Implication of a Rare Deletion at Distal 16p11.2 in Schizophrenia

Saurav Guha, PhD; Elliott Rees, MRes; Ariel Darvasi, PhD; Dobril Ivanov, PhD; Masashi Ikeda, MD, PhD; Sarah E. Bergen, PhD; Patrik K. Magnusson, PhD; Paul Cormican, PhD; Derek Morris, PhD; Michael Gill, MRCPsych, PhD; Sven Cichon, PhD; Jeffrey A. Rosenfeld, PhD; Annette Lee, PhD; Peter K. Gregersen, PhD; John M. Kane, MD; Anil K. Malhotra, MD; Marcella Rietschel, MD; Markus M. Nöthen, MD; Franziska Degenhardt, PhD; Lutz Priebe, MSc; René Breuer, MSc; Jana Strohmaier, Dipl Psych; Douglas M. Ruderfer, MSc; Jennifer L. Moran, PhD; Kimberly D. Chambert, MSc; Alan R. Sanders, MD; Jianxin Shi, PhD; Kenneth Kendler, MD; Brien Riley, PhD; Tony O'Neill, MD; Dermot Walsh, MD; Dheeraj Malhotra, PhD; Aiden Corvin, MRCPsych, PhD; Shaun Purcell, PhD; Pamela Sklar, MD, PhD; Nakao Iwata, MD, PhD; Christina M. Hultman, PhD; Patrick F. Sullivan, MD, FRANZCP; Jonathan Sebat, PhD; Shane McCarthy, PhD; Pablo V. Gejman, MD; Douglas F. Levinson, MD; Michael J. Owen, PhD, FRCPsych; Michael C. O'Donovan, PhD, FRCPsych; Todd Lencz, PhD; George Kirov, PhD, MRCPsych; for the Molecular Genetics of Schizophrenia Consortium and the Wellcome Trust Case Control Consortium 2

Context: Large genomic copy number variations have been implicated as strong risk factors for schizophrenia. However, the rarity of these events has created challenges for the identification of further pathogenic loci, and extremely large samples are required to provide convincing replication.

Objective: To detect novel copy number variations that increase the susceptibility to schizophrenia by using 2 ethnically homogeneous discovery cohorts and replication in large samples.

Design: Genetic association study of microarray data.

Setting: Samples of DNA were collected at 9 sites from different countries.

Participants: Two discovery cohorts consisted of 790 cases with schizophrenia and schizoaffective disorder and 1347 controls of Ashkenazi Jewish descent and 662 parent-offspring trios from Bulgaria, of which the offspring had schizophrenia or schizoaffective disorder. Replication data sets consisted of 12 398 cases and 17 945 controls.

Main Outcome Measures: Statistically increased rate of specific copy number variations in cases vs controls.

Results: One novel locus was implicated: a deletion at distal 16p11.2, which does not overlap the proximal 16p11.2 locus previously reported in schizophrenia and autism. Deletions at this locus were found in 13 of 13 850 cases (0.094%) and 3 of 19 954 controls (0.015%) (odds ratio, 6.25 [95% CI, 1.78-21.93]; $P = .001$, Fisher exact test).

Conclusions: Deletions at distal 16p11.2 have been previously implicated in developmental delay and obesity. The region contains 9 genes, several of which are implicated in neurological diseases, regulation of body weight, and glucose homeostasis. A telomeric extension of the deletion, observed in about half the cases but no controls, potentially implicates an additional 8 genes. Our findings add a new locus to the list of copy number variations that increase the risk for development of schizophrenia.

JAMA Psychiatry. 2013;70(3):253-260.
Published online January 16, 2013.
doi:10.1001/2013.jamapsychiatry.71

Author Affiliations are listed at the end of this article.
Group Information: Pablo V. Gejman, MD, takes authorship responsibility for the Molecular Genetics of Schizophrenia Consortium, and Aiden Corvin, MRCPsych, PhD, takes authorship responsibility for the Wellcome Trust Case Control Consortium 2. Lists of the Molecular Genetics of Schizophrenia Consortium and Wellcome Trust Case Control Consortium 2 members appear in eMaterial 4 and eMaterial 5 at http://www.jamapsych.com.

Uncovering the genetic factors underlying schizophrenia (SZ) has proven difficult despite heritability estimates of up to 80%.[1] Copy number variations (CNVs) at several loci show consistently replicated evidence for association with SZ.[2,3] These CNVs are individually very rare, are not fully penetrant, and are found cumulatively in approximately 2% of SZ cases; therefore, large samples were required to establish their association. Given their low baseline frequency, further CNV susceptibility loci likely have yet to be discovered.

In the present study, we report the identification of a CNV locus at distal 16p11.2 that increases the risk for SZ. Findings pointing to a possible association between this locus and SZ were obtained independently by 2 teams of investigators. During the process of obtaining replication data, the 2 groups became aware of each other's work and decided to combine results from their discovery and replication cohorts. Using high-resolution microarrays, one group (from New York and Israel) examined an SZ case-control cohort from the Ashkenazi Jewish (AJ) population, whereas the other group

Table. CNVs at the Distal 16p11.2 Region in the Discovery and Replication Sets[a]

Study	Platform	No. of Subjects Undergoing Testing		No. With 16p11.2 Deletion		No. With 16p11.2 Duplication	
		Cases	Controls	Cases	Controls	Cases	Controls
Discovery samples							
Bulgarian trios	Affymetrix 6.0[b]	662	662	2	0	1	0
Ashkenazi Jewish subjects	HumanOmni1-quad[c]	790	1347	2	0	0	1
Replication							
GAIN/MGS EA	Affymetrix 6.0[b]	2671	2648	5	0	1	4
GAIN/MGS AA	Affymetrix 6.0[b]	1274	963	0	0	0	0
ISC	Affymetrix 6.0/5.0[b]	3045	3185	0	0	2	0
Vacic et al[7]	Array CGH HD2[d]	757	742	0	0	1	0
WTCCC2/Irish	Affymetrix 6.0[b]/Illumina 1m[c]	1269	6175	0	2	0	4
Japanese	Affymetrix 5.0[b]	490	516	1	0	0	0
Swedish	Affymetrix 6.0[b]	1506	2089	2	1	0	2
German	HH550, 610, 660[c]	1386	1627	1	0	1	2
Total Replication Samples, No. or No. (%)		**12 398**	**17 945**	**9** (0.073)	**3** (0.017)	**5** (0.040)	**12** (0.067)
Total (Discovery + Replication), No. or No. (%)		**13 850**	**19 954**	**13** (0.094)	**3** (0.015)	**6** (0.043)	**13** (0.065)

Abbreviations: AA, African Americans; EA, European Americans; CNVs, copy number variations; ISC, International Schizophrenia Consortium; GAIN/MGS, Genetic Association Information Network/Molecular Genetics of Schizophrenia Consortium; WTCCC2, Wellcome Trust Case Control Consortium 2.

[a] Details on all samples are presented in eMaterial 1. For significance for the rate of deletions in the replication sample by 1-tailed Fisher exact test, $P = .018$; by Cochran Mantel-Haenszel test, $P = .04$. For significance for the total sample (discovery + replication) by 2-tailed Fisher exact test, $P = .0014$; by Cochran Mantel-Haenszel test, $P = .0031$. Odds ratio for replication samples were 4.35 (95% CI, 1.18-16.06); for the total sample, 6.25 (95% CI, 1.78-21.93). The rate of duplications in the region is not significantly different between cases and controls.

[b] Available from the Broad Institute.

[c] Available from Illumina, Inc.

[d] Available from NimbleGen.

(from Cardiff, Wales) examined a cohort of parent-offspring trios from Bulgaria (BG). Because of the need for large-scale replication, we contacted research groups worldwide who were willing to share raw data from microarray-based CNV studies in cohorts of SZ cases and controls and obtained data from a total of approximately 34 000 individuals.

METHODS

BG TRIOS SAMPLE

Sample Description

The final sample (after quality control [QC]) consisted of 662 BG offspring with all their parents in 638 families (615 families with 1 offspring, 22 families with 2, and 1 family with 3). Details on this cohort have been previously described,[4] but that previous publication reported on de novo CNVs only; here we report on the transmitted CNVs in this cohort. This cohort does not include patients with severe developmental disorders (all probands had attended mainstream schools, from which people with known intellectual disabilities were excluded). Diagnoses were made according to *DSM-IV* criteria,[5] using an interview from the Schedules for Clinical Assessment in Neuropsychiatry[6] and a review of hospital discharge summaries. We included patients with SZ or schizoaffective disorder. Concomitant medical conditions were not systematically assessed except as related to psychiatric diagnosis. The CNVs found in the parents of each trio but not transmitted to the affected offspring constituted the pseudocontrol population listed in the **Table** under the control headings.

Genotyping and QC

All samples were genotyped on a single-nucleotide polymorphism array (Affymetrix 6.0; Broad Institute). Analysis was performed using software from the manufacturer (Affymetrix Genotyping Console 4.0 software; Broad Institute) one batch at a time, with each batch containing 70 to 90 arrays. The QC process included removal of CNVs if they were from the X or Y chromosome, were less than 15 kilobases (kb), were covered by fewer than 15 probes, or had a probe density (size/probe number) greater than 7500 base pairs. We used the PLINK tool, version 1.07,[8] to exclude CNVs if 50% or more of their length was covered by a segmental duplication (SD). Copy number variation loci with a frequency of greater than 1% were excluded. Individuals with multiple large duplicate CNVs on the same chromosome were excluded because these duplications are likely to be artifacts.[9] Samples were also removed if their total number of CNVs was very high and constituted an outlier for the distribution within that sample (>50 CNVs for this experiment).

For additional QC of the BG trios, we used a modification of the median *z*-score outlier detection algorithm proposed by McCarthy et al[10] in 2009 and described in detail by Kirov et al[4] in 2012. A *z* score is the median of the standardized $\log_2$ ratios for all probes within a specified chromosomal region. Through comparison of all individual *z* scores for a given region, true CNVs are represented as outliers from the normal distribution of the *z* scores. We show the distribution of the *z* scores for the 16p11.2 distal region in eFigure 1 (http://www.jamapsych.com), which demonstrates that the only outliers for this region are the 2 probands with deletions and their parents.

AJ SAMPLE

Sample Description

Case (n = 1156) and control (n = 2279) samples were selected from an AJ repository (Hebrew University Genetic Resource; http://hugr.huji.ac.il). Patients for the discovery analysis were recruited from hospitalized inpatients at 7 medical centers in Israel. All diagnoses were assigned after direct interview using the Structured Clinical Interview,[11] a questionnaire with inclusion and exclusion criteria, and cross-references to medical records. Chronic medical disorders and conditions were recorded on the basis of patient report and hospital records. The inclusion criteria specified that subjects had to be diagnosed as having SZ or schizoaffective disorder by *DSM-IV* criteria,[5] that all 4 grandparents of each subject were reported to be of AJ ethnic origin, and that each subject or the subject's legal representative had signed the informed consent form. Exclusion criteria included any psychotic disorder due to a general medical condition, substance-induced psychotic disorder, pervasive developmental disorders, or any cluster A (ie, schizotypal, schizoid, or paranoid) personality disorder. Samples from healthy AJ individuals were collected from volunteers at the Israeli Blood Bank; these subjects did not undergo psychiatric screening but reported no chronic disease and were taking no medication at the time of the blood draw. Corresponding institutional review boards and the National Genetic Committee of the Israeli Ministry of Health approved the studies. All samples were fully anonymized immediately after collection, and genomic DNA was extracted subsequently from blood samples through the use of a commercially available kit (Nucleon; Pharmacia). Genotyping and analyses were performed under protocols approved by the institutional review board of the North Shore–Long Island Jewish Health System.

Genotyping and QC

Genotyping was performed with commercially available arrays (HumanOmni1-Quad; Illumina, Inc) according to the manufacturer's specifications for approximately 1.4 million genome-wide markers (approximately 900 000 single-nucleotide polymorphisms [SNPs] and 500 000 CNV intensity probes). The SNPs were filtered on the following basis: call rate of less than 98%, minor allele frequency of less than 0.02, and Hardy-Weinberg exact test results of $P < .000001$ in controls. Samples were filtered on the basis of genotype QC filtration (sample call rate, <97%, sex mismatch) and examined for cryptic identity and first- or second-degree relatedness using the pairwise identity-by-descent estimation in the PLINK tool set[8] with 128 403 linkage disequilibrium pruned ($r^2 > 0.2$) genome-wide SNPs. Samples were excluded based on a pairwise identity-by-descent estimation of greater than 0.125; the individual with the lower call rate from each control-control or case-case pair was excluded, and controls were excluded from case-control pairs. The remaining samples were further examined for underlying population stratification using principal component analysis with ancestry informative markers specific for the AJ population.[12] Samples with principal component analysis results suggestive of 1 or more non-AJ grandparents were identified as outliers on the basis of a first principal component score greater than 0.01 and were excluded from further analysis (eFigure 2). After QC based on SNP markers, the data set contained 2544 samples consisting of 904 cases (573 male and 331 female) and 1640 controls (1216 male and 424 female) who underwent genotyping on 762 372 high-quality SNPs with a 99.8% overall call rate.

CNV Calls and Validation

Normalization and log ratio data calculation for 904 cases and 1640 controls were performed using the manufacturer's software platform (GenomeStudio; Illumina, Inc). The resulting log R ratios and B-allele frequencies were used to identify CNVs on autosomes for each subject. We used variations of 3 algorithms for CNV detection: PennCNV,[9] QuantiSNP,[13] and cnvPartition (http://www.illumina.com). QuantiSNP and PennCNV are based on the hidden Markov model, and cnvPartition is based on bivariate Gaussian distribution as implemented on the manufacturer's software platform (http://www.illumina.com).

Following the methods of Need et al[14] and Sanders et al,[15] we excluded any individuals with a PennCNV threshold of log R standard deviation of at least 0.30, B-allele frequency drift of at least 0.002, and/or waviness factor deviating from 0 by greater than 0.04. Individuals containing more than 500 CNVs (before filtration described in the following paragraphs) were also excluded from the analysis. The final data set contains 790 cases and 1347 controls.

We further excluded CNV calls on the basis of the QC thresholds recommended by each of the respective algorithms. Thus, CNV calls were excluded from further analysis if the log Bayes factor was no greater than 10 in QuantiSNP, if the confidence threshold was no greater than 35 in cnvPartition, or if default QC parameters in PennCNV were not obtained.

After the QC steps, all the CNV calls were merged using the CNVision program.[15] The final rare CNV calls were based on consensus calls from all 3 algorithms (with no more than 25% of the length drawn from 1 algorithm only), with the following filtration criteria: at least 20 probes, at least 100 kb in size, and less than 1% frequency in the total sample. Copy number variations of the same type (ie, deletion or duplication) that were separated by at least 3 probes were merged into 1 contiguous segment as recommended by Vacic et al.[7] All CNVs were annotated using the CNVision program. Based on previous findings in SZ and other neuropsychiatric disorders,[16] purely intergenic CNVs were excluded.

REPLICATION SAMPLES

Evidence for replication of the findings was obtained from 8 case-control samples recruited and genotyped by other teams from the United States, Europe, and Japan. These samples consisted of 12 398 cases and 17 945 controls who underwent genotyping with high-resolution arrays. Details on the samples, genotyping platforms, and QC used by these teams are specified in eMaterial 1. The minimally affected region was covered well by all arrays used by the other teams (eFigure 3).

All coordinates in this study are based on the Human Genome Build NCBI36/hg18.

RESULTS

After stringent QC procedures, 790 cases and 1347 controls from the AJ cohort and 662 probands from 638 BG families were examined for rare large CNVs. Replication was sought in other case-control data sets for any CNVs that were observed in at least 2 cases and no controls (for the AJ cohort) and transmitted at least twice with no non-transmissions in the BG trios cohort. Several relevant CNVs were found in the 2 discovery data sets at loci already reported to increase the risk for development of SZ; because they are already known susceptibility factors, however, we only list them in eTable 4. In the AJ cohort, CNVs

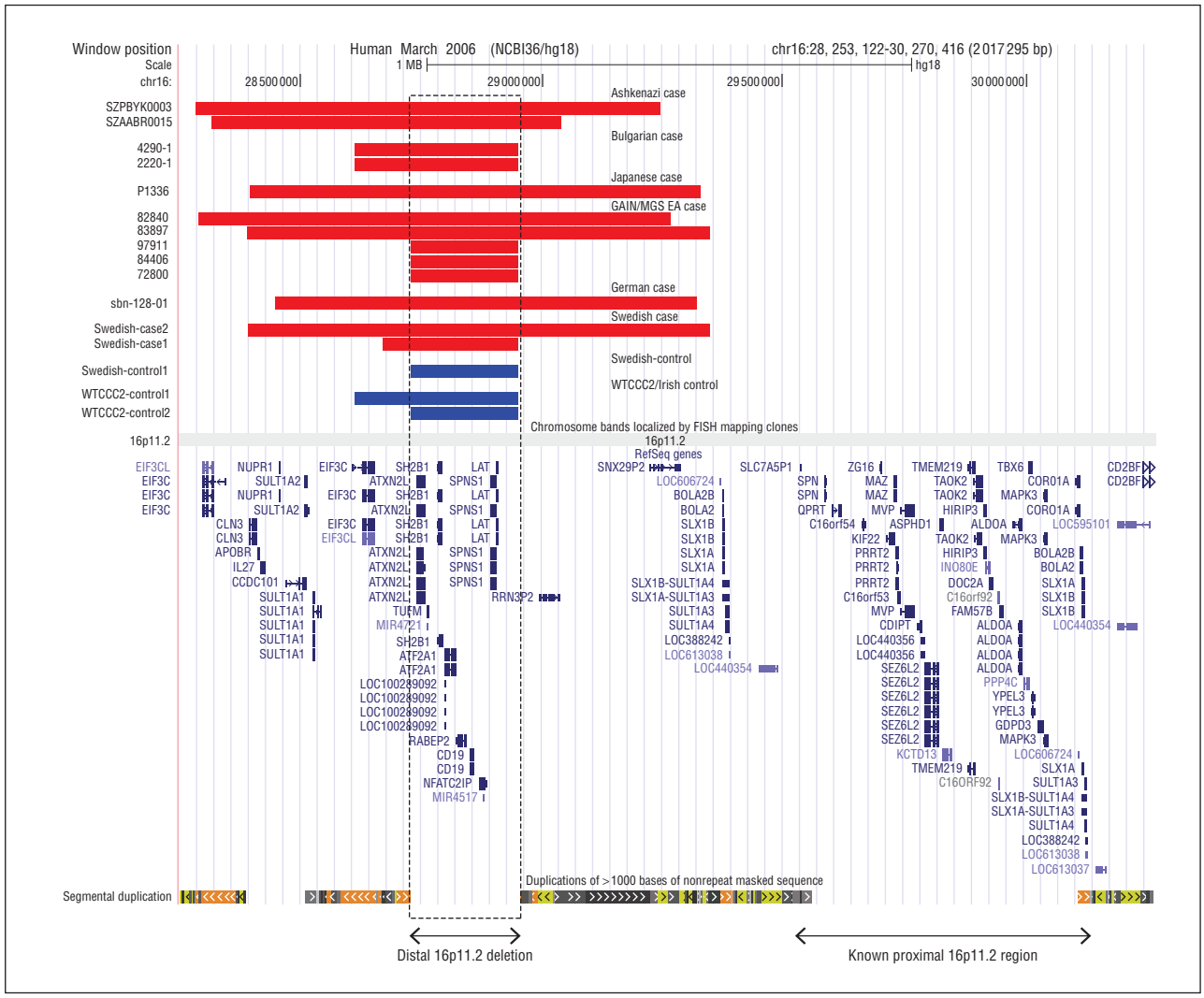

Figure. Microdeletions at the distal 16p11.2 region in the present study. The region intersects 9 genes and is flanked by 2 blocks of segmental duplications (SDs). Red lines represent deletions in schizophrenia cases and blue lines represent deletions in control subjects. The double-headed arrows indicate the intervals implicated in the present study (distal 16p11.2 deletion; minimal common region, 28.73-28.95 megabases [Mb], hg18) and the known 16p11.2 copy number variation locus (known proximal 16p11.2 region; minimal common region, 29.56-30.11 Mb). Bp indicates base pair; chr, chromosome; EA, European American; FISH, fluorescent in situ hybridization; GAIN/MGS, Genetic Association Information Network/Molecular Genetics of Schizophrenia Consortium; hg, human genome; RefSeq, Reference Sequence collection; and WTCCC2, Wellcome Trust Case Control Consortium 2.

at 2 additional loci were observed in 2 cases and no controls. These loci were at chromosome 6q14.3 (hg18 coordinates: 85.25-85.58 megabases [Mb]) and 7q33 (133.39-133.50 Mb), but replication evidence was not observed. No other CNV of this type was supported by replication evidence in the BG data (apart from the 16p11.2 deletion). The lists of all rare and large (>100 kb) CNVs in the 2 samples that intersected genes are available as files on request from the authors (AJ SZ CNVs >100 kb.xls and BG SZ trio CNVs >100 kb.xls).

The only CNV of interest that overlapped between the 2 discovery samples was a deletion at the distal region of 16p11.2, with a minimal common region extending from 28.73 to 28.95 Mb (build 36, hg18). This region intersects 9 genes and is flanked by 2 SD blocks (**Figure**). The region does not overlap the known 16p11.2 locus at 29.56 to 30.11 Mb that has been implicated in SZ,[10,16] autism,[15,17] and developmental delay.[18] Deletions at this locus were found in 2 cases (and no controls) from the AJ cohort and 2 offspring from the BG samples, both transmitted from mothers (parents transmitted this CNV in every trio). Duplications at this locus were observed in 1 AJ control and in 1 BG parent (who transmitted it to an affected offspring).

We sought evidence for association between this deletion with SZ in 8 independent case-control cohorts (12 398 cases and 17 945 controls) for which we had access to the raw data (Table and eMaterial 1). Deletions overlapping this region were observed in an additional 9 cases and 3 controls (P = .018, 1-tailed Fisher exact test for the replication sample; odds ratio, 4.35 [95% CI, 1.18-16.06]). Combining the discovery and replication cohorts, we found 13 deletions among 13 850 cases (0.094%) and 3 among 19 954 controls (0.015%) (P = .0014, 2-tailed Fisher exact test; odds ratio, 6.25 [95% CI, 1.78-21.93]). The positions of the CNVs are shown in the Figure. No excess of duplications in cases was found at distal 16p11.2.

The minimal common region for all deletions reported in the Table encompasses 9 genes within a 220-kb interval flanked by blocks of SD (Figure). Some CNVs extend over the SDs (however, we noted that no CNV in the 16p11.2 region was excluded on the basis of >50% overlap with an SD). Different break points that extend over the flanking SD regions (but do not reach the telomeric region that is free of SDs) are more likely to reflect the different coverage of arrays (eMaterial 2 and eFigure 3) and/or the problems of calling CNVs over repetitive regions rather than different pathogenicity, especially because these regions have fewer genes. Seven deletions cover an additional region of unique DNA sequence at the telomeric side (Figure, the interval free of SDs) that contains further genes. Evidence of pathogenicity of the 7 CNVs that extended over the telomeric region was nearly as strong as for the implicated critical region (7 of 13 850 cases vs 0 of 19 954 controls, $P = .0019$, 2-tailed Fisher exact test). However, the critical distal 16p11.2 region remains the more likely candidate owing to its confirmed involvement in other neurodevelopmental disorders discussed in the "Comment" section and the lack of isolated CNVs in the smaller telomeric region. Of the 3 controls with deletions, 1 (in the Swedish data set) was recruited at the age of 45 years and had type 2 diabetes mellitus and high blood pressure but no other medical or psychiatric problems. No further information is available on the 2 anonymized controls from the Wellcome Trust Case Control Consortium 2/Irish data set; one is from the British Blood Transfusion service (therefore presumably healthy), and the other is from the 1958 cohort.

The new distal locus is approximately 600 kb telomeric from the previously implicated proximal 16p11.2 CNV (29.56-30.11 Mb).[10] The CNVs at proximal 16p11.2 have been shown to increase risk for SZ, autism, and developmental delay when duplicated[5,16] and for autism and developmental delay when deleted.[10,15,18] None of the CNVs in our study extend over the proximal region (Figure).

We have previously demonstrated that the known SZ-associated CNVs have high mutation rates and that strong selection pressure operates against them.[19] We can estimate the de novo rate for this deletion at 25% based on the present study (2 transmitted deletions and no information on inheritance in the other subjects) and 4 available data sets with a total of 5 de novo occurrences of 20 events with a known inheritance[20-23] (eMaterial 3 and eTable 3). This rate approximates to a selection pressure of 0.25. In line with this, we observe the 2 BG proband deletions to be found on different haplotypes and therefore very likely to be independent mutations.

Phenotypic data, where available, indicate a spectrum of typical presentations of SZ with no evidence of intellectual disability or a specific clinical profile (eTable 2). This indication is similar to the lack of specific clinical presentations reported for the other large CNVs implicated in neurodevelopmental disorders.[2,15,20] The possible exception is the presence of 2 individuals with obesity and 2 with type 2 diabetes mellitus (plus 1 control with type 2 diabetes mellitus) in line with previous reports. In addition to the 13 cases listed in the Table and eTable 2, the brother of 1 case (in the Japanese sample) carries the same deletion and also has SZ. Further probands had positive family histories of SZ, but we do not know whether their affected relatives also carry the deletion. Although the transmission status of the CNVs is only available for the BG cohort, we further note that both deletions were transmitted maternally.

COMMENT

Several lines of evidence from the literature support the distal 16p11.2 deletion as a true SZ-associated CNV locus. The deletion has been implicated in developmental delay and other clinical phenotypes[18,20,21] (details are given in eTable 1), similar to other SZ-associated CNVs.[2,3,10] Briefly, Cooper et al[18] reported a very similar increased rate of 0.1% (15 of 15 767) for this deletion in children with intellectual disability, autism spectrum disorders, and congenital malformations who were referred for genetic testing compared with a control rate of 0.01% (1 of 8329). Similar rates were found in another large study on patients with developmental delay and a range of other abnormal phenotypes[20]: 31 of 23 084 cases (0.13%) and 1 of 7700 controls (0.01%). Of the 6 cases in that study for whom detailed clinical information was available, one had autism, behavioral problems, attention-deficit/hyperactivity disorder, and SZ; another had behavioral problems, attention-deficit/hyperactivity disorder, and bipolar disorder; and a third had autism. Four of these 6 cases were overweight and all 6 had developmental delay. Moreover, additional telomeric extension of the deletion (to approximately 28.4 Mb) was present in 9 of the 31 cases and was never observed in the controls. Similarly, in our study, 7 of the 13 cases demonstrated this telomeric extension, which was not seen in the controls. The controls used in these studies partially overlap ours, so these control rates are not independent (eTable 1). Additional published reports of distal 16p11.2 deletions include 5 patients from 2 separate families,[22] all of whom have developmental delay and behavioral problems, and 1 child of 4284 patients with mental retardation.[23]

Distal 16p11.2 deletions have also been shown to be enriched in patients with severe early-onset obesity (3 of 300 [1%]) compared with unscreened population controls (2 of 7366 [0.03%]),[21] consistent with the findings in the study by Bachmann-Gagescu et al.[20] The most likely obesity candidate within the distal 16p11.2 region was postulated to be *SH2B1* because this gene plays a role in the regulation of body weight and glucose homeostasis in mice.[24] Two of our cases were obese or overweight, and 2 cases and 1 control had type 2 diabetes mellitus (consistent with being overweight, although this information is not available). However, 1 carrier (from Japan) had documented evidence of normal weight, and several did not have recorded evidence of obesity despite being drawn from cohorts that were assessed for this and other medically relevant phenotypes.

Considerable heterogeneity of phenotypic expression has been reported for most large rare CNVs implicated in SZ, with carriers often manifesting nonpsychotic phenotypes, including intellectual disability, autism, epilepsy, obesity, and cardiac disorders.[2,16] Plei-

otropy appears to also be the case for distal 16p11.2 deletions, possibly owing to the presence of multiple genes within the deleted region.

Clinical presentations for distal 16p11.2 deletion carriers are unremarkable for SZ, with diagnoses ranging across all major subtypes: paranoid, catatonic, undifferentiated, and schizoaffective. The age at SZ onset for deletion carriers ranges from 15 to 30 (mean age, 23.4) years, with no clear evidence of early onset. Two parents who transmitted the deletions to probands did not have psychotic disorders, although one had a mood disorder. Of the 3 controls who carry the deletion, one (from Sweden) did not report psychiatric problems at age 45 years, which is past the usual accepted age for the period of risk for SZ. Of the other 2 controls, one had also passed through the risk period (from the 1958 cohort, examined at age 44-45 years [eMaterial 1]) and the third one, a blood donor, is presumably healthy and not receiving any medication. These observations indicate that this CNV does not have full penetrance, similar to most other CNVs implicated in SZ. None of the carriers had any other SZ-associated CNVs.

Mutations in several of the 9 genes within the critical region of distal 16p11.2 have been implicated in neurological diseases. Homozygous mutations in the gene *TUFM* have been described in infants with fatal encephalopathy[25]; *ATP2A1* is implicated in Brody disease, in which patients are unable to relax their muscles during exercise,[26] and its homologue, *ATP2A2*, has been implicated in neuropsychiatric phenotypes.[27] In addition, *ATXN2L* (although unknown in function) encodes a protein belonging to the spinocerebellar ataxia family. The remaining genes are involved in immunity or in insulin and leptin signaling (*SH2B1*) or are of unknown function. In addition to the 9 genes in the minimal critical region, the larger CNVs with telomeric extensions include 8 additional deleted genes (7 of them in a DNA region that is free of SDs [Figure]), possibly increasing the pathogenicity of these larger CNVs. Most notable among these 8 genes is *CLN3*, for which recessive mutations are associated with Batten disease, characterized by childhood-onset neurodegeneration.[28] Moreover, *CLN3* is the only gene in the minimal or the extended region that is implicated in synaptic function on the basis of Gene Ontology annotation. Our previous study of de novo CNVs indicated an enrichment of such genes in SZ-related events; however, *CLN3* is not among the postsynaptic density genes implicated in that study.[4] Additional evidence from animal knockout models may help disentangle the contributions of each of these genes to the observed range of phenotypes.

In conclusion, we have obtained strong evidence of the role of a new CNV locus in SZ. Similar to other such loci, it is very rare and increases the risk for other neurodevelopmental phenotypes.

Submitted for Publication: May 31, 2012; accepted July 19, 2012.
Published Online: January 16, 2013. doi:10.1001/2013.jamapsychiatry.71
Author Affiliations: Division of Psychiatry Research, The Zucker Hillside Hospital, Glen Oaks, New York (Drs Guha, Kane, A. K. Malhotra, and Lencz); Medical Research Council Centre for Neuropsychiatric Genetics and Genomics and Neuroscience and Mental Health Research Institute, Cardiff University, Cardiff, Wales (Mr Rees and Drs Ivanov, Owen, O'Donovan, and Kirov); Department of Genetics, The Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel (Dr Darvasi); Fujita Health University School of Medicine, Toyoake, Japan (Drs Ikeda and Iwata); Psychiatric and Neurodevelopmental Genetics Unit, Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School (Drs Bergen and Purcell), and Center for Human Genetics Research, Massachusetts General Hospital (Drs Bergen and Purcell), Boston, Massachusetts; Stanley Center for Psychiatric Research, Broad Institute, Cambridge, Massachusetts (Drs Bergen, Moran, and Purcell and Ms Chambert); Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden (Drs Magnusson and Hultman); Department of Psychiatry, Trinity College Dublin (Drs Cormican, Morris, Gill, and Corvin), and Health Research Board (Dr Walsh), Dublin, Ireland; Institute of Human Genetics (Drs Cichon, Nöthen, and Degenhardt and Mr Priebe) and Department of Genomics, Life and Brain Center (Drs Cichon and Degenhardt and Mr Priebe), University of Bonn, and German Center for Neurodegenerative Diseases (Dr Nöthen), Bonn, Germany; Institute of Neuroscience and Medicine, Research Center Juelich, Juelich, Germany (Dr Cichon); IST/High Performance and Research Computing, University of Medicine and Dentistry of New Jersey, Newark (Dr Rosenfeld); Robert S. Boas Center for Human Genetics and Genomics (Drs Lee and Gregersen) and Center for Psychiatric Neuroscience (Drs Kane, A. K. Malhotra, and Lencz), The Feinstein Institute for Medical Research, Manhasset, New York; Department of Psychiatry, Hofstra University School of Medicine, Long Island, New York (Drs Kane, A. K. Malhotra, and Lencz); Central Institute of Mental Health, Medical Faculty Mannheim/Heidelberg University, Mannheim, Germany (Dr Rietschel, Mr Breuer, and Ms Strohmaier); Division of Psychiatric Genomics, Department of Psychiatry, Mount Sinai School of Medicine, New York, New York (Mr Ruderfer and Drs Purcell and Sklar); Department of Psychiatry and Behavioral Sciences, NorthShore University HealthSystem Research Institute, Evanston, Illinois (Drs Sanders and Gejman); Department of Psychiatry and Behavioral Sciences, University of Chicago, Chicago, Illinois (Drs Sanders and Gejman); National Cancer Institute, Bethesda, Maryland (Dr Shi); Virginia Institute for Psychiatry and Behavioral Genetics (Dr Kendler) and Department of Psychiatry (Dr Riley), School of Medicine, Virginia Commonwealth University, Richmond; Department of Psychiatry, Queens University, Belfast, Ireland (Dr O'Neill); Beyster Center for Genomics of Psychiatric Diseases, University of California, San Diego, La Jolla (Drs D. Malhotra and Sebat); Departments of Genetics, Psychiatry, and Epidemiology, University of North Carolina at Chapel Hill (Dr Sullivan); Stanley Institute for Cognitive Genomics, Woodbury Genome Center, Cold Spring Harbor Laboratory, Cold Spring Harbor, Cold Spring, New York (Mr McCarthy); and Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, California (Dr Levinson).
Correspondence: Todd Lencz, PhD, Division of Psychiatry Research, The Zucker Hillside Hospital, 75-59 263rd

St, Glen Oaks, NY 11004 (tlencz@NSHS.edu); George Kirov, PhD, MRCPsych, Medical Research Council Centre for Neuropsychiatric Genetics and Genomics and Neuroscience and Mental Health Research Institute, Cardiff University, Heath Park, Cardiff CF4 4XN, Wales (Kirov @cardiff.ac.uk).

Author Contributions: Dr Guha and Mr Rees contributed equally to the research. Drs Lencz and Kirov contributed equally to the research. Drs Guha, Lencz, and Kirov and Mr Rees take full responsibility for the integrity of the data.

Conflict of Interest Disclosures: None reported.

Funding/Support: Work at Cardiff University was supported by grant G0800509 from the Medical Research Council Programme and Centre Grants, grant 2 P50 MH066392-05A1 from the National Institute of Mental Health (NIMH), and grant EU-GEI from the European Union. Genotyping was funded by multiple grants to the Stanley Center for Psychiatric Research at the Broad Institute from the Stanley Medical Research Institute, the Merck Genome Research Foundation, and the Herman Foundation. The Zucker Hillside Hospital work was supported by the North Shore–Long Island Jewish Health System Foundation and grants RC2 MH089964 (Dr Lencz), R01 MH084098 (Dr Lencz), P50 MH080173 (Dr A. K. Malhotra), and P30 MH090590 (Dr Kane) from the National Institutes of Health. Molecular Genetics of Schizophrenia Consortium data collection was supported by R01 grants MH67257, MH59588, MH59571, MH59565, MH59587, MH60870, MH59566, MH59586, MH61675, MH60879, and MH81800 and U01 grants MH46276, MH46289, MH46318, MH79469, and MH79470 from the NIMH. We acknowledge support through Young Investigator Awards (Dr Sanders and Jubao Duan, PhD) from the National Alliance for Research on Schizophrenia and Depression and from the Genetic Association Information Network, the Paul Michael Donovan Charitable Foundation, and the Walter E. Nichols, MD, and Eleanor Nichols endowments (Stanford University). Genotyping performed by the Center for Genotyping and Analysis at the Broad Institute of Harvard and Massachusetts Institute of Technology was supported by grant U54 RR020278 from the National Center for Research Resources. The German study was supported by grants 01GS08144 and 01GS08147 from the German Federal Ministry of Education and Research (BMBF), within the context of the Integrated Genome Research Network MooDS (Systematic Investigation of the Molecular Causes of Major Mood Disorders and Schizophrenia). Control data were provided by the community-based studies of PopGen (Population-based Recruitment of Patients and Controls for the Analysis of Complex Genotype-Phenotype Relationships), Cooperative Health Research in the Region Augsburg, the Heinz Nixdorf Recall study, and MooDS. Genotyping of Wellcome Trust Case Control Consortium 2 samples was funded by grants 085475/B/08/Z and 085475/Z/08/Z from the Wellcome Trust. The British 1958 Birth Cohort DNA collection was funded by grant G0000934 from the UK Medical Research Council and grant 068545/Z/02 from the Wellcome Trust, and the UK National Blood Service controls by the Wellcome Trust. The Japanese cohort was supported in part by research grants from the Japan Ministry of Education, Culture, Sports, Science and Technology; the Ministry of Health, Labor and Welfare; the Core Research for Evolutional Science and Technology; and the Health Sciences Foundation (Research on Health Sciences focusing on Drug Innovation).

Role of the Sponsors: The sponsors had no role in the design and conduct of the study; in the collection, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript.

Online-Only Material: The eMaterial, eTables, and eFigures are available at http://www.jamapsych.com.

REFERENCES

1. Owen MJ, Williams NM, O'Donovan MC. The molecular genetics of schizophrenia: new findings promise new insights. *Mol Psychiatry*. 2004;9(1):14-27.
2. Kirov G. The role of copy number variation in schizophrenia. *Expert Rev Neurother*. 2010;10(1):25-32.
3. Levinson DF, Duan J, Oh S, Wang K, Sanders AR, Shi J, Zhang N, Mowry BJ, Olincy A, Amin F, Cloninger CR, Silverman JM, Buccola NG, Byerley WF, Black DW, Kendler KS, Freedman R, Dudbridge F, Pe'er I, Hakonarson H, Bergen SE, Fanous AH, Holmans PA, Gejman PV. Copy number variants in schizophrenia: confirmation of five previous findings and new evidence for 3q29 microdeletions and VIPR2 duplications. *Am J Psychiatry*. 2011;168(3):302-316.
4. Kirov G, Pocklington AJ, Holmans P, Ivanov D, Ikeda M, Ruderfer D, Moran J, Chambert K, Toncheva D, Georgieva L, Grozeva D, Fjodorova M, Wollerton R, Rees E, Nikolov I, van de Lagemaat LN, Bayés A, Fernandez E, Olason PI, Böttcher Y, Komiyama NH, Collins MO, Choudhary J, Stefansson K, Stefansson H, Grant SG, Purcell S, Sklar P, O'Donovan MC, Owen MJ. De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Mol Psychiatry*. 2012;17(2):142-153.
5. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. ed 4. Washington, DC: American Psychiatric Association; 1994.
6. Wing JK, Babor T, Brugha T, Burke J, Cooper JE, Giel R, Jablenski A, Regier D, Sartorius N; Schedules for Clinical Assessment in Neuropsychiatry. SCAN: Schedules for Clinical Assessment in Neuropsychiatry. *Arch Gen Psychiatry*. 1990;47(6):589-593.
7. Vacic V, McCarthy S, Malhotra D, Murray F, Chou HH, Peoples A, Makarov V, Yoon S, Bhandari A, Corominas R, Iakoucheva LM, Krastoshevsky O, Krause V, Larach-Walters V, Welsh DK, Craig D, Kelsoe JR, Gershon ES, Leal SM, Dell Aquila M, Morris DW, Gill M, Corvin A, Insel PA, McClellan J, King MC, Karayiorgou M, Levy DL, DeLisi LE, Sebat J. Duplications of the neuropeptide receptor gene *VIPR2* confer significant risk for schizophrenia. *Nature*. 2011;471(7339):499-503.
8. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-575.
9. Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF, Hakonarson H, Bucan M. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res*. 2007;17(11):1665-1674.
10. McCarthy SE, Makarov V, Kirov G, Addington AM, McClellan J, Yoon S, Perkins DO, Dickel DE, Kusenda M, Krastoshevsky O, Krause V, Kumar RA, Grozeva D, Malhotra D, Walsh T, Zackai EH, Kaplan P, Ganesh J, Krantz ID, Spinner NB, Roccanova P, Bhandari A, Pavon K, Lakshmi B, Leotta A, Kendall J, Lee YH, Vacic V, Gary S, Iakoucheva LM, Crow TJ, Christian SL, Lieberman JA, Stroup TS, Lehtimäki T, Puura K, Haldeman-Englert C, Pearl J, Goodell M, Willour VL, Derosse P, Steele J, Kassem L, Wolff J, Chitkara N, McMahon FJ, Malhotra AK, Potash JB, Schulze TG, Nöthen MM, Cichon S, Rietschel M, Leibenluft E, Kustanovich V, Lajonchere CM, Sutcliffe JS, Skuse D, Gill M, Gallagher L, Mendell NR, Craddock N, Owen MJ, O'Donovan MC, Shaikh TH, Susser E, Delisi LE, Sullivan PF, Deutsch CK, Rapoport J, Levy DL, King MC, Sebat J; Wellcome Trust Case Control Consortium. Microduplications of 16p11.2 are associated with schizophrenia. *Nat Genet*. 2009;41(11):1223-1227.
11. Spitzer RL, Williams JBW, Gibbon M, First MB. The Structured Clinical Interview for *DSM-III-R* (SCID), I: history, rationale, and description. *Arch Gen Psychiatry*. 1992;49(8):624-629.
12. Guha S, Rosenfeld JA, Malhotra AK, Lee AT, Gregersen PK, Kane JM, Pe'er I, Darvasi A, Lencz T. Implications for health and disease in the genetic signature of the Ashkenazi Jewish population. *Genome Biol*. 2012;13(1):R2 doi:10.1186/gb-2012-13-1-r2.

13. Colella S, Yau C, Taylor JM, Mirza G, Butler H, Clouston P, Bassett AS, Seller A, Holmes CC, Ragoussis J. QuantiSNP: an objective Bayes hidden-Markov model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Res*. 2007;35(6):2013-2025.
14. Need AC, Ge D, Weale ME, Maia J, Feng S, Heinzen EL, Shianna KV, Yoon W, Kasperaviciūte D, Gennarelli M, Strittmatter WJ, Bonvicini C, Rossi G, Jayathilake K, Cola PA, McEvoy JP, Keefe RS, Fisher EM, St Jean PL, Giegling I, Hartmann AM, Möller HJ, Ruppert A, Fraser G, Crombie C, Middleton LT, St Clair D, Roses AD, Muglia P, Francks C, Rujescu D, Meltzer HY, Goldstein DB. A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genet*. 2009;5(2):e1000373. doi:10.1371/journal.pgen.1000373.
15. Sanders SJ, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D, Chu SH, Moreau MP, Gupta AR, Thomson SA, Mason CE, Bilguvar K, Celestino-Soper PB, Choi M, Crawford EL, Davis L, Wright NR, Dhodapkar RM, DiCola M, DiLullo NM, Fernandez TV, Fielding-Singh V, Fishman DO, Frahm S, Garagaloyan R, Goh GS, Kammela S, Klei L, Lowe JK, Lund SC, McGrew AD, Meyer KA, Moffat WJ, Murdoch JD, O'Roak BJ, Ober GT, Pottenger RS, Raubeson MJ, Song Y, Wang Q, Yaspan BL, Yu TW, Yurkiewicz IR, Beaudet AL, Cantor RM, Curland M, Grice DE, Günel M, Lifton RP, Mane SM, Martin DM, Shaw CA, Sheldon M, Tischfield JA, Walsh CA, Morrow EM, Ledbetter DH, Fombonne E, Lord C, Martin CL, Brooks AI, Sutcliffe JS, Cook EH Jr, Geschwind D, Roeder K, Devlin B, State MW. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron*. 2011;70(5):863-885.
16. Malhotra D, Sebat J. CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell*. 2012;148(6):1223-1241.
17. Levy D, Ronemus M, Yamrom B, Lee YH, Leotta A, Kendall J, Marks S, Lakshmi B, Pai D, Ye K, Buja A, Krieger A, Yoon S, Troge J, Rodgers L, Iossifov I, Wigler M. Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron*. 2011;70(5):886-897.
18. Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu TH, Baker C, Williams C, Stalker H, Hamid R, Hannig V, Abdel-Hamid H, Bader P, McCracken E, Niyazov D, Leppig K, Thiese H, Hummel M, Alexander N, Gorski J, Kussmann J, Shashi V, Johnson K, Rehder C, Ballif BC, Shaffer LG, Eichler EE. A copy number variation morbidity map of developmental delay. *Nat Genet*. 2011;43(9):838-846.
19. Rees E, Moskvina V, Owen MJ, O'Donovan MC, Kirov G. De novo rates and selection of schizophrenia-associated copy number variants. *Biol Psychiatry*. 2011;70(12):1109-1114.
20. Bachmann-Gagescu R, Mefford HC, Cowan C, Glew GM, Hing AV, Wallace S, Bader PI, Hamati A, Reitnauer PJ, Smith R, Stockton DW, Muhle H, Helbig I, Eichler EE, Ballif BC, Rosenfeld J, Tsuchiya KD. Recurrent 200-kb deletions of 16p11.2 that include the *SH2B1* gene are associated with developmental delay and obesity. *Genet Med*. 2010;12(10):641-647.
21. Bochukova EG, Huang N, Keogh J, Henning E, Purmann C, Blaszczyk K, Saeed S, Hamilton-Shield J, Clayton-Smith J, O'Rahilly S, Hurles ME, Farooqi IS. Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature*. 2010;463(7281):666-670.
22. Barge-Schaapveld DQ, Maas SM, Polstra A, Knegt LC, Hennekam RC. The atypical 16p11.2 deletion: a not so atypical microdeletion syndrome? *Am J Med Genet A*. 2011;155A(5):1066-1072.
23. Bijlsma EK, Gijsbers AC, Schuurs-Hoeijmakers JH, van Haeringen A, Fransen van de Putte DE, Anderlid BM, Lundin J, Lapunzina P, Pérez Jurado LA, Delle Chiaie B, Loeys B, Menten B, Oostra A, Verhelst H, Amor DJ, Bruno DL, van Essen AJ, Hordijk R, Sikkema-Raddatz B, Verbruggen KT, Jongmans MC, Pfundt R, Reeser HM, Breuning MH, Ruivenkamp CA. Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals. *Eur J Med Genet*. 2009;52(2-3):77-87.
24. Morris DL, Cho KW, Rui L. Critical role of the Src homology 2 (SH2) domain of neuronal SH2B1 in the regulation of body weight and glucose homeostasis in mice. *Endocrinology*. 2010;151(8):3643-3651.
25. Valente L, Tiranti V, Marsano RM, Malfatti E, Fernandez-Vizarra E, Donnini C, Mereghetti P, De Gioia L, Burlina A, Castellan C, Comi GP, Savasta S, Ferrero I, Zeviani M. Infantile encephalopathy and defective mitochondrial DNA translation in patients with mutations of mitochondrial elongation factors EFG1 and EFTu. *Am J Hum Genet*. 2007;80(1):44-58.
26. Odermatt A, Taschner PE, Khanna VK, Busch HF, Karpati G, Jablecki CK, Breuning MH, MacLennan DH. Mutations in the gene-encoding SERCA1, the fast-twitch skeletal muscle sarcoplasmic reticulum Ca^{2+} ATPase, are associated with Brody disease. *Nat Genet*. 1996;14(2):191-194.
27. Gordon-Smith K, Jones LA, Burge SM, Munro CS, Tavadia S, Craddock N. The neuropsychiatric phenotype in Darier disease. *Br J Dermatol*. 2010;163(3):515-522.
28. Lerner TJ, Boustany RN, Anderson JW, D'Arigo KL, Schlumpf K, Buckler AJ, Gusella JF, Haines JL; International Batten Disease Consortium. Isolation of a novel gene underlying Batten disease, *CLN3*. *Cell*. 1995;82(6):949-957.