УДК 575.224.2:616.8-056.7

ГЕНОМНЫЕ ТЕХНОЛОГИИ ДИАГНОСТИКИ УМСТВЕННОЙ ОТСТАЛОСТИ: РЕЗУЛЬТАТЫ ПРОЕКТА CHERISH

Игорь Николаевич ЛЕБЕДЕВ[1], Людмила Павловна НАЗАРЕНКО[1], Анна Александровна КАШЕВАРОВА[1], Николай Алексеевич СКРЯБИН[1], Ольга Александровна САЛЮКОВА[1], Наталья Николаевна ЧЕЧЁТКИНА[1], Екатерина Николаевна ТОЛМАЧЁВА[1], Елена Александровна САЖЕНОВА[1], Светлана Леонидовна ВОВК[1], Алексей Анатольевич РУДКО[1], Валерий Павлович ПУЗЫРЁВ[1], Памела МАГИНИ[2], Клаудио ГРАЦИАНО[2], Джованни РОМЕО[2]

[1] *ФГБУ НИИ медицинской генетики СО РАМН*
Россия, 634050, Томск, Набережная р. Ушайки, 10

[2] *Alma Mater Studiorum – Universita di Bologna*
Italy, 40126, Bologna, Via Zamboni, 33

Умственная отсталость – чрезвычайно гетерогенное заболевание, как с генетической, так и с клинической точек зрения. В большинстве случаев ее диагностика по-прежнему остается основанной на клиническом обследовании пациента и не включает молекулярно-генетическое тестирование. Получение новых знаний о наследственных причинах умственной отсталости и улучшение ее диагностики за счет применения современных высокопроизводительных геномных технологий явились основными целями исследовательского проекта «Улучшение диагностики умственной отсталости у детей в Центральной Восточной Европе и Центральной Азии через генетическую характеристику, биоинформатику и статистику» (CHERISH), выполненного при поддержке 7 Рамочной программы Европейского союза в 2009–2012 гг. научными коллективами 9 европейских стран, включая Россию. В настоящей статье представлена информация о данном проекте и его основных результатах.

Ключевые слова: умственная отсталость, врожденное слабоумие, матричная сравнительная геномная гибридизация, микроделеционные/микродупликационные синдромы, вариации числа копий ДНК, секвенирование нового поколения.

GENOMIC TECHNOLOGIES FOR DIAGNOSIS OF INTELLECTUAL DISABILITY: RESULTS OF THE «CHERISH» PROJECT

Igor Nikolaevich LEBEDEV[1], Lyudmila Pavlovna NAZARENKO[1], Anna Aleksandrovna KASHEVAROVA[1], Nikolay Alekseevich SKRYABIN[1], Olga Aleksandrovna SALYUKOVA[1], Natalia Nikolaevna CHECHETKINA[1], Ekaterina Nikolaevna TOLMACHEVA[1], Elena Aleksandrovna SAZHENOVA[1], Svetlana Leonidovna VOVK[1], Aleksei Anatol'evich RUDKO[1], Valeriy Pavlovich PUZYREV[1], Pamela MAGINI[2], Claudio GRAZIANO[2], Giovanni ROMEO[2]

[1] *Institute of Medical Genetics of SB RAMS*
Russia, 634050, Tomsk, Ushaika str., 10

[2] *Alma Mater Studiorum – Universita di Bologna*
Italy, 40126, Bologna, Via Zamboni, 33

Intellectual disability is a highly heterogeneous disorder both from genetic and clinical point of view. In most cases of idiopathic mental retardation the diagnosis is still based on clinical examination only and not involved molecular genetic testing. Generation of new knowledge about hereditary causes of intellectual disability and improving its diagnosis

through the application of modern high-throughout whole genome technologies was the main goal of the CHERISH project implemented within the 7th Framework Program of the European Union in 2009–2012 in 9 European countries, including Russia. Information about the project and its key results is given in this article.

Key words: intellectual disability, mental retardation, array-CGH, microdeletion/microduplication syndromes, copy number variation (CNV), next generation sequencing.

INTRODUCTION

Intellectual Disability (ID) is a neuro-developmental disorder characterized by a substantial limitation in cognitive functioning (IQ < 70), disturbances of adaptive and social behavior that manifests before the age of 18 years. It is one of the main disabling conditions in children, and is estimated to affect 2–3 % of the population. The frequency of the most severe forms (IQ less than 30) consists of 15 % among all cases of ID [32].

ID is a highly heterogeneous disorder, with the majority of cases (25–45 %) recognizing a genetic etiology. Chromosomal aberrations, such as aneuploidy, deletions, duplications and translocations, cause ID in approximately 30–35 % of all patients, in particular in syndromic cases, which accompany other clinical features. Tens of new microdeletion/microduplication syndromes associated with ID were described during last years due to application of high-resolution arrays technologies in cytogenetic studies [33]. However, the most cases of ID in patients with normal karyotype remain undiagnosed.

As regards to single gene bases of ID, it is known about 140 forms of syndromic X-linked ID, with Martin-Bell syndrome is the most frequent, and 95 variants of non-syndromic ID. Primary gene defect is found for approximately 66 syndromic forms, whereas only specific X-chromosome regions are mapped for 50 forms [30]. Gene mutations are known for less than 50 % of non-syndromic X-linked ID. It is reported about mutations in 20 genes, which are associated with non-syndromic forms. They are genes of 5 protein's groups – transmembrane receptors, small GTP-regulators and effectors, enzymes and translation regulators [3]. A total of 80 genes were associated with X-linked ID [30].

Less definite situation is still remaining for autosomal forms of ID. It was no reports about gene mutations responsible for autosomal dominant ID. As concerns to autosomal recessive ID, which accounts for 25 % of mental retardation cases, 14 regions of linkage only were known several years ago before the introduction of the high-throughput technologies of next generation sequencing. Previously it was reported about mutations in 6 genes, which products belong to families of serine proteases (*PRSS12* – neurotrypsin) [24]. ATP-dependent proteases (*CRBN* – cereblon) [11], calcium-dependent transcription repressors (*CC2D1A*) [4], transmembrane glutamate

Лебедев И.Н. *– д.б.н., зав лабораторией цитогенетики, e-mail: igor.lebedev@medgenetics.ru*
Назаренко Л.П. *– д.м.н., проф., зам. директора по научной работе, e-mail: ludmila.nazarenko@medgenetics.ru*
Кашеварова А.А. *– к.б.н., научный сотрудник лаборатории цитогенетики, e-mail: anna.kashevarova@medgenetics.ru*
Скрябин Н.А. *– к.м.н., научный сотрудник лаборатории цитогенетики, e-mail: skryabinn@mail.ru*
Салюкова О.А. *– к.м.н., научный сотрудник лаборатории наследственной патологии, e-mail: olga.salukova@mail.ru*
Чечеткина Н.Н. *– аспирант, e-mail: valete@mail.ru*
Толмачева Е.Н. *– к.б.н., научный сотрудник лаборатории цитогенетики, e-mail: kate.tolmacheva@medgenetics.ru*
Саженова Е.А. *– к.б.н., научный сотрудник лаборатории цитогенетики, e-mail: elena.sazhenova@medgenetics.ru*
Вовк С.Л. *– зав. клинико-диагностической лабораторией генетической клиники, e-mail: svetlana.vovk@medgenetics.ru*
Рудко А.А. *– к.м.н., главный врач генетической клиники, e-mail: aleksey.rudko@medgenetics.ru*
Пузырев В.П. *– д.м.н., проф., академик РАМН, директор, e-mail: valery.puzyrev@medgenetics.ru*
Magini Pamela *– PhD, Laboratory of medical genetics, department of gynecology, obstetrics and pediatrics sciences, e-mail: pamela.magini@unibo.it*
Graziano Claudio *– MD, Laboratory of medical genetics, Department of gynecology, obstetrics and pediatrics sciences, e-mail: claudio.graziano@unibo.it*
Romeo Giovanni *– professor of medical genetics, medical school, e-mail: romeo@eurogene.org*

receptors (*GRIK2*) [25], NF-κB signaling proteins (*TRAPPC9*) [23, 28], and enzymes of long chain fatty acids conversion (*TECR*) [5].

Introduction of next generation sequencing rapidly increases the number of genes responsible for ID. At present this list contains approximately 50 genes [6, 26]. Mutations in more than 450 genes have been associated with ID and related cognitive disorders, such as autism [32]. However, majority of these mutations are reported in single patients or families. This means, that molecular epidemiology of new genetic variants associated with ID in different populations remains undefined. As before, the identification of primary gene defect responsible for ID in given patient is a hard task due to significant clinical and genetic heterogeneity as well as usual lack of national specific registers. In most cases the diagnosis of ID is still based on clinical examination only and not involved molecular genetic testing. Recent advances in gene identification techniques have revolutionized the clinical approach to ID. Thus, the search of new hereditary mutations involved in ID pathogenesis and development of standard for molecular genetic diagnosis of ID in childhood is a challenge for current basic and clinical research.

THE CHERISH PROJECT: MAIN TASKS

This aim was pursued during implementation of the research project within 7th Framework Program of the European Union «Improving diagnosis of mental retardation in children in Eastern Europe and Central Asia through genetic characterization and bioinformatics/statistics (CHERISH)» (grant agreement № 223692). The overall goal of the CHERISH project was to establish an interdisciplinary Eastern Europe and Central Asia consortium of experts to perform a research program of clinical, scientific and public activities for generation of new knowledge about genetic causes of ID. The main objectives of the project were: 1) the development of a standardized approach for clinical diagnosis of ID; 2) the creation of large database and bio-bank of DNA samples from patients with clinically well defined ID of unknown genetic etiology; 3) the identification of cryptic chromosomal rearrangements by molecular cytogenetic analysis; 4) the identification of new mutations and genes responsible for ID by linkage analysis in familial cases, sequencing of known genes and whole exome sequencing of the patients with ID; 5) the dissemination of knowledge about the project and its results in scientific publications, meetings and workshops for researchers, clinicians and society.

Ten Universities and research Institutes from 9 European and former USSR countries were involved in the investigations:

1) Alma Mater Studiourum – Universita di Bologna, Italy (project coordinator);

2) University of Tartu (Institute of Molecular and Cell Biology), Estonia;

3) Vilnius University (Department of Human and Molecular Genetics), Lithuania;

4) Charles University of Prague (Department of Biology and Medical Genetics), Czech Republic;

5) Poznan University of Medical Sciences (Department of Medical Genetics), Poland;

6) Institute of Molecular Biology and Genetics of the National Academy of Sciences of Ukraine, Kiev, Ukraine;

7) The Cyprus Institute of Neurology and Genetics (Department of Cytogenetics), Nicosia, Cyprus;

8) Institute of Medical Genetics, Siberian Branch of Russian Academy of Medical Sciences, Tomsk, Russia;

9) Center of Medical Genetics and Primary Health Care, Erevan, Armenia;

10) European Genetic Foundation, Bologna, Italy.

PATIENTS SELECTION FOR THE STUDY: CLINICAL AND GENETIC CRITERIA

During the first Work Package of the CHERISH project an interdisciplinary Eastern European and Central Asia consortium of experts was established in order to lay down a basis for a significant improvement of clinical, educational and laboratory diagnostic developments in the field of genetics of ID. A standardized approach for the diagnosis of ID with common criteria for patient selection was defined, and allowed the creation of a large sample's collection of patients with clinically well defined ID. The consortium identified a web-based «Cartagenia» database tool [16] for the management of all the information regarding the phenotypic and genetic characteristics of the included patients, in order to obtain a secure system to share data to which the partners only could have access; this tool also allowed sample anonymization to be a routine process. Partners collected a blood sample from every patient enrolled in the project. For sporadic cases, samples from parents were collected whenever possible, while for familial cases, blood samples collection included patients and other family members who gave their informed consent.

One thousand and four hundred fifty seven patients with idiopathic ID (including 206 children recruited by Genetic Clinics of the Institute of Medical Genetics, SB RAMS, Tomsk) were sampled and analyzed during the project, aiding the detection of new genetic causes of neuro-developmental disorders. Thus, the CHERISH project was the largest investi-

Table

Summary of results of genetic and biochemical tests in patients with ID

Tests	Number of analyses	Number of patients with abnormal results
Karyotyping (G-banding)	1457	51 (4 %)
FISH for known microdeletion/microduplication syndromes	153	1 (0.7%)
FISH for subtelomeric rearrangements	67	5 (7.5 %)
MLPA for subtelomeric rearrangements	164	12 (7.3 %)
FRAXA molecular genetic testing	647	2 (0.3 %)
Other molecular genetic tests including Rett, Prader-Willi and Angelman syndromes testing	373	11 (2.9 %)
Amino acids (tandem mass-spectrometry)	442	7 (1.6 %)
Urinary organic acids	243	3 (1.2 %)
Lysomal enzymes	25	0
Lactic acid	152	6 (3.9 %)
Creatine defect (creatine/creatinine index from urine)	32	0
Other metabolic defects	156	11 (7 %)

gation of ID in European countries and Russia. Previously in Western Europe (Netherlands, Belgium, Germany, and France) the EURO-MRX project was performed focusing on genetics of X-linked ID [17]. More than 600 families were examined and 17 new X-linked genes were discovered.

In order to be enrolled, preliminary testing of all patients included standard karyotyping and, where indicated, Fragile X, Rett, Prader-Willi, Angelman syndrome molecular testing, FISH for known microdeletion/microduplication syndromes, FISH and MLPA analysis for subtelomeric rearrangements, neuro-metabolic screening. Summary results of these tests are presented in the table. As a rule, the data obtained for the incidence of the most frequent genetic and biochemical defects associated with ID were in agreement with literature data. Patients with normal results were enrolled in the main research part of the project.

WHOLE GENOME ARRAY-BASED COMPARATIVE GENOMIC HYBRIDIZATION IN PATIENTS WITH INTELLECTUAL DISABILITY AND NORMAL KARYOTYPE

This part contained three Work Packages – array-based comparative genomic hybridization (array-CGH) testing, analysis of families with ID, and whole exome sequencing. Analyses of chromosomal microdeletions/microduplications and copy number variations (CNV) were performed by array-CGH in University of Bologna, The Cyprus Institute of Neurology and Genetics, and Institute of Medical Genetics (Tomsk, Russia). This study was performed on 44K, 105K, and 400K whole genome arrays (Agilent Technologies, USA). A total of 378 patients were investigated and 563 CNVs were detected. The frequency of pathogenic CNVs varied from 10 to 20 % in different observed populations that is common for patients with ID.

Eighty Russian patients were selected for array-CGH studies in Institute of Medical Genetics (Tomsk, Russia). Eleven of these patients were 3–5 years old and showed developmental delay; the other 69 patients were older than 5 years of age and had ID defined as an IQ lower than 70. Balanced karyotype was confirmed after array-CGH study for 35 children. Twenty-two patients carried only benign CNVs, which previously were observed in healthy individuals according to information of Database of Genomic Variants (DGV) [15]. A total of 27 pathogenic or likely pathogenic CNVs were detected in the other 23 affected children. Among these CNVs, microdeletions associated with known syndromes were identified in seven patients, allowing for a definite diagnosis in 9 % of the patients in the Russian group. These diagnoses were Cri du Chat syndrome, 15q24 microdeletion syndrome (two cases), 16p11.2 microdeletion syndrome (two cases), 16p12.2 microdeletion syndrome, and 22q11.2 microdeletion syndrome. Information about clinical and genetic findings in this group of patients can be obtained from our previously published study [21]. Nineteen CNVs that had not previously been associated with any known syndrome were detected in 16 patients (20 %). Below, we will focus on some interesting findings.

Case #1. A 6 years old boy with mild mental retardation (IQ = 57), hyperdynamia, attention deficit hyperactivity disorder (ADHD) and speech delay. Array-CGH revealed a 1.155 Mb microdeletion at 11p13 (Fig. 1a). Although the deleted region seems to correspond to the location indicated in Wilms

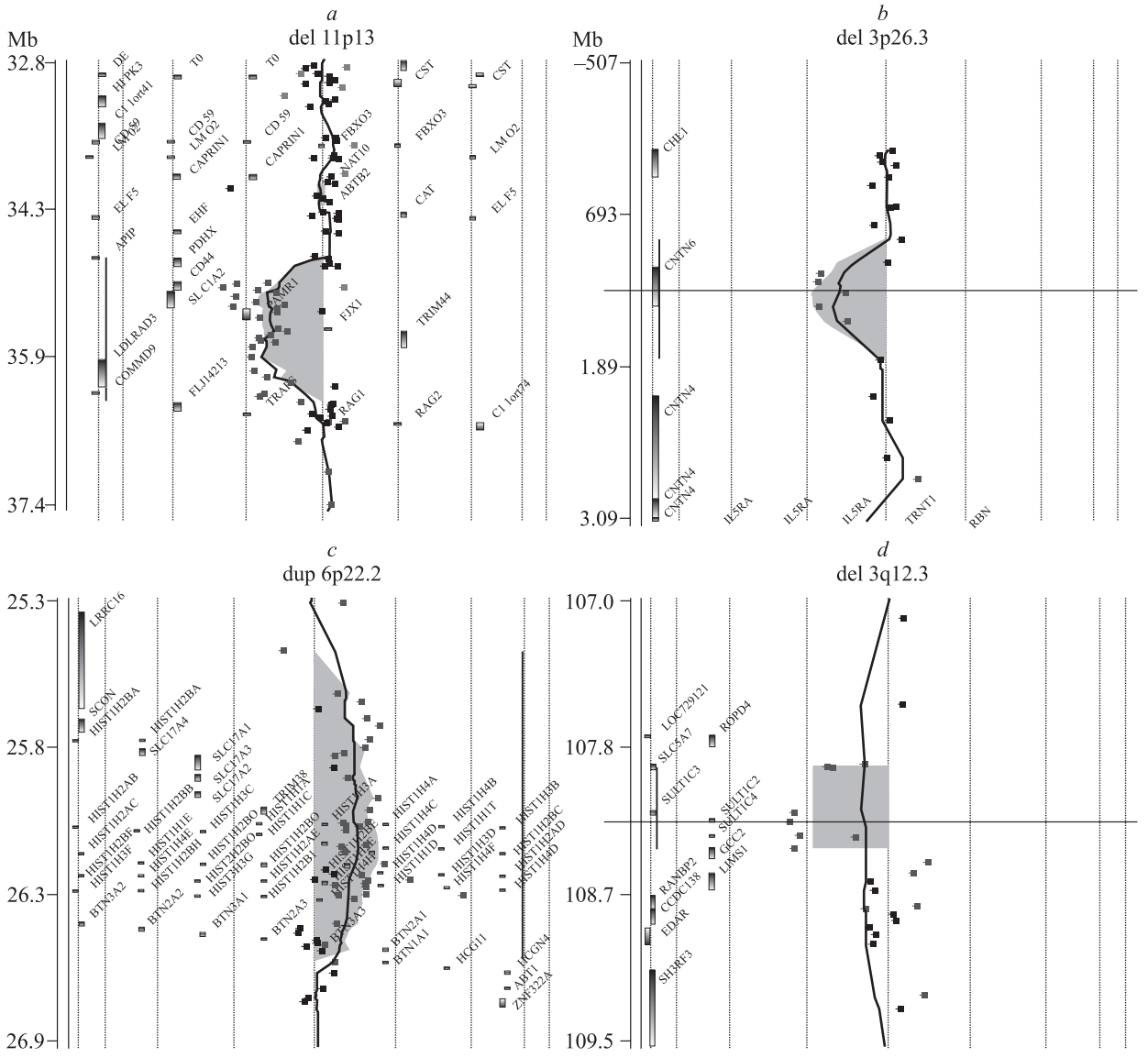

Fig. 1. *Chromosomal microdeletions and microduplications detected by array-CGH in Russian patients with ID*

tumor, aniridia, genitourinary anomalies and mental retardation (WAGR) 11p13 deletion syndrome (MIM 194072) [14], the precise location of the microdeletion in our patient does not overlap with the classical deletion given in DECIPHER database [13] (11:35180424-36335424 vs. 11:31806339-32457087). Thus, it is not surprising that the boy does not have the typical WAGR symptoms. However, the proband is still mentally retarded, and there are several genes within the region of the microdeletion that may be associated with ID. For example, the SLC1A2 protein is responsible for glutamate transport. The accumulation of extracellular glutamate causes a dysfunction in calcium homeostasis, increases production of NO, free radicals, and leads to activation of proteases. These factors can result in neuronal damage, leading to neurodegenerative disease, inflammation, or ischemia development [22]. The function of FJXI, the product of another gene in this region, remains unknown in humans; however, this gene product has been shown to regulate dendrite extension in rodents [2]. The product of a third gene, TRIM44, may play a role in neuronal differentiation and maturation [18]. The product of a fourth gene in this region, LDLRAD3, participates in the proteolysis of an amyloid precursor protein, leading to beta amyloid formation; this fiber form is the primary component of amyloid plaques found in the brain of Alzheimer's disease patients [29]. The 11p13 microdeletion observed in our patient was confirmed by real-time PCR.

Case #2. A 8 years old girl with mild mental retardation (IQ = 65), scrambled speech, ADHD and some clinical features: low-set ears, short philtrum,

wide-spaced nipples, shoulder pigmentation, and overweight. Array-CGH revealed a 112 kb microdeletion at 3q22.1. Among three genes located within affected region, *DNAJC13* and *ACAD11* are especially interesting. Variation in DNAJC13 is involved in Tourette syndrome (MIM 137580) [14]. ACAD11 belongs to a family of enzymes that participate in mitochondrial fatty acid beta-oxidation. High levels of expression of *ACAD11* have been observed in adult brains, and this protein participates in generating energy as well as likely having a role in the degradation of unique cellular lipids and catabolism of aromatic amino acids, vital compounds for neuron function [10].

Case #3. A 9 years old boy with moderate mental retardation (IQ = 47), delay of motor milestones and fine motor skills deficiency, dysarthria, ADHD, some dysmorphic features: dolichocephalic skull, mongoloid slant, divergent left strabism, protrunding ears, hypoplastic antihelix and antiragus. Array-CGH identified a 766 kb microduplication at 3p26.3. The only gene located within this region is *CNTN6*, which encodes contactin 6. Significantly, contactin 6 is a neuronal membrane protein that functions as a cell adhesion molecule and may take part in the formation of axon connections in the developing nervous system. *CNTN6* and other members of this family have been identified as potentially pathogenic genes in neuro-developmental disorders. These genes have been suggested to participate in pathways important for correct brain development [34]. In contrast to case #3, a 369 kb microdeletion at 3p26.3 overlapping with *CNTN6* was identified in another patient (case #4, Fig. 1b). He was a 10 years old boy with mild mental retardation (IQ = 55). He had tower skull, frontal bossing, antimongoloid slant, epicanthus, wide nasal bridge, low-set ears, hypertrichosis, hair nevus, clinodactyly, dysarthria, ADHD. Magnetic resonance imaging (MRI) revealed lateral ventricles asymmetry as a result of outflow from left ventricles, and neurodystrophic foci due to hypoxia. It is interesting, that although the chromosomal aberrations in these cases were the reverse of one another, both patients had intellectual disability, dysarthria, and ADHD. The presence of microduplication and microdeletion of 3p26.3 in our patients was confirmed by real-time PCR.

Case #5. A 6 years old boy with mild mental retardation (IQ = 50), hydrocephaly, large protruding ears, wide-spaced nipples, abnormal hair growth, and hypertrichosis. A patient pronounces single syllables only, has an autistic signs and food fastidiousness. Array-CGH revealed a 716 kb microduplication at 6p22.2 (Fig. 1c). One of the genes within this region that has gained attention in neuroscience is *HFE*, which regulates ion homeostasis. Mutations in this gene are commonly associated with hereditary hemochromatosis – an autosomal recessive disorder of ion metabolism (MIM 235200) [14]. It has been recently shown that the H67D polymorphism in *Hfe* in mice impacts brain ion homeostasis, creates an environment favorable for oxidative stress, and predisposes mice to neurodegenerative disorders [27]. The other gene of interest in this region is *SCGN*, which encodes secretagonin. *SCGN* is expressed in neurons in the embryonic nervous system and is confined to differentiated cells in the adult brain. Secretagonin may be implicated in the control of neuronal turnover and differentiation. This protein is also re-expressed in brain tumors. In addition, secretagonin may functions as a Ca2+ sensor regulating the exocytosis of neurotransmitters, neuropeptides, and hormones [1]. Microduplication of 6p22.2 in our patient was confirmed by real-time PCR.

Case #6. A 8 years old boy with mild mental retardation (IQ = 50), microcephaly, small ears, gothic palate, irregular teeth, abdominal obesity, hypogonadism, speech delay, undue fatigability, aggressiveness, auto-aggressiveness, and mood swings. Array-CGH identified a 1.95 Mb microduplication at 14q11.2. There are two genes attracting particular attention in this region: *SLC7A7* and MMP14. The SLC7A7 protein transfers cationic and large neutral amino acids from the cell to the extracellular space. Mutations in the *SLC7A7* gene are found in patients with lysinuric protein intolerance (MIM 222700) [14], a multi-organ disease with a variety of clinical symptoms, including moderate intellectual disability [9]. Increased *MMP14* expression has been observed in different cancers, including brain tumors [18].

Case #7. A 12 years old boy with moderate mental retardation (IQ = 45), triangular face, protruding ears, delay in motor milestones, absent of speech (certain sounds only), seizures, aggressiveness, auto-aggressiveness, autistic signs, and ADHD. Array-CGH analysis revealed a 464 kb microdeletion at 2q12.3 (Fig. 1d). One gene in this region, *SLC5A7*, is an Na(+)- and Cl(-)-dependent high-affinity transporter that mediates the uptake of choline for acetylcholine synthesis. Acetylcholine is a neurotransmitter in the central and peripheral nervous system that regulates a variety of autonomic, cognitive, and motor functions. This gene was shown to be associated with ADHD [8].

After microarray studies, a cohort of 584 patients with normal array-CGH results was selected to undergo high-resolution CNVs analysis through Single Nucleotide Polymorphism (SNP)-array. These studies were performed at the University of Bologna in Italy and at the University of Tartu in Estonia using «Genome-wide human SNP Array 6.0» (Affymetrix), which contains 950000 SNP, «HumanCNV370

BeadChip» (Illumina) with 5 kb resolution over the whole genome, and «HumanCytoSNP-12 BeadChip» (Illumina) with 300 kb resolution. Selection criteria were substantially the same as for the array-CGH analysis. SNP-arrays have the advantage of a higher resolution with respect to array-CGH, and allows the identification of stretches of homozygosity, that can suggest a region with uniparental disomy or of true homozygosity in the case of parental consanguinity, as well as the regions of loss of heterozigosity. On the other hand, due to its resolution, SNP-array analysis makes interpretation of results more complex because a high number of CNVs are identified, most of them being of benign origin. In order to exclude such neutral variations, identified CNVs were compared with those recurrently present in the Database of Genomic Variants [15] and in the databases of national general populations. The potential clinical significance of CNVs was evaluated using OMIM and DECIPHER databases and peer-reviewed literature searches in the PubMed database. As a result 228 new probably pathogenic CNVs were identified. The presence of genomic aberrations with potential clinical relevance was confirmed by quantitative PCR.

IDENTIFICATION OF NEW GENES RESPONSIBLE FOR THE DEVELOPMENT OF INTELLECTUAL DISABILITIES

One of the main objectives of the project was the identification of new genes involved in the pathogenesis of ID. The classical approach to the identification of genes is based on recognition of families with multiple affected family members, linkage analysis through the study of short tandem repeat polymorphisms, and sequencing analysis of candidate genes inside the regions showing positive linkage scores. Only a few families with sufficient size were collected by the CHERISH consortium.

Consanguineous marriages are infrequent in Italy and eastern European countries, but we identified a very interesting Italian family where 2 male siblings, born from first cousin parents (Fig. 2a), showed a peculiar and similar phenotype: early development was normal, but loss of ambulatory skills due to lower limb spasticity led them to be wheel-chair bound in the first decade of life. Motor problems were followed by progressive cognitive decline and both brothers showed severe mental delay after a few years. Brain MRI showed progressive leukoencephalopathy with cerebellar, brainstem and cervical cord atrophy, and mild hypointensity of the globi pallidi on T2 weight images. Electromyography / nerve conduction studies and muscle biopsy were normal. Neurometabolic investigations were negative, except for the presence of trace amounts of urinary sulfatides. They have a normal sister and brother.

On a first basis, X-linked Pelizaeus Merzbacher Disease (MIM 312080) [14] was excluded through mutation analysis of the *PLP1* gene. The asset of this family seemed excellent in order to perform linkage analysis. Blood was collected from both the affected and healthy sibs and from their parents. DNA was extracted from the blood samples and autozygosity mapping was performed by means of a 6090 SNPs Illumina Infinium-12 genome wide scan. Three regions, which were heterozygous in the parents, were inherited in homozygosity by the affected brothers but not by the healthy offspring: on chromosome 1 (7.4 Mb between SNPs rs1619856 and rs1342872), on chromosome 3 (14.1 Mb between SNPs rs1017967 and rs1039524) and on chromosome 16 (18.4 Mb between SNPs rs30882 and rs1037973). Evaluating the genes contained in the longest autozygous region on chromosome 16q22-23, the most striking candidate was Fatty Acid 2-Hydroxylase (*FA2H*), since homozygous mutations in *FA2H* were recently shown to be associated with leukodystrophy with spastic para-

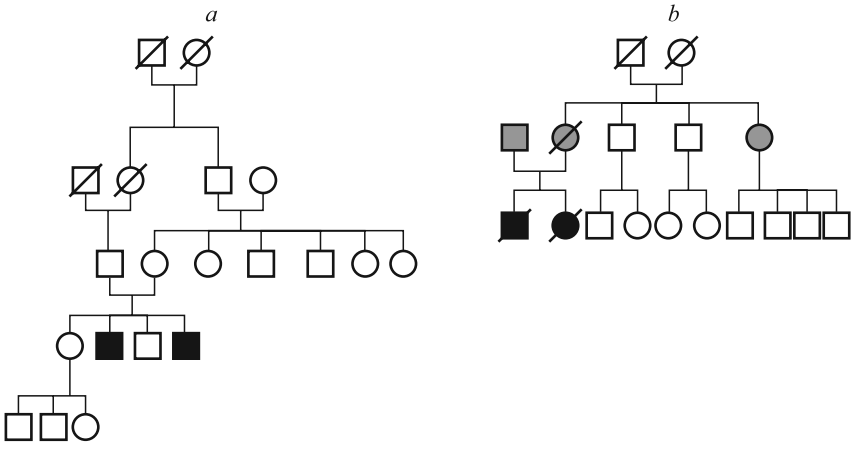

Fig. 2. *Pedigrees of Italian families with mutation in FA2H (a) and deletion of 7q31.32 (b)*

paresis and dystonia, complicated by neurodegeneration with brain iron accumulation [7].

Sequencing of *FA2H* coding region led to the identification of a homozygous c.270+3A>T mutation in the affected individuals. Parents were heterozygous while the healthy brother and sister, as well as 150 ethnically matched control individuals, showed the reference wild type genotype. We then retrotranscribed and amplified the *FA2H* specific cDNA with PCR primers on exons 1 and 2. Compared to the wild type control, a longer cDNA fragment was amplified in one of the homozygous patients and in his heterozygous mother, resulting in a band of higher molecular weight in the agarose gel run. Sequencing of the cDNA demonstrated the insertion of a 97 base pairs intronic sequence with consequent frame shift and creation of a premature stop codon after 42 amino acids in exon 2. Overall, these findings indicate that we have identified the causative *FA2H* mutation in this family.

In another family a 250 kb deletion of 7q31.32 involving one gene, *CADPS2*, was identified by array-CGH (44K, Agilent Technologies) in patient with normal karyotype, affected by mild mental ID and generalized epilepsy with normal brain imaging. The same deletion was identified in his affected sister, who shows highly similar clinical features, but was absent in their father and aunt (Fig. 2b). It is thus likely to be maternally inherited: the mother died with breast cancer but was reported to show behavioural abnormalities. The deletion was characterized in detail by quantitative real-time PCR and maps between exon 4 and exon 28 of this large gene, resulting in a complete disruption of the gene product. Sequencing of *CADPS2* was performed in the proband, in the hypothesis of a recessive disorder, but no point mutations were present.

This gene encodes a member of the calcium-dependent activator of secretion (CAPS) protein family, which are calcium binding proteins that regulate the exocytosis of synaptic and dense-core vesicles in neurons and neuroendocrine cells. Several lines of evidence suggest that *CADPS2* is good candidate gene for ID and autism spectrum disorders (ASD), given that: 1) it is involved in the release of neurotrophins such as BDNF/NT3, as shown from studies on animal models (mice) where *Cadps2* was "knocked out"; these mice are also reported to show autistic traits [31]; 2) alternative spliced forms lacking important domains have been reported to be more frequent in individuals with autism or with lower IQ compared to controls, although data from different groups are controversial; 3) *CADPS2* maps to the AUTS1 locus, one of the few linkage intervals for autism susceptibility identified from many independent groups. Therefore, a mutation screening of *CADPS2* was performed in 120 ID/ASD sporadic patients, in order to identify other putative deleterious variants. All exon and exon-intron boundaries were analyzed by direct sequencing of the PCR products. Novel variants M630T and F645V in exon 13 and V1137M in exon 26 were identified. All variants were not found in 500 control chromosomes (250 Italian healthy individuals). But it is important that M630T and F645V mutations were identified once in the group of 46 patients with ASD, whereas V1137M was detected in one from 36 individuals from the additional cohort of Italian patients with ID.

The segregation of all variants was evaluated in the corresponding pedigrees and an excess of maternal transmission was identified. In particular, the change in exon 26 was inherited from the mother and was not found in an unaffected brother. Conversely, the mutation in exon 13 was of paternal origin. These data suggested possible parent-of-origin and/or imprinting effects of these coding changes. Considering that *CADPS2* maps to chromosome 7q31, where a cluster of imprinted genes has already been reported, the expression analysis of the different parental alleles of *CADPS2* was performed in controls and affected individuals of whom blood RNA was available, and in some samples from archive of different brain tissues regions. Monoallelic expression from maternal chromosome in the brain, but biallelic in lymphocytes was found.

Technological development that took place after the beginning of the project in 2009 gave the opportunity to analyze almost all the known exons of an individual in one single experiment at a reasonable cost. This technology, termed Whole-Exome Sequencing (WES), allowed the identification of disease-causing genes of various genetic disorders in the last years, but has also opened the possibility to speed up the molecular diagnosis of heterogeneous genetic disorder such as ID. The main issue with WES is the analysis of the hundreds of genetic variants identified in every single patient. This technology revolutionized the way to disease-causing gene identification, rapidly making the classical approach obsolete. The knowledge of linkage regions in a specific family (or of homozygosity stretches in consanguineous families) can nevertheless help the process of gene prioritization, given the extremely high number of variants identified with WES.

Affected individuals from families with ID were selected for WES in the CHERISH project. Only families with multiple ID affected members were selected: 3 families with an X-linked pattern of inheritance, 1 family with an autosomal dominant pattern, 6 families with autosomal recessive pattern and 2 families with an unclear pattern (2 affected brothers, compatible with either autosomal recessive and

X-linked inheritance). A total of 21 affected patients from 12 families were sequenced. The results obtained indicate, that there are about 18000–19000 single-nucleotide variants (SNV) and 300–500 insertions/deletions per each exome. The most of SNVs (>98 %) and insertions/deletions (>70 %) are presented in the database of normal variants – «dbSNP» [19]. Whereas a new, not previously reported SNVs and insertions/deletions affect at the average 260–270 and 105–110 genes in each patients, respectively. Analysis and validation of WES results are under implementation.

POTENTIAL IMPACT OF THE CHERISH PROJECT AND MAIN DISSEMINATION ACTIVITIES

Dissemination of the obtained results was one of the major goal of the CHERISH project. The project had an obvious potential impact for the scientific community, as well as for the participating families and for society. First of all the participation in the consortium provided the possibility of a professional growth for young scientists, PhD students and clinicians, who had the opportunity to travel to the laboratories of the involved partners to gain specific knowledge and expertise. The participating countries have strongly profited from the project in several ways. The first period of the project, that involved training and increasing the awareness of ID, allowed the participation of young scientists and clinicians in courses that addressed different aspects of ID diagnosis and management. This knowledge was then further disseminated in the local scientific and clinical communities, and among the general public. Each team could profit from newly established scientific and clinical contacts, which were realized through mutual visits among partner laboratories. The project also supported the participation of the partner teams in scientific conferences where they could present their findings, learn about the progress in the field, further discuss the methods and interpretation of the results, and find new collaborations. Young researchers and clinicians were preferentially supported to attend scientific meetings.

During the stage of sample collection and clinical characterization, many local professional collaborators were involved in each country, further increasing the outreach of the project and the public awareness of ID. This preliminary part of the project brought a clear benefit to some families who were identified to carry visible cytogenetic aberrations. These families were excluded from further studies, but nevertheless received a diagnosis explaining their child phenotype. The CHERISH project helped to introduce array-based analysis as the first-tier diagnostic test for patients with ID (and multiple congenital anomalies) in participating countries, as suggested by internationally recognized authoritative sources in the field of cytogenetics, e.g. ISCA Consortium [20].

Between 10 and 20 % of patients in the different national cohorts received a specific diagnosis thanks to array analysis. Every single patient who receives such a diagnosis brings a potential benefit for his/her family, at least in terms of genetic counselling and evaluation of recurrence risks for family members. In other families new variants potentially causative for ID were identified: the immediate clinical relevance is limited in these families, but the findings are of scientific interest as they can point to novel syndromes or novel candidate genes for ID.

If array analysis was the best tool for diagnosis of idiopathic ID at the beginning of this project, research and diagnosis in medical genetics underwent a real revolution in most recent years with the advent of new technologies for DNA sequencing, allowing analysis of all the known genes of an individual in one single experiment (Whole Exome Sequencing). This approach has been the object of a specific amendment submitted during the last year of implementation of the project. We thus had the opportunity to introduce this new technology in Eastern European countries, contributing to its widespread use. The use of such advanced molecular diagnostics tools in the definition of genetic disorders will eventually open a new phase in the prevention of the manifestations of many monogenic as well as complex genetic disorders.

In order to raise awareness on ID in general and on the project's results specifically among the scientific community and the general public, the Consortium has developed a dissemination strategy based on the use of different channels: scientific publications, a Scientific Symposium, a CHERISH stand during the annual ESHG meeting, 3 newsletters, an international workshop, 3 targeted events for researchers, a «Meet the Experts» Symposium, the realization of a multilingual brochure for families and patients' associations, and Press Conference.

The last Newsletter, containing information concerning the Consortium dissemination activities, scientific publications, portal updates and project meetings, was sent on July 2012 to 44.000 e-mail contacts. All CHERISH Newsletters are freely available on the project website [12]. Educational brochures and informative leaflets were prepared in English, translated in the languages of participating countries and provided to a wide range of audiences including families, parent groups, health care professionals and the general public.

Finally, all the partners were involved in dissemination of the project objectives and results at Inter-

national meetings, such as the yearly conferences of the European Society of Human Genetics (ESGH), but also in the local organization of meetings with families and patients' support groups. Along this line, towards the end of the project a «Meet the Experts» Symposium addressed to patients, patients' families and associations was organized in Bologna. The CHERISH project was presented and patients and their representatives had a chance to discuss their personal experiences. During the ESHG Annual Conference in Amsterdam (the Netherlands) on May 28-31, 2011, the Cherish project had a stand in the international exhibition area, where copies of dissemination materials (handouts, leaflets, posters) were distributed to more than 500 participants. During the meeting the EGF staff also recorded some interviews with Consortium members. All the interviews are available on the project portal [12].

A series of one full day scientific sessions dedicated to the CHERISH Project were organized in Bologna within the frame of 3 residential courses (Basic and Advanced Course in Genetic Counseling in Practice, Course on Molecular and Statistical Genetics of Consanguinity, and the 25th Course in Medical Genetics) planned by the European Genetics Foundation. The Consortium had the opportunity to present the results of the project to a wide international audience and to underline the strength of the project and the importance of creating international networks for sharing genetic practices in a cross-cultural setting.

Dissemination activities reached their culmination on May 24th 2012, when University of Bologna (project coordinator) organized in collaboration with the City Hall of Bologna and the European Genetics Foundation, a Public Awareness of Genetics event with the presence of Nobel Laureate Professor Mario Capecchi. The event was introduced by the Rector of the University of Bologna, Prof. Ivano Dionigi and by the Vice-Mayor of Bologna Prof. Silvia Giannini. During this unique occasion, the CHERISH partners briefly presented their work and introduced the new powerful diagnostic tools that medical genetics offers for the diagnosis of ID, just before Professor Capecchi's lecture for the general public. CHERISH leaflets were distributed to 200 participants. The event has also been web-casted and a recorded version of the partners' presentation is available on the project website [12].

In conclusion, it should be noted, that participation in the CHERISH project has also a valuable significance for Institute of Medical Genetics. Two hundred and six Russian patients with ID were clinically examined according to international standard, and 80 of them received results of array-CGH studies. Known microdeletion/microduplication syndromes associated with ID were diagnosed in 7 patients. For 16 patients new, previously not reported chromosomal microdeletions/microduplications were detected and candidate genes responsible for ID within affected chromosomal regions were delineated.

During the project implementation training courses for young scientists and clinicians were organized. As a result, new research and diagnostic technologies became available. For example, diagnostic of chromosomal aberrations by array-CGH is now performed at the Genetic Clinic of the Institute of Medical Genetics for patients with intellectual disability and developmental delay. This technology can be applied also for prenatal and preimplantation diagnosis of chromosomal diseases.

Participation in the CHERISH project allowed establish and improve scientific cooperation between partners, providing possibilities to continue research in the field of genetics of mental retardation. At present, these studies are supported partially by the grant of Federal Program of the Ministry of Education and Science of Russian Federation «Scientific and Educational personnel of innovative Russia» (№ 8727). Moreover, other studies within this Program in the areas of human reproduction (grants № 8276 and 8720), common diseases (grant № 8062) and cancer genetics (grants № 8595, 8602, and 8719) based on array-CGH technology have been performed in Institute of Medical Genetics with cooperation of Institute of Complex Issues of Cardiovascular Diseases of SB RAMS (Kemerovo) and Institute of Oncology of SB RAMS (Tomsk).

REFERENCES

1. *Alpár A., Attems J., Mulder J. et al*. The renaissance of Ca2+-binding proteins in the nervous system: secretagogin takes center stage // Cell Signal. 2012. 24. 378–387.

2. *Ashery-Padan R., Alvarez-Bolado G., Klamt B. et al*. *Fjx1*, the murine homologue of the Drosophila four-jointed gene, codes for a putative secreted protein expressed in restricted domains of the developing and adult brain // Mech. Dev. 1999. 80. 213–217.

3. *Bahl-Buisson N., Chelly J., des Portes V*. Update on the genetics of X-linked mental retardation // Rev. Neurol. (Paris). 2006. 162. 952–963.

4. *Basel-Vanagaite L., Attia R., Yahav M. et al.* The *CC2D1A*, a member of a new gene family with C2 domains, is involved in autosomal recessive non-syndromic mental retardation // J. Med. Genet. 2006. 43. 203–210.

5. *Caliskan M., Chong J.X., Uricchio L. et al.* Exome sequencing reveals a novel mutation for autosomal recessive non-syndromic mental retardation in the *TECR* gene on chromosome 19p13 // Hum. Mol. Genet. 2011. 20. 1285–1289.

6. *de Ligt J., Willemsen M.H., van Bon B.W.M. et al*. Diagnostic exome sequencing in persons with severe intellectual disability // N. Engl. J. Med. 2012. 367. 1921–1929.

7. *Edvardson S., Hama H., Shaag A. et al*. Mutations in the fatty acid 2-hydroxylase gene are associated with leukodystrophy with spastic paraparesis and dystonia // Am. J. Hum. Genet. 2008. 83. 643–648.

8. *English B.A., Hahn M.K., Gizer I.R. et al.* Choline transporter gene variation is associated with attention-deficit hyperactivity disorder // J. Neurodev. Disord. 2009. 1. 252–263.

9. *Font-Llitjos M., Rodriguez-Santiago B., Espino M. et al*. Novel *SLC7A7* large rearrangements in lysinuric protein intolerance patients involving the same AluY repeat // Eur. J. Hum. Genet. 2009. 17. 71–79.

10. *He M., Zhengtong P., Mohsen A. et al.* Identification and characterization of new long chain Acyl-CoA dehydrogenases // Mol. Genet. Metab. 2011. 102. 418–429.

11. *Higgins J.J., Hao J., Kosofsky B., Rajadhyaksha A.M*. Dysregulation of large-conductance Ca(2+)-activated K(+) channel expression in nonsyndromal mental retardation due to a cereblon p.R419X mutation // Neurogenetics. 2008. 9. 219–223.

12. http://cherishproject.eu – an official web-site of the CHERISH Consortium.

13. http://decipher.sanger.ac.uk – Database of Chromosomal Imbalance and Phenotype in Humans Using Ensemble Resources (DECIPHER).

14. http://omim.org – Online Mendelian Inheritance in Man (An Online Catalog of Human Genes and Genetic Disorders).

15. http://projects.tcag.ca/variation – Database of Genomic Variants (DGV).

16. http://www.cartagenia.com – Cartagenia database.

17. http://www.euromrx.com – an official web-site of the EURO-MRX Consortium.

18. http://www.genecards.org – GeneCards Database.

19. http://www.ncbi.nlm.nih.gov/projects/SNP – Database of Short Genetic Variations (dbSNP).

20. https://www.iscaconsortium.org – an official web-site of the International Standards for Cytogenomic Arrays Consortium (ISCA).

21. *Kashevarova A.A., Skryabin N.A., Cheremnykh A.D. et al*. Clinical and genetic analysis of idiopathic intellectual disability based on array comparative genomic hybridization // Korsakov J. Neurol. Psychiatry (Neuroscience and Behavioral Physiology). 2013. 9. 70–74.

22. *Kim K., Lee S., Kegelman T.P. et al*. Role of excitatory amino acid transporter-2 (EAAT2) and glutamate in neurodegeneration: opportunities for developing novel therapeutics // J. Cell Physiol. 2011. 226. 2484–2493.

23. *Mir A., Kaufman L., Noor A. et al*. Identification of mutations in *TRAPPC9*, which encodes the NIK- and IKK-beta-binding protein, in nonsyndromic autosomal-recessive mental retardation // Am. J. Hum. Genet. 2009. 85. 905–915.

24. *Molinari F., Rio M., Meskenaite V. et al.* Truncating meurotrypsin mutation in autosomal recessive nonsyndromic mental retardation // Science. 2002. 298. 1779–1781.

25. *Motazacker M.M., Rost B.R., Hucho T. et al.* A defect in the ionotropic glutamate receptor 6 gene (*GRIK2*) is associated with autosomal recessive mental retardation // Am. J. Hum. Genet. 2007. 81. 792–798.

26. *Najmabadi H., Hu H., Garshasbi M. et al.* Deep sequencing reveals 50 novel genes for recessive cognitive disorders // Nature. 2011. 478. 57–63.

27. *Nandar W., Neely E.B., Unger E., Connor J.R*. A mutation in the *HFE* gene is associated with altered brain iron profiles and increased oxidative stress in mice // Biochim. Biophys. Acta. 2013. 1832. 729–741.

28. *Philippe O., Rio M., Carioux A. et al.* Combination of linkage mapping and microarray-expression analysis identifies NF-kappa-B signaling defect as a cause of autosomal-recessive mental retardation // Am. J. Hum. Genet. 2009. 85. 903–908.

29. *Ranganathan S., Noyes N.C., Migliorini M. et al*. LRAD3, a novel low-density lipoprotein receptor family member that modulates amyloid precursor protein trafficking // J. Neurosci. 2011. 31. 10836–10846.

30. *Raymond F.L*. Monogenic causes of mental retardation // Genetics of Mental Retardation / Ed. S.J.L. Knight. Monogr. Hum. Genet. Basel: Karger, 2010. 18. 89-100.

31. *Sadakata T., Washida M., Iwayama Y. et al.* Autistic-like phenotypes in Cadps2-knockout mice and aberrant *CADPS2* splicing in autistic patients // J. Clin. Invest. 2007. 117. 931–943.

32. *van Bokhoven H*. Genetic and epigenetic networks in intellectual disabilities // Annu Rev Genet. 2011. 45. 81–104.

33. *Vissers L.E.L.M., Stankiewicz P*. Microdeletion and microduplication syndromes // Genomic Structural Variants: Methods and Protocols / Ed. L. Feuk L. Methods in Molecular Biology. Springer Science-Business Media, LLC, 2012. 838. 29–75.

34. *Zuko A., Bouyain S., van der Zwaag B., Burbach J.P*. Contactins: structural aspects in relation to developmental functions in brain disease // Adv. Protein Chem. Struct. Biol. 2011. 84. 143–180.

Lebedev I.N. *– doctor of biological sciences, head of the laboratory of cytogenetics, e-mail: igor.lebedev@medgenetics.ru*
Nazarenko L.P. *– doctor of medical sciences, professor, vice-head, e-mail: ludmila.nazarenko@medgenetics.ru*
Kashevarova A.A. *– candidate of biological sciences, researcher of the laboratory of cytogenetics, e-mail: anna.kashevarova@medgenetics.ru*
Skryabin N.A. *– candidate of medical sciences, researcher of the laboratory of cytogenetics, e-mail: skryabinn@mail.ru*
Salyukova O.A. *– candidate of medical sciences, researcher of the laboratory of hereditary pathology, e-mail: olga.salukova@mail.ru*
Chechetkina N.N. *– postgraduate student, e-mail: valete@mail.ru*
Tolmacheva E.N. *– candidate of biological sciences, researcher of the laboratory of cytogenetics, e-mail: kate.tolmacheva@medgenetics.ru*
Sazhenova E.A. *– candidate of biological sciences, researcher of the laboratory of cytogenetics, e-mail: elena.sazhenova@medgenetics.ru*
Vovk S.L. *– head of the clinical and diagnostic laboratory of Genetic clinic, e-mail: svetlana.vovk@medgenetics.ru*
Rudko A.A. *– candidate of medical sciences, head of Genetic clinic, e-mail: aleksey.rudko@medgenetics.ru*
Puzyrev V.P. *– doctor of medical sciences, professor, academician of RAMS, head, e-mail: valery.puzyrev@medgenetics.ru*
Magini Pamela *– PhD, Laboratory of medical genetics, Department of gynecology, obstetrics and pediatrics sciences, e-mail: pamela.magini@unibo.it*
Graziano Claudio *– MD, Laboratory of medical genetics, Department of gynecology, obstetrics and pediatrics sciences, e-mail: claudio.graziano@unibo.it*
Romeo Giovanni *– professor of medical genetics, e-mail: romeo@eurogene.org*