Published in final edited form as:

J Neural Transm. 2015 September ; 122(9): 1289–1301. doi:10.1007/s00702-015-1366-8.

Genomic Structural Variants Are Linked With Intellectual Disability

Kazima Bulayeva^{1,*}, Klaus-Peter Lesch², Oleg Bulayev¹, Christopher Walsh³, Stephen Glatt³, Farida Gurganova¹, Jamilja Omarova¹, Irina Berdichevets¹, and Paul M. Thompson^{4,5}

¹ N.I. Vavilov Institute of General Genetics, Russian Academy of Sciences, Gubkin str. 3, Moscow, 119991, Russia.

² Division of Molecular Psychiatry, Department of Psychiatry, Psychosomatics and Psychotherapy, University of Wuerzburg, Fuechleinstr. 15, 97080 Wuerzburg, Germany.

³ Department of Pathology, Children's Hospital Boston, 300 Longwood Avenue, Boston, MA 02115, USA. Department of Psychiatry and Behavioral Sciences, and Medical Genetics Research Center; SUNY Upstate Medical University; 750 East Adams Street; Syracuse, NY 13210, US.

⁴ Imaging Genetics Center, University of Southern California, Los Angeles, CA 90095-7332

⁵ Departments of Neurology, Psychiatry, Radiology, Pediatrics, Engineering, & Ophthalmology, Keck/USC School of Medicine, Los Angeles, CA, USA.

Keywords

mental retardation; genetic isolate; genome-wide linkage scans; SNP; consanguinity; genome variability; LOH analysis

INTRODUCTION

Intellectual disability (ID), earlier known as mental retardation, is a generalized neuropsychiatric disorder, characterized by impaired cognitive functioning and difficulties with adaptive behaviors. ID is one of the most genetically heterogeneous complex disorders, and is an important healthcare problem worldwide; it can also result from genetic mutation. ID occurs in 2-3% of newborns in the general population and mutations in more than 500 genes are thought to impact risk. Even so, its causes have remained elusive in most cases (Roeleveld et al. 1997; Yeargin-Allsopp et al. 1997; Kaufman et al. 2010; Najmabadi et al. 2011). Over 90 different gene defects have been identified for X-chromosome-linked intellectual disability alone. To expedite the molecular elucidation of autosomes related to

*To whom correspondence should be addressed: Tel: 81 82 2575895; Fax: +7 499 1355076; bulaeva@vigg.ru. kplesch@mail.uni-wuerzburg.de, christopher.walsh@childrens.harvard.edu, glatts@upstate.edu, Phone: 323-44-BRAIN; 206-2101(?), FAX: 310-2065518, pthomp@usc.edu

ETHICAL STATEMENT

Written informed consent was obtained from each participant prior to clinical interviews and blood sample collections. The study was approved by the Dagestan IRB (Dagestan Center of the Russian Academy of Sciences).

All authors declare that have no any conflict of interests in submitting this original research article.

intellectual disability, here we performed genome-wide linkage scans and analyzed molecular aberrations within the identified linkage regions in highly consanguineous families with autosomal-recessive intellectual disability from Dagestan.

Specifically, the hippocampal formation - a crucial brain structure for cognitive processes such as learning and memory - has consistently been implicated in ID and in several other neuropsychiatric disorders (Maguire et al. 2000; Burgess et al. 2002; Snyder et al. 2011; Freitag et al. 2009). Hippocampal volume is progressively reduced in Alzheimer's disease (Jack et al. 2011; Simic et al. 1997; Shenton et al. 2001; Theo et al. 2014) and meta-analyses show that it can also be reduced in schizophrenia (SCZ) (Theo et al. 2014) and major depressive disorder (MDD) (Videbech and Ravnkilde 2004).

The heritability of many volumetric brain imaging phenotypes, such as hippocampal volume, is in the range 62-74% (Stein et al. 2012; Peper et al. 2007; Kremen and Jacobson 2010). Even so, identifying specific genetic mechanisms underlying volumetric brain differences has been challenging, as each genetic variant has only a small effect; magnetic resonance imaging (MRI) is also hard to conduct on a scale large enough to achieve sufficient statistical power for genetic association studies (Stein et al. 2011). Recently, however, we and others published the first very large genome-wide association studies (GWAS; N=21,151) for human brain measures, including hippocampal and intracranial volumes measured using MRI from infancy to old age (Stein et al. 2012; Ikram et al. 2012; Bis et al. 2012; Taal et al. 2012).

Earlier we reported our results from a genome-wide linkage scan in a pedigree ascertained in a genetic isolate of Dagestan (Northern Caucasus, Russia) with aggregation of non-specific ID. We analyzed chromosome 12 for linkage with ID (mental retardation) and found LOD>3 at 12q24 (Bulayeva et al. 2002; 2005). Such genetically homogeneous isolates offer exceptional resources to detect susceptibility genes for complex human diseases (Alkuraya 2013). Consanguinity is common in such genetic isolates and autosomal recessive diseases are more common. Recent progress in molecular genetics and bioinformatics make it possible to identify the role of consanguinity in detecting causal mutations in autosomal recessive diseases in an efficient and cost-effective way, as compared to genetically heterogeneous outbred populations. Here we report our follow-up linkage analysis results obtained in an extended pedigree from a Dagestan genetic isolate with aggregation of ID. We focused on chromosomal regions where we previously found significant associations with volumetric differences in brain structures that underlie cognition and have been implicated in neuropsychiatric disorders in large cohorts scanned with MRI (Videbech and Ravnkilde 2004; Lee et al. 2012; Brouwer et al. 2012).

MATERIALS AND METHODS

Description of populations

Dagestan contains 26 small indigenous ethnic groups that have existed for over 10,000 years in the same highland region, according to archaeological data. The ethnic groups are subdivided into numerous remote highland villages; according to population type classifications, some of these are designated as genetic isolates. We previously reported on a

large multigenerational pedigree from one of the isolates with a dense aggregation of ID (designated as DGH011). The isolate and pedigree members were ascertained as part of our long-term research program on the 'Dagestan Genetic Heritage (DGH)' (Bulayeva et al. 2000; 2002; 2003; 2005; 2007).

Clinical Assessment and Diagnostic Procedure

Historically, ID was defined based on an 'Intelligence Quotient' (IQ) score below 70 (ICD-10, version 2010; F70-F79). In the general population, the average IQ is 100 and an individual is considered intellectually disadvantaged if he or she has an IQ of less than 70 to 75. Unspecific ID affected subjects had been diagnosed in Dagestan psychiatric hospitals in early childhood. In a total of 21 observed ID cases, two probands displayed co-morbidity with schizoaffective disorder and three with neurosensory deafness. We re-diagnosed all affected probands using a structured psychiatric interview - the Diagnostic Interview for Genetic Studies (DIGS), based on the DSM-IV diagnostic criteria (Nurnberger et al. 1994) - as well as using ICD-10.

We previously translated the DIGS into Russian for our research and adapted it to ethnic and cultural differences of Dagestan regions, in both computerized and paper forms. Diagnostic assessments were conducted by two Dagestan psychiatrists (R. Kurbanov and U. Guseinova) trained on the DIGS and experienced with clinical evaluations using ICD-10, during their long-term work assignment in regional psychiatric hospitals. The affected subjects who met ID criteria were included as probands in our linkage study, while two probands with schizoaffective co-morbidities were excluded from the analysis.

Pedigree information was corroborated and verified across interviews of different family members within each isolate. For drawing of pedigrees and data management, we used Progeny software, which combines the power of a pedigree drawing tool with an easy-to-use back-end database.

Written informed consent was obtained from each participant prior to clinical interviews and blood sample collections. The study was approved by the Dagestan IRB (Dagestan Center of the Russian Academy of Sciences).

Genotyping

Blood samples were collected from four current generations of affected and healthy pedigree members. Genomic DNA was extracted from peripheral-blood leukocytes with standard protocols. Using standard methods, approximately 400 µg of DNA was isolated from each subject's blood samples. This DNA was sent to the high-throughput genotyping facilities at the Mammalian Genotyping Service of the Marshfield Medical Research Foundation of the National Institutes of Health (US) (Weber et al. 1993). The Weber/CHLC map contains approximately 400 STR markers (short tandem repeats; mostly tri- and tetranucleotide repeats), covering the genome at 10 cM resolution.

Computation of CNVs and ROH

Affymetrix SNP 5.0 array sets were used with standard protocols at the Harvard Medical Center (C. Walsh's lab) for genotyping affected and un-affected members of the kindred with ID aggregation. A CNV was defined when there were at least three consecutive SNPs showing consistent deletion or duplication. As it was not possible to delineate the exact boundaries of each CNV with genome-wide SNP-genotyping arrays, we used the positions of SNPs to approximate boundaries. If two individual CNVs overlapped, we merged them as a CNVR, using as boundaries the maximum-interval SNPs selected from these two CNVs. The boundary and size of the CNV intervals are defined based on the positions of the first and last array probes identified as lying within the CNV.

Loss of homozygosity (LOH) measures the degree of allelic similarity in a genome. LOH can occur either in uniparental disomy (UPD) - when both copies of a gene or genomic region are inherited from the same parent - or by bi-parental autozygosity, when alleles are derived from both parents. The generic term "Runs of Homozygosity" (ROH) may be used to denote either autozygous segments or UPD. ROHs were defined, for each subject in the present study, as any window of 100 or more consecutive SNPs on a single chromosome not receiving a heterozygous call. The minimum number of SNPs that constituted a ROH (l) was calculated with a method similar to that proposed by Lencz et al. (2007). The main goal of this study was to find overlapping ROH in affected cases within linked genomic regions as a possible pathogenic cause of the clinical phenotype studied. We identified overlapping regions in ROH segments, using the sliding-window approach mentioned above. We then calculated for each SNP the proportion of homozygous windows among ID-affected participants and their healthy relatives that overlap that same position. The percentage of members that had the region with the most overlapping ROH on each chromosome was plotted, and the candidate genes located in ROH segments were identified.

Statistical analysis

For linkage analyses, we designated pedigree members with ID as 'affected'; pedigree members were considered 'unaffected' if they were determined to have no mental illness based on all available clinical information. Other individuals, including pedigree members with unclear clinical symptoms, were considered 'unknown'. For the genome-wide STRs linkage study we used Simwalk2.91 based on the Markov Chain/Monte Carlo (MCMC) algorithm (Sobel and Lange 1996), which can be used to analyze large pedigrees as it considers the underlying configurations in proportion to their likelihood. Using multipoint parametric linkage analyses, we tested both dominant and recessive models with a 90% penetrance model of inheritance, the disease allele frequency of 0.02, and the assumption of genetic heterogeneity. A detailed description of using Simwalk2 for linkage analysis in the Dagestan genetic isolates is described in our prior publications (Bulayeva et al. 2005; 2007).

Following this linkage scan, samples were analyzed using Affymetrix 5.0 SNP microarrays in order to refine the information on the contribution of CNV and ROH in the genomic regions linked with ID. Such molecular aberrations were calculated with the Golden Helix SVS7.7.3 software, which implements a Hidden Markov Model (HMM)-based algorithm to identify chromosomal gains and losses by comparing the signal intensity of each SNP probe

set against a reference set. The arrays that passed the QC call rate threshold were analyzed using the Birdseed algorithm with the default setting of 0.1; a minimum segment length of 100 kb was required for calling a CNV. To minimize the risk of invalid samples passing the QC Call Rate metric, samples with outlier values for heterozygosity rate were rejected. A total of three samples that failed CNV quality control were excluded from further analysis. All arrays had a call rate of $\geq 95\%$ and were thus included in subsequent analyses. The final mean BRLMM call rates reached a high level of 99.0% and 99.1% for the ID samples, respectively. All subjects that passed SNP QC procedures were entered into the CNV and LOH analysis.

RESULTS

Demographic and clinical characteristics for the isolate ID kindred are presented in **Table 1**. The pedigree was reconstructed for 21 affected cases (available for our expedition study from a total of 27 cases) and from a total of 35 individuals. The reconstructed pedigree showed that all affected ID cases are located in one extended pedigree with limited numbers of common ancestors. Interestingly, for the founder effects, many of the neighboring highland villages of the same ethnic background and population size in Dagestan do not have *any* diagnosed cases of ID. The pedigree has 187 members, across 11 generations (**Table 1**). The sex ratio across the isolate and in the pedigree reconstructed is similar and close to 1:1.

We analyzed the marriage structure in the reconstructed extended pedigree and found a high rate of 88% of traditional endogamy (within village marriages) and among them 57% are consanguineous marriages. We calculated the coefficient of inbreeding, F , using several methods: (a) based on traditional population-genetic studies, three-generation marriage analyses among assessed subjects ($F_{\text{pop}}=0.0066$); b) based on marriage structure within the reconstructed pedigree ($F_{\text{ped}}=0.0346$), i.e., the population-based F_{pop} is five times is lower than pedigree-based F_{ped} . We obtained similar differences between affected and healthy probands in the isolate as well: the mean coefficient of inbreeding among ID cases was 2.7 times higher, compared to healthy subjects ($\chi^2=30.5$; $df=1$; $P=0.000$).

We performed genome-wide multipoint linkage analyses in this pedigree using the statistical genetics application, SimWalk2. This employed Markov chain Monte Carlo and simulated annealing algorithms that are suitable for analyzing large complex pedigrees with inbred loops (**Table 2**). Parametric linkage analysis results in the pedigree with an adequately specified level of penetrance and disease alleles frequencies. We detected a number of statistically significant ($p \leq 0.05$) suggestive linkage peaks in a total of 10 genomic regions: 1q41, 2p25.3-p24.2, 3p13-p12.1, 4q13.3, 10p11, 11q23, 12q24.22-q24.31, 17q24.2-q25.1, 21q22.13 and 22q12.3-q13.1 (**Table 2**). Three significant linkages with LOD >3 we obtained at 2p25.3-p24.2 under the dominant model, a peak at 21 cM, flanked loci D2S2976 and D2S2952; at 12q24.22-q24.31 under the recessive model, a peak at -120 cM, flanked by loci D12S2070 and D12S395 and at 22q12.3 under the dominant model, a peak -32 cM, flanked loci D22S683 and D22S445 (**Table 2**). These linked regions contain genes implicated in the pathogenesis of ID and/or other neurodevelopment psychiatric disorders. The linked region 2p25.3-p24.2 with LOD = 3.77 contains *MYT1L* (myelin transcription

factor 1-like) was shown as associated with ID, as well as with SCZ, MDD and ADHD. Other genes, such as *SNTG2* which was reported as a cause of neuronitis, muscular dystrophy, ASD and ID, as well as genes *TPO* associated with hypothyroidism *COLEC11* related to learning disability, and *SOX11* linked to neuronitis and ID also are located in this linked region (**Table 2, Fig. 1**).

The genomic region 12q24.22-q24.32 with LOD=3.87 under a recessive mode of inheritance, peak -120 cM and flanking loci D12S2070- D12S395, contains genes related to cognition and psychoses - *TESC*, *HRK*, *FBXW8*, *SBNO1*, *MED13*, *CDK2AP1*, *CIT*, *DYNLL1*, *P2RX7*, *NOS1*, *PLA2G1B*, *CCDC60* (**Table 2, Fig. 1**). The gene *CIT* has been shown to be related to schizophrenia and gene *NOS1* was reported to be a candidate gene for impulsivity and aggressiveness.

Region 22q12.3-q13.1 where we obtained LOD=3.4 under the dominant mode, peak=32 and flanking loci D22S683-D22S445 contains a well known associated with muscular dystrophy and ID-associated gene, *LARGE* (like-glycosyltransferase) that is a protein-coding gene.. Mutation in the *LARGE* gene can cause 2 different forms of muscular dystrophy (<http://omim.org/entry/603590>): dystroglycanopathy (MDDG): a severe congenital form with brain and eye anomalies (type A6; MDDGA6, [613154](#)), formerly designated Walker-Warburg syndrome (WWS) or muscle-eye-brain disease (MEB), and a less severe congenital form with mental retardation (type B6; MDDGB6; [608840](#)), formerly designated congenital muscular dystrophy type 1D (MDC1D). Other important gene associated with cognition and psychosis in this genomic region are *SYN3* (Synapsin III). Synapsins encode neuronal phosphoproteins, which associate with the cytoplasmic surface of synaptic vesicles. Family members are characterized by common protein domains, and they are implicated in synaptogenesis and the modulation of neurotransmitter release, suggesting a potential role in several neuropsychiatric diseases that were supported in previous studies of associations with schizophrenia, bipolar disorder, multiple sclerosis and ADHD. Gene *TOM1* (Target Of Myb Protein 1), located in same linked region, is a protein-coding gene, and is affiliated with the lncRNA class. Diseases associated with *TOM1* include bipolar disorder, chronic wasting disease, cystic fibrosis. Other genes here are *CSNK1E*, *PVALB*, *APOL1-APOL4*, *RBM9*, *MYH9*, *IL2RB*, *CACNG2*, *SOX10*, *DDX17*, *RPSAP52* (**Table 2, Fig. 1**). All these genes have been previously reported as associated or linked with cognitive impairment or with aspects of neurodevelopment or neurodegenerative diseases.

After specific chromosomal regions had been identified by linkage analyses and certain genes were identified as possible candidates, we used exploratory examination of CNV and ROH based on Affymetrix 5.0 microarray data to detect structural variations within the ID-linked regions. In regions linked with ID, we obtained 60 LOH and 27 CNV. "Hot spots" for ROH and CNV were observed in linked regions 12q24.31 (11 ROH and 7 CNV) and in 17q21.31 (14 ROH and 11 CNV). The results obtained suggest that at least some of the observed ROH are related to deletions in CNVs. In five of the selected linked regions with LODs=2-3, ROH and CNV segments length sizes were larger among ID cases than among healthy their relatives (**Supplemental Fig. 1 and 2**). Our results showed that summarized for all linked regions ROH length size among ID cases was larger by a factor of 1.8 times,

compared to healthy pedigree members: in ID, cases' mean length for ROH segments was 1263 kb, but among healthy members it was only 671 kb ($t=2.9$, $df=170$, $p=0.008$).

Both groups of ID cases and healthy pedigree members demonstrated statistically significant higher rates of heterozygosity losses and higher rate of genomic structural variations among offspring of consanguineous marriages, compared offspring of unrelated marriages (**Fig 2 a,b**). It is well known that inbreeding with close relatives leads to long ROH stretches among offspring, as spouses share much of their DNA obtained from common ancestor. Endogamous marriages — within local populations — also leads to increased ROH stretches but they are less long (medium) than is the case for inbred marriages between close relatives. In our genetic isolate, we studied the offspring of two types of marriages: consanguineous and non-consanguineous. We found statistically significant differences between these two groups of offspring in the lengths of the ROH segments: segmental mean length size among offspring of consanguineous marriages was 2652 ± 1163 kb, and among offspring of non-consanguineous marriages, it was 702 ± 122 kb ($t=2.87$, $p=0.008$).

Given the nonuniform patterns of ROHs across the genome, we were able to uncover regions with different lengths between groups. Results obtained show that the most significant difference in the sizes of ROH segments between ID and healthy groups was at the 12q24 linked region: in the ID group mean size length was 2.5 Mb, but in healthy subjects 682 kb (**Fig. 3**). Within ROH segments at 12q24.31, 15 of the 21 ID cases in the pedigree demonstrated LOH in the gene *SBNO1* (*Supplemental Fig. 3*); four cases showed LOH in this gene between *rs62774561* and *rs10773005*, while 11 cases showed LOH between *rs62774561* and *rs11268916*. In the gene *CDK2AP1*, we found LOH between *rs4759414* and *rs10846494* in mother with ID affection and in her four heavily affected ID offspring. These four children also had LOH in the gene *MED13L*, between *rs34802971* and *rs3851642*. Three affected cases demonstrated LOH between *rs60354985* and *rs10774903* the in region 12q24.22 where the genes *HRK*, *FBXW8* and *TESC* are located (**Table 3a,b**).

“Hot spots” – or higher frequencies of CNV and ROH segments – we observed in the linked regions 12q24.21-q24.31 (**Fig. 4**). Eleven cases in 12q24.22 have CNV (deletions) and 15 cases of ROH segments were found in the genes *MED13L*, *HRK*, *FBXW8* and *TESC* (*rs11609741*, *rs79667445* and *rs4767490*) (**Figure 4**). In the same 12q24.22 intergenic region with *rs7294919* (between *HRK* and *FBXW*) three of the most severely affected ID cases demonstrated ROH (**Fig 4**). We found ROH also in genes *CDK2AP1* and *SBNO1* (12q24.31) for 15 CN segments in 7 ID cases (**Fig. 5**). Among unaffected members of the pedigree, we found 4 ROH and CN segments, three of them at 12q24.31 and one in the 12q24.22 region (**Fig. 4, Table 4**). In the linked region 2p25.3-p24.2, we found deletions among 5 of 21 affected participants, within genes *SNTG2* and *TPO*. And in linked region 22q12.2-q13.1 CNV variation analyses showed common for four ID affected deletions in gene *LARGE* with *rs8141384* and *rs8140012*.

We also checked genomic regions where recent studies reported associations with key brain structures related to cognition. In 8 ID cases, we found CNV deletion within gene *HMGA2* located in the 12q14.3 region. Seven of affected ID cases also showed loss of heterozygosity in the same genomic region (**Table 4**).

In region 6q22.32, we observed 3 affected ID cases with deletions and 9 cases with LOH in the NCOA7 gene (Table 4). This gene enhances the transcriptional activities of several nuclear receptors and is involved in the co-activation of different nuclear receptors, such as ESR1, THRB, PPARG and RARA.

We also obtained CNV and ROH 'hot spots' at 17q21.31 -11 in affected ID cases demonstrated deletions and duplications and loss of heterozygosity in genes CRHR1, KIAA1267 and MAPT that recently has been shown associated with brain measures derived from MRI scans (Stein et al, 2012) (**Table 4, Supplemental Fig 4 and 5**).

DISCUSSION

Overall, we obtained significant LOD scores for ID in 10 genomic regions in kindred studied within isolate DGH011. We obtained a LOD >3 for ID in the regions 2p25.3-p24.2, 12q24.22-q24.31 and 22q12.3-q13.1 (**Table 2**). One of the genes, *MYT1L* is a protein-coding gene, located in the 2p25.1-p24.2 region (chr2:1,887,649-1,927,035). It is a strong candidate implicated in the ID phenotype and has previously been associated with SCZ, MDD and ADHD (Lee et al. 2012; Li et al. 2012; Stevens et al. 2011; Lesch et al. 2008). The gene is functionally implicated in neuronal differentiation and in the development of neurons and oligodendroglia in the CNS. Other ID candidate genes located in this region are *COLEC11*, *SNTG2*, *TPO*, *SOX11*, *KIDINS220* and *ASAP2*.

KCNJ2, located in the ID region 17q24.2-q25.1, is important in the development of cleft lip and palate, as well as *KCNJ16* and *USH2A*, which are associated with sensory neural hearing loss. We found a combination of these co-morbidities in three ID cases, in the isolate studied.

Three genomic regions we found linked to ID: 11q23, 12q24.23-q24.32 and 22q12.3-q13.1, have previously been reported to be associated with SCZ and MDD (Bulayeva et al, 2007; 2011; 2012). The genomic regions linked to this spectrum of clinical phenotypes were also confirmed by MRI findings in SCZ (Simic et al. 1997). Other studies demonstrated a reduced size of the amygdala-hippocampal complex, particularly in the posterior portion, in first-episode bipolar disorder (BPD) patients - e.g. (Hirayasu et al. 1998; Velakoulis et al. 1999) and in chronic MDD patients, e.g. (Altshuler et al. 1998; Drevets et al. 1992; Mayberg 1993; Derrek et al. 2014). This risk factor for developing psychiatric disorders across diagnostic boundaries is also related to cognitive impairments, including working memory deficits, in patients with SCZ, as suggested by several investigators, e.g. (Goldberg et al. 1994; Nestor et al. 1993; Weinberger 1999). This further supports the notion that genes involved in the development and maintenance of hippocampal circuitry or in the expression of molecules that mediate processes of neural plasticity in the hippocampus may play a critical role in the genetic predisposition to SCZ.

Of particular interest is the ID-linked region 12q24 generating the highest LOD=3.87. This is the same region where our recent studies using brain MRI scans of 21,151 individuals found associations with the volumes of key brain structures related to cognition that may also be involved in the pathogenesis of ID. The pedigree's ID cases share a haplotype

block with identical-by-descent (IBD) alleles in two loci flanked by D12S2070-D12S395 with the following alleles combined in genotype -77-44-, except three affected sibs with genotypes -76-43- that most likely are related to recombination events. Further study of the 12q24 region is warranted based on recent GWAS findings where we and others demonstrated replicated associations between genomic variants at 12q24 and key brain structures related to cognition (Videbech and Ravnkilde 2004; Brouwer et al. 2012; Li et al. 2012; Pahari et al. 2013). This region contains about 30 genes; several of them have been implicated in cognition and ID, including genes *ESC*, *HRK*, *FBXW8*, *SBNO1* and *CDK2AP1*, as well as with neuropsychiatric disorders (SCZ, MDD and BPD), including genes *CIT*, *DYNLL1*, *P2RX7*, *NOS1*, *PLA2G1B* and *CCDC60*. Those genes related to both cognition and ID are also known to harbor SNPs associated with hippocampal volume and overall brain volume. In a recent meta-analysis of seven GWAS (Nobs = 10,768 of European ancestry), two top signals in 6 replication studies were obtained: SNP *rs7980687* in *SBNO1* on chromosome 12q24.23 and *rs1042725* on chromosome 12q14. These SNPs were robustly associated with head circumference in infancy (Lee et al. 2012). A study in *zebrafish larvae* found that knockout of *Sbno1* specifically affects regionalization along the anterior-posterior axis of the brain suggesting essential roles of *SBNO* genes in vertebrate brain development (Takano et al. 2011). In the present study we found that all our ID cases had a loss of heterozygosity within gene *SBNO1* located within linked and ROH segment region.

In our own study as part of the ENIGMA Consortium (<http://enigma.ini.usc.edu>), we found that the intergenic variant *rs7294919* was associated with hippocampal volume (12q24.22; $N = 21,151$; $P = 6.70 \times 10^{-16}$) and with the expression levels of the positional candidate *TESC* in brain tissue (Stein et al. 2012; Bis et al. 2012). *TESC* expression is strongly regulated during cell differentiation in a cell lineage-specific fashion and plays a significant role in cell proliferation and differentiation (Levy and Slepak 2007; 2010) which is likely to be relevant for hippocampal volume and brain development (Stein et al. 2012). In Dagestan genetic isolate most severe affected ID cases demonstrated ROH in this intergenic between *HRK* and *FBXW* variant *rs7294919* that confirm with our ENIGMA consortium findings.

Results obtained in present paper suggest that at least some of the observed ROH are related to deletions in CNVs. For example in gene *FBXW8* LOH and deletions were detected. Further screening for structural variations in same genomic region linked with ID in our isolate DGH011 indicated that, in this genomic region, eight ID cases have CNV in genes *FBXW8* and *TESC* (*rs11609741*, *rs79667445* and *rs4767490*).

In conclusion, we performed a linkage and genome structural variation analyses in a pedigree ascertained in a highland genetic isolate, with aggregations of ID. We were able to implicate genomic regions that recent studies have associated with brain measures derived from MRI scans, especially in region 12q24.

ACKNOWLEDGMENTS

This work was supported in part by research grants from the RFBR and ‘Dynamics of Gene Pools’ of Russian Academy of Sciences Council. The study was supported also in part by the European Community (EC: AGGRESSOTYPE FP7/No. 602805)

REFERENCES

1. Roeleveld N, Zielhuis GA. The prevalence of mental retardation: a critical review of recent literature. *J Developmental Medicine & Child Neurology*. 1997; 39(2):125–32.
2. Yeargin-Allsopp M, Murphy CC, Cordero JF, Decouflé P, Hollowell JG. Reported biomedical causes and associated medical conditions for mental retardation among 10-year-old children, metropolitan Atlanta, 1985 to 1987. *Dev Med Child Neurol*. 1997; 39(3):142–9. [PubMed: 9112961]
3. Kaufman L, Ayub M, Vincent JB. The genetic basis of non-syndromic intellectual disability: a review. *J Neurodev Disord*. 2010; 2(4):182–209. doi: 10.1007/s11689-010-9055-2. [PubMed: 21124998]
4. Najmabadi H, Hu H, Garshasbi M, et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. 2011 doi:10.1038/nature10423.
5. Maguire EA, Gadian DG, Johnsrude IS, Good CD, Ashburner J, Frackowiak RS, Frith CD. Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences of the United States of America*. 2000; 97:4398–4403. [PubMed: 10716738]
6. Burgess N, Maguire EA, O'Keefe J. The human hippocampus and spatial and episodic memory. *Neuron*. 2002; 35:625–641. [PubMed: 12194864]
7. Snyder JS, Soumier A, Brewer M, Pickel J, Cameron HA. Adult hippocampal neurogenesis buffers stress responses and depressive behaviour. *Nature*. 2011; 476:458–461. [PubMed: 21814201]
8. Freitag CM, Luders E, Hulst HE, Narr KL, Thompson PM, Toga AW, Krick C, Konrad C. Total brain volume and corpus callosum size in medication-naïve adolescents and young adults with autism spectrum disorder. *Biol Psychiatry*. 2009; 66:316–319. [PubMed: 19409535]
9. Jack CR Jr, Barkhof F, Bernstein MA, Cantillon M, Cole PE, Decarli C, Dubois B, Duchesne S, Fox NC, Frisoni GB, Hampel H, Hill DL, Johnson K, Mangin JF, Scheltens P, Schwarz AJ, Sperling R, Suhy J, Thompson PM, Weiner M, Foster NL. Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer's disease. *Alzheimers Dement*. 2011; 7:474–485. e474. [PubMed: 21784356]
10. Simic G, Kostovic I, Winblad B, Bogdanovic N. Volume and number of neurons of the human hippocampal formation in normal aging and Alzheimer's disease. *J Comp Neurol*. 1997; 379:482–494. [PubMed: 9067838]
11. Shenton ME, Dickey CC, Frumin M, McCarley RW. A review of MRI findings in schizophrenia. *Schizophr Res*. 2001; 49:1–52. [PubMed: 11343862]
12. Theo GM, van Erp, Derrek PH, Jerod MR, David CG, Godfrey DP, et al. Subcortical Brain Volume Abnormalities in 2,028 Patients with Schizophrenia and 2,540 Healthy Controls. via the ENIGMA Consortium, submitted to *Nature Neuroscience*. Mar 22.2014 2014.
13. Videbech P, Ravnkilde B. Hippocampal volume and depression: a meta-analysis of MRI studies. *Am J Psychiatry*. 2004; 161:1957–1966. [PubMed: 15514393]
14. Stein JL, Medland SE, Vasquez AA, Hibar DP, Senstad RE, Winkler AM, et al. Identification of common variants associated with human hippocampal and intracranial volumes. *Nat Genet*. 2012; 44:552–561. [PubMed: 22504417]
15. Peper JS, Brouwer RM, Boomsma DI, Kahn RS, Hulshoff Pol HE. Genetic influences on human brain structure: a review of brain imaging studies in twins. *Hum Brain Mapp*. 2007; 28:464–473. [PubMed: 17415783]
16. Kremen WS, Jacobson KC. Introduction to the special issue, pathways between genes, brain, and behavior. *Behav Genet*. 2010; 40:111–113. [PubMed: 20155393]
17. Stein JL, Hibar DP, Madsen SK, et al. Discovery and replication of dopamine-related gene effects on caudate volume in young and elderly populations (N=1198) using genome-wide search. *Mol. Psychiatry*. 2011; 16:927–937. 10.1038/mp.2011.32. [PubMed: 21502949]
18. Ikram MA, Fornage M, Smith AV, Seshadri S, Schmidt R, et al. Common variants at 6q22 and 17q21 are associated with intracranial volume. *Nat Genet*. 2012; 44:539–544. [PubMed: 22504418]

19. Bis JC, DeCarli C, Smith AV, van der Lijn V, Crivello F, Fornage M, et al. Common variants at 12q14 and 12q24 are associated with hippocampal volume. *Nat Genet.* 2012; 44:545–551. [PubMed: 22504421]
20. Taal HR, St Pourcain B, Thiering E, Das S, Mook-Kanamori DO, Warrington NM, et al. Common variants at 12q15 and 12q24 are associated with infant head circumference. *Nat Genet.* 2012; 44:532–538. [PubMed: 22504419]
21. Bulayeva KB, Pavlova TA, Kurbanov RM, Bulayev OA. Mapping genes of complex disease in genetic isolates of Daghestan. *Russian Journal of Genetics.* 2002; 38:1304–1311.
22. Bulayeva KB, Leal SM, Pavlova TA, Kurbanov RM, Glatt SJ, Bulayev OA, et al. Mapping genes of complex psychiatric diseases in Daghestan genetic isolates. *Am J Med Genet B Neuropsychiatr Genet.* 2005; 132:76–84. [PubMed: 15389762]
23. Bulayeva KB, Glatt SJ, Bulayev OA, Pavlova TA, Tsuang MT. Genome-wide linkage scan of schizophrenia: a cross-isolate study. *Genomics.* 2007; 89:167–177. [PubMed: 17140763]
24. Alkuraya FS. Impact of new genomic tools on the practice of clinical genetics in consanguineous populations: the Saudi experience. *Clinical Genetics.* 2013; 84:203–208. [PubMed: 23451714]
25. Lee Y, Mattai A, Long R, Rapoport JL, Gogtay N, Addington AM. Microduplications disrupting the MYT1L gene (2p25.3) are associated with schizophrenia. *Psychiatr Genet.* 2012; 22:206–209. [PubMed: 22547139]
26. Brouwer RM, Mandl RC, Schnack HG, van Soelen IL, van Baal GC, Peper JS, et al. White matter development in early puberty: a longitudinal volumetric and diffusion tensor imaging twin study. *PLoS One.* 2012; 7:e32316. [PubMed: 22514599]
27. Bulayeva K, Devlin B, Bacanu SA, Bennett P, Roeder K, Byerley WF. Linkage analysis of schizophrenia in isolated Daghestanian kindreds. *American Journal of Human Genetics.* 2000; 67:309–309.
28. Bulayeva KB, Pavlova TA, Kurbanov RM, Leal S, Bulayev OA. Genetic and epidemiological studies in Daghestan highland isolates. *Russian Journal of Genetics.* 2003; 39:333–341.
29. Nurnberger J, Blehar M, Kaufman C, York-Cooler C, Simpson S, Harkavy-Friedman J, et al. Diagnostic interview for genetic studies: rationale, unique features and training. *Arch Gen Psychiatry.* 1994; 51:849–862. [PubMed: 7944874]
30. Weber JL, Wang Z, Hansen K, Stephenson M, Kappel C, Salzman Sh, Wilkie PJ, Keats B, Dracopoli NC, Brandriff BF, Olsenet AS. Evidence for Human Meiotic Recombination: Interference Obtained through Construction of a Short Tandem Repeat-Polymorphism Linkage Map of Chromosome 19. *Am J Hum Genet.* 1993; 53:1079–1095. [PubMed: 8213834]
31. Lencz T, Lambert C, DeRosse P, Burdick KE, Morgan TV, Kane JM, et al. Runs of homozygosity reveal highly penetrant recessive loci in schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America.* 2007; 104:19942–19947. [PubMed: 18077426]
32. Sobel E, Lange K. Descent Graphs in Pedigree Analysis: Applications to Haplotyping, Location Scores, and Marker Sharing Statistics. *Am J Hum Genet.* 1996; 58:1323–1337. [PubMed: 8651310]
33. Li W, Wang X, Zhao J, Lin J, Song XQ, Yang Y, et al. Association study of myelin transcription factor 1-like polymorphisms with schizophrenia in Han Chinese population. *Genes Brain Behav.* 2012; 11:87–93. [PubMed: 21923761]
34. Stevens SJ, van Ravenswaaij-Arts CM, Janssen JW, Klein Wassink-Ruiter JS, van Essen AJ, Dijkhuizen T, et al. MYT1L is a candidate gene for intellectual disability in patients with 2p25.3 (2pter) deletions. *Am J Med Genet A.* 2011; 155A:2739–2745. [PubMed: 21990140]
35. Lesch KP, Timmesfeld N, Renner TJ, Halperin R, Röser C, Nguyen TT, Craig DW, Romanos J, Heine M, Meyer J, Freitag C, Warnke A, Romanos M, Schäfer H, Walitza S, Reif A, Stephan DA, Jacob C. Molecular genetics of adult ADHD: converging evidence from genome-wide association and extended pedigree linkage studies. *J Neural Transm.* 2008; 115:1573–1585. [PubMed: 18839057]
36. Bulayeva KB, Lencz T, Glatt S, Takumi T, Gurganova FR, Bulayev OA. Genome-wide linkage scan of major depressive disorder in two Dagestan genetic isolates. *Central European Journal of Medicine.* 2011; 6:616–624.

37. Bulayeva K, Lencz T, Glatt S, Takumi T, Gurganova F, Kawakami H, Bulayev O. Mapping genes related to early onset major depressive disorder in dagestan genetic isolates. *Turk Psikiyatri Derg.* 2012; 23(3):161–70. [PubMed: 22949285]
38. Hirayasu Y, Shenton ME, Salisbury DF, Dickey CC, Fischer IA, Mazzoni P, et al. Lower left temporal lobe MRI volumes in patients with first-episode schizophrenia compared with psychotic patients with first-episode affective disorder and normal subjects. *The American journal of psychiatry.* 1998; 155:1384–1391. [PubMed: 9766770]
39. Velakoulis D, Pantelis C, McGorry PD, Dudgeon P, Brewer W, Cook M, et al. Hippocampal volume in first-episode psychoses and chronic schizophrenia: a high-resolution magnetic resonance imaging study. *Arch Gen Psychiatry.* 1999; 56:133–141. [PubMed: 10025437]
40. Altshuler LL, Bartzokis G, Grieder T, Curran J, Mintz J. Amygdala enlargement in bipolar disorder and hippocampal reduction in schizophrenia: An MRI study demonstrating neuroanatomic specificity. *Archives of General Psychiatry.* 1998; 55:663–664. [PubMed: 9672058]
41. Drevets WC, Videen TO, Price JL, Preskorn SH, Carmichael ST, Raichle ME. A functional anatomical study of unipolar depression. *J Neurosci.* 1992; 12:3628–3641. [PubMed: 1527602]
42. Mayberg, HS. Neuroimaging studies of depression in neurologic disease.. In: Starkstein, SERRG., editor. *Depression in Neurological Disease.* Johns Hopkins University Press; 1993. p. 186–216. 1993 Johns Hopkins University Press, pp 186–216
43. Derrek PH, Lars TW, Theo GM, van Erp, for the ENIGMA Bipolar Disorder Working Group. Robust subcortical volumetric reductions in bipolar disorder: findings from the ENIGMA bipolar disorder working group including 1,745 cases and 2,613 controls. submitted to *Nature Neuroscience.* Mar 3.2014 2014.
44. Goldberg TE, Torrey EF, Berman KF, Weinberger DR. Relations between neuropsychological performance and brain morphological and physiological measures in monozygotic twins discordant for schizophrenia. *Psychiatry Res.* 1994; 55:51–61. [PubMed: 8047629]
45. Nestor PG, Shenton ME, McCarley RW, Haimson J, Smith RS, O'Donnell B, et al. Neuropsychological correlates of MRI temporal lobe abnormalities in schizophrenia. *The American journal of psychiatry.* 1993; 150:1849–1855. [PubMed: 8238641]
46. Weinberger DR. Cell biology of the hippocampal formation in schizophrenia. *Biological psychiatry.* 1999; 45:395–402. [PubMed: 10071707]
47. Pahari DR, Gu YJ, van Oeveren W, El-Essawi A, Harringer W, Brouwer RM. Effect of minimized perfusion circuit on brain injury markers carnosinase and brain-type Fatty binding protein in coronary artery bypass grafting patients. *Artif Organs.* 2013; 37:128–135. [PubMed: 23020859]
48. Takano A, Zochi R, Hibi M, Terashima T, Katsuyama Y. Function of strawberry notch family genes in the zebrafish brain development. *Kobe J Med Sci.* 2011; 56:E220–230. [PubMed: 21937870]
49. Levay K, Slepak VZ. Tescalcin is an essential factor in megakaryocytic differentiation associated with Ets family gene expression. *J Clin Invest.* 2007; 117:2672–2683. [PubMed: 17717601]
50. Levay K, Slepak VZ. Up- or downregulation of tescalcin in HL-60 cells is associated with their differentiation to either granulocytic or macrophage-like lineage. *Exp Cell Res.* 2010; 316:1254–1262. [PubMed: 20060826]

Fig. 1.

Characterization of multipoint Lod scores 2.1–3.87 and genes located in linked regions across chromosomes 2 (a), 11 (b), 12 (c) and 22 (d) in the pedigree with aggregation of intellectual disability (ID)

Fig. 2.

a Box-plot in CN segment length sizes between healthy (N) and ID groups. Differences are statistically significant: Mann–Whitney $Z_{adj} = 1.93$, $p = 0.050$. b Box-plot in ROH segment length sizes between healthy (N) and ID groups. Differences are statistically significant: Mann–Whitney $Z_{adj} = 2.37$, $p = 0.018$

Fig. 3.
ROH mean segment length sizes: differences between ID and healthy (N) groups in five linked regions

Fig. 4.
Family aggregation and transmission of CN segments in 12q24.22 (white triangles) and at 12q24.31 (black triangle); same region ROH segments are marked by white and black square

Table 1

Demographic structure of the isolate DGH011 and the pedigree reconstructed there.

Parameters	Numbers
Nt *	2000
Endogamy rate (%)	83
Nt of pedigree	187
Fpop ** /Fped ***	0.0066/ 0.0346
Male/Female ratio in pedigree (%)	54/46
Total # of affected alive	27
Number of observed subjects	35(21)

Notice:

* Nt -a total size of the isolate studied

** F -average coefficient of inbreeding evaluated using analyses of traditional for population genetic studies three generations marriages structures among selected unrelated subjects (Fpop).

*** Fped -coefficient of inbreeding evaluated using extended pedigree reconstructed for every affected ID cases up to 12 generations.

Table 2

Genome-wide Multipoint linkage scan and candidate genes in linked regions in kindred with aggregation of MRT ascertained in genetic isolate

Map	LOD, D/M-R/M	Flanking loci, peak (cM)	Start-end (bp)	Genes
<i>1q41</i>	1.4, D/M	D1S2141- D1S549, 231	215,095,283-220,743,932	<i>KCNK2, USH2A, RAB3GAP2, MARK1</i>
<i>2p25.3-p24.2</i>	3.77, D/M	D2S2976-D2S2952, 21	835,510-11,707,195	<i>MYT1L, SNTG2, TPO, COLEC11, SOX11, KIDINS220, ASAP2</i>
<i>3p13-p12.1</i>	1.7, D/M	D3S2406-D3S4529, 90	73,158,376-85,952,736	<i>MITF, RYBP, ROBO1, ROBO2, FOXP1, GBE1</i>
<i>4q13</i>	1.5, D/M	D4S3243-D4S1647, 68	68,182,802-81,032,553	<i>GNRHR, HUNK, GC, SLC444, AFP</i>
<i>10p11.23-p11.21</i>	2.5, D/M	D10S1426-D10S1208, 32	30,435,660-35,397,920	<i>NRP1, EPC1, ARHGAP12, REM,</i>
<i>11q22.3-q23.1</i>	2.1, R/M	D11S1986-D11S1998, 107	111,123,472-117,798,030	<i>DRD2, HTR3A, HTR3B, FXYD6, PTS, ANKK1, ZBTB16, SLC35F2</i>
<i>12q24.22-q24.31</i>	3.87, R/M	D12S2070-D12S395, 120	114,466,941-126,627,408	<i>MED13L, HRK, TESC, FBXW8, SBNOL, CDK2AP1, CIT, NOS1, PLA2G1B, CCDC60,</i>
<i>17q24.2-q25.1</i>	1.6, D/M	D17S2193-D17S1301, 96	66,447,685-72,781,109	<i>HNI, SLC25A19, ACOX1, SRP68, KCNJ2, KCNJ16, SSTR2, KCTD2, SLC9A3R1, GRIN2C, ATP5H, SLC25A19, ACOX1, SRP68, EXOC7</i>
<i>21q22.13</i>	1.9, D/M	D21S1440-D21S2055, 36	39,041,451-41,291,662	<i>KCNJ6, CLDN14, PCP4, DSCR4, KCNJ15, ERG, WRB, HMGNI, DYRK1A</i>
<i>22q12.3-q13.1</i>	3.4, D/M	D22S683-D22S445, 32	36,413,691-37,666,244	<i>LARGE, APOL1-APOL4, CSNK1E, RBM9, MYH9, IL2RB, CACNG2, SOX10, RPSAP52, IL2RB, DDX17</i>

Notices: Genes associated with ID according previous studies, and found as linked with ID in our study, were bolded.

Table 3a

ROH segments in linked 12q24 regions with genes located

DGH #	Start (kb)	End (kb)	Length (kb)	Gene Name
DGH001	1157800	1159272	1472	MED13L , MIR620, MIR4472-2, NCRNA00173, MAP1LC3B2, C12orf49, RNFT2, HRK , FBXW8 , TESC , FBXO21 , NOS1, KSR2, RFC5, WSB2, VSIG10, PEBP1, TAOK3, SUDS3, SRRM4, HSPB8, LOC144742, CCDC60, TMEM233, PRKAB1, CIT, MIR1178, CCDC64, RAB35,
DGH002	1157800	1159272	1472	MED13L , MIR620, MIR4472-2, NCRNA00173, MAP1LC3B2, C12orf49, RNFT2, HRK , FBXW8 , TESC , FBXO21 , NOS1, KSR2, RFC5, WSB2, VSIG10, PEBP1, TAOK3, SUDS3, SRRM4, HSPB8, LOC144742, CCDC60, TMEM233, PRKAB1, CIT, MIR1178, CCDC64, RAB35,
DGH003	1157800	1159272	1472	MED13L , MIR620, MIR4472-2, NCRNA00173, MAP1LC3B2, C12orf49, RNFT2, HRK , FBXW8 , TESC , FBXO21 , NOS1, KSR2, RFC5, WSB2, VSIG10, PEBP1, TAOK3, SUDS3, SRRM4, HSPB8, LOC144742, CCDC60, TMEM233, PRKAB1, CIT, MIR1178, CCDC64, RAB35,
DGH004	123201	123822	621	HCAR3, HCARI, DENR, CCDC62, HIP1R, VPS37B, ABCB9, OGFOD2, ARL6IP4, PITPNM2, MIR4304, LOC100507091, C12orf65, CDK2AP1 , SBNO1
DGH004	116000	117387	1387	MED13L , MIR620, MIR4472-2, NCRNA00173, MAP1LC3B2
DGH005	123105	123822	717	KNTC1, HCAR2, HCAR3, HCARI, DENR, CCDC62, HIP1R, VPS37B, ABCB9, OGFOD2, ARL6IP4, PITPNM2, MIR4304, LOC100507091, C12orf65, CDK2AP1 , SBNO1
DGH005	117000	117500	500	C12orf49, RNFT2, HRK , FBXW8 , TESC , FBXO21 , NOS1, KSR2, RFC5, WSB2, VSIG10, PEBP1, TAOK3, SUDS3, SRRM4, HSPB8, LOC144742, CCDC60, TMEM233, PRKAB1, CIT, MIR1178, CCDC64, RAB35,
DGH006	125089	128647	3559	SCARBI, UBC, DHX37, BRI3BP, AACs, TMEM132B, LOC400084, LOC100128554, LOC387895, LOC100507206, LOC440117, FLJ37505
DGH007	122804	123790	986	CLIP1, ZCCHC8, RSR2, KNTC1, HCAR2, HCAR3, HCARI, DENR, CCDC62, HIP1R, VPS37B, ABCB9, OGFOD2, ARL6IP4, PITPNM2, MIR4304, LOC100507091, MPHOSPH9, C12orf65, CDK2AP1 , SBNO1
DGH008	122804	123790	986	CLIP1, ZCCHC8, RSR2, KNTC1, HCAR2, HCAR3, HCARI, DENR, CCDC62, HIP1R, VPS37B, ABCB9, OGFOD2, ARL6IP4, PITPNM2, MIR4304, LOC100507091, MPHOSPH9, C12orf65, CDK2AP1 , SBNO1
DGH009	122804	123790	986	CLIP1, ZCCHC8, RSR2, KNTC1, HCAR2, HCAR3, HCARI, DENR, CCDC62, HIP1R, VPS37B, ABCB9, OGFOD2, ARL6IP4, PITPNM2, MIR4304, LOC100507091, MPHOSPH9, C12orf65, CDK2AP1 , SBNO1
DGH010	122804	123790	986	CLIP1, ZCCHC8, RSR2, KNTC1, HCAR2, HCAR3, HCARI, DENR, CCDC62, HIP1R, VPS37B, ABCB9, OGFOD2, ARL6IP4, PITPNM2, MIR4304, LOC100507091, MPHOSPH9, C12orf65, CDK2AP1 , SBNO1
DGH011	123201	123822	621	HCAR3, HCARI, DENR, CCDC62, HIP1R, VPS37B, ABCB9, OGFOD2, ARL6IP4, PITPNM2, MIR4304, LOC100507091, MPHOSPH9, C12orf65, CDK2AP1 , SBNO1
DGH012	122804	123790	986	CLIP1, ZCCHC8, RSR2, KNTC1, HCAR2, HCAR3, HCARI, DENR, CCDC62, HIP1R, VPS37B, ABCB9, OGFOD2, ARL6IP4, PITPNM2, MIR4304, LOC100507091, MPHOSPH9, C12orf65, CDK2AP1 , SBNO1
DGH013	123105	123822	717	KNTC1, HCAR2, HCAR3, HCARI, DENR, CCDC62, HIP1R, VPS37B, ABCB9, OGFOD2, ARL6IP4, PITPNM2, MIR4304, LOC100507091, MPHOSPH9, C12orf65, CDK2AP1 , SBNO1
DGH014	123105	123743	639	KNTC1, HCAR2, HCAR3, HCARI, DENR, CCDC62, HIP1R, VPS37B, ABCB9, OGFOD2, ARL6IP4, PITPNM2, MIR4304, LOC100507091, MPHOSPH9, C12orf65, CDK2AP1 , SBNO1
DGH015	123105	123822	717	KNTC1, HCAR2, HCAR3, HCARI, DENR, CCDC62, HIP1R, VPS37B, ABCB9, OGFOD2, ARL6IP4, PITPNM2, MIR4304, LOC100507091, MPHOSPH9, C12orf65, CDK2AP1 , SBNO1
DGH016	123201	123822	621	HCAR3, HCARI, DENR, CCDC62, HIP1R, VPS37B, ABCB9, OGFOD2, ARL6IP4, PITPNM2, MIR4304, LOC100507091, MPHOSPH9, C12orf65, CDK2AP1 , SBNO1

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

DGH_#	Start (kb)	End (kb)	Length (kb)	Gene Name
DGH017	123201	123822	621	HCAR3, HCARI, DENR, CCDC62, HIP1R, VPS37B, ABCB9, OGFOD2, ARL6IP4, PITPNM2, MIR4304, LOC100507091, MPHOSPH9, C12orf65, CDK2API , SBN01

Notices: Gene-candidate associated with ID according previous publications are bolded

Table 3b

Copy number segments in linked 12q24.21-24.31 regions with genes located

DGH #	Position	End Position	Segment Mean	Length (kb)	Gene Name
DGH001	116396381	78410558	-0,16622	9416	MED13L, HRK, FBXW8, TESC, NOS1
DGH002	97801884	125931905	-0,04391	28130	MED13L, HRK, FBXW8, TESC, CDK2AP1, SBN01
DGH003	115688402	117344276	0,743	1656	MED13L, HRK
DGH004	116396381	78410558	-2,14765	9416	MED13L, HRK, FBXW8, TESC, NOS1
DGH007	116396381	77450413	-1,02624	8455	MED13L, HRK, FBXW8, TESC, NOS1
DGH010	117617646	124558791	-2,09709	6941	CDK2AP1, SBN01
DGH010	115688402	117344276	0,743	1656	MED13L, HRK
DGH012	116192880	124558791	0,509505	8366	MED13L, HRK, FBXW8, TESC, CDK2AP1, SBN01
DGH013	118227544	119376800	-2,8537	1149	CDK2AP1, SBN01
DGH015	118227544	119376800	-2,345	1149	CDK2AP1, SBN01
DGH016	123757861	124558791	-2,71676	801	SBN01
DGH017	116396381	122613000	-0,1891	53715	MED13L, HRK, FBXW8, TESC, NOS1

Table 4

ROH and CNV ascertained among MRT cases from genetic isolate in genomic region associated with adult intracranial volume.

Cyband	SNP ID	Effect allele	Gene	Flanking genes	Map	Authors	DGH LOD	DGH/ROH cases	DGH/CNV cases
1q23.3	rs10494373 *	C	DDR2	UAP1, HSD17B7(SCZ)	chr1:162,619,112-162,619,612	Stein et al, 2011		1	1 -del
2q24	rs6741949	-	DPP4	SLC4A10, LOC100131604	chr2:162,909,973-162,910,473	Bis et al, 2012		1	
6q22.32	rs4273712	G	intergenic	FABP7, NCOA7	chr6:126,964,260-126,964,760	Ikram et al, 2012		9	3 -del
	rs1042725	C	HMG2	RPSAP52, GRIP1	chr12:66,358,097-66,358,597	Taal et al, 2012		1	4 -del
12q14.3	rs17178006	-	MSRB3	LEM3, LOC100129644	chr12:65,718,049-65,718,549	Bis et al, 2012		1	
	rs10784502	C	HMG2	IRAK3, GRIP1	chr12:66,343,560-66,344,060	Stein et al, 2011		5	4 -del
12q24.22	rs7294919	C	intergenic	HRK, FBXW8, TESC	chr12:117,327,342-117,327,842	Stein et al, 2011		5	
	rs7294919	C	intergenic	HRK, FBXW8, TESC	chr12:117,327,342-117,327,842	Bis et al, 2012	MRT: Lod=3.87, R/M	5	11del/gain
12q24.31	rs7980687 *	A	SBNO1	CDK2AP1	chr12:123,822,461-123,822,961	Taal et al, 2012		22	8del/gain
	rs11655470	C	CRHR1	c17orf69	chr17:43,795,183-43,795,683	Taal et al, 2012		2	
17q21.31	rs9303525	G	KIAA1267	MART, ARL178	chr17:44,187,007-44,187,507	Ikram et al, 2012		2	11del/gain
	rs9915547 **		KIAA1267	MART, ARL178	chr17:44,212,532-44,213,032	Ikram et al, 2012		2	
								56	43

rs11655470 (17q21.31) showed suggestive evidence of association with head circumference ($P = 3.9 \times 10^{-6}$). SNPs correlated to the 17q21 signal have shown genome-wide association with adult intracranial volume, Parkinson's disease and other neurodegenerative diseases, indicating that a common genetic variant in this region might link early brain growth with neurological disease in later life.

* rs7980687 (12q24.31) & rs1042725 (12q14.3) were robustly associated with head circumference in infancy (Taal et al, 2012). Although these loci have previously been associated with adult height, their effects on infant head circumference were largely independent of height ($P = 3.8 \times 10^{-7}$) for rs7980687 and $P = 1.3 \times 10^{-7}$ for rs1042725 after adjustment for infant height).

** found association with Parkinson D (Tobin et al, 2008)