DETECTION OF GENOMIC IMBALANCES BY ARRAY-BASED COMPARATIVE GENOMIC HYBRIDIZATION IN BULGARIAN PATIENTS WITH AUTISM SPECTRUM DISORDERS

Daniela Avdjieva-Tzavella[1], Savina Hadjidekova[2], Blaga Rukova[2], Desislava Nesheva[2], Ivan Litvinenko[1], Dimitrina Hristova-Naydenova[3], Emil Simeonov[4], Radka Tincheva[1], Draga Toncheva[2]

[1]Medical University of Sofia, University Pediatrics Hospital, Sofia, Bulgaria

[2]Medical University of Sofia, Medical Faculty, Sofia, Bulgaria

[3]Tokuda Hospital, Department of Pediatrics, Sofia, Bulgaria

[4]University Hospital Alexandovska, Clinic of Pediatrics, Sofia, Bulgaria

Correspondence to: Daniela Avdjieva-Tzavella

E-mail: davdjieva@yahoo.com

ABSTRACT

Autism spectrum disorders (ASDs) constitute a group of behaviorally-defined conditions whose main features are qualitative changes in social interactions, defect in communication abilities, and repetitive and stereotyped interests and activities. ASDs are disorders which can be either isolated, or syndromic. The exact etiology of autism remains unknown, although it is likely to result from a complex combination of genetic and nongenetic factors. One genetic mechanism known to be associated with ASDs is submicroscopic chromosomal imbalances that are undetectable at the level of traditional cytogenetic analysis. Array-based comparative genomic hybridization (array CGH) is a powerful and high-resolution approach for detection of DNA copy number variants (CNVs). Forty-seven autistic patients were investigated using a whole-genome oligo-based array CGH, covering the genome at an average distance of 35 kb. Four clinically significant rearrangements, ranging from 494 kb to 3.47 Mb in size, were identified in 3 patients. Confirmation studies were performed on array CGH results using FISH. These data strongly support the idea that only a whole-genome high-resolution analysis such as array CGH is able to provide an accurate diagnosis for chromosomal imbalance in patients with ASDs.

Keywords: array-based comparative genomic hybridization (array CGH), autism spectrum disorders (ASDs), genomic imbalances

Abbreviations: Array CGH: Array-based comparative genomic hybridization; ASDs: Autism spectrum disorders; CNVs: copy number variants; DGV: Database of Genomic Variants; DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, 4th Ed.; FISH: fluorescence *in situ* hybridization; FMRP: fragile X mental retardation protein; PBLs: peripheral blood lymphocytes; PDD-NOS: pervasive developmental disorder not otherwise specified; SNP: single nucleotide polymorphism

Introduction

Autism is a neurodevelopmental condition defined by impairments in three core domains: social interaction, language and range of interests. In addition to autism, autism spectrum disorders (ASDs) include Asperger syndrome, pervasive developmental disorder not otherwise specified (PDD-NOS), childhood disintegrative disorder, and Rett syndrome (1, 42). The prevalence of autism is 1:500, while the prevalence of ASDs is 1:160 (11). ASDs can be either isolated or syndromic, i.e. associated with other clinical features. About 10 % to 20 % of the individuals with ASDs have an identifiable genetic etiology corresponding to known rare single-gene disorders or chromosomal rearrangements (13). The other cases remain unexplained, although they are likely to result from a complex combination of genetic and nongenetic factors. One genetic mechanism known to be associated with ASDs is submicroscopic chromosomal imbalances that are undetectable at the level of traditional cytogenetic analysis. The development of microarray based technologies for comparative genomic hybridization analysis (array CGH) has enabled the detection of copy number variants (CNVs) throughout the whole genome (6, 16, 23, 32). The detection rate in different studies is variable and it is highly influenced by the resolution of the applied array platform, the clinical selection of patients and the criteria used to define the CNVs as pathogenic. One of the first studies of array CGH analysis in ASDs was published by Jacquemont et al. in 2006 (16). They identified eight clinically relevant CNVs among a cohort of 29 patients with syndromic ASDs (27.5 %). Subsequently, a larger series of 427 unrelated ASDs cases have been investigated by single nucleotide polymorphism (SNP) microarray. This analysis identified 277 CNVs in 44 % of the patients and CNVs were classified as pathogenic in 14 % of the cases (23). In the same year, Christian et al. (6) showed an overall CNVs rate of 11.6 % in 397 autistic patients. More recently, Qiao et al. (32) investigated 100 subjects with syndromic ASDs using BAC array CGH and identified 9 % pathogenic CNVs. In the above reported literature, only a fraction of CNVs ranging from 9 % to 27 % have been classified as pathogenic.

This paper reports the oligo-based array CGH investigation of a Bulgarian group of children with ASDs that was performed in an attempt to further delineate the contribution of microstructural genomic abnormalities to the etiology of these diseases.

Materials and Methods

Subjects

We studied a group of 47 patients (35 boys and 12 girls) with ASDs. The clinical diagnosis was made in accordance with the criteria of DSM-IV (1). All patients were evaluated by a multidisciplinary professional staff including: clinical geneticist, psychiatrist, and psychologists. Patients with associative medical conditions including chromosomal aberrations, fragile-X syndrome, neurocutaneous diseases, and metabolic disorders were excluded from the study.

The research protocol was approved by The Ethics Committee for Research Investigations to The Medical University of Sofia, Bulgaria. Informed consent was obtained from the guardians of all patients.

Methods

We used two groups of methods: oligo-based array CGH and FISH. DNA was extracted from peripheral blood lymphocytes (PBLs) of the patients by the phenolchloroform method. The DNA pool samples used for reference were extracted from PBLs of six physically and clinically normal males and females without family history of psychiatric disorders. DNA concentration was measured by Nanodrop 2000c, and the purity of DNA was also estimated. As an additional quality check DNA, DNA was run in a 1 % agarose gel: DNA of high molecular weight (more than 10 kbp to 20 kbp) met the quality criteria and was suitable for use.

Oligo-based arrays

We used BlueGnome CytoChip oligo 2X105K microarray, v1.1, with 35 kb backbone resolution. Test and sex-matched reference genomic DNA (500 ng) was labeled by random
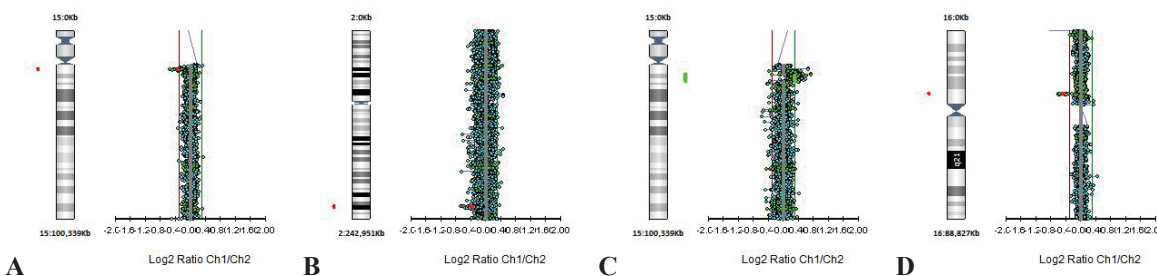

Fig. 1. Genomic profiles of the involved chromosomal regions obtained from oligo array: (**A**) Patient 1 chip idiogram of chromosome 15 showing deletion, arr15q11.2(20,317,022-20,811,926)x1; (**B**) Patient 2 chip idiogram of chromosome 2 showing deletion, arr2q36.3 (227,041,210-228,497,433)x1; (**C**) Patient 2 chip idiogram of chromosome 15 showing duplication, arr15q12q13.1(23,602,087-27,075,266)x3; (**D**) Patient 3 chip idiogram of chromosome 16 showing deletion, arr16p11.2(29,592,950-30,106,078)x1.
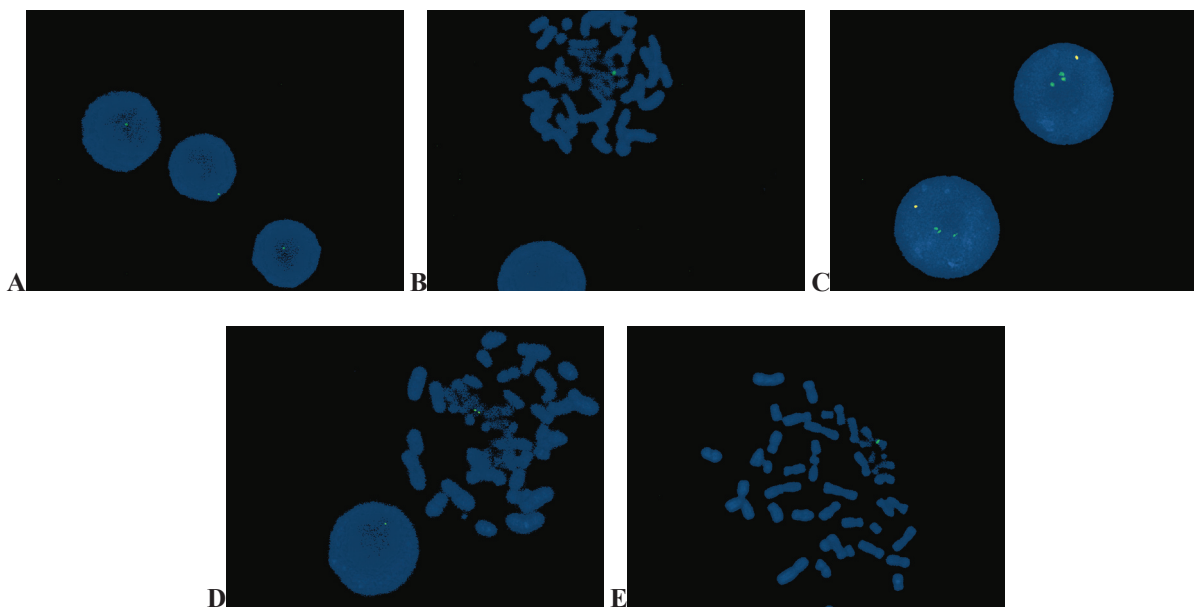

Fig. 2. FISH validation with BAC clones: (**A**) Patient 1 heterozygous deletion in 15q11.2 (one green signal) – ish 15q11.2 (RP11-26F2x1); (**B**) Mother of patient 1– ish 15q11.2 (RP11-26F2x1); (**C**) Patient 2 heterozygous deletion in 2q36.3 (one orange signal) – ish 2q36.3 (RP11-185H21x1) and duplication in 15q12 (three green signals) – ish 15q12 (RP11-142M24x3); (**D**) Patient 3 heterozygous deletion in 16p11.2 (one green signal) – ish 16p11.2 (RP11-301D18x1); (**E**) Father of patient 3 – ish 16p11.2 (RP11-301D18x1).

priming, using BlueGnome Fluorescent dUTP Labeling System. Arrays were scanned on a GenePix 4100A two-color fluorescent scanner (Axon Instruments, Union City, CA, U.S.A.). The images were analyzed by BlueFuse Multi, version 2.2. (BlueGnome, Cambridge, UK). Normal copy numbers were considered in log2 ratio between -0.3 and +0.3, values above +0.3 were evaluated as duplications and the ones under -0.3, as losses (deletions).

FISH analysis

To confirm the regions of copy number variation that are not present in the general population, we performed fluorescence *in situ* hybridization (FISH) with pre-labeled BlueFish BAC clones from the BlueGnome database. The FISH was carried out on interphases and metaphases from cultured PBLs according to the standard procedure.

Results and Discussion

Although ASDs are well known to be strongly genetic, no single genetic cause has been identified that contributes to this group of disorders in a significant proportion of patients. With the advent of genome-wide screening methods for copy number changes, a significant amount of new cytogenetic abnormalities, involving all chromosomes, have been reported in association with ASDs. The majority, however, have been found in single cases and their importance remains unknown. Chromosome abnormalities are more prevalent in syndromic (27.5 %) ASD cases than in simple cases (7.2 %) (16, 35). We used the technique of oligo array CGH to screen for micro-structural whole genome copy number changes in 47 autistic patients. We detected four potentially clinically relevant genomic abnormalities in 3 patients, which corresponds to about 6 % of the cohort. This is lower than the 9 % and 11.6 % reported by Qiao et al. (32) and Christian et al. (6) respectively. The pathogenic role of CNVs was determined on the basis of the following criteria: already reported in ASDs, size, genes content and knowledge on gene function/expression (7, 9). The four CNVs reported in this study contain 22 RefSeq genes which could be considered as candidate genes for autism (**Table 1**). The size of the aneusomic segments varied from 494 kb to 3.47 Mb. One girl (patient 2) was identified with

two rearrangements (**Fig. 1B, C**) and the two remaining cases (patients 1 and 3) exhibited one rearrangement (**Fig. 1A, D**). The array CGH results were confirmed by FISH analysis (**Fig. 2**). To further show the clinical relevance of these chromosomal aberrations, clinically normal parents of patients 1 and 3 were tested by FISH to discriminate between *de novo* and inherited anomalies. The two rearrangements were shown to be inherited from a normal parent. The first is a 15q11.2 microdeletion in a boy (patient 1), inherited from his healthy mother (**Fig. 2B**). The second (patient 3) is a 16p11.2. microdeletion inherited from the healthy father (**Fig. 2E**). For patient 2 *de novo* occurrence could not be examined because her parents refused to be tested. While the occurrence of specific CNVs in healthy controls is generally indicative of a lack of pathogenicity, issues such as mosaicism, incomplete penetrance, variable expression, positional effect, and gene–gene interactions could modify biological functions and need to be considered. Lee et al. (21) proposed that apparently inherited CNVs sometimes have different breakpoints when analyzed with higher resolution assays and thus might result in different functions or phenotypes. For these reasons and the availability of promising candidate genes in 15q11.2 and 16p11.2 regions we suggest that these CNVs, although inherited from healthy parents, are a cofactor of the disease.

The microdeletion 15q11.2 in patient 1 involves four genes: *TUBGCP5, NIPA1, NIPA2,* and *CYFIP1*; the latter three are widely expressed in the central nervous system, while *TUBGCP5* is expressed in the subthalamic nuclei. Multiple cases with microdeletions of 15q11.2 with different behavioral problems were described by Doornbos et al. (8). They found that the deletion often was inherited from a normal or mildly affected parent as a result of an incomplete penetrance. According to Napoli et al. (27) *CYFIP1* is the best candidate gene in this region for a causal role in ASDs. This gene is a part of a complex that includes fragile-X mental retardation protein (FMRP) that regulates mRNA transport and translation at the synapse, controlling critical aspects of activity-dependent protein synthesis. Other studies highlighted *CYFIP1* and *NIPA1* as autism risk genes functioning in axonogenesis and synaptogenesis (39). A third gene, *TUBGCP5*, could be

TABLE 1

Clinical data and array CGH testing results of the patients with identified microaberrations

Patient No	Chromosome	Rearrangement	Inheritance	Size (bp)	Candidate genes	Clinical data
1	15q11.2	loss	maternal	494,904	*NIPA1, NIPA2, CYFIP1*	Autism, moderate mental retardation
2	2q36.3	loss	unknown	1,456,223	*MFF, SLC19A3, TM4SF20, WDR69*	Autism, severe mental retardation, dysmorphic face
	15q12- q13.1	gain	unknown	3,473,180	*ATP10A, GABRB3, GABRA5, GABRG3, APBA2*	
3	16p11.2	loss	paternal	513,128	*SEZ6L2, TAOK2, MAZ, DOC2A, PPP4C, ALDOA, HIRIP3, MAPK3, QPRT*	Autism, moderate mental retardation, dysmorphic face

involved in behavioral disorders, such as ADHD and obsessive-compulsive behavior (5). In a large study in schizophrenia patients deletions of the exact same region on chromosome 15 were found to be associated with schizophrenia (37). It is interesting that our patient's aunt suffered from schizophrenia.

A genome-wide analysis of structural variations ruled by Marshall et al. (23) identified an autism susceptibility locus on 16p11.2 involving deletions and duplications. They reported that this CNV region was found at near to 1 % in the studied autistic cohort and not in controls. Kumar et al. (19) suggested that the 16p11.2 microdeletion is one of the most common recurrent genomic disorders associated with autism. Several studies proposed that duplication of the same region similarly influences susceptibility to developmental delay with variable features of autism and that one or more genes in the region may be particularly dosage sensitive with significant developmental manifestations when the region is either duplicated or deleted (10, 41). The microdeletion 16p11.2 in patient 3 contains several genes that represent promising candidates for autism based on known expression and functional data (**Table 1**). *SEZ6L2* expression in human fetal brain was highest in post-mitotic cortical layers, hippocampus, amygdala, and thalamus (20). *TAOK2* encodes a serine/threonine kinase that has a role in cell signaling, microtubule organization and stability, and in apoptosis. Highest *TAOK2* expression was detected in fetal brain and in all specific adult brain regions examined (20). *MAZ* is expressed in several tissues, with the highest expression in brain found in motor and midfrontal cortex. *MAZ* directly regulates genes involved in GABA signaling, neuronal differentiation and the serotonin pathway (26, 29). *DOC2A* is expressed predominantly in the brain and localizes to synaptic vesicles. It has been hypothesized that *DOC2A* regulates synaptic activity though Ca-dependent mechanisms (12). *PPP4C* encodes a serine/threonine phosphatase that has a role in motor neuron survival (4). *ALDOA* is expressed in brain and other tissues. Martins-de-Souza et al. (24) showed that male and female schizophrenia patients presented different patterns of aldolase C activity in brain. *HIRIP3* is a widely expressed gene that interacts directly with *HIRA*, a major candidate for the DiGeorge syndrome and related developmental disorders (2). *MAPK3* is expressed in human fetal and adult brains and MAPK3 acts as an intermediary across several pathways that have been implicated in ASDs pathogenesis, including: serotonin, oxytocin and IL-6/immune signaling (31). QPRT encodes a key enzyme in the catabolism of quinolinate that acts as a most potent endogenous excitotoxin to neurons. Elevation of quinolinate levels in the brain has been associated with the pathogenesis of several neurodegenerative disorders (38). Although the subjects with 16p11.2 deletion and duplication do not have clinically recognisable phenotypes, Shinawi et al. (36) found that certain facial features are shared among them. Subjects with the deletions shared the following features: broad forehead, micrognathia, hypertelorism, and a flat midface which were different from the dysmorphic features in our patient: up-slanting palpebral fissures, synophrys, large

prominent ears, strabismus, hypertelorism, and anteverted nares. Patient 2 (girl with autism, severe mental retardation, and dysmorphic face: prognathism, forward-placed large incisors, posteriorly rotated ears) had two relatively large rearrangements (**Fig. 1B, C**). Aberrations of 15q11-q13 (deletions, duplications and inversions) in ASDs have been reported in several studies with frequencies ranging from 1.0 % to 3.0 % (22, 40). In our patient the duplicated 15q12-q13.1 region harbored several genes with expression in the central nervous system and in synapses and others that code for protein subunits of neurotransmitters (**Table 1**). *ATP10C* is preferentially expressed from the maternal chromosome in the brain. The product of the *ATP10C* gene is thought to function as a phospholipid transporter protein that may be involved in CNS signaling (14). Linkage to this gene has been detected in ASDs but no mutations have been identified in a small cohort of subjects (18, 28). Three of the γ-aminobutyric acid (GABA) receptor subunit genes (*GABRB3, GABRA5* and *GABRG3*) form a cluster on chromosome 15q11-q13. GABA is an inhibitory neurotransmitter in the mammalian brain (34). Several studies identified evidence for linkage or association with the GABA receptor genes on chromosome 15q in autism patients (15, 17). The protein encoded by *APBA2* interacts and inhibits production of proteolytic Alzheimer's disease amyloid precursor protein (APP) (25). Babatz et al. (3) identified novel nonsynonymous coding variants on *APBA2* in ASD subjects. The microdeletion 2q36 in patient 2 involves four possible candidate genes for autism and mental retardation. *MFF* functions as an essential factor in mitochondrial fission (30). No data has been reported on the relation between *MFF* mutations and autism, but it is well known that mitochondrial dysfunction can lead to ASDs (33). Homozygous or compound heterozygous mutations in *SLC19A3* cause two distinct clinical phenotypes, biotin-responsive basal ganglia disease and Wernicke's-like encephalopathy (43). *TM4SF20* and *WDR69* are expressed in the central nervous system, however their functions have not been well investigated. Our patient's phenotype was caused by two different aberrations: 2q36 deletion and 15q12-q13.1 duplication, which makes the interpretation of results more difficult.

No recurrent abnormality was found in this cohort. These results underlie the genetic heterogeneity of ASDs.

Conclusions

Our study confirmed the benefits of whole-genome array CGH as an essential diagnostic tool for assessing autistic patients. Our results are in support of the participation of candidate genes localized at 15q11-q13 and 16p11.2 in susceptibility to ASDs. Moreover, our findings underline a challenge in genetic counselling in interpreting array CGH data in patients with inherited CNVs from healthy parents. From a research standpoint, chromosomal microaberrations offer opportunities for the identification of candidate regions for gene discovery. However, we are unable to differentiate for each specific CNV whether it is pathogenic or has no phenotypic effect without

analysis of much larger sample sizes or using other methods of investigation.

Acknowledgements

REFERENCES

1. **American Psychiatric Association** (1994) Diagnostic and Statistical Manual of Mental Disorders, 4th Ed., American Psychiatric Association, Washington, D.C.

2. **Assrir N., Filhol O., Galisson F., Lipinski M.** (2007) Biol. Chem., **388**, 391-398.

3. **Babatz T.D., Kumar R.A., Sudi J., Dobyns W.B., et al.** (2009) Autism Res., **2**, 359-364.

4. **Carnegie G.K., Sleeman J.E., Morrice N., Hastie C.J., et al.** (2003) J. Cell Sci., **116**, 1905-1913.

5. **Chai J.H., Locke D.P., Greally J.M., Knoll J.H., et al.** (2003) Am. J. Hum. Genet., **73**, 898-925.

6. **Christian S.L., Brune C.W., Sudi J., Kumar R.A., et al.** (2008) Biol. Psychiatry, **63**, 1111-1117.

7. **Database of Genomic Variants** <http://projects.tcag.ca/variation/>

8. **Doornbos M., Sikkema-Raddatz B., Ruijvenkamp C.A., Dijkhuizen T., et al.** (2009) Eur. J. Med. Genet., **52**, 108-115.

9. **Fan Y.S., Jayakar P., Zhu H., Barbouth D., et al.** (2007) Hum. Mutat., **28**, 1124-1132.

10. **Finelli P., Natacci F., Bonati M.T., Gottardi G., et al.** (2004) J. Med. Genet., **41**, e90.

11. **Fombonne E.** (2005) J. Clin. Psychiatry, **66**, 3-8.

12. **Groffen A.J., Friedrich R., Brian E.C., Ashery U., et al.** (2006) J. Neurochem., **97**, 818-833.

13. **Gropman A.L., Batshaw M.L.** (2010) J. Dev. Behav. Pediatr., **31**, 582-591.

14. **Guffanti G., Strik Lievers L., Bonati M.T., Marchi M., et al.** (2011) Psychiatry Res., **185**, 33-38.

15. **Guilmatre A., Dubourg C., Mosca A.L., Legallic S., et al.** (2009) Arch. Gen. Psychiatry, **66**, 947-956.

16. **Jacquemont M.-L., Sanlaville D., Redon R., Raoul O., et al.** (2006) J. Med. Genet., **43**, 843-849.

17. **Kelemenova S., Schmidtova E., Ficek A., Celec P., et al.** (2010) Psychiatr. Genet., **20**, 137-139.

18. **Kim S.J., Cox N., Courchesne R., Lord C., et al.** (2002) Am. J. Med. Genet., **114**, 137-143.

19. **Kumar R.A., KaraMohamed S., Sudi J., Conrad D.F., et al.** (2008) Hum. Mol. Genet., **17**, 628-638.

20. **Kumar R.A., Marshall C.R., Badner J.A., Babatz T.D., et al.** (2009) PloS ONE., **4**, e4582.

21. **Lee C., Iafrate A.J., Brothman A.R.** (2007) Nat. Genet., **39**, 48-54.

22. **Locke D.P., Segraves R., Nicholls R.D., Schwartz S., et al.** (2004) J. Med. Genet., **41**, 175-182.

23. **Marshall C.R., Noor A., Vincent J.B., Lionel A.C., et al.** (2008) Am. J. Hum. Genet., **82**, 477-488.

24. **Martins-de-Souza D., Schmitt A., Roder R., Lebar M., et al.** (2010) J. Psychiatr. Res., **44**, 989-991.

25. **Mitchell J.C., Ariff B.B., Yates D.M., Lau K.F., et al.** (2009) Hum. Mol. Genet., **18**, 4492-4500.

26. **Murakami A., Matsuda M., Nakasima A., Hirata M.** (2006) Gene, **382**, 129-139.

27. **Napoli I., Mercaldo V., Boyl P.P., Eleuteri B., et al.** (2008) Cell, **134**, 1042-1054.

28. **Nurmi E.L., Amin T., Olson L.M., Jacobs M.M., et al.** (2003) Mol. Psychiatry, **8**, 624-634.

29. **Okamoto S., Sherman K., Bai G., Lipton S.A.** (2002) Brain Res. Mol. Brain Res., **107**, 89-96.

30. **Otera H., Wang C., Cleland M.M., Setoguchi K. et al.** (2010) JCB, **191**, 1141-1158.

31. Page D.T. (2011) Anat. Rec. (Hoboken), **294**, 1671-1684.

32. **Qiao Y., Riendeau N., Koochek M., Liu X., et al.** (2009) J. Med. Genet., **46**, 680-688.

33. **Rossignol D.A., Frye R.E.** (2011) Mol. Psychiatry, 1-25.

34. **Saus E., Brunet A., Armengol L., Alonso P., et al.** (2010) J. Psychiatr. Res., **44**, 971-978.

35. **Sebat J., Lakshmi B., Malhotra D., Troge J., et al.** (2007) Science., **316**, 445-449.

36. **Shinawi M., Liu P., Kang S.H., Shen J., et al.** (2010) J. Med. Genet., **47**, 332-341

37. **Stefansson H., Rujescu D., Cichon S., Pietiläinen O.P., et al.** (2008) Nature, **455**, 232-236.

38. **Vamos E., Pardutz A., Klivenyi P., Toldi J., et al.** (2009) J. Neurol. Sci., **283**, 21-27.

39. **van der Zwaag B., Staal W.G., Hochstenbach R., Poot M., et al.** (2010) Am. J. Med. Genet. B. Neuropsychiatr. Genet., **153B**, 960-966.

40. **Vorstman J.A., Staal W.G., van Daalen E., van Engeland H., et al.** (2006) Mol. Psychiatry, **11**, 18-28.

41. **Weiss L.A., Shen Y., Korn J.M., Arking D.E., et al.** (2008) N. Engl. J. Med., **358**, 667-675.

42. **World Health Organization** (1993) International Classification of Diseases: 10th revision, Chapter 5, Mental and Behavioural Diseases, Diagnostic Criteria for Research, Geneva.

43. **Yamada K., Miura K., Hara K., Suzuki M. et al.** (2010) A wide spectrum of clinical and brain MRI findings in patients with SLC19A3 mutations. BMC Med Genet., **11**, 171.