency of 50; interval spacing at 2 bp (mean) and 3 bp (median). We also employed a fine tiling commercial design for chr16 from human genome (hg18; NCBI Build 36) designed as above except that interval spacing was 158 bp (mean) and 165 bp (median). All hybridizations were performed as previously described (42). Our estimates of the telomeric breakpoint boundaries were consistent with the data generated from a lower density Chromosome 16 microarray (Nimble-Gen); however, the centromeric boundary could not be accurately recovered due to a paucity of probes in this region on the lower resolution array (data not shown). A Gaussian mixture model approach was used to identify change points between two adjacent segments of different copy number. Briefly, a large window was manually designed that spanned an equal number of probes in each segment. The model was then constructed as a mixture of two probe sets, with the mixture proportions determined by the (unknown) location of the breakpoint. Data from the probes spanning the interval from the left edge of the design window to the breakpoint parameterize one of the mixture distributions; the other set of probes, the other distribution. The change point was then identified by evaluating the likelihood of the data across a grid of points (one point per interval between probes), re-estimating the other parameters in the model at each point. The maximum likelihood estimates for the three change points apparent in Figure 1H are: 29 557 173–29 559 112 (change point 1); 30 107 181 (change 2); 30 253 847–30 253 852 (change 3). All coordinates are relative to NCBI 36. Strikingly, these estimated change points correspond closely to the edges of known SDs: a 147-kb duplication on the centromeric flank of the deletion covers the sequence from 30 107 356–30 254 364. The other half of this pair is located at 29 368 017–29 514 353; however, there is a third, 22-kb duplication from 29 534 625–29 557 497, whose centromeric edge corresponds exactly with change point 1 identified here. It is possible that a CNV within this 22-kb SD is creating additional changes between HI2467 and the reference, thus complicating the analysis, or that the 22-kb SD is directly involved in NAHR with the centromeric 147-kb SD. Moving averages of the Nimblegen data were constructed with the ‘caTools’ package written in the ‘R’ programming language (http://cran.r-project.org/src/contrib/Descriptions/caTools.html).

Quantitative real-time PCR

qPCR was performed using two TaqMan® primer/probe amplicon sets designed against the 16p11.2 microdeletion region (Assays-by-Design; Applied Biosystems) as well as a reference primer/probe set (TaqMan® RNaseP gene, Applied Biosystems). Sequences for the 16p11.2-region primers and FAM-labeled probes are as follows: SEZ6L2-F, 5′-CCTTCTCTTCCCCACAAAGG-3′; SEZ6L2-R, 5′-TG GACAGCCTGGTTCTCTCT-3′; SEZ6L2-probe, 5′- CCTC TGCTCAACCTCTA-3′; C16orf54-F, 5′-ACTGCCCTCA GCGAAGATG-3′; C16orf54-R, 5′-GCCAACCCCG ACC-3′; C16orf54-probe, 5′-CTTCCCAGGCTCCC-3′. Reactions were run in triplicate in 6 ul total volume containing 40 ng genomic DNA, 2X TaqMan Universal Master Mix (ABI part number 4304437) and 20X primer/probe mix. Each assay included a no-template control and three controls subjects each validated by FISH for 16p11.2 copy number (i.e. one, two and three copies, respectively). We repeated qPCR for every subject showing copy number alterations at 16p11.2 by re-running each sample an additional six times for the target and reference probes. Each experiment was performed using a 384-well optical PCR plate and the 7900 HT Real-Time PCR System (Applied Biosystems) under the following cycling conditions: initial step at 50°C for 2 min, denaturation at 95°C for 10 min, followed by 40 cycles at 95°C for 15 s. Data analysis was performed using the SDS v2.1 software’s Relative Quantitation feature. Copy number of the 16p11.2 target region was defined as 2T −ΔCT, where ΔCT is the difference threshold cycle number for the test and reference loci. For each subjects showing copy number alterations at the target loci, we confirmed our finding by re-running each sample six times for the target and reference probes.

Fluorescence in situ hybridization

Lymphoblastoid cell lines for each proband and selected family members were acquired from the Rutgers University Cell and DNA Repository and cultured using standard techniques. RPCI-11 BACs that defined the boundaries of the 16p11.2 microdeletion were hybridized on metaphase spreads from human genome (hg18; NCBI Build 36) covering 29 318 000–30 304 400 Mb using either Spectrum Green or Spectrum Orange fluorescent dyes (Abbott Labs). FISH was performed using standard techniques. Slides were analyzed with a Zeiss Axioplan 2 fluorescent microscope with a cooled CCD camera and Applied Imaging CytoVision v3.7 software.

Microsatellite analysis

Microsatellites were selected from the UCSC Genome Browser microsatellite or simple repeat tracks and primers were designed using the MIT Primer3 program. For a single reaction, a master mix of 1 μl 10× PCR buffer with 15 mM MgCl (2), 1 μl 10 mm dNTP, 0.1 μl Ampli Taq Gold enzyme, 0.8 μl 10 μM primers (forward and reverse) and...
Pathway analyses

Gene accession numbers for our ‘focus’ genes (i.e. the 25 genes residing within the 16p11.2 microdeletion) were imported into IPA software (Systems, Mountain View, CA, USA; https://analysis.ingenuity.com/pa), a web-delivered application that enables discovery, visualization and exploration of molecular interaction networks. Genetic networks were ranked by a score that takes into account the number of focus genes and size of networks. The probability that a collection of genes equal to or greater than the number in a network could be achieved by chance alone was calculated. A score of three indicates that there is a 1/1,000 chance that the focus gene(s) are in a network due to random chance. Scores of three or higher have a 99.9% confidence level of not being generated by random chance alone and was used as the cut-off for identifying gene networks.

Statistical analysis

To assess the significance of the difference of the frequency of the 16p11.2 microdeletion between autistic subjects and controls, the Fisher Exact Test was used, as implemented at http://www.matforsk.no/ola/fisher.htm.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENTS

We gratefully acknowledge the resources provided by the Autism Genetic Resource Exchange ( AGRE ) Consortium and the participating AGRE patients and families. Additional autism families and the control samples were acquired as part of the NIMH Center for Collaborative Genetic Studies on Mental Disorders.

Conflict of Interest statement. None declared.

FUNDING

This study was supported by the National Institutes of Health, National Institute of Neurological Diseases and Stroke (1R01NS51812 to S.L.C.) and National Institute of Mental Health (1R01MH64457-01 to T.C.G), the National Alliance for Autism Research to S.L.C., and an Autism Speaks Post-Doctoral Fellowship Award to E.H.C. Jr. and C.B. AGRE is supported, in part, by a grant to D.H.G. from the National Institute of Mental Health (NIMH) (MH64547).

REFERENCES

19. Engelen, J.J., de Die-Smulders, C.E., Direcks, R., Verhoeven, W.M., Tuinier, S., Curs, L.M. and Hamers, A.J. (2002) Duplication of 7.1 μl sterile H2O was prepared. One microliter DNA (10 ng/μl) was added to each reaction. The PCR reaction was run in ABI 9700 thermocyclers using the following conditions: hot start at 94°C for 10 min, 94°C for 30 s, 55°C for 30 s, 72°C for 30 s for 35 cycles followed by a final extension step at 72°C for 10 min. Samples were analyzed on an ABI 3730 XL DNA sequencing analyzer and processed using GeneMapper 3.7 software (Applied Biosystems).


