Copy Number Variation Characteristics in Subpopulations of Patients With Autism Spectrum Disorders

Anna Bremer,1,2* MaiBritt Giacobini,1,2,3 Mats Eriksson,4 Peter Gustavsson,1,2,5 Viviann Nordin,6 Elisabeth Fernell,4,7 Christopher Gillberg,8 Ann Nordgren,1,2,5 Asa Uppströmer,5 Britt-Marie Anderlid,1,2,4,5 Magnus Nordenskjöld,1,2,5* and Jacqueline Schoumans1,2,5,9

1Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden
2Center for Molecular Medicine, Karolinska Institutet, Stockholm, Sweden
3PRIMA Child and Adolescent Psychiatry AB, Stockholm, Sweden
4Department of Neuropediatrics, Astrid Lindgren Children’s Hospital, Stockholm, Sweden
5Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden
6Department of Neuropediatrics, Sachs’ Children’s Hospital, Södersjukhuset, Stockholm, Sweden
7Autism Centre for Young Children, Karolinska University Hospital, Stockholm, Sweden
8Institute of Neuroscience and Physiology, Child and Adolescent Psychiatry, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
9Department of Medical Genetics, Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland

Received 10 August 2010; Accepted 19 October 2010

Autism spectrum disorders (ASDs) are a heterogeneous group of disorders with a complex genetic etiology. We used high-resolution whole genome array-based comparative genomic hybridization (array-CGH) to screen 223 ASD patients for gene dose alterations associated with susceptibility for autism. Clinically significant copy number variations (CNVs) were identified in 18 individuals (8%), of which 9 cases (4%) had de novo aberrations. In addition, 20 individuals (9%) were shown to have CNVs of unclear clinical relevance. Among these, 13 cases carried rare but inherited CNVs that may increase the risk for developing ASDs, while parental samples were unavailable in the remaining seven cases. Classification of all patients into different phenotypic and inheritance pattern groups indicated the presence of different CNV patterns in different patient groups. Clinically relevant CNVs were more common in syndromic cases compared to non-syndromic cases. Rare inherited CNVs were present in a higher proportion of ASD cases having first- or second-degree relatives with an ASD-related neuropsychiatric phenotype in comparison with cases without reported heredity ($P = 0.0096$). We conclude that rare CNVs, encompassing potential candidate regions for ASDs, increase the susceptibility for the development of ASDs and related neuropsychiatric disorders giving us further insight into the complex genetics underlying ASDs.

Key words: ASD; array CGH; CNV; microdeletion; microduplication

How to Cite this Article:
INTRODUCTION

Autism spectrum disorders (ASDs) [Mendelian inheritance in man (MIM) 209850] comprise a heterogeneous group of disorders. A large proportion of individuals with ASD also have intellectual disabilities, physical/visible malformations, and/or dysmorphic features. The patient group is very heterogeneous with regard to cognitive abilities and daily life skills. The prevalence for these disorders is now estimated at 1% when the whole autism spectrum is included [Gillberg and Wing, 1999; Fombonne, 2003; Baron-Cohen et al., 2009; Kogan et al., 2009]. The concordance rate for autism has been estimated to be approximately 90% in monozygotic twins and 2–10% in dizygotic twin pairs [Folstein and Rosen, 2001], making autism one of the most genetically influenced disorders of all developmental neuropsychiatric disorders [Kumar and Christian, 2009].

Despite the strong genetic component, support for associated regions and candidate genes has been weak and it has been difficult to obtain consistent results in independent samples. However, a few recurrent aberrations have been reported, such as the maternally derived duplication of chromosome band 15q11.2–13 identified in 0.5–3% of ASD cases [Hogart et al., 2008]. In addition, autism or autistic features often occur in single gene disorders such as Tuberous Sclerosis, Fragile X syndrome, and Rett syndrome [Gillberg and Coleman, 2000] but these disorders only explain around 2–5% of the autism cases [Baker et al., 1998; Carney et al., 2003; Kielinen et al., 2004; Volkmar et al., 2005; Hatton et al., 2006]. Furthermore, both copy number variations (CNVs) and minor mutations in several genes, such as SHANK3, NLGN3, NLGN4, and NRXN1, have been reported in a small number of individuals with ASDs [Jamain et al., 2003; Laumonnier et al., 2004; Moessner et al., 2007; Ching et al., 2010]. Taken together, an underlying genetic diagnosis is identified in around 10–15% of ASD cases [Kumar and Christian, 2009], while cytogenetically visible chromosomal rearrangements are found in 2–6% of individuals with ASDs [Wassink et al., 2001; Kielinen et al., 2004; Shen et al., 2010].

Until recently, karyotyping has been the standard method for the detection of cytogenetic aberrations in patients with developmental disorders. The development of whole-genome screening methodologies for the detection of CNVs, such as array-based comparative genomic hybridization (array-CGH), provides a much higher resolution than karyotyping leading to the identification of novel microdeletion- and microduplication syndromes, such as deletions and duplications in chromosome band 15q13.2q13.3, 16p11.2, and 17p11.2, often associated with an autism phenotype [Ballif et al., 2007; Potocki et al., 2007; Weiss et al., 2008; Miller et al., 2009]. The constantly increasing resolution of the arrays has further improved the detection of copy number abnormalities down to single genes and is likely to provide new advances in the autism genetics field. In the present study, we used genome-wide high resolution array-CGH in order to identify copy number aberrations in a cohort of 223 ASD patients. All patients were classified into different subpopulations regarding gender, phenotypic features, and inheritance, which revealed specific patterns of CNVs within different ASD subgroups.

MATERIALS AND METHODS

Patient Material

One hundred fifty-one ASD patients with a normal karyotype were recruited to the study. An additional patient with a non-pathogenic, inherited balanced translocation identified by karyotyping (case 16, Table 1), was included. Seventy-two patients in which karyotyping had not been carried out were recruited to the study for investigation with array-CGH alone. Of these 224 patients, 147 had been tested negative for Fragile X syndrome. The remaining cases, except one patient with a duplication comprising the MECP2 gene known to be associated with ASD [Van Esch et al., 2005], were screened after the array-CGH analysis. In one case a Fragile X syndrome diagnosis was revealed excluding the patient from the study resulting in a final cohort of 223 patients. DNA isolated from peripheral blood samples using standard procedures was used for array-CGH and Fragile X syndrome analysis.

The 223 ASD patients recruited to the study included 159 males and 64 females. One hundred sixty-four were sporadic cases with no first or second degree relative with ASD or any other related neuropsychiatric phenotype and 27 were familial cases including individuals with one or several first degree relatives with ASD. Twenty-five individuals had one or several first and/or second degree relatives with broader neuropsychiatric phenotypes (BNPs) including autistic traits and related neuropsychiatric disorders such as attention-deficit/hyperactivity disorder (ADHD), mental retardation (MR), dyslexia, and/or other speech and language disorders and/or one or several second degree relatives with ASD, referred to as BNP-familial cases. Of the seven remaining cases, three were adopted, and the family history was unavailable for the remaining four. The patients were recruited from neuropsychiatric units in Stockholm County and from child and adolescent psychiatric outpatient clinics in Sweden. They were offered to participate in the study after having undergone a neuropsychiatric workup done by the standard procedures used at that clinic resulting in an ASD diagnosis according to the criteria listed in the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV).

The patients were classified into four major phenotype groups with regard to syndromic features and cognitive ability. Patients with dysmorphic features and/or growth disorders and/or malformations were classified as syndromic, and mental retardation was defined as an IQ below 70 (measured by the Wechsler scales sometimes in combination with non-verbal scales such as Leiter or Wechsler non-verbal scale) in conjunction with significant limitations in the adaptive functioning. Of the 223 patients participating, 25 cases were syndromic with an IQ within the normal range, 45 cases were syndromic and had MR, 60 cases were non-syndromic with a normal IQ, and 93 patients were non-syndromic but had MR.

BAC-Tiling Array-CGH

BAC 33K and 38K arrays with an average practical resolution of approximately 300 kb were used to screen the whole genome for chromosomal imbalances in a total of 63 patients. Twenty-eight were screened with 33K BAC arrays consisting of 33, 370 BAC clones.
## TABLE I. Clinically Relevant CNVs Identified With Array-CGH in Patients With ASDs

<table>
<thead>
<tr>
<th>Case</th>
<th>Sex</th>
<th>Age</th>
<th>Inheritance pattern</th>
<th>Phenotype</th>
<th>Array-CGH platform</th>
<th>Chromosome band</th>
<th>Type</th>
<th>Size</th>
<th>Genes of interest</th>
<th>Total of gene</th>
<th>Position</th>
<th>Inheritance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>6</td>
<td>Sporadic</td>
<td>Non-syndromic, MR</td>
<td>X</td>
<td>2p16.3</td>
<td>Loss</td>
<td>430 kb</td>
<td>NRXN1</td>
<td>1</td>
<td>50,963,194–51,381,978</td>
<td>374 kb paternal, 50,975,595–51,144,527</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>13</td>
<td>Sporadic</td>
<td>Syndromic, MR</td>
<td>X</td>
<td>3q27.2q29</td>
<td>Loss</td>
<td>11.3 Mb</td>
<td>CHRNP, KLF13</td>
<td>73</td>
<td>180,453,169–180,571,851</td>
<td>180,466,994–197,826,819</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>16</td>
<td>Sporadic, IQ &gt; 70</td>
<td>Syndromic, MR</td>
<td>X</td>
<td>15q13.2q13.3</td>
<td>Gain</td>
<td>2 Mb</td>
<td>CREBBP</td>
<td>10</td>
<td>28,441,169–30,423,251</td>
<td>28,441,169–30,423,251</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>6</td>
<td>Sporadic</td>
<td>Syndromic, MR</td>
<td>X</td>
<td>16p11.3</td>
<td>Gain</td>
<td>2.4 Mb</td>
<td>MAZ, SEZ6L2, TBX6, MAPK3</td>
<td>30</td>
<td>28,500,284–30,240,082</td>
<td>28,500,284–30,240,082</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>6</td>
<td>Sporadic</td>
<td>Syndromic, IQ &gt; 70</td>
<td>X</td>
<td>16p11.2</td>
<td>Gain</td>
<td>580 kb</td>
<td>MAZ, SEZ6L2, TBX6, MAPK3</td>
<td>30</td>
<td>29,500,284–30,240,082</td>
<td>29,500,284–30,240,082</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>3</td>
<td>Sporadic</td>
<td>Non-syndromic, IQ &gt; 70</td>
<td>X</td>
<td>17p11.3</td>
<td>Gain</td>
<td>1.1 Mb</td>
<td>YWHAE, CRK</td>
<td>16</td>
<td>286,785–1,371,703</td>
<td>286,785–1,371,703</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>9</td>
<td>Familial</td>
<td>Non-syndromic, IQ &gt; 70</td>
<td>X</td>
<td>22q11.21</td>
<td>Gain</td>
<td>2.6 Mb</td>
<td>PKD2, SNAP29, Tbk1, SLC25A3</td>
<td>55</td>
<td>17,270,271–19,891,514</td>
<td>17,270,271–19,891,514</td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>9</td>
<td>Sporadic</td>
<td>Syndromic, MR</td>
<td>X</td>
<td>22q11.21q11.22</td>
<td>Gain</td>
<td>1.6 Mb</td>
<td>HIC2, UBE2L3, YPEL1, MAPK1</td>
<td>48</td>
<td>19,800,000–21,400,000, Balanced translocation</td>
<td>19,800,000–21,400,000, Balanced translocation</td>
</tr>
<tr>
<td>11</td>
<td>F</td>
<td>7</td>
<td>Familial</td>
<td>Non-syndromic, MR</td>
<td>X</td>
<td>22q13</td>
<td>Loss</td>
<td>25 kb</td>
<td>SHANK3</td>
<td>2</td>
<td>49,499,711–49,525,130</td>
<td>49,499,711–49,525,130</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>4</td>
<td>Sporadic</td>
<td>Non-syndromic, MR</td>
<td>X</td>
<td>9p24.1p24.2</td>
<td>Gain</td>
<td>420 kb</td>
<td>SLC1A1</td>
<td>2</td>
<td>4,118,618–4,538,970</td>
<td>4,118,618–4,538,970</td>
</tr>
<tr>
<td>13</td>
<td>F</td>
<td>6</td>
<td>Familial</td>
<td>Non-syndromic, MR</td>
<td>X</td>
<td>9q38</td>
<td>Loss</td>
<td>6.4 Mb</td>
<td>MEC2</td>
<td>16</td>
<td>152,933,677–153,158,866</td>
<td>152,933,677–153,158,866</td>
</tr>
<tr>
<td>14</td>
<td>F</td>
<td>6</td>
<td>Familial</td>
<td>Syndromic, MR</td>
<td>X</td>
<td>3p25.3-pter</td>
<td>Gain</td>
<td>10 Mb</td>
<td></td>
<td>40</td>
<td>1–10,068,676</td>
<td>1–10,068,676</td>
</tr>
<tr>
<td>15</td>
<td>M</td>
<td>38</td>
<td>Sporadic</td>
<td>Syndromic, IQ &gt; 70</td>
<td>X</td>
<td>7p22.1</td>
<td>Gain</td>
<td>1.4 Mb</td>
<td>SLC29A4</td>
<td>19</td>
<td>5,100,000–6,500,000</td>
<td>5,100,000–6,500,000</td>
</tr>
<tr>
<td>16</td>
<td>M</td>
<td>10</td>
<td>Sporadic</td>
<td>Non-syndromic, MR</td>
<td>X</td>
<td>9q34.13q34.31</td>
<td>Loss</td>
<td>10.9 Mb</td>
<td></td>
<td>50</td>
<td>68,429,077–79,001,357</td>
<td>68,429,077–79,001,357</td>
</tr>
<tr>
<td>17</td>
<td>M</td>
<td>19</td>
<td>Sporadic</td>
<td>Syndromic, MR</td>
<td>X</td>
<td>17p11.22</td>
<td>Loss</td>
<td>3 Mb</td>
<td>SLC25A1, FXR2, NGLN2</td>
<td>91</td>
<td>4,500,000–7,500,000</td>
<td>4,500,000–7,500,000</td>
</tr>
<tr>
<td>18</td>
<td>M</td>
<td>23</td>
<td>Sporadic</td>
<td>Syndromic, IQ &gt; 70</td>
<td>X</td>
<td>18q22.2-qter</td>
<td>Loss</td>
<td>10.1 Mb</td>
<td></td>
<td>30</td>
<td>65,820,701–75,365,502</td>
<td>65,820,701–75,365,502</td>
</tr>
</tbody>
</table>

F, female; M, male. aGene numbers were obtained from the Endeavour web server [http://homes.esat.kuleuven.be/~bioiuser/endeavour/tool/endeavourweb.php].

and the remaining 35 patients were screened with 38K BAC arrays including an additional 5000 BAC clones resulting in a denser coverage of the subtelomeric regions (Swegene DNA Microarray Resource Center, Department of Oncology, Lund University, Sweden, http://swegene.onk.lu.se). All experiments and analysis were performed as previously described by Schoumans et al. [2006].

In order to evaluate the significance of the CNVs identified, the Database of Genomic Variants (http://projects.tcag.ca/variation/) (DGV), cataloguing putatively benign variants, and the database DECIPHER (https://decipher.sanger.ac.uk/), cataloguing deletion and duplication syndrome regions and causative aberrations, were used. In addition, the literature was searched for more detailed information. Aberrations larger than 3 Mb were presumed more likely to be of clinical relevance [Lee et al., 2007]. Large de novo aberrations and both de novo and inherited CNVs in microdeletion and microduplication syndrome regions were interpreted as clinically significant CNVs. CNVs smaller than 1 Mb were noted as unclear unless they comprised regions or genes highly associated with ASDs, such as SHANK3 and NRXN1. Unclear and small CNVs, comprising potential ASD- and BNP candidate regions and genes, were noted and when possible investigated for inheritance.

Oligonucleotide Array-CGH, 244K and 180K

The Agilent 244K and 180K oligonucleotide array platforms (Agilent Technologies, Wilmington, DE) with a complete genome coverage and an average practical resolution of around 30–50 kb was used for screening 160 patients. Patient and reference DNA, the latter consisting of a pool of 10 healthy individuals (Promega GmbH, Mannheim, Germany), were labeled using a kit for oligonucleotide arrays (Enzo Life Sciences, Inc., New York, NY). All experiments and analysis were performed according to the manufacturer’s protocol and as previously described by Zhang et al. [2008]. Evaluation of the significance of identified CNVs is described above.

Multiplex Ligation-Dependent Probe Amplification (MLPA)

MLPA was used for confirmation of aberrant CNVs identified in our ASD cohort and for investigation of parental samples. Commercial MLPA-kits from MRC-Holland (Amsterdam, Netherlands) were used for testing genomic regions when available. For other genomic regions of interest, synthetic MLPA probes were designed as previously described [Stern et al., 2004] in UCSC Genome Browser (http://genome.ucsc.edu/) and synthesized by Sigma–Aldrich (St. Louis, MO). The MLPA reactions and analysis were performed as previously described [Bremer et al., 2010].

Fluorescence In Situ Hybridization (FISH)

FISH analysis on metaphase chromosomes was used in order to confirm aberrations identified by array and for investigating parental samples. All clones were obtained from CHORI BACPAC resources (http://bapac.chori.org) and were labeled with Fluorescein-12-dUTP and Spectrum Orange-11-dUTP using random priming labeling according to the manufacturer’s protocol (Enzo Life Sciences, Inc.). A labeled chromosome control probe was added to each labeling and inverted DAPI staining was used for chromosome identification.

Statistical Analysis

All statistical analyses were performed in StatXact-4 (Cytel, Inc., Cambridge, MA). Fisher exact P-tests were used for estimations of the significance of differences in proportions between several patient groups simultaneously. For the estimation of the significance of differences between two patient groups, two-tailed Fisher exact P-tests were used. P-values below 0.05 were considered statistically significant. StatXact-4 was also used for calculations of binomial confidence intervals.

RESULTS

CNVs of Clinical Significance

CNVs considered to be of clinical significance were identified in 18 out of 223 ASD patients (8%) of which 9 cases (4%) had de novo aberrations (Fig. 1, Table I). Twelve cases (5.4%) had CNVs located in previously described microdeletion and microduplication syndrome regions, or in regions containing genes in which deletions or duplications have been reported to cause ASDs (case 1–12 in Table I). These genomic alterations were located in chromosome bands 2p16.3, 3q27.2q29, 15q13.2q13.3, 16p11.2, 16p13.3, 17p11.2, 17p13.3, 22q11.2, 22q13.3, and Xq28 [Van Esch et al., 2005; Ballif et al., 2007; Moessner et al., 2007; Potocki et al., 2007; Ou et al., 2008; Weiss et al., 2008; Bi et al., 2009; Miller et al., 2009; Ramocki et al., 2009; Thienpont et al., 2009]. The deletions and duplications found in the six remaining cases (2.7% of all cases) were not located within any specific genetic syndrome regions, but within regions with previously described larger or partly overlapping aberrations. Four of these aberrations were shown to be de novo, while one or both parental samples were unavailable for the two remaining cases (cases 14 and 17, Table I). These six aberrations appeared in
One of the 18 cases identified with a clinically significant alteration was case 13, identified with a duplication in chromosome band 1q25.3q31.1, 3p25.3-pter, 7p22.1, 9q13q21.31, 17p13.2, and 18q22.2-qter (Table I). In case 14, identified with a duplication in chromosome band 3p25.3, an unbalanced translocation between chromosome 3 and 13, 46,XX,der(13)t(3;13)(p25.3;qter) was identified by FISH analysis.

CNVs Potentially Increasing ASD Risk

CNVs of unclear clinical significance were identified in seven (3.1%) cases (case 19–25 in Table II). These, although small in size (<1 Mb), spanned regions encompassing potential ASD-related genes or regions where no or only a few variations had been reported in the DGV possibly representing rare benign variants or risk factors for ASD development. In this group of patients, one or both parental samples were unavailable for testing.

In an additional 13 (5.8%) cases (case 26–38 in Table III), CNVs were detected in similar regions, but were shown to be inherited from a healthy parent. These CNVs may increase the risk for ASD due to the possibility of incomplete penetrance, imprinting effects, or that the other copy of a gene in the same region may be affected by a minor mutation leading to a homozygous alteration in the child.

All CNVs described above correspond to a proportion of 17% of the 223 ASD patients representing 38 individuals with a total of 44 genomic variations (Fig. 1).

Fisher exact P-tests (3×2 tables) were used to estimate significant differences in the proportion of cases with different types of CNVs between the three inheritance subgroups. Cases with other CNVs, than the type at the time being tested, were excluded. The results indicated a statistically significant difference in proportion of cases with rare inherited CNVs between the different subgroups (P = 0.0139). Separate comparison of frequencies between the three inheritance subgroups (excluding all cases with another type of CNV, Tables I and II) using two-tailed Fisher exact P-tests (2×2 tables) showed that rare inherited CNVs were significantly more common in BNP-familial cases and sporadic cases compared with non-familial cases. Since the BAC-arrays have a lower resolution compared with the Agilent arrays, the statistical estimation was in addition performed for the BAC-arrays as with the Agilent arrays [Zhang et al., 2008]. Therefore, the statistical estimation was in addition performed for the BAC-arrays as with the Agilent arrays [Zhang et al., 2008].

<table>
<thead>
<tr>
<th>Case</th>
<th>Sex</th>
<th>Age</th>
<th>Inheritance pattern</th>
<th>Phenotype</th>
<th>Array-CGH platform</th>
<th>Identified abnormalities of unclear significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>M</td>
<td>10</td>
<td>BNP-familial</td>
<td>Non-syndromic, IQ &gt; 70</td>
<td>Agilent 244K</td>
<td>X</td>
</tr>
<tr>
<td>20</td>
<td>M</td>
<td>6</td>
<td>Adopted</td>
<td>Non-syndromic, MR</td>
<td>Agilent 244K</td>
<td>X</td>
</tr>
<tr>
<td>21</td>
<td>F</td>
<td>29</td>
<td>Sporadic</td>
<td>Non-syndromic, MR</td>
<td>Agilent 244K</td>
<td>X</td>
</tr>
<tr>
<td>22</td>
<td>F</td>
<td>8</td>
<td>Sporadic</td>
<td>Non-syndromic, IQ &gt; 70</td>
<td>Agilent 244K</td>
<td>X</td>
</tr>
<tr>
<td>23</td>
<td>M</td>
<td>18</td>
<td>Sporadic</td>
<td>Non-syndromic, MR</td>
<td>Agilent 244K</td>
<td>X</td>
</tr>
<tr>
<td>24</td>
<td>M</td>
<td>10</td>
<td>Familial</td>
<td>Non-syndromic, MR</td>
<td>Agilent 244K</td>
<td>X</td>
</tr>
<tr>
<td>25</td>
<td>M</td>
<td>11</td>
<td>Sporadic</td>
<td>Syndromic, MR</td>
<td>Agilent 244K</td>
<td>X</td>
</tr>
</tbody>
</table>

M, male; F, female. *Gene numbers were obtained from the Endeavour web server (http://homes.esat.kuleuven.be/bioiuser/endeavour/tool/endeavourweb.php).

| Case | Sex | Age | Inheritance | Phenotype | Array-CGH platform | Identified abnormalities | Total of genes | Genes of interest | Chromosome | band | Type | Size | Genes of interest |
|------|-----|-----|-------------|-----------|-------------------|------------------------|----------------------|-----------------|-------------|-----------|------|------|--------|---------|
| 26   | M   | 4   | BNP-familial | Non-syndromic, IQ > 70 | Agilent 244K | 800 kb | FBXL7 | 1 | 15,118,837 | X | 5p15 | Gain | 800 kb | 1 |
| 27   | F   | 18  | BNP-familial | Non-syndromic, IQ > 70 | Agilent 244K | 270 kb | NQO2 | 7 | 2,777,486 | X | 6p25.2 | Gain | 270 kb | 2 |
| 28   | M   | 18  | Sporadic     | Non-syndromic, MR | Agilent 180K | 120 kb | DKFZp564N2472 | 1 | 103,472,391 | X | 7p12.1 | Gain | 2 Mb | 10 |
| 29   | M   | 8   | BNP-familial | Non-syndromic, IQ > 70 | Agilent 244K | 270 kb | CNTNAP2 | 1 | 139,115,238 | X | 9q34.3 | Gain | 90 kb | 6 |
| 30   | M   | 19  | Sporadic     | Non-syndromic, MR | Agilent 244K | 270 kb | GRIN1 | 6 | 5,406,064 | X | 11p14.2 | Gain | 160 kb | 2 |
| 31   | F   | 7   | Familial     | Non-syndromic, MR | Agilent 244K | 270 kb | IMPA2 | 2 | 11,925,571 | X | 9q31.1 | Gain | 140 kb | 3 |
| 32   | M   | 6   | BNP-familial | Non-syndromic, MR | Agilent 244K | 300 kb | GRIN3A | 1 | 103,472,391 | X | 11p13.31 | Gain | 110 kb | 2 |
| 33   | M   | 13  | Familial     | Non-syndromic, MR | Agilent 244K | 270 kb | SLFN11,SLFN12,SLFN13 | 3 | 30,701,599 | X | 12p12.1 | Gain | 270 kb | 1 |
| 34   | M   | 48  | Sporadic     | Non-syndromic, IQ > 70 | Agilent 244K | 270 kb | TSPAN7 | 1 | 139,115,238 | X | 9q34.3 | Gain | 270 kb | 1 |
| 35   | M   | 17  | BNP-familial | Non-syndromic, MR | Agilent 244K | 120 kb | GRIN3A | 1 | 103,472,391 | X | 11p14.2 | Gain | 120 kb | 1 |
| 36   | M   | 48  | Sporadic     | Non-syndromic, IQ > 70 | Agilent 244K | 120 kb | GRIN3A | 1 | 103,472,391 | X | 11p14.2 | Gain | 120 kb | 1 |
| 37   | F   | 6   | Sporadic     | Non-syndromic, IQ > 70 | Agilent 244K | 120 kb | GRIN3A | 1 | 103,472,391 | X | 11p14.2 | Gain | 120 kb | 1 |

aGene numbers were obtained from the Endeavour web server (http://homes.esat.kuleuven.be/bioiuser/endeavour/tool/endeavourweb.php).


M, male; F, female.
the oligonucleotide arrays only \((P = 0.0128)\) (Supplementary Tables I and II).

**CNVs in Different Patient Subgroups**

Out of the 18 patients with clinically significant CNVs, 10 were males (6.3% of the male cases) and 8 were females (12.5% of the female cases) giving a male/female ratio of 1.25:1 compared with the initial whole patient cohort male/female ratio of 2.48:1 [non-significant (n.s.)].

When comparing the frequency of clinically relevant aberrations across phenotypically different patient groups, non-syndromic patients without MR had a lower number of pathogenic CNVs compared with the other phenotypic subgroups, especially compared with the syndromic patient groups (Table V). However, no difference in frequency of aberrations was statistically significant.

Dividing the groups with regard to presence of syndromic features showed clinically relevant CNVs in 13% in the syndromic patients versus 6% in the non-syndromic (n.s.) (Table V). Clinically relevant CNVs did not discriminate patients with intellectual disabilities from those without (Table V).

**Deletions Versus Duplications**

The clinically relevant CNVs consisted of eight deletions and ten duplications. When phenotypic characteristics were compared between patients with deletions and duplications, all but one of the patients with a deletion had MR, while five of the ten patients with a duplication had normal IQ (n.s.) (Table I).

**DISCUSSION**

We have screened the whole genome of 223 patients diagnosed with ASD in order to identify deletions and duplications contributing to their behavioral and physical phenotype. Pathogenic CNVs were identified in 18 (8%) individuals comprising 8 losses and 10 gains (Table I). This is in contrast to what has been reported in cohorts of patients with mental retardation where deletions seem to be the main pathogenic mechanism [Menten et al., 2006; Wincent et al., 2010]. In agreement with our results, Marshall et al. [2008] also found an even distribution of gains and losses in a cohort of patients with ASD. Within all the six clinically relevant CNVs that did not involve any well-recognized ASD associated regions, overlapping or minor anomalies had been reported in DECIPHER. However, the individuals reported in DECIPHER were seldom described to have autism but a phenotype including cognitive delay, behavioral problems, head and facial dysmorphisms, and/or malformations.

In addition to the clinically relevant aberrations found in 8% of our patients, another 9% had potentially pathogenic aberrations categorized into two CNV groups, shown in Tables II and III. Of the individuals referred to as BNP-familial cases, 20.0% carried rare inherited CNVs including potential ASD or BNP candidate regions (Table III), while in sporadic cases such CNVs were present in 3.7% \((P = 0.0096)\) (Table IV). According to the current knowledge, CNVs are the most common structural variants present in the

---

**TABLE IV. Distribution of Different Types of CNVs Identified in the Patient Subgroups Recognized by Having Different Inheritance Patterns Regarding First and Second Degree Relatives**

<table>
<thead>
<tr>
<th>Inheritance pattern</th>
<th>Sporadic cases</th>
<th>Familial cases</th>
<th>BNP-familial cases</th>
<th>Cases with unknown inheritance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients [n]</td>
<td>164</td>
<td>27</td>
<td>25</td>
<td>7</td>
</tr>
<tr>
<td>Patients with clinically relevant CNVs % [n]</td>
<td>8.5% [14]</td>
<td>14.8% [4]</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

---

**TABLE V. Distribution of the Patients, Male/Female Ratio, and CNV Findings into Different Phenotypic Subgroups**

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Syndromic, IQ &gt; 70</th>
<th>Syndromic, MR</th>
<th>Non-syndromic, IQ &gt; 70</th>
<th>Non-syndromic, MR</th>
<th>Syndromic</th>
<th>Non-syndromic</th>
<th>MR</th>
<th>IQ &gt; 70</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients [n]</td>
<td>25</td>
<td>45</td>
<td>60</td>
<td>93</td>
<td>70</td>
<td>153</td>
<td>138</td>
<td>85</td>
<td>223</td>
</tr>
<tr>
<td>Male/female ratio</td>
<td>2.13:1</td>
<td>1.65:1</td>
<td>3.29:1</td>
<td>2.72:1</td>
<td>1.80:1</td>
<td>2.92:1</td>
<td>2.29:1</td>
<td>2.86:1</td>
<td>2.48:1</td>
</tr>
<tr>
<td>Patients with clinically relevant CNVs [n]</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>12</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>Patients with clinically relevant CNVs [%]</td>
<td>16.0</td>
<td>11.1</td>
<td>3.3</td>
<td>7.5</td>
<td>12.9</td>
<td>5.9</td>
<td>8.7</td>
<td>7.1</td>
<td>8.1</td>
</tr>
</tbody>
</table>
human genome and several different mechanisms have been described for different kinds of CNV events [Stankiewicz and Lupski, 2010]. De novo CNVs are more common in ASD patients compared with healthy individuals [Sebat et al., 2007]. In addition, it has been shown that the overall number of CNVs in the genome seems to be similar between ASD cases and healthy controls, while rare CNVs containing ASD or BNP-associated genes are enriched in ASD cases [Marshall et al., 2008; Bucan et al., 2009; Pinto et al., 2010] suggesting that it is the location and the nature of the genes that the CNVs include that cause or predispose the individual for the development of ASDs. Recent studies suggest that two or several CNVs could interact in predisposing individuals for neuropsychiatric disorder [Stankiewicz and Lupski, 2010]. Girirajan et al. [2010] reported a two-hit model in which a second CNV in patients with 16p12.1 deletions was shown to lead to a more severe phenotype outcome of the syndrome. Notably, the majority of the 16p12.1 deletions were inherited with a significantly increased risk of manifestations of neuropsychiatric traits in carrier parents compared with non-carrier parents [Girirajan et al., 2010]. Our result of a significantly higher presence of rare inherited, potentially neuropsychiatric disorder related CNVs in BNP-familial cases than in sporadic cases, may indicate that these types of CNVs, rather than being directly causative (they were inherited from a healthy parent), increase susceptibility for the development of ASDs and ASD-related phenotypes. This may, at least in part, explain the complex genetics underlying ASDs and the difficulties we face in finding recurrent genetic causes.

Among the 18 cases with CNVs considered causative, 14 were sporadic (8.5% of sporadic cases) and four were familial (14.8% of familial cases). In previous studies, higher rates of aberrations have been found in sporadic than in familial cases with frequencies of pathogenic CNVs reported at 7–10% and 2–3%, respectively [Sebat et al., 2007; Marshall et al., 2008]. However, the results from Szatmari et al. [2007], identifying pathogenic CNVs in 7.5–11.5% of cases in three familial ASD patient cohorts, and those from Pinto et al. [2010], identifying equal rates of ~5.5% de novo CNVs in both simplex and multiplex cases, contrast with other earlier reports. Even though our cohort is small our results are similar to the findings by Szatmari et al. [2007], supporting that causative CNVs could be at least as common in familial cases as in sporadic cases.

As mentioned above, similar numbers of pathogenic deletions and duplications were identified which is in contrast to what is found in patients with multiple congenital anomalies (MCA) and MR, where causative CNVs generally consist of deletions [Menten et al., 2006; Wincent et al., 2010]. Nearly all our patients with deletions had intellectual disabilities, while half of the patients with duplications had IQ in the normal range. A stronger selection against deletions has been shown in RefSeq and disease related gene regions, in ultra-conserved elements and for deletions of larger sizes in the normal population than for genomic duplications [Redon et al., 2006]. In addition, reciprocal duplications in previously known submicroscopic deletion syndrome regions have usually been associated with milder phenotypes [Thienpont et al., 2009]. Our findings support that duplications could be more common in cases with milder phenotypes. However, the deletions and duplications were different regarding location, size, and gene content and the findings should therefore be interpreted with caution. In addition, there have been studies in which causative deletions have been identified in a higher proportion than duplications in ASD patients [Sebat et al., 2007]. Future studies in larger populations may give us a better understanding of CNV patterns in ASD patients with regard to different phenotypic expressions such as the intellectual ability status.

The low male/female ratio of 2.48 in our patient cohort may reflect a high number of individuals with cognitive disabilities in addition to their ASD which would lead to an equalization of male and female cases [Fombonne, 2003]. However, in our patient cohort the presence of syndromic features seemed to have more influence on equalization of gender distribution than the presence of intellectual disabilities (n.s.) (Table V). Nevertheless, the male/female ratio of 1.25:1 in patients with clinically relevant CNVs was lower compared to the initial whole patient cohort male/female ratio of 2.48:1. The probability of finding a clinically significant CNV in a female was almost twice that of a male in our ASD cohort. Similar findings have been reported in previous studies but with even larger differences [Sebat et al., 2007; Qiao et al., 2009]. The increased risk in males for developing ASDs [Fombonne, 2003] together with the higher presence of clinically relevant CNVs in females with ASDs may point toward factors other than CNVs being responsible for the increased susceptibility for ASDs in males. The skewed male/female ratio has been argued to depend on epigenetic effects such as an increased vulnerability for dysregulation of methylation of brain-expressed genes on the X-chromosome and sex-specific responses to different hormones [Carter, 2007; Jones et al., 2008]. Males may also be more vulnerable to minor variations in ASD susceptibility genes located on the X-chromosome [Noor et al., 2010].

More causative CNVs were found in syndromic patients than in non-syndromic patients (Table V). This finding was not statistically significant but is in line with previous studies [Jacquemont et al., 2006; Sebat et al., 2007]. However, as genomic abnormalities are also found in a significant number of non-syndromic patients with ASDs it is advisable to clinically investigate all ASD patients with whole-genome screening methodologies. When patients were categorized according to intellectual disability, no clear difference in pathogenic CNV frequency was noted between the two patient groups (Table V).

The significant frequency of findings in our study further confirms that screening of the whole genome by array-CGH should become standard in all clinical investigations of ASD patients. Notably, one case, a male patient with a maternally inherited 2.2 Mb chromosome band 15q13.2-q13.3 duplication, was shown to have a CGG-repeat expansion in the *FMR1* gene revealing a Fragile X syndrome diagnosis and was therefore excluded from the study. This additionally illustrates the importance of carrying out all genetic tests relevant for ASDs [Shen et al., 2010], especially concerning patients with genetic syndromes with incomplete penetrance.

We conclude that rare but inherited CNVs may have a significant influence in increasing the risk for development of ASDs and related neuropsychiatric disorders together with other CNVs and genetic alterations as well as epigenetic and environmental factors. Recently, it has been reported that rare CNVs containing ASD- and BNP-associated genes are accumulated in ASD cases compared with
healthy controls. In addition, it has been shown that two CNVs can interact increasing the severity of a genomic syndrome. These findings further support our conclusion. CNVs currently hypothesized to be clinically benign, may play an important role in the development of ASDs and ASD-related phenotypes as well as in other complex diseases and should be judged with caution.

Our study illustrates the complex genetics underlying ASDs and the importance of reporting rare variants concerning these disorders. Future larger scale studies of both healthy and affected individuals are needed in order to elucidate the significance and effects of these types of genetic alterations. The increased resolution of array-CGH in combination with new technologies, such as whole genome sequencing and bioinformatics programs, will play an important role in helping us further understand the complex genetic basis of autism. The implementation of these high resolution techniques in the genetic research of ASDs may lead to the unveiling of specific genotypes and subtypes of ASDs for which new diagnostic and therapeutic strategies can be developed.

ACKNOWLEDGMENTS

We would like to thank the patients and their families for participating in this study. We would also like to thank Professor Åke Borg (Department of Oncology, Lund University Hospital, Sweden) for providing the BAC arrays. Special thanks to Margareta Lagerberg, Sigrid Sahlén, Irene White, Anna Hammarsjö, and Lena Forsberg for technical assistance. This work was supported by funds from the Swedish Research Council, Karolinska Institutet, The Independent Order of Odd Fellows, Frimurarna Barnhuset Foundation, Sven Jerring Foundation, Åke Wiberg Foundation, Kronprinsessan Lovisas Foundation, and Linnea and Josef Carlsson Foundation.

REFERENCES


