**MLPA is a practical and complementary alternative to CMA for diagnostic testing in patients with autism spectrum disorders and identifying new candidate CNVs associated with autism**

**Pavlina Capkova**, **Josef Srovnal** Corresp., **Zuzana Capkova**, **Katerina Staffova**, **Vera Becvarova**, **Marie Trkova**, **Katerina Adamova**, **Alena Santava**, **Vaclava Curtisova**, **Marian Hajduch**, **Martin Prochazka**

1. Department of Medical Genetics, University Hospital Olomouc, Olomouc, Czech Republic
2. Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacky University Olomouc, Olomouc, Czech Republic
3. Gennet, s.r.o., Prague, Czech Republic

Corresponding Author: Josef Srovnal
Email address: josef.srovnal@upol.cz

**Background:** Autism spectrum disorder (ASD) is a complex heterogeneous developmental disease with a significant genetic background that is frequently caused by rare copy number variants (CNV). Microarray-based whole-genome approaches for CNV detection are widely accepted. However, the clinical significance of most CNV is poorly understood, so results obtained using such methods are sometimes ambiguous. We therefore evaluated a targeted approach based on multiplex ligation probe-dependent amplification (MLPA) using selected probemixes to detect clinically relevant variants for diagnostic testing of ASD patients. We compare the reliability and efficiency of this test to those of CMA and other tests available to our laboratory. In addition, we identify new candidate genes for ASD identified in a cohort of ASD-diagnosed patients.

**Method:** We describe the use of multiplex ligation probe-dependent amplification (MLPA), chromosomal microarray analysis (CMA), and karyotyping to detect CNV in 92 ASD patients and evaluate their clinical significance.

**Result:** Pathogenic and likely pathogenic mutations were identified by CMA in 8 (8.07 % of the studied cohort) and 12 (13.04 %) ASD patients, respectively, and in 8 (8.07 %) and 4 (4.35 %) patients, respectively, by MLPA. The detected mutations include the 22q13.3 deletion, which was attributed to ring chromosome 22 formation based on karyotyping. CMA revealed a total of 91 rare CNV in 55 patients: 8 pathogenic, 15 designated VOUS - likely pathogenic, 10 VOUS - uncertain, and 58 VOUS - likely benign or benign. MLPA revealed 18 CNV in 18 individuals: 8 pathogenic, 4 designated as VOUS- likely pathogenic, and 6 designated as VOUS - likely benign/benign. Rare CNV were detected in 17 (58.62 %) out of 29 females and 38 (60.32 %) out of 63 males in the cohort. Two genes, DOCK8 and PARK2, were found to be overlapped by CNV designated pathogenic, VOUS - likely pathogenic, or VOUS - uncertain in multiple patients. Moreover, the studied ASD cohort exhibited significant ($p < 0.05$) enrichment of duplications encompassing DOCK8.

**Conclusion:** MLPA and CMA yielded concordant results for 12 patients bearing CNV designated pathogenic or VOUS-likely pathogenic. Unambiguous diagnoses were achieved for 8 patients (corresponding to 8.7 % of the total studied population) by both MLPA and CMA, for one (1.09 %) patient by karyotyping, and for one (1.09 %) patient by FRAXA testing. MLPA and CMA thus achieved identical reliability with respect to clinically relevant findings. As such, MLPA could be useful as a fast and inexpensive test in patients with syndromic autism. The detection rate of potentially pathogenic variants (VOUS - likely pathogenic) achieved by CMA was higher than that for MLPA (13.04 % vs. 4.35 %).
However, there was no corresponding difference in the rate of unambiguous diagnoses of ASD patients. In addition, the results obtained suggest that DOCK8 may play a role in the etiology of ASD.
MLPA is a practical and complementary alternative to CMA for diagnostic testing in patients with autism spectrum disorders and identifying new candidate CNVs associated with autism.

Pavlina Capkova\textsuperscript{1\#}, Josef Srovnal\textsuperscript{1,2*\#}, Zuzana Capkova\textsuperscript{1}, Katerina Staffova\textsuperscript{2}, Vera Becvarova\textsuperscript{3}, Marie Trkova\textsuperscript{3}, Katerina Adamova\textsuperscript{1}, Alena Santava\textsuperscript{1}, Vaclava Curtisova\textsuperscript{1}, Marian Hajduch\textsuperscript{2} and Martin Prochazka\textsuperscript{1}

\textsuperscript{1}Department of Medical Genetics, University Hospital Olomouc, Czech Republic

\textsuperscript{2}Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacky University and University Hospital in Olomouc, Czech Republic

\textsuperscript{3}Gennet, s. r. o., Prague, Czech Republic

*Corresponding author: Josef Srovnal, MD, PhD, Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacky University and University Hospital in Olomouc, Hnevotinska 5, 77900 Olomouc, Czech Republic, Telephone: +420-58-5632137, Fax: +420-58-5632180, e-mail: josef.srovnal@upol.cz, www.imtm.cz

\#Both authors contributed equally to the work
Abstract

Background: Autism spectrum disorder (ASD) is a complex heterogeneous developmental disease with a significant genetic background that is frequently caused by rare copy number variants (CNV). Microarray-based whole-genome approaches for CNV detection are widely accepted. However, the clinical significance of most CNV is poorly understood, so results obtained using such methods are sometimes ambiguous. We therefore evaluated a targeted approach based on multiplex ligation probe-dependent amplification (MLPA) using selected probemixes to detect clinically relevant variants for diagnostic testing of ASD patients. We compare the reliability and efficiency of this test to those of CMA and other tests available to our laboratory. In addition, we identify new candidate genes for ASD that were detected in a cohort of ASD-diagnosed patients.

Method: We describe the use of multiplex ligation probe-dependent amplification (MLPA), chromosomal microarray analysis (CMA), and karyotyping to detect CNV in 92 ASD patients and evaluate their clinical significance.

Result: Pathogenic and likely pathogenic mutations were identified by CMA in 8 (8.07% of the studied cohort) and 12 (13.04%) ASD patients, respectively, and in 8 (8.07%) and 4 (4.35%) patients, respectively, by MLPA. The detected mutations include the 22q13.3 deletion, which was attributed to ring chromosome 22 formation based on karyotyping. The FMR1 mutation was detected in one patient. CMA revealed a total of 91 rare CNV in 55 patients: 8 pathogenic, 15 designated VOUS - likely pathogenic, 10 VOUS - uncertain, and 58 VOUS - likely benign or benign.
MLPA revealed 18 CNV in 18 individuals: 8 pathogenic, 4 designated as VOUS- likely pathogenic, and 6 designated as VOUS - likely benign/benign. Rare CNV were detected in 17 (58.62 %) out of 29 females and 38 (60.32 %) out of 63 males in the cohort. Two genes, DOCK8 and PARK2, were found to be overlapped by CNV designated pathogenic, VOUS - likely pathogenic, or VOUS - uncertain in multiple patients. Moreover, the studied ASD cohort exhibited significant ($p < 0.05$) enrichment of duplications encompassing DOCK8.

**Conclusion:** MLPA and CMA yielded concordant results for 12 patients bearing CNV designated pathogenic or VOUS-likely pathogenic (corresponding to 60 % of the study population exhibiting such CNV), including a ring chromosome 22. Unambiguous diagnoses were achieved for 8 patients (corresponding to 8.7 % of the total studied population) by both MLPA and CMA, for one (1.09 %) patient by karyotyping, and for one (1.09 %) patient by FRAXA testing. MLPA and CMA thus achieved identical reliability with respect to clinically relevant findings. As such, MLPA could be useful as a fast and inexpensive test in patients with syndromic autism. The detection rate of potentially pathogenic variants (VOUS – likely pathogenic) achieved by CMA was significantly higher than that for MLPA (13.04 % vs. 4.35 %). However, there was no corresponding difference in the rate of unambiguous diagnoses of ASD patients. In addition, the results obtained suggest that DOCK8 may play a role in the etiology of ASD.

**Introduction**

Autistic spectrum disorders (ASD) are complex heterogeneous developmental diseases with a significant genetic background and include three closely related diagnoses: autistic disorder, Asperger syndrome, and pervasive developmental disorder - not otherwise specified (PDD-NOS). They are characterized by simultaneous deficits in 3 domains of behavior: reciprocal social
interaction, communication, and stereotyped and restricted behaviors. Their estimated prevalence in the population is 1:68, with males being predominantly affected (McCarthy, 2014; de la Torre-Ubieta et al., 2016; Schaefer, 2016; Schaefer & Mendelsohn, 2013; Miller et al., 2010).

ASD occurs frequently with the following comorbidities: a motor deficit, sleep abnormalities, gastrointestinal disturbances, epilepsy, and intellectual disability (de la Torre-Ubieta et al., 2016; Schaefer, 2016). These comorbidities can also overlap with phenotypes observed in other neuropsychiatric disorders.

Submicroscopic copy number variants (CNV) may have a causal or susceptibility–related role in the heritability of ASD. However, the causality and/or pathogenicity of these CNV is largely unknown due to their variable expression and incomplete penetrance, which can result in a spectrum of phenotypes ranging from asymptomatic to intellectual disability (ID)/developmental delay (DD)/ASD. Many studies have sought to identify new candidate genes for ASD but fewer have sought to clarify their clinical significance for patients and their families. Chromosomal microarrays (CMAs) were recently identified as a first-tier method for testing in patients with ID/DD/ASD (Schaefer, 2016; Schaefer & Mendelsohn, 2013; Miller et al., 2010; South et al., 2013).

However, little is currently known about the clinical significance of most CNV, which can hamper the interpretation of test results and complicate genetic counseling. Unlike CMA, multiplex ligation-dependent probe amplification (MLPA) is fast and provides easily interpreted results. It can therefore serve as a clinically effective targeted test to detect recurrent CNV associated with ASD, and is currently used as a preliminary test to exclude recurrent pathogenic CNV in our department.
This study compares the efficiency and reliability of MLPA and CMA in diagnostic testing of ASD patients. A secondary objective was to identify new candidate genes for ASD. We present results obtained by using a combination of MPLA and CMA to analyze CNV in a cohort of 92 children diagnosed with ASD.

**Materials and Methods**

**Participants**

The study involved 92 individuals of Caucasian ethnicity - 63 males and 29 females - with autism (54), PDD-NOS (35) and Asperger syndrome (3), predominantly from simplex families (89 %). These patients were referred to genetic counselling solely on the basis of an ASD, PDD-NOS, or Asperger syndrome diagnosis, or based on some level of neurodevelopmental impairment together with ASD or PDD-NOS. The size of the study population is approximately 1/5th of the number of child patients that undergo genetic testing in our department each year. Patients with known monogenic syndromes (e.g. familial cases and clear syndromic cases) associated with ASD (tuberose sclerosis, neurofibromatosis, etc.) were excluded. Peripheral blood samples were collected after genetic counseling in the Department of Medical Genetics at the University Hospital in Olomouc, Czech Republic, during the years 2012-2016. For 17 patients, the DNA of both parents was obtained to verify the origin of the detected variants. The study was approved by the Institutional Review Board of the University Hospital and Faculty of Medicine and Dentistry, Palacky University, Olomouc (IRB number 96/17). All procedures were conducted in accordance with the Declaration of Helsinki. Written informed consent for the use of personal particulars and genetic information for research purposes was collected from the patients’ parents or guardians during genetic counselling. The cohort’s mean age at evaluation was 5.0 ± 2.9 years.
The patients underwent rigorous examinations by pediatricians, neurologists, psychiatrists and geneticists, including metabolic tests and brain imaging. Individuals were diagnosed with ASD by clinicians after performing the Autism Diagnostic Interview-Revised and Autism Diagnosis Observation Schedule. The subjects with pervasive developmental disorders and varying levels of impairment were diagnosed with broad-spectrum disorder, which includes conditions such as pervasive developmental disorder not otherwise specified (PDD-NOS) and Asperger syndrome. Phenotype descriptions for patients with pathogenic/VOUS - likely pathogenic and VOUS - uncertain findings are presented in the Supplemental file. The frequency of CNV overlapping the DOCK8 gene in the healthy population was determined by MLPA analysis (using the SALSA MLPA P385-A2 DOCK8 probemix) of 40 male and 40 female control individuals exhibiting no related health conditions.

**Study design**

The study used a retrospective observational design.

**Methods**

Systematic screening for pathogenic mutations was performed by karyotyping, fragile X syndrome testing, screening for metabolic disorders, targeted MLPA testing, and CMA. Cytogenetic analysis was performed using cultured lymphocytes by conventional G-banding with a resolution of 550 bphs. DNA was isolated from peripheral blood by the saline method. DNA samples were tested for the FMR1 mutation by PCR using fluorescently labelled primers as described previously (Zhou, 2006). MLPA tests were performed with SALSA® MLPA® sets P070-B3 and P036-E3 (the
Subtelomeres Mix 1 and Mix 2B probemixes), P245-B1 and P297-B2/C1 (the Microdeletions 1A and 2 probemixes), P343-C3 (the Autism1 probemix), and P106-C1 (the Mental retardation X-linked probemix) in accordance with the manufacturer’s protocol. The Coffalyser program was used for CNV calling (MRC-Holland, Amsterdam, Netherlands). PCR products were identified and quantified by capillary electrophoresis on an ABI 3130 genetic analyzer, using the Gene Mapper software from Applied Biosystems, Foster City, CA. Other MLPA probemixes (P051/P052-D1 and P385-A2) were used to verify CMA findings relating to PARK2 and DOCK8.

Chromosomal microarray analyses were performed using a Cytoscan HD (Affymetrix, Santa Clara, CA, USA) or CytoSNP-12(Illumina, San Diego, CA, USA) instrument according to the manufacturer’s protocol. The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus database (Edgar, Domrachev & Lash, 2002) and are accessible using GEO Series accession number GSE114870 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE114870). The programs CHAS v. 1.2.2 and Illumina KaryoStudio 1.3 from Genome Studio V2011.1 were used for CNV calling. The observed CNV were compared to CNV recorded in the DGV to assess their frequency in the population. CNV encompassing coding regions with frequencies of <1 % in the population were considered rare variants. The clinical significance of individual CNV was evaluated according to the ACMG Standard and Guidelines (Kearney et al., 2011). Where possible, parental samples were collected for patients exhibiting pathogenic CNV and variants of unknown significance (VOUS) to determine the CNV’s origin. Detected CNV were systematically compared to CNV recorded in curated databases (ISCA, DECIPHER, SFARI, and DGV) to determine their clinical significance. Fisher’s exact test was used to calculate the statistical significance of the frequency of DOCK8.
duplication in the ASD population (www.socscistatistics.com/tests/fisher/Default2.aspx). As a control, we used data from population sequencing studies held by the 1000 Genomes Consortium (Mills et al., 2011).

Results

Karyotyping revealed chromosomal aberrations in 3 (3.26%) patients from the cohort: 1 pathogenic r(22)(q13.3)(1.08%) and 2 likely benign t(10;11)(q26;p13)pat, inv(Y)(p11.2q11.23) (Table 1). The FMR1 mutation was identified in 1 (1.09%) individual. In total, MLPA revealed 18 CNV in 18 (19.57%) individuals: 8 (44.44%) pathogenic (including a terminal deletion in ring chromosome 22), 4 (22.22%) designated as VOUS - likely pathogenic and confirmed by CMA, and 6 (33.33%) designated as VOUS - likely benign/benign. All but 3 of these detected variants were confirmed by CMA. The variants not confirmed by CMA were small deletions (MAPK3 - pat inherit., SNRPN – mat inherit., FRG1 - pat inherit.) identified using one probe; these deletions probably correspond to SNV that were inherited from healthy parents and lie in the probe’s hybridization region. This phenomenon has been described previously and represents an inherent limitation of MLPA (Cai et al., 2008).

CMA identified 91 rare CNV (60 duplications and 31 deletions) that contained coding regions and had MAF values of < 1% in 55 (59.78%) patients. Among these were 8 (8.79%) pathogenic CNV, 15 (16.48%) CNV designated as VOUS - likely pathogenic, 10 (10.99%) designated as VOUS - uncertain, and 58 (63.73%) designated as VOUS - likely benign or benign (Table 2). The percentages of rare CNV in the males (60.32%) and females (58.62%) of our cohort were similar. MLPA and CMA yielded concordant results for all eight patients (8.7% of the study cohort) exhibiting pathogenic mutations. CNV designated VOUS - likely pathogenic were identified as
primary CMA findings in 12 individuals (13.04 % of the cohort), but only 4 (4.35 %) of these were also discovered by MLPA (Table 1). The rate of detection for CNV designated as VOUS – likely pathogenic when using CMA was thus 8.69 percentage points higher than that achieved using MLPA. This difference was significant (p = 0.039). These CNV represent potentially harmful mutations but there is currently insufficient evidence to classify them as being causal of the patient’s disorder. The rate of pathogenic variant detection using MLPA and CMA was 7.61 percentage points higher than that achieved by karyotyping alone in the cohort of ASD patients. A terminal deletion of chromosome 22 affecting the gene \textit{SHANK3} was detected by all methods in the patient with ring chromosome 22. However, without karyotyping, this deletion’s mechanism of occurrence would not have been determined. The duplication 15q11.2q13.1 (patient 1522/16) was identified as a chromosome heteromorphism during cytogenetic assessment, but both MLPA and CMA revealed the duplication.

Eight recurrent CNV known to be associated with ASD were found in our cohort – deletions 7q11.2 (Williams-Beuren Syndrome; WBS), 22q11.2 (Velocardiofacial Syndrome; VCFS), 22q13 (Phelan-McDermid Syndrome), 16p11.2, and Xp21.2-p21.1 (Becker muscular dystrophy; BMD); and duplications 1q21.1, 16p11.2, and 15q11-q13 (Table 2A). Several CNV encompassed genes reported to be important in the etiology of autism or schizophrenia (\textit{APOO, ARX, TSPAN7, NRXN1, CSMD1, CTNNA3, RBFOX1, MACROD2, ASMT, DISC1, PARK2, DOCK8}); these CNV were designated “likely pathogenic” because they are listed in curated databases (SFARI) or have repeatedly been identified as being involved in the etiology of autism or intellectual disability (Table 2B). Recurrent duplications (9p24.3) overlapping the \textit{DOCK8} gene (specifically, spanning exons 1-2, 2-43, and 1-26) were detected in 3 unrelated patients by MLPA and CMA (Figure 1). All of these
duplications are currently recorded as variants of unknown clinical significance. A similar MLPA analysis of a control cohort of 80 healthy individuals revealed no individuals with CNV encompassing this gene. Moreover, DOCK8 gain variants were found in only 12 of the 2504 healthy individuals whose genetic data was published by the 1000 Genomes Consortium (Mills et al., 2011). DOCK8 gain variants are thus significantly enriched in the ASD cohort relative to the population as a whole ($p < 0.05$).

We also analyzed the phenotypes of patients with similar duplications encompassing the gene DOCK8 that have been reported in the literature and the ISCA and Decipher databases (Table 3). The analyzed cases exhibited similar phenotypes involving ASD, DD, speech delay seizures, dysmorphic features and behavioral abnormalities. Additional phenotypes observed in our patients with DOCK8 duplications include hearing impairment and attention hyperactivity disorder (ADHD), which was also reported previously by Glessner (Glessner et al., 2017). Our patients also exhibited undescended testes (2) and atrophy of one testis (1).

Ten CNV with uncertain significance were found in the cohort (Table 2C). This category included variants that lack entries in the DGV but contained brain-expressed and/or dosage-sensitive genes, such as dup11q14-q21, which spans melatonin receptor 1B (MTNR1B) and fat tumor suppressor 3 (FAT3); del 6q22, which spans protein-tyrosine phosphatase, receptor-type, kappa (PTPRK); dup 17q21.31, which spans granulin (GRN) and run domain-containing protein 3A (RUNDC3A); and dup 9q33.2, which spans stomatin (STOM) and gelsolin (GSN). The category also included variants with entries in the DGV that span brain-expressed dosage-sensitive genes for which there is little or no evidence of involvement in the etiology of autism. CNV in this latter group were del 2p12, which spans catenin, alpha-2 (CTNNA2); dup 16q11.2, which spans vacuolar
protein sorting 35 (VSP35); and dup Xp11.22, which spans (SMC1A) and 17-beta-hydroxysteroid dehydrogenase X (HSD17B10). Other CNV in this group included variants overlapping OMIM morbid genes associated with etiologies for non-neuropsychiatric disorders, namely dup 5q32.3, del 2q13, and dup 16p12.2, spanning muscle segment homeobox 2 (MSX2), nephrocystin 1 (NPHP1), and otoancorin (OTOA), respectively.

Three patients exhibited multiple CNV designated as pathogenic, likely pathogenic, or VOUS-uncertain: D1377/15, D1190/16 and D1094/16 (Tab 2B, 2C).

Interestingly, secondary findings, i.e. CNV encompassing genes of interest (here defined as genes associated with autism or expressed in the brain), were discovered in the patient with the FMR1 mutation. One of these CNV encompassed the genes ASMT, and the other encompassed HSD17B10 and SMC1A. A paternally inherited deletion/SNV involving the MAPK3 gene (16p11.2 region) detected by MLPA was also identified in this patient.

CNV encompassing genes involved in the pathophysiology of parkinsonism were detected in 3 (3.26 %) patients. A deletion (covering exons 2 - 3) and a duplication (covering exon 2) of the gene PARK2 (6q26) were discovered in two unrelated patients, and duplication of 16q11.2 including the gene VSP35 (PARK17) was discovered in the third.

Discussion

The detection rate for pathogenic CNV in ASD patients achieved using MLPA was identical to that achieved using CMA. We therefore suggest that MLPA is sufficient to diagnose unambiguously pathogenic variants under some circumstances - for instance, when CMA is unavailable or in cases where the need to interpret VOUS variants or incidental findings would be problematic and patients would be unwilling to deal with the implications of such variants being detected.
especially in cases where a prenatal genetic diagnosis would have to be followed up with further investigations within the family). The patients would retain the ability to refuse to be informed of such findings despite their possible pathogenic impact; in such cases, CMA would be redundant because we have confirmed that clinically significant variants can be detected by both methods. This example demonstrates that the benefits of CMA in clinical applications differ from those in research contexts, and shows that CNV analysis by CMA in individuals with ASD can enable diagnosis and appropriate genetic counseling in a small number of cases. CMA is unavailable to some laboratories because it requires specialized equipment. Therefore, many labs would have to make a large capital outlay to perform CMA but will be readily able to perform MLPA with existing equipment to determine the causes of known syndromes. These laboratories can rely on targeted methods such as MLPA to exclude possible causes of syndromic autism. However, MLPA would not be sufficient to detect CNV associated with nonsyndromic ASD. Therefore, CMA remains an invaluable tool for studying the causes of ASD. Our results confirm that CMA outperforms MLPA at detecting CNV classified as VOUS – likely pathogenic. These variants can help reveal new genes involved in the emergence of ASD. Most ASD patients with detected pathogenic CNV are classified as having syndromic autism because the diagnosis of ASD is usually secondary to developmental delay or intellectual disability with further comorbidities such as dysmorphic features or growth delay that may suggest the involvement of a syndrome based on the procedures adopted here (see Supplemental files). Although all the patients in the studied cohort had been assessed by a genetic counsellor before the study was conducted, none of them had been suggested to have any syndrome prior to our testing. This could be partly due to the phenotypic variability of some syndromes.
Even in the boy with BMD, the diagnosis of the syndrome was based on MLPA testing and subsequent confirmation by targeted DNA analysis at 12 years of age. However, the primary reasons given when referring this patient for genetic investigation were severe growth delay, dystrophia together with autistic features. Because the patient’s dystrophia was milder than in DMD, the case was classified as BMD. The typical VCFS phenotype did not manifest in the patient with microdeletion 22q11.2 because the deleted interval did not include the *TBX1* or *HIRA* genes – the deletion was rather distal, spanning the ASD candidate gene *LZTR1* (Krumm et al., 2015).

The distal microdeletion 22q11.2 has been linked to behavioral and psychiatric impairments (Burnside, 2015). In the patient with ring chromosome 22, MLPA confirmed the suspected loss of the terminal part of chromosome 22 and the loss of the *SHANK3* gene, which has been associated with ASD (Durand et al., 2007; Nemirovsky et al., 2015). These findings explained the patient’s phenotype and resulted in a diagnosis of Phelan-McDermid syndrome. The most notable aspects of the patient’s phenotype were severe developmental delay and neuropsychiatric impairment (recently described as low functioning autism). However, at the age (24 months) when the girl was tested, the syndrome’s hallmarks had not fully manifested. Further testing was required in 4 patients with CNV VOUS – likely pathogenic and 6 patients with CNV VOUS - likely benign or benign because there may have been pathogenic CNV outside the loci covered by the tested probemixes in these cases. The detection rate of MLPA depends somewhat on the chosen probemix, but is comparable to that for CMA if one restricts one’s focus to clinically well described recurrent pathogenic CNV. This makes MLPA a convenient method for fast, reliable, and inexpensive targeted exclusion of CNV involved mostly in syndromic autism. The detection rate for pathogenic or likely pathogenic CNV by CMA in ASD patients ranges from 3 – 30 % depending
on the cohort and acceptance criteria (Nava et al., 2013; Wang et al., 2016; Guo et al., 2017; Shen et al., 2010; Bremer et al., 2011; Cappuccio et al., 2016; Leppa et al., 2016). In this work, CMA revealed 8 index cases with pathogenic CNV. This result is comparable to previous reports (Nava et al., 2013; Shen et al., 2010; Bremer et al., 2011). CMA achieved a higher detection rate for potentially pathogenic variants than MLPA or karyotyping in this work. However, without karyotyping it would have been very difficult to determine the mechanism of occurrence of deletion 22q13.33 (which was due to ring 22). Nor would we have detected the balanced chromosomal rearrangements in two patients. The possibility that these rearrangements may have contributed to the etiology of ASD in these patients cannot be completely excluded.

We observed a relatively high frequency of CNV encompassing genes associated with parkinsonism in our group of ASD patients. Variants encompassing PARK2 have previously been detected in ASD patients (Nava et al., 2013; Yin et al., 2016). However, we also detected a CNV involving VSP35 (PARK17) in one of our patients whose genome contained multiple CNV. This gene has been suggested to play a role in parkinsonism (Zimprich et al., 2011). The relatively high frequency of CNV overlapping genes associated with this disease raises the possibility that these patients may have an elevated risk of developing parkinsonism in adulthood. An increased frequency of parkinsonism among ASD patients has previously been reported (Starkstein et al., 2015). However, the role of parkinsonism-related genes in the etiology of ASD is currently unclear.

It is possible that different kinds of genomic changes affecting certain genes can lead to different phenotypes (Scheuerle & Wilson, 2011).

Two patients exhibited variants encompassing two melatonin-related genes: acetylserotonin O-methyltransferase (ASMT) and the melatonin receptor (MNTR1B). Both genes have been
identified as potentially affecting the risk of ASD (Cai et al., 2008; Nava et al., 2013; Jonsson et al., 2010; Anderson et al., 2009). The duplication region 11q14.3-q21 encompassing MNTR1B and FAT3 cooccurred with duplication Xq27.3 encompassing the SOX3 gene, which was previously linked to the etiology of ID, hypopituitarism, and speech disorders, but not ASD (Solomon, 2004; Stankiewicz et al., 2005).

Two patients exhibited variants overlapping genes encoding catenins (cadherin-associated proteins): deletion 10q21.3 overlaps catenin alpha 3 (CTNNA3), and deletion 2p12 overlaps catenin alpha 2 (CTNNA2). The first CTNNA3 deletion cooccurred with other CNV – deletion 8p23.2 (CSMD1), duplication 16q11.2 (VSP35), and duplication 1q42.2 (DISC1). While variations in CTNNA3 have been described in patients with ASD, variations in CTNNA2 have not (Folmsbee et al., 2016; Bacchelli et al., 2014). CTNNA2 is predominantly expressed in the brain and helps regulate the stability of synaptic contacts and axogenesis, brain morphogenesis, dendrite morphogenesis, and synapse structural plasticity (www.uniprot.org), making it a plausible candidate for involvement in the etiology of ASD.

Duplications covering the gene DOCK8 were identified in multiple patients by both MLPA and CMA. DOCK8 encodes a member of the Dock protein family of atypical Rho guanine nucleotide exchange factors (GEFs) for Rac and/or Cdc42 GTPases that play pivotal roles in various processes of brain development. To date, 11 members of the Dock family have been identified in mammals. Dock proteins regulate the actin cytoskeleton, cell adhesion and dendritic migration (Gadea & Blangy, 2014). There is also evidence that members of the Dock family are associated with several neurodegenerative and neuropsychiatric diseases, including Alzheimer’s disease and ASD (Shi, 2014). Homozygous loss of function of the DOCK8 gene causes Autosomal Recessive Hyper-IgE
Recurrent Infection Syndrome (Zhang et al., 2009). In addition, evidence was recently presented supporting a causal relationship between heterozygous disruption of DOCK8 and mental retardation, pervasive developmental disorders, autism, and bipolar disorders (Nava et al., 2013; Wang et al., 2016; Griggs et al., 2008; Glessner et al., 2017; Krgovic et al., 2018;). Our results support the findings of Glessner et al., who identified DOCK8/KANK1 as novel significant loci for ASD and ADHD (Glessner et al., 2017). We observed significant enrichment of CNV involving gains of the DOCK8 gene in the studied ASD cohort. This may indicate that the region of the DOCK8/KANK1 locus associated with ASD and ADHD is likely to be within or proximal to the gene DOCK8. Our patients with DOCK8 gains had all been diagnosed with ADHD. We detected no individuals with any CNV overlapping with the DOCK8 gene in our control cohort, so we regard this gene as an interesting candidate for further study on the etiology of ASD. Because the DOCK8 duplication was inherited from a healthy father in one case, we assume that variant increases the risk of ASD or other neuropsychiatric conditions but that its phenotypic impact may be limited by incomplete penetrance or/and variable expressivity. Both of these factors are known to complicate genetic counselling in patients with CNV encompassing neurosusceptibility loci.

The greatest limitation of this study, aside from the relatively small cohort, is the lack of information about the inheritance of most of the identified CNV. It seems that CNV with variable expressivity (del/dup16p11.2, NRXN1) are frequently inherited. An analysis of ASD patients’ parental genomes could thus help to explain the patients’ phenotypes. A segregational analysis was performed for the family with a heterozygotic loss in NRXN1, revealing that this CNV exhibited incomplete penetrance (Figure 2).
Conclusion

Diagnosis was achieved in only 8 index cases (8.7 % of the studied cohort), all of which involved patients classified as having syndromic autism. The rate of pathogenic CNV detection by CMA was identical to that achieved with MLPA using probemixes targeted towards losses associated with ID, DD, and ASD. Consequently, our results indicate that MLPA and CMA are equally reliable methods for obtaining clinically relevant findings. Therefore, MLPA can be used as a quick alternative to CMA for excluding syndromes associated with ASD. MLPA is also frequently used to confirm CMA findings, and for targeted verification of the origins of CNV during parental testing.

However, many genes are involved in the etiology of ASD, creating a clear need for whole-genome screening to identify genes associated with ASD and to clarify the clinical impact of VOUS. Moreover, increasing knowledge of known candidate genes in ASD provided by CMA (or NGS) enables to develop the new targeted tests (MLPA probemixes, targeted arrays) for quick exclusion known pathogenic mutations in particular patients. Our results confirm that traditional karyotyping is an indispensable tool for deciphering the origins of specific CNV and detecting balanced chromosomal changes, the clinical significance of which cannot be totally disregarded in ASD patients. This work presents further evidence that genes identified in the etiology of various genetic conditions can be linked to the pathophysiology of ASD (DOCK8). However, the exact pathophysiological mechanism underlying the functions of these genes in the development of phenotypes such as ASD remains unknown.

Abbreviations

ASD Autism spectrum disorders
PDD-NOS  Pervasive developmental disorder - not otherwise specified
MLPA  Multiplex ligation probe-dependent amplification
CMA  Chromosomal microarray analysis
CNV  Copy number variants
VOUS  Variant of unknown significance
ID  Intellectual disability
DD  Developmental delay

References

Anderson BM., Schnetz-Boutaud NC., Bartlett J., Wotawa AM., Wright HH., Abramson RK.,
Cuccaro ML., Gilbert JR., Pericak-Vance MA., Haines JL. 2009. Examination of association of genes
in the serotonin system to autism. *Neurogenetics* 10:209-216 DOI: 10.1007/s10048-009-0171-7.

Bacchelli E., Ceroni F., Pinto D., Lomartire S., Giannandrea M., D'Adamo P., Bonora E., Parchi P.,
Tancredi R., Battaglia A., Maestrini E. 2014. A CTNNA3 compound heterozygous deletion
implicates a role for αT-catenin in susceptibility to autism spectrum disorder. *Journal of Neurodevelopmental Disorders* 6:17 DOI: 10.1186/1866-1955-6-17.

Bremer A., Giacobini MB., Eriksson M., Gustavsson P., Nordin V., Fernell E., Gillberg C., Nordgren
A., Uppströmer Å., Anderlid B-M., Nordenskjöld M., Schoumans J. 2011. Copy number variation
characteristics in subpopulations of patients with autism spectrum disorders. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 156:115-124 DOI: 10.1002/ajmg.b.31142.

Burnside RD. 2015. 22q11.21 Deletion Syndromes: A Review of Proximal, Central, and Distal Deletions and Their Associated Features. *Cytogenetic and Genome Research* 146:89-99 DOI: 10.1159/000438708.

Cai G., Edelmann L., Goldsmith JE., Cohen N., Nakamine A., Reichert JG., Hoffman EJ., Zurawiecki DM., Silverman JM., Hollander E., Soorya L., Anagnostou E., Betancur C., Buxbaum JD. 2008. Multiplex ligation-dependent probe amplification for genetic screening in autism spectrum disorders: Efficient identification of known microduplications and identification of a novel microduplication in ASMT. *BMC Medical Genomics* 1:50 DOI: 10.1186/1755-8794-1-50.

Cappuccio G., Vitiello F., Casertano A., Fontana P., Genesio R., Bruzese D., Ginocchio VM., Mormile A., Nitsch L., Andria G., Melis D. 2016. New insights in the interpretation of array-CGH: autism spectrum disorder and positive family history for intellectual disability predict the detection of pathogenic variants. *Italian Journal of Pediatrics* 42:39 DOI: 10.1186/s13052-016-0246-7.

Chiocchetti AG., Bour HS., Freitag CM. 2014. Glutamatergic candidate genes in autism spectrum disorder: an overview. *Journal of Neural Transmission* 121:1081-1106 DOI: 10.1007/s00702-014-1161-y.
Deardorff MA., Kaur M., Yaeger D., Rampuria A., Korolev S., Pie J., Gil-Rodríguez C., Arnedo M.,
Loeys B., Kline AD., Wilson M., Lillquist K., Siu V., Ramos FJ., Musio A., Jackson LS., Dorsett D.,
Krantz ID. 2007. Mutations in Cohesin Complex Members SMC3 and SMC1A Cause a Mild Variant of Cornelia de Lange Syndrome with Predominant Mental Retardation. The American Journal of Human Genetics 80:485-494 DOI: 10.1086/511888.

Decipher https://decipher.sanger.ac.uk

Durand CM., Betancur C., Boeckers TM., Bockmann J., Chaste P., Fauchereau F., Nygren G.,
Rastam M., Gillberg IC., Anckarsäter H., Sponheim E., Goubran-Botros H., Delorme R., Chabane N., Mouren-Simeoni M-C., de Mas P., Bieth E., Rogé B., Héron D., Burglen L., Gillberg C., Leboyer M., Bourgeron T. 2007. Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. Nature Genetics 39:25-27 DOI: 10.1038/ng1933.

Edgar R., Domrachev M., Lash AE. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Research 30:207-10 DOI: 10.1093/nar/30.1.207.

Folmsbee SS., Wilcox DR., Tyberghein K., De Bleser P., Tourtellotte WG., van Hengel J., van Roy F., Gottardi CJ. 2016. AT-catenin in restricted brain cell types and its potential connection to autism. Journal of Molecular Psychiatry 4:2 DOI: 10.1186/s40303-016-0017-9.
Froyen G., Corbett M., Vandewalle J., Jarvela I., Lawrence O., Meldrum C., Bauters M., Govaerts K., Vandeleur L., Van Esch H., Chelly J., Sanlaville D., van Bokhoven H., Ropers H-H., Laumonnier F., Ranieri E., Schwartz CE., Abidi F., Tarpey PS., Futreal PA., Whibley A., Raymond FL., Stratton MR., Fryns J-P., Scott R., Peippo M., Sipponen M., Partington M., Mowat D., Field M., Hackett A., Marynen P., Turner G., Gécz J. 2008. Submicroscopic Duplications of the Hydroxysteroid Dehydrogenase HSD17B10 and the E3 Ubiquitin Ligase HUWE1 Are Associated with Mental Retardation. *The American Journal of Human Genetics* 82:432-443 DOI: 10.1016/j.ajhg.2007.11.002.

Gadea G., Blangy A. 2014. Dock-family exchange factors in cell migration and disease. *European Journal of Cell Biology* 93:466-477 DOI: 10.1016/j.ejcb.2014.06.003.

Glessner JT., Li J., Wang D., March M., Lima L., Desai A., Hadley D., Kao C., Gur RE., Cohen N., Sleiman PMA., Li Q., Hakonarson H. 2017. Copy number variation meta-analysis reveals a novel duplication at 9p24 associated with multiple neurodevelopmental disorders. *Genome Medicine* 9:106 DOI: 10.1186/s13073-017-0494-1.

Griggs BL., Ladd S., Saul RA., DuPont BR., Srivastava AK. 2008. Dedicator of cytokinesis 8 is disrupted in two patients with mental retardation and developmental disabilities. *Genomics* 91:195-202 DOI: 10.1016/j.ygeno.2007.10.011.
Guo H., Peng Y., Hu Z., Li Y., Xun G., Ou J., Sun L., Xiong Z., Liu Y., Wang T., Chen J., Xia L., Bai T., Shen Y., Tian Q., Hu Y., Shen L., Zhao R., Zhang X., Zhang F., Zhao J., Zou X., Xia K. 2017. Genome-wide copy number variation analysis in a Chinese autism spectrum disorder cohort. *Scientific Reports* 7:44155 DOI: 10.1038/srep44155.

ISCA http://dbsearch.clinicalgenome.org/search/

Jonsson L., Ljunggren E., Bremer A., Pedersen C., Landén M., Thuresson K., Giacobini MB., Melke J. 2010. Mutation screening of melatonin-related genes in patients with autism spectrum disorders. *BMC Medical Genomics* 3:10 DOI: 10.1186/1755-8794-3-10.

Kearney HM., Thorland EC., Brown KK., Quintero-Rivera F., South ST. 2011. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genetics in Medicine* 13:680-685 DOI: 10.1097/GIM.0b013e3182217a3a.

Krgovic D., Kokalj Vokac N., Zagorac A., Gregoric Kumperscak H. 2018. Rare structural variants in the DOCK8 gene identified in a cohort of 439 patients with neurodevelopmental disorders. *Scientific Reports* 8:9449 DOI: 10.1038/s41598-018-27824-0.
Krumm N., Turner TN., Baker C., Vives L., Mohajeri K., Witherspoon K., Raja A., Coe BP., Stessman HA., He Z-X., Leal SM., Bernier R., Eichler EE. 2015. Excess of rare, inherited truncating mutations in autism. Nature Genetics 47:582-588 DOI: 10.1038/ng.3303.

Leppa V M., Kravitz S N., Martin C L., Andrieux J., Le Caignec C., Martin-Coignard D., DyBuncio C., Sanders S J., Lowe J K., Cantor R M., Geschwind D H. 2016. Rare Inherited and De Novo CNVs Reveal Complex Contributions to ASD Risk in Multiplex Families. The American Journal of Human Genetics 99:540-554 DOI: 10.1016/j.ajhg.2016.06.036.

McCarthy M. 2014. Autism diagnoses in the US rise by 30%, CDC reports. BMJ 348:g2520-g2520 DOI: 10.1136/bmj.g2520.

Miller DT., Adam MP., Aradhya S., Biesecker LG., Brothman AR., Carter NP., Church DM., Crolla JA., Eichler EE., Epstein CJ., Faucett WA., Feuk L., Friedman JM., Hamosh A., Jackson L., Kaminsky EB., Kuhn RM., Lee C., Ostell JM., Rosenberg C., Scherer SW., Spinner NB., Stavropoulos DJ., Tepperberg JH., Thorland EC., Vermeesch JR., Waggoner DJ., Watson MS., Martin CL., Ledbetter DH. 2010. Consensus Statement: Chromosomal Microarray Is a First-Tier Clinical Diagnostic Test for Individuals with Developmental Disabilities or Congenital Anomalies. The American Journal of Human Genetics 86:749-764 DOI: 10.1016/j.ajhg.2010.04.006.
Mills RE., Walter K., Stewart C., Handsaker RE., Chen K., Alkan C., Abyzov A., Yoon SC., Ye K.,
Cheetham RK., Chinwalla A., Conrad DF., Fu Y., Grubert F., Hajirasouliha I., Hormozdiari F.,
Iakoucheva LM., Iqbal Z., Kang S., Kidd JM., Konkel MK., Korn J., Khurana E., Kural D., Lam HYK.,
Leng J., Li R., Li Y., Lin C-Y., Luo R., Mu XJ., Nemesh J., Peckham HE., Rausch T., Scally A., Shi X.,
Stromberg MP., Stütz AM., Urban AE., Walker JA., Wu J., Zhang Y., Zhang ZD., Batzer MA., Ding
L., Marth GT., McVean G., Sebat J., Snyder M., Wang J., Ye K., Eichler EE., Gerstein MB., Hurles
ME., Lee C., McCarroll SA., Korbel JO. 2011. Mapping copy number variation by population-scale
genome sequencing. Nature 470:59-65 DOI: 10.1038/nature09708.

Nakanishi M., Deardorff MA., Clark D., Levy SE., Krantz I., Pipan M. 2012. Investigation of autistic
features among individuals with mild to moderate Cornelia de Lange syndrome. American Journal
of Medical Genetics Part A 158A:1841-1847 DOI: 10.1002/ajmg.a.34014.

Nava C., Keren B., Mignot C., Rastetter A., Chantot-Bastaraud S., Faudet A., Fonteneau E., Amiet
C., Laurent C., Jacquette A., Whalen S., Afenjar A., Périsse D., Doummar D., Dorison N., Leboyer
M., Siffroi J-P., Cohen D., Brice A., Héron D., Depienne C. 2013. Prospective diagnostic analysis of
copy number variants using SNP microarrays in individuals with autism spectrum
disorders. European Journal of Human Genetics 22:71-78 DOI: 10.1038/ejhg.2013.88.

Nemirovsky SI., Córdoba M., Zaiat JJ., Completa SP., Vega PA., González-Morón D., Medina NM.,
Fabbro M., Romero S., Brun B., Revale S., Ogara MF., Pecci A., Marti M., Vazquez M., Turjanski A.,
Kauffman MA., Hu VW. 2015. Whole Genome Sequencing Reveals a De Novo SHANK3 Mutation
in Familial Autism Spectrum Disorder. *PLOS ONE* 10:e0116358 DOI:
10.1371/journal.pone.0116358.

Schaaf CP., Sabo A., Sakai Y., Crosby J., Muzny D., Hawes A., Lewis L., Akbar H., Varghese R.,
Boerwinkle E., Gibbs RA., Zoghbi HY. 2011. Oligogenic heterozygosity in individuals with high-
functioning autism spectrum disorders. *Human Molecular Genetics* 20:3366-3375 DOI:
10.1093/hmg/ddr243.

Schaefer GB., Mendelsohn NJ. 2013. Clinical genetics evaluation in identifying the etiology of
autism spectrum disorders: 2013 guideline revisions. *Genetics in Medicine* 15:399-407 DOI:
10.1038/gim.2013.32.

Schaefer G. 2016. Clinical Genetic Aspects of Autism Spectrum Disorders. *International Journal of
Molecular Sciences* 17:180 DOI: 10.3390/ijms17020180.

Scheuerle A., Wilson K. 2011. PARK2 copy number aberrations in two children presenting with
autism spectrum disorder: Further support of an association and possible evidence for a new
microdeletion/microduplication syndrome. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 156:413-420 DOI: 10.1002/ajmg.b.31176.

Shen Y., Dies KA., Holm IA., Bridgemohan C., Sobeih MM., Caronna EB., Miller KJ., Frazier JA.,
Silverstein I., Picker J., Weissman L., Raffalli P., Jeste S., Demmer LA., Peters HK., Brewster SJ.,
Kowalczyk SJ., Rosen-Sheidley B., McGowan C., Duda AW., Lincoln SA., Lowe KR., Schonwald A., Robbins M., Hisama F., Wolff R., Becker R., Nasir R., Urion DK., Milunsky JM., Rappaport L., Gusella JF., Walsh CA., Wu B-L., Miller DT. 2010. Clinical Genetic Testing for Patients With Autism Spectrum Disorders. *Pediatrics* 125:e727-e735 DOI: 10.1542/peds.2009-1684.

Shi L. 2014. Dock protein family in brain development and neurological disease. *Communicative & Integrative Biology* 6:e26839 DOI: 10.4161/cib.26839.

Solomon NM. 2004. Array comparative genomic hybridisation analysis of boys with X linked hypopituitarism identifies a 3.9 Mb duplicated critical region at Xq27 containing SOX3. *Journal of Medical Genetics* 41:669-678 DOI: 10.1136/jmg.2003.016949.

South ST., Lee C., Lamb AN., Higgins AW., Kearney HM. 2013. ACMG Standards and Guidelines for constitutional cytogenomic microarray analysis, including postnatal and prenatal applications: revision 2013. *Genetics in Medicine* 15:901-909 DOI: 10.1038/gim.2013.129.

Stankiewicz P., Thiele H., Schlicker M., Cseke-Friedrich A., Bartel-Friedrich S., Yatsenko SA., Lupski JR., Hansmann I. 2005. Duplication of Xq26.2-q27.1, including SOX3, in a mother and daughter with short stature and dyslalia. *American Journal of Medical Genetics Part A* 138A:11-17 DOI: 10.1002/ajmg.a.30910.
Starkstein S., Gellar S., Parlier M., Payne L., Piven J. 2015. High rates of parkinsonism in adults with autism. *Journal of Neurodevelopmental Disorders* 7:29 DOI: 10.1186/s11689-015-9125-6.

de la Torre-Ubieta L., Won H., Stein JL., Geschwind DH. 2016. Advancing the understanding of autism disease mechanisms through genetics. *Nature Medicine* 22:345-361 DOI: 10.1038/nm.4071.

UniProt: the universal protein knowledgebase. The UniProt Consortium. 2017. www.uniprot.org. Accessed 21 January 2017.

Wang J-C., Mahon LW., Ross LP., Anguiano A., Owen R., Boyar FZ. 2016. Enrichment of small pathogenic deletions at chromosome 9p24.3 and 9q34.3 involving DOCK8, KANK1, EHMT1 genes identified by using high-resolution oligonucleotide-single nucleotide polymorphism array analysis. *Molecular Cytogenetics* 9:82. DOI: 10.1186/s13039-016-0291-3.

Yin C-L., Chen H-I., Li L-H., Chien Y-L., Liao H-M., Chou MC., Chou W-J., Tsai W-C., Chiu Y-N., Wu Y-Y., Lo C-Z., Wu J-Y., Chen Y-T., Gau SS-F. 2016. Genome-wide analysis of copy number variations identifies PARK2 as a candidate gene for autism spectrum disorder. *Molecular Autism* 7:23 DOI: 10.1186/s13229-016-0087-7.
Zhang Q., Davis JC., Lamborn IT., Freeman AF., Jing H., Favreau AJ., Matthews HF., Davis J., Turner ML., Uzel G., Holland SM., Su HC. 2009. Combined Immunodeficiency Associated with DOCK8 Mutations. *New England Journal of Medicine* 361:2046-2055 DOI: 10.1056/NEJMoa0905506.

Zhou Y. 2006. Simplified Molecular Diagnosis of Fragile X Syndrome by Fluorescent Methylation-Specific PCR and GeneScan Analysis. *Clinical Chemistry* 52:1492-1500 DOI: 10.1373/clinchem.2006.068593.

Zimprich A., Benet-Pagès A., Struhal W., Graf E., Eck S H., Offman M N., Haubenberger D., Spielberger S., Schulte E C., Lichtner P., Rossle S C., Klopp N., Wolf E., Seppi K., Pirker W., Presslauer S., Mollenhauer B., Katzenschlager R., Foki T., Hotzy C., Reinthaler E., Harutyunyan A., Kralovics R., Peters A., Zimprich F., Brücke T., Poewe W., Auff E., Trenkwalder C., Rost B., Ransmayr G., Winkelmann J., Meitinger T., Strom T M. 2011. A Mutation in VPS35, Encoding a Subunit of the Retromer Complex, Causes Late-Onset Parkinson Disease. *The American Journal of Human Genetics* 89:168-175 DOI: 10.1016/j.ajhg.2011.06.008.
### Table 1 (on next page)

Detection rates of rare CNV in patients with ASD for different methods

* detected by karyotyping, MLPA and CMA: r(22)(q13.3)

** detected by MLPA and CMA

***15 confirmed by CMA, 3 not confirmed

**** CNVs of coding region with frequency <1% in population

2a t(10;11)(q26;p13)pat; inv(Y)(p11.2q11.23)
| Method | pathogenic CNV | %     | VOUS likely pathogenic | %     | VOUS uncertain | %     | VOUS likely benign | %     | patients with rare variants totally | %     | negative |
|--------|----------------|-------|------------------------|-------|----------------|-------|-------------------|-------|-------------------------------|-------|----------|
| karyo  | 1              | 1.09  | 0                      | 0     | 0              | 0     | 2                 | 2.17  | 2^a + 1^*                      | 3.26  | 89       |
| FMR1   | 1              | 1.09  | -                      | -     | -              | -     | -                 | -     | 1                             | 1.09  | 91       |
| MLPA   | 7+1^*          | 8.69  | 4                      | 4.35  | 0              | 0     | 6                 | 6.52  | 18***                         | 19.57 | 74       |
| CMA    | 7+1^*          | 8.69  | 8+4**                  | 13.04 | 6              | 6.52  | 26+3**            | 31.52 | 55****                        | 59.78 | 37       |
|        | 9              | 9.78  | 12                     | 13.04 | 6              | 6.52  | 34                | 36.96 | 61                           | 66.3  | 31       |
Table 2 *(on next page)*

The list of detected CNV

*not in DGV

**patient with FMR1 mutation

***patients with multiple CNV
### A. Pathogenic CNV

| Patient ID | Band      | CNV status | Region GRCh37/hg19 | Length (kb) | Inheritance | Gender | Method of detection                     | Syndrome (phenotype MIM number) |
|------------|-----------|------------|--------------------|-------------|-------------|--------|-----------------------------------------|---------------------------------|
| D980/11    | 1q21.1-q21.2 | gain       | 146476526-147825662 | 1349        | de novo     | F      | MLPA(P297), CMA                        | dup 1q21.1(612475)              |
| D1277/08   | 7q11.22   | loss       | 72701018-74143060  | 1442        | de novo     | M      | MLPA(P245), CMA                        | WBS(194050)                     |
| D1522/16   | 15q11.2-q13.1 | gain     | 20737094-31293264  | 10556       | de novo     | M      | MLPA(P297,P343, P070,P036, P245), CMA | dup 15q11q13(608636)            |
| D731/15    | 16p11.2   | loss       | 29432212-30190029  | 758         | paternal    | M      | MLPA(P297,P343), CMA                   | del 16p11.2 (611913)            |
| D767/14    | 16p11.2   | gain       | 29600878-30177240  | 576         | paternal    | M      | MLPA(P297,P343), CMA                   | dup 16p11.2 (614671)            |
| D1981/12   | 22q11.21  | loss       | 20733667-21460220  | 727         | NA          | M      | MLPA(P245), CMA                        | del 22q11 (192430)              |
| 1764/16    | 22q13.31-q13.33 | loss | 47349588-51197838  | 3848        | de novo     | F      | MLPA(P070,P036,P343, P245), karyorr (22)(q13.3), CMA | Phelan-McDermid sy (606230)     |
| 770/16     | Xp21.1-q21.2 | loss     | 31518523-31948537  | 430         | NA          | M      | MLPA(P245), CMA                        | BMD (300376)                    |
## B. VOUS – likely pathogenic CNV

| Patient ID | Band       | CNV status | Region GRCh37/hg19 | Length (kb) | Inheritance | Gender | Method of detection     | Gene of interest |
|------------|------------|------------|--------------------|-------------|-------------|--------|-------------------------|----------------|------------------|
| D1377/15*** | 1q42.2     | gain       | 231712603-231816159 | 104         | NA          | M      | CMA                     | DISC1          |
| D1377/15*** | 10q21.3    | loss       | 68312378-68445989  | 134         | M           | CMA    | CSMD1                   |                |
| D1377/15*** | 8p23.2     | loss       | 3879391-4100961    | 222         | NA          | M      | CMA                     | NRXN1          |
| D357/12     | 2p16.3     | loss       | 50943528-51041472  | 98          | maternal    | M      | CMA                     | PARK2          |
| D1320/14    | 6q26       | loss       | 162631070-162982289| 351         | paternal    | M      | CMA                     | PARK2          |
| D1123/15    | 6q26       | gain       | 162716322-162912832| 197         | NA          | M      | CMA                     |                |
| D714/10     | 9p24.3     | gain       | 1-271132           | 271         | NA          | M      | CMA                     | DOCK8          |
| D684/13     | 9p24.3     | gain       | 203861-398865      | 195         | NA          | M      | CMA, MLPA (P070)        | DOCK8          |
| D1748/15    | 9p24.3     | gain       | 271533-440683      | 169         | maternal    | M      | CMA, MLPA (P070)        |                |
| 840/15      | 16p13.3    | loss       | 7212403-7459707    | 247         | NA          | M      | CMA                     | RBFOX1         |
| D1327/13    | 20p12.1    | loss       | 14806577-15042599  | 236         | NA          | F      | CMA                     | MACROD2        |
| D781/16     | Xp22.11p21.3 | gain       | 23794728-28429500  | 4635        | de novo     | F      | MLPA (P106-MRX), CMA   | APOO,POLA1,ARX* |
| D1094/16**  | Xp22.33    | gain       | 1755742-1821317    | 66          | NA          | M      | CMA                     | ASMT*          |
| D1194/15    | Xp11.4     | gain       | 38486618-38634614  | 148         | NA          | F      | MLPA (P106-MRX), CMA   | TSPAN7         |
| D1190/16*** | Xq27.3     | gain       | 138614319-143156110| 4541        | NA          | M      | CMA                     | SOX3, F9, ATP11C, CDR1, LDOC1* |

## C. VOUS – uncertain CNV
| Patient ID | Band       | CNV status | Region GRCh37/hg19 | Length (kb) | Inheritance | Gender | Method of detection | Gene of interest |
|-----------|------------|------------|--------------------|-------------|-------------|--------|---------------------|-------------------|
| D1474/15  | 2p12       | loss       | 80271129-80377518  | 106         | NA          | F      | CMA                 | CTNNA2            |
| D538/16   | 2q13       | loss       | 110874326-111365996| 492         | maternal    | F      | CMA                 | NPHP1             |
| D1471/14  | 6q22.33    | loss       | 128633613-128811348| 178         | NA          | M      | CMA                 | PTPRK*            |
|           |            |            |                    |             |             |        |                     | GSN, STOM, DAB2IP, EPB72* |
| D970/14   | 9q33       | gain       | 124034190-12444190 | 408         | NA          | NA(maternal excluded) | CMA | MTNR1B (melatonin receptor), FAT3, SLC36A4* |
| D1190/16*** | 11q14.3-q21 | gain       | 90076743-93118662  | 3040        | NA          | M      | CMA                 | MSX2              |
| D1190/16*** | 5q32.2   | gain       | 174151663-174157517| 5           | NA          | M      | CMA                 | MYLK3, ORC6, SHCBP1, VPS35 |
| D1377/15*** | 16q11.2       | gain       | 46600773-46830637  | 230         | NA          | M      | CMA                 | OTOA, METTL9, IGSF6, GRN, ITGA2B, SLC4A1, ASB16, GPATCH8, RUNDC3A, SLC25A39, UBTF, SHC1P2* |
| D1116/12  | 16p12.2    | gain       | 21591157-21951415  | 360         | NA          | M      | CMA                 | SMC1A, HSD17B10  |
| D2121/13  | 17q21.32   | gain       | 42212006-42542693  | 331         | NA          | F      | CMA                 |                   |
| D1094/16** | Xp11.22    | gain       | 53444924-53459515  | 15          | NA          | M      | CMA                 |                   |
**Table 3** (on next page)

Clinical characterisation of the patients with DOCK8 gains.

*Glessner et al., 2017; CC – corpus callosum*
| resource | ISCA         | Decipher     | Krogovic     | N   | %    | D714/10 | D684/13 | D1748/15 |
|----------|--------------|--------------|--------------|-----|------|---------|---------|----------|
| GRCh37/hg19 | 52389-416351 | 41587-489842 | 204193-271316 | 58  | 100.0|         |         |          |
| number of assessed cases | 37 | 19 | 2 | 58 | 100.0 | + | - | - |
| DD | 8 | 4 | 1 | 13 | 22.41 | + | - | - |
| ASD* | 3 | 6 | 1 | 10 | 17.24 | + | + | + |
| ID | 1 | 6 | 1 | 8 | 13.79 | + | - | - |
| behavioural abn. | 0 | 4 | 2 | 6 | 10.34 | - | - | - |
| dysmorphic | 3 | 3 | 1 | 7 | 12.06 | + | - | - |
| speech delay/disorder | 0 | 2 | 2 | 4 | 6.9 | + | + | + |
| seizures | 2 | 2 | 0 | 4 | 6.9 | - | - | - |
| ambiguous genitalia | 1 | 1 | 0 | 2 | 3.45 | - | - | + |
| obesity | 0 | 2 | 0 | 2 | 3.45 | - | - | + |
| sleep disturbance | 0 | 1 | 0 | 1 | 1.72 | - | - | + |
| microcephaly | 0 | 0 | 1 | 1 | 1.72 | - | - | - |
| dolichocephaly | 0 | 1 | 0 | 1 | 1.72 | - | - | - |
| plagiocephaly | 0 | 1 | 0 | 1 | 1.72 | - | - | - |
| craniosynostosis | 0 | 1 | 0 | 1 | 1.72 | - | - | - |
| cardiac abnorm. | 1 | 0 | 0 | 1 | 1.72 | - | - | - |
| short stature | 0 | 1 | 0 | 1 | 1.72 | - | - | - |
| agenesis CC | 1 | 0 | 0 | 1 | 1.72 | - | - | - |
| feeding problems | 0 | 1 | 0 | 1 | 1.72 | - | - | - |
| undescended testes | 0 | 0 | 0 | 0 | 0 | + | - | + |
| ADHD* | 0 | 0 | 0 | 0 | 0 | + | + | + |
| hypacusis | 0 | 0 | 0 | 0 | 0 | - | + | + |
Figure 1

Duplication of the gene DOCK8 in 3 unrelated patients.
Figure 2

Segregation analysis for the family of a patient bearing a CNV causing loss of NRXN1.