Molecular Psychiatry (2011) 16, 17–25

www.nature.com/mp

ORIGINAL ARTICLE

Copy number variations of chromosome 16p13.1 region associated with schizophrenia

A Ingason[1,2], D Rujescu[3], S Cichon[4,5], E Sigurdsson[6], T Sigmundsson[6], OPH Pietiläinen[7], JE Buizer-Voskamp[8,9], E Strengman[9], C Francks[10], P Muglia[10], A Gylfason[1], O Gustafsson[1], PI Olason[1], S Steinberg[1], T Hansen[2], KD Jakobsen[2], HB Rasmussen[2], I Giegling[3], H-J Möller[3], A Hartmann[3], C Crombie[11], G Fraser[11], N Walker[12], J Lonnqvist[13], J Suvisaari[13], A Tuulio-Henriksson[13], E Bramon[14], LA Kiemeney[15], B Franke[16], R Murray[14], E Vassos[14], T Toulopoulou[14], TW Mühleisen[4], S Tosato[17], M Ruggeri[17], S Djurovic[18,19], OA Andreassen[18,19], Z Zhang[20], T Werge[2], RA Ophoff[9,21], GROUP Investigators[22], M Rietschel[23], MM Nöthen[4,5], H Petursson[6], H Stefansson[1], L Peltonen[7,24,25], D Collier[14], K Stefansson[1] and DM St Clair[11]

[1]*deCODE genetics, Reykjavík, Iceland;* [2]*Research Institute of Biological Psychiatry, Mental Health Centre Sct. Hans, Copenhagen University Hospital, Roskilde, Denmark;* [3]*Division of Molecular and Clinical Neurobiology, Department of Psychiatry, Ludwig-Maximilians-University and Genetics Research Centre GmbH, Munich, Germany;* [4]*Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany;* [5]*Institute of Human Genetics, University of Bonn, Bonn, Germany;* [6]*Department of Psychiatry, National University Hospital, Reykjavík, Iceland;* [7]*Department for Molecular Medicine, National Public Health Institute, Helsinki, Finland;* [8]*Department of Psychiatry, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, Utrecht, The Netherlands;* [9]*Department of Medical Genetics and Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, Utrecht, The Netherlands;* [10]*Medical Genetics, GlaxoSmithKline R&D, Verona, Italy;* [11]*Institute of Medical Sciences, University of Aberdeen, Aberdeen, Scotland;* [12]*Ravenscraig Hospital, Greenock, Scotland;* [13]*Department of Mental Health and Addiction, National Public Health Institute, Helsinki, Finland;* [14]*Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College, London, UK;* [15]*Department of Epidemiology & Biostatistics (133 EPIB)/Department of Urology (659 URO), Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands;* [16]*Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands;* [17]*Section of Psychiatry and Clinical Psychology, University of Verona, Verona, Italy;* [18]*Institute of Psychiatry, University of Oslo, Oslo, Norway;* [19]*Departments of Medical Genetics and Psychiatry, Ulleval University Hospital, Oslo, Norway;* [20]*Department of Statistics, UCLA, Los Angeles, CA, USA;* [21]*UCLA Center for Neurobehavioral Genetics and Department of Human Genetics, Los Angeles, CA, USA;* [22]*See footnote;* [23]*Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health Mannheim, University of Heidelberg, Mannheim, Germany;* [24]*Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK and* [25]*The Broad Institute, Cambridge, MA, USA*

Deletions and reciprocal duplications of the chromosome 16p13.1 region have recently been reported in several cases of autism and mental retardation (MR). As genomic copy number variants found in these two disorders may also associate with schizophrenia, we examined 4345 schizophrenia patients and 35 079 controls from 8 European populations for duplications and deletions at the 16p13.1 locus, using microarray data. We found a threefold excess of duplications and deletions in schizophrenia cases compared with controls, with duplications present in 0.30% of cases versus 0.09% of controls ($P = 0.007$) and deletions in 0.12 % of cases and 0.04% of controls ($P > 0.05$). The region can be divided into three intervals defined by flanking low copy repeats. Duplications spanning intervals I and II showed the most significant ($P = 0.00010$) association with schizophrenia. The age of onset in duplication and deletion carriers among cases ranged from 12 to 35 years, and the majority were males with a family history of psychiatric disorders. In a single Icelandic family, a duplication spanning intervals I and II was present in two cases of schizophrenia, and individual cases of alcoholism, attention deficit hyperactivity disorder and dyslexia. Candidate genes in the region include *NTAN1* and *NDE1*. We conclude that duplications and perhaps also deletions of chromosome 16p13.1, previously reported to be associated with autism and MR, also confer risk of schizophrenia.
Molecular Psychiatry (2011) **16,** 17–25; doi:10.1038/mp.2009.101; published online 29 September 2009

Keywords: 16p13.1; CNV; schizophrenia; duplication

Correspondence: Professor DM St Clair, Institute of Medical Sciences, University of Aberdeen, Aberdeen AB25 2ZH, Scotland, UK.
E-mail: d.stclair@abdn.ac.uk
[22]Members of the GROUP Investigators are listed in the Appendix.
Received 11 June 2009; revised 18 August 2009; accepted 21 August 2009; published online 29 September 2009

Introduction

Copy number variants (CNVs) are emerging as an important genomic cause of autism, mental retardation (MR) and schizophrenia.[1–8] There is no simple pattern: several CNVs are associated with a broad range of neuropsychiatric phenotypes with no respect for traditional clinical diagnostic boundaries,[4–6,8] and no agreed explanation is found for this clinical and genetic overlap.[9–11] Despite high mutation rates,[6] some recurrent CNVs seem to maintain themselves at low frequencies in human populations through the reduced fecundity and negative selection associated with severe neuropsychiatric disorders such as autism, MR and schizophrenia.

Ullmann *et al*.[12] performed a high-resolution, tiling path BAC array comparative genomic hybridization on DNA from 70 autistic individuals from an Australian cohort. Identical submicroscopic 16p13.1 duplications were found in two unrelated patients. The duplication encompassed an interval of 1.5 Mb, ranging from chromosome 16p, 14.89–16.39 Mb (Human Genome Build 35). A third duplication was identified by quantitative PCR in a second Australian cohort of 112 patients. Two of the duplications were familial, and in one family, a severely autistic brother of the index case also carried the duplication. One of the brothers was continuously hyperactive, destructive and aggressive, whereas the younger brother was passive and easy to manage. Other carriers included a sister, who had learning difficulties and mother who had learning difficulties coupled with obsessive–compulsive disorder. They also reported two deletion patients with severe MR. The former was *de novo*; the latter had a mildly affected carrier mother. Hannes *et al*.[13] in a study of 1027 patients with MR and/or multiple congenital anomalies found six deletions and seven duplications in the chromosome 16p13 region. Five deletions and five duplications corresponded roughly to the CNVs found by Ullmann *et al*.[12] with distal breakpoints at 14.7–14.75 Mb and proximal breakpoints at 16.3–16.77 Mb (Human Genome Build 36). Two further duplications had distal breakpoints at 15.1–15.4 Mb and proximal breakpoints at 18.05–18.45 Mb, and one deletion had breakpoints at 16.3–16.77 and 18.3–18.4 Mb. Despite phenotypic variability, common features included three deletions with MR, microcephaly and epilepsy, and three duplications with pronounced behavioral problems in addition to MR and/or multiple congenital anomalies. However, when they examined the region in roughly 2000 controls, they found five duplications (no estimate of breakpoints provided) and no deletions, leading them to the conclusion that, although the deletion was probably pathogenic, the reciprocal duplication might be benign.[13]

The chromosome 16p13.1 duplication/deletion interval is located in a region with reported linkage to bipolar disorder[14–17] and to puerperal psychosis.[18] In addition, analysis of 458 Finnish schizophrenia families reported linkage to the *DISC1* locus;[19] however, when these families were later conditioned for a risk haplotype spanning intron 1 and exon 2 of the *DISC1* gene, linkage was found in 16p13.1 region.[20] The duplicated/deleted region contains the gene coding the DISC1-binding protein, NDE1, and the authors found significant allelic association between *NDE1* and schizophrenia. However, this has not been confirmed in a recent Japanese study.[21] Finally, another significant association was recently reported between *NDE1* and schizophrenia, when schizophrenia cases and controls were conditioned for the presence of Cys residue at the non-synonymous polymorphism Ser704Cys encoded by the *DISC1* gene.[22]

In the present study, we assessed association of CNVs in the 16p13.1 region with schizophrenia as part of a genome-wide scan using the Illumina microarrays HumanHap300, HumanHap550v3 and Human 610-Quad (Illumina, San Diego, CA, USA); and the Affymetrix SNP 6.0 microarray (Affymetrix, Santa Clara, CA, USA) in a sample of 4345 schizophrenia patients and 35 079 controls from 8 European populations, namely, Iceland, Finland, Germany, the Netherlands, Norway, Italy, Denmark and the United Kingdom. Preliminary details on the Scottish sample and part of the German sample have already been documented.[7,23]

Materials and methods

Samples

A total of 4345 schizophrenia affected individuals and 35 079 screened controls from eight European populations were successfully examined at deCODE for CNVs at the locus studied here; they comprised 1435 schizophrenia patients and 28 554 control individuals from Iceland, Scotland, Germany, England, Italy and Finland (The SGENE sample; http://www.SGENE.eu); to these were added a further 866 patients and 856 controls from Aberdeen, Scotland and Munich, Germany; 491 patients and 881 controls from Bonn, Germany; 502 patients and 477 controls from Denmark; 806 Dutch cases and 4039 controls; and 245 patients and 272 controls from Norway. For a full description of samples see Supplementary Information, part one. Ethical approval was obtained from the local Ethics Committees. All participants gave a written informed consent.

The SGENE samples were typed on the HumanHap300 BeadArray (Illumina) at deCODE genetics. The additional samples from Aberdeen and Munich were typed at Duke University in collaboration with GlaxoSmithKline on HumanHap550v1 and HumanHap300 BeadArray (Illumina, respectively). The samples from Bonn were typed at Bonn University on the HumanHap550v3 BeadArray (Illumina). The Dutch samples from Utrecht University were genotyped at the University of California, Los Angeles, on HumanHap550v3 BeadArray (Illumina). The remaining Dutch samples were genotyped at deCODE genetics on HumanHap300 BeadArray (Illumina). The Danish

samples were genotyped at deCODE genetics on the Human 610-Quad BeadArray (Illumina). The Norwegian samples were genotyped on Affymetrix GeneChip(r) GenomeWide SNP 6.0 array and analyzed using the Affymetrix Power Tools 1.8.0.

CNV detection

Dosage Miner software developed at deCODE genetics and QuantiSNP software developed at Wellcome Trust Centre for Human Genetics and the University of Oxford (http://www.well.ox.ac.uk/QuantiSNP/)[24] were used to identify deletions and duplications within the region reported by Ullmann *et al*. and Hannes *et al*.[12,13] Dosage Miner, described in detail elsewhere,[6] uses the intensities from single nucleotide polymorphism (SNP) probes on the Illumina microarrays to estimate copy number of genomic regions and models factors, such as SNP effect, sample effect and GC-content the in the neighboring regions, to normalize the intensities. The software then automatically registers SNP loci in which intensities fall above or below an empirical threshold.

The QuantiSNP program relies on an Objective Bayes Hidden-Markov Model to estimate copy number variations. In this model, the hidden states denote the unknown copy number at the inspected SNPs. Genotype data was used to compute different states. The algorithm computes a Bayes factor that is used to calibrate the model to a fixed type I (false-positive) error rate. A Bayes factor threshold of 10 is considered as a promising value for the possible presence of a CNV. Usually, such values occur when 5–10 consecutive SNPs are deleted/duplicated. Differences in GC base pairs may result in biased hybridization behavior of SNP probes bearing the risk of miscalling genotypes. To normalize this, QuantiSNP assigns a locus-specific GC value to each probe. All potential CNVs that were detected by both softwares, and spanned at least 10 consecutive SNP probes, were subsequently visually inspected and confirmed.

Association analysis

A two-tailed Cochrane–Mantel–Haentzel analysis assuming common odds ratios was performed, stratifying samples by country of origin to take account of the possible effect of geographical variation on the results of the analysis.

Results

We limited our search on 16p13.1 to the region between 14.66 and 18.70 Mb (Human Genome Build36). We subdivided the region into three single copy sequence intervals, which we called intervals I, II and III; each interval is flanked by sequences rich in low copy repeats with 99% sequence homology (Figure 1). The duplications and deletions previously reported[12,13] are all contained within this region, with the most common breakpoints in the low copy repeat clusters distal to interval I and proximal to interval II (Figure 1).

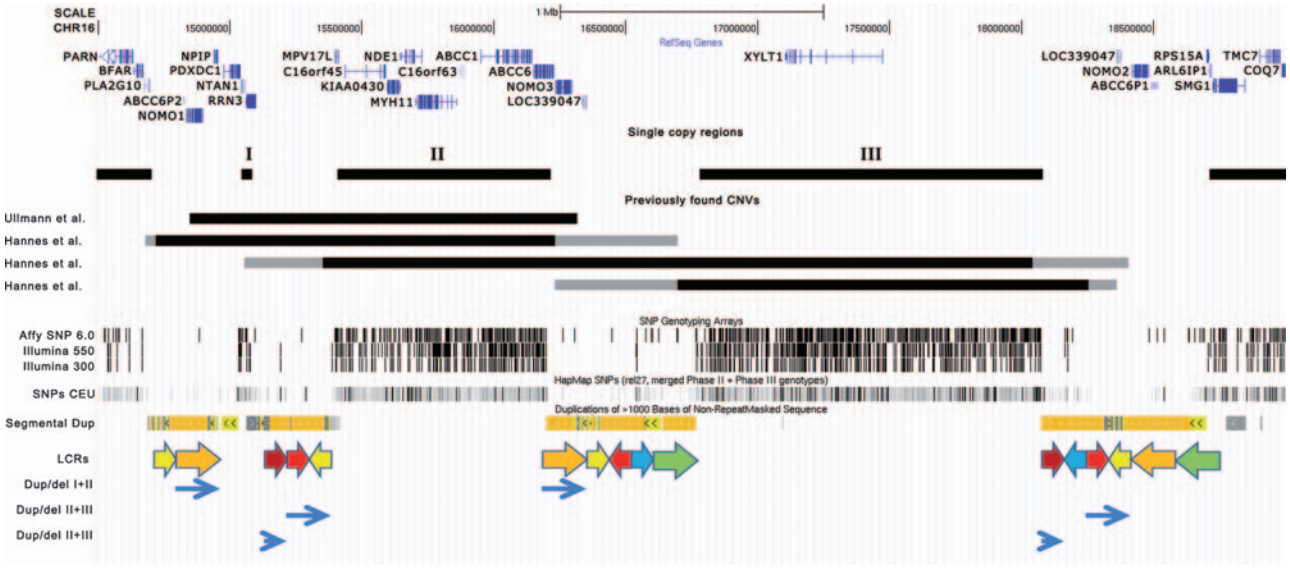

Figure 1 The 16p13.11-p12.3 region. Chromosomal coordinates according to the Human Genome Build 36 are shown at top, followed by genes from the Reference Sequence database. The locus is marked by regions of single copy sequence interspersed with stretches of segmental duplications. The single copy regions, referred to in the article as I, II and III are shown as black bars. Copy number variants (CNVs) identified by Ullmann *et al*.[12] and Hannes *et al*.[13] are also indicated with black bars, gray bars reflect uncertainty of exact breakpoints. The genomic coverage of microarrays used in this study is also shown, and reflects the division of the locus into single copy regions and segmental duplications, as does the coverage of HapMap single nucleotide polymorphisms (SNPs) from the CEU trios and the track displaying segmental duplications. At the bottom, we have depicted the largest low copy repeats in the region (<50 kb) with high sequence homology (>98%), different colors denote different repeats and arrows show directionality. Finally, we show the most likely sites for NaHR causing four of the six CNVs found in the study.

Table 1 Duplications and deletions of 16p 13.1 in European populations

	Iceland[a]		*Scotland*[b]		*Germany*		*Holland*[b]		*Denmark*	
Status	*Scz*	*Ctrl*	*Scz*	*Ctrl*	*Scz*	*Ctrl*	*Scz*	*Ctrl*	*Scz*	*Ctrl*
No. of cases	648	27 747	661	665	1102	1493	806	4039	502	477
Percentage of male	63	39	72	58	57	49	76	60	58	58
All Dup	2	24	6	1	1	0	3	6	1	1
All Del	0	12	1	0	3	3	0	0	1	0
Dup I + II	2	18	6	0	0	0	3	2	1	0
Dup II	0	3	0	1	0	0	0	1	0	1
Dup II + III	0	3	0	0	1	0	0	3	0	0
Del I + II	0	10	0	0	3	1	0	0	1	0
Del II	0	0	0	0	0	0	0	0	0	0
Del II + III	0	2	1	0	0	2	0	0	0	0

Abbreviations: ADHD, attention deficit hyperactivity disorder; CNV, copy number variant; Ctrl, control; Del, deletion; Dup, duplication; Scz, schizophrenia.
[a]Further duplications and deletions were detected in Icelandic samples that were excluded from the case–control analysis. These include one autism case with Del II, one alcoholism case with Dup II, two alcoholism cases with Del I + II, one ADHD and one alcoholism case with Dup II + III, and finally, four first-degree relatives of schizophrenia cases, two ADHD, two alcoholism and one dyslexia case with Dup I + II (some of the Dup I + II carriers counted here are shown in figure 3).
[b]Genome-wide CNV results for 54 of the Dutch cases were published recently[42] but these do not include carriers of CNVs at 16p13.1. The Scottish sample has been independently genotyped on the Affymetrix platform by the International Schizophrenia consortium[7], and 16p13.1 CNVs are listed with other CNV findings from that study at http://pngu.mgh.harvard.edu/isc/isc-r1.cnv.bed. Furthermore, genome-wide CNV results for 866 patients and 856 controls of the Scottish and German sample, including carriers of three of the Dup I + II as well as one of the Del I + II, the Del II + III and Dup II found in this study, have been published elsewhere.[43]

Table 1 lists the duplications and deletions found in our sample, together with the population of origin. None were found in cases or controls from England ($n = 104$ and 95), Italy ($n = 86$ and 92), Finland ($n = 191$ and 200) and Norway ($n = 245$ and 272). Accordingly these samples are not included in the table or in the Cochran–Mantel–Haentzel statistical analysis.

Owing to varying geographical origin of the samples, we analyzed the data for association using Cochran–Mantel–Haentzel algorithm to correct for stratification. Within the total population examined, we found a threefold excess of duplications and deletions in cases compared with controls (see Table 2). We found six distinct CNV categories in the region, corresponding to deletions and reciprocal duplications of three overlapping regions (Figure 2). Duplications were present in 0.30% of schizophrenia cases versus 0.09% controls ($P = 0.007$). When analysis was restricted to duplications containing intervals I and II (Dup I + II, Figure 2), the significance increased further ($P = 0.00010$). The Dup I + II was present in four male and two female Scottish cases, two male Icelandic, one male Danish and three male Dutch cases, as well as in 12 female and six male Icelandic controls, in none of the Scottish controls and in two Dutch male controls. Common odds ratio was 8.79 (males) and 3.63 (females). The two Icelandic cases were independently ascertained and are included in the analysis as separate probands. However, when genealogical analysis was later performed, we found that the two were second-degree relatives. Other carriers in the family included single cases of alcoholism (under treatment), dyslexia and attention deficit hyperactivity disorder (See Supplementary Figure 1).

Deletions were present in 0.12% cases and 0.04% controls ($P > 0.05$). A deletion of interval I and II (Del I + II, Figure 2) was present in 3 German and 1 Danish schizophrenia cases, and 1 German and 10 Icelandic controls ($P > 0.05$), whereas a deletion of intervals II and III (Del II + III, Figure 2) was present in 1 Scottish schizophrenia case, and 2 German and 2 Icelandic controls ($P > 0.05$).

We also tested allelic association for all SNP markers on the Illumina microarrays that spanned the 16p13.1 region in 2687 schizophrenia cases and 13 484 controls. Although none of the markers in the interval were genome-wide significant, one marker, rs2283508, was associated ($P = 1.5E\text{-}05$) and remained significant after locus-wide correction $P = 0.0043$ (117 markers tested). This marker is located within an intron of the *ABCC6* gene. No other markers spanning the region were significant after correcting for the 117 markers on the HumanHap300 chip in the interval spanning the duplication (see Supplementary Information, part 3).

In view of the report[22] of significant associations between *NDE1* and schizophrenia when schizophrenia cases were stratified by the presence of a Cys residue at Ser704Cys of the *DISC1* gene, we also conditioned our schizophrenia cases. The *DISC1* Ser704Cys SNP, rs821616, is not on the HumanHap300 chip. However, a SNP that is in complete linkage disequilibrium ($r^2 = 1$) with rs821616 in

Table 2 *P*-values, common odds ratios (ORs) and 95% confidence intervals (CIs) for 16p13.1 duplications and deletions

	All samples			*Male only*			*Female only*		
16p13.11 CNV	P-*values*[a]	*Common OR*	*95% CI*	P-*values*[a]	*Common OR*	*95% CI*	P-*values*[a]	*Common OR*	*95% CI*
All duplications	0.0071[b]	3.27	1.29–7.94	0.014	3.54	1.20–10.31	0.59	2.05	0.17–12.89
All deletions	0.51	1.65	0.36–7.06	1	1.34	0.17–9.18	0.60	2.03	0.13–22.98
Duplications intervals I and II	0.00010[c]	7.27	2.49–19.82	0.00030	8.79	2.39–32.38	0.24	3.63	0.22–28.74
Deletions intervals I and II	0.23	2.71	0.42–15.35	0.28	3.66	0.18–45.46	0.60	2.10	0.13–25.65
Duplications intervals II and III	1	1.01	0.02–11.47	1	1.86	0.02–60.28	1	0	0–29.09
Deletions intervals II and III	1	0.59	0.01–9.48	1	0.50	0.01–8.59	1	0	0–2623.69

Abbreviation: CNV, copy number variant.
[a]Uncorrected.
[b]Significant ($P = 0.014$) after correction for multiple testing.
[c]Significant ($P = 0.00060$) after correction for multiple testing.

the Hapmap CEU population, rs821596, was present. We therefore used rs821596 to divide the schizophrenia cases into Cys704-carrier and non-carrier groups, and then looked for allelic association with SNPs at the *NDE1* locus in the two groups. None were significantly associated after correcting for the number of tested markers. The data regarding 51 SNPs in, or within, 200 Kb of *NDE1* for the Cys704 carrier and non-carrier groups are given in Supplementary Information, part 3.

We also performed exploratory analyses. As the majority of the duplication cases were Scottish in origin, we examined the haplotype background of the duplicated regions. The CNV occurred on a different haplotype background in each individual. In addition, none of the non-Icelandic carriers, for whom we had genotype data, had a CNV on the same haplotype background as any of the Icelandic carriers. This suggests that there was no founder mutation, and each of the events is likely to have arisen independently. Within Iceland itself, for each of the CNV duplication and deletion subtypes (Figure 2) found in more than one individual, there was no founder mutation. There were also enough individuals in the Icelandic population with Dup I+II to look at clustering patterns. Clustering occurred at a rate of three- to fourfold less than expected if the duplications were selectively neutral. See Supplementary Information, part 4 for discussion of cluster analysis.

Discussion

We have found a statistically significant ($P = 0.\ 007$) threefold overrepresentation of duplications in the chromosome 16p13.1 region in schizophrenia cases compared with controls. We found a similar threefold overrepresentation of deletions in cases, but this was not significant ($P > 0.05$). The great majority of duplications and deletions we found using Illumina microarrays are identical to those reported by two groups using array comparative genomic hybridization.[12,13] They span the same 1.5 Mb region that includes intervals that we refer to as intervals I and II; we have also identified novel duplications and a single deletion involving interval II only.

The distinction between CNVs involving intervals I and II and interval II alone is dependent on a group of seven SNP probes at 15.0–15.1 Mb (Figure 1). However, we are confident that our classifications are correct; every CNV detected by the independent Dosage Miner and QuantiSNP algorithms was then inspected by eye, and in all cases, we could make a clear distinction between those involving both intervals I and II (Dup/Del I+II) and interval II only (Dup/Del II). In addition, genealogical analysis of the Icelandic data set, performed *post hoc*, showed a clear separation of CNV types. In total, 36 out of 54 carriers were clustered in 11 families (ancestral clustering, max depth three generations), including seven families with Dup I+II (nine, three and five times two carriers), two families with Del I+II (six and two

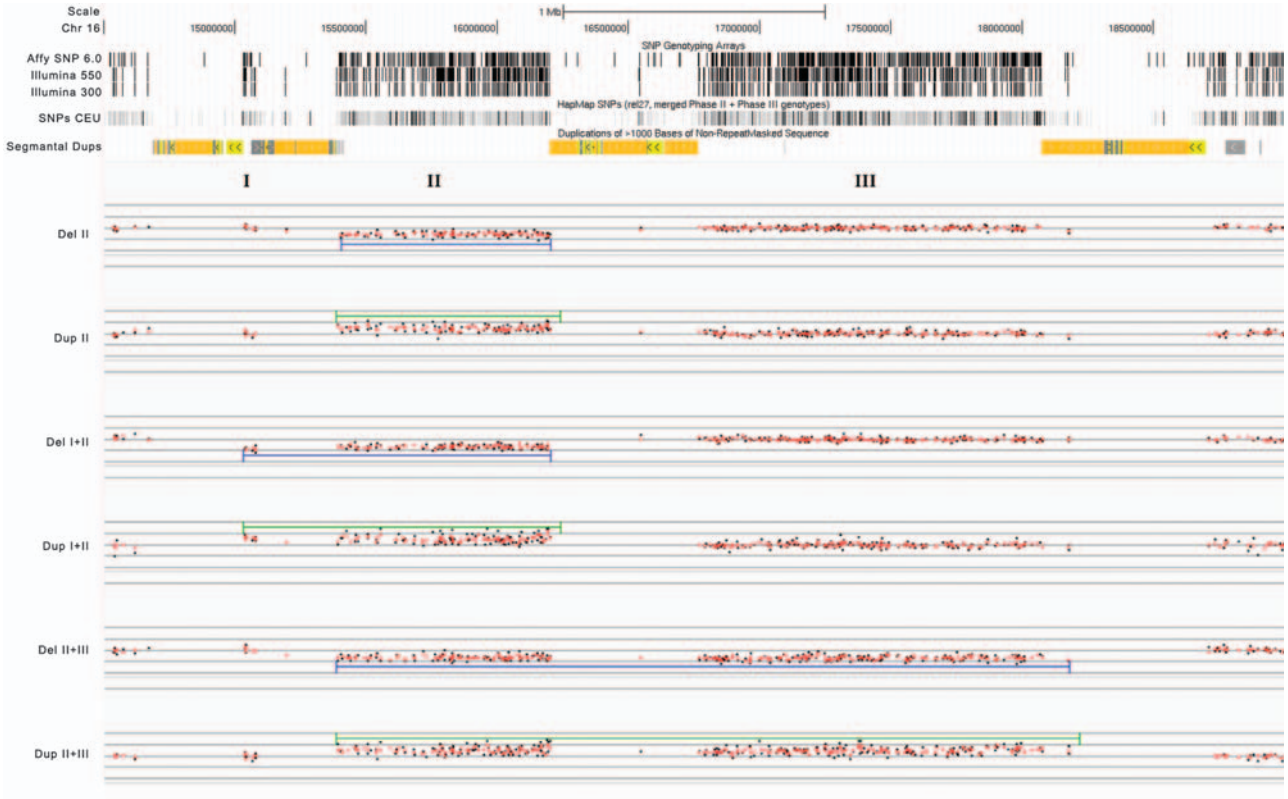

Figure 2 Intensity plots for the Copy number variants (CNVs) identified in this study in the 16p13.11-p12.3 region. Chromosomal coordinates according to the Human Genome Build 36 are shown at top, followed by markers from the microarrays used in this study, HapMap single nucleotide polymorphisms (SNPs) from the CEU population and the track displaying known segmental duplications. At the bottom, intensity plots from Dosage Miner are shown for one carrier of each of the identified CNVs. The blue and green bars show the intervals of Dosage Miner automatic calls for deletions and duplications, respectively.

carriers), one family with Dup II (four carriers) and one family with Dup II + III (two carriers). In no instance did we observe two different CNVs segregating within the same family.

The breakpoints for all three types of deletion/duplication are located in areas with high low copy repeat content, reflecting apparent genomic instability of the region.[25] The repeats are in the same orientation, and non-allelic homologous recombination (NaHR) between these low copy repeats seems to be the most likely explanation for the recurrence of the rearrangements and for their identical size. Three inversion polymorphisms have previously been described in the 16p13.1 region.[26,27] A large duplication in a patient with MR has also been reported,[28] as has a smaller *de novo* duplication.[29] A much larger duplication (8 Mb) of the region has also been reported in two unrelated patients with autistic features.[30]

Our most striking finding is the increased risk of schizophrenia associated with duplications at the 16p13.1 locus. Recurrent deletions at several loci have now been reported to be significantly associated with schizophrenia, but, to date, duplications associated with schizophrenia have mostly been isolated case reports. It is more difficult to decide whether the duplications are genuinely pathogenic or benign. The problem has already been encountered by Hannes *et al*.[13] with duplications in MR at this locus. They found duplications in five out of 1027 cases, three of whom had striking behavioral problems similar to those reported by Ullmann *et al*.[12] However, because the duplication was present in 5 out of 2000 controls, they could not decide whether it was genuinely pathogenic. The different sizes of duplications and deletions at the 16p13.1 locus also present difficulties when it comes to assessing statistical association. Statistically, we have used the straightforward approach of counting all duplications and then deletions as equivalent events, and only then tried to condition on those duplications or deletions that have the same breakpoints as those reported previously in autism and MR/multiple congenital anomalies.[12,13]

Although caution must be exercised when interpreting results from such a small number of cases, there are several grounds supporting that our findings are indeed genuine. First, given the rarity of the duplications, the overall association with schizophrenia is statistically ($P = 0.0071$) significant, and increases ($P = 0.00010$) when the Dup I + II is considered separately. In addition, identical duplications at the 16p13.1 locus have already been found

associated with autism.[12] Second, four of the schizophrenia duplication cases had an early onset of illness (12, 17, 19 and 19 years) and in this respect resembled the 16p13.1 deletion cases in which three of the five schizophrenia cases also had early onset of illness (15, 17 and 18 years, see Supplementary Information, part 2). With a mean age at onset of 24 years, this reflects an overrepresentation of early onset cases, although not reaching significance. Third, in an Icelandic family, the duplication co-segregates with neuropsychiatric disorders, including two cases of schizophrenia, and individual cases of attention deficit hyperactivity disorder, dyslexia and alcoholism. Owing to the small size of the family and the diversity of the neuropsychiatric phenotypes, calculating a logarithm of the odds (LOD) score did not seem appropriate. This range of phenotypes we observe in a single family is not unexpected, as an overlap of phenotypic features between autism and attention deficit hyperactivity disorder has been extensively reported, and individuals with attention deficit hyperactivity disorder are at increased risk of schizophrenia.[31–33] Fourth, the duplications at this locus seem to be under negative selection. Cluster analysis of the Dup I+II events in the Icelandic population finds that the carriers cluster less than expected in families, that is, the genealogical clusters are smaller, and sporadic carriers more numerous than would be expected if the duplications were selectively neutral. The arguments and methods are laid out in Supplementary Information, part 4 and are also given in Stefansson *et al*.[6] These clustering results, however, are informal, and so any conclusions must be provisional. Fifth, Kirov *et al*.[34] have identified three chromosome 16p13.11 duplications in 471 schizophrenia cases and 6 out of 2792 controls. In addition, the International Schizophrenia Consortium finds overrepresentation of large 16p13.11 duplications (13 of 3391 cases versus 7 of 3181 controls have duplications encompassing interval II, see at http://pngu.mgh.harvard.edu/isc/isc-r1.cnv.bed),[7] but this overrepresentation is mostly explained by the Aberdeen sample, which overlaps between our study and the International Schizophrenia Consortium study (although separately genotyped and analyzed), and therefore, does not provide further support to our finding.

Finally, with regard to the Dup I+II duplications, although the Scottish sample seems to be driving the association observed, as 6 out of 12 carriers among affected cases belong to that sample, the association of Dup I+II in this study remains significant even after removing the Scottish samples from the analysis ($P = 0.0033$, OR = 5.94).

The duplicated region contains two strong candidate genes (*NTAN1* and *NDE1*), over- or underexpression of either or both of which at key stages of neurodevelopment could predispose to autism, MR and/or schizophrenia.

The *NTAN1* gene is located in the small island of non-repeated sequence called interval one. It encodes an N-terminal asparagine amidase that has been implicated in social behavior and memory. Overexpression of *NTAN1* leads to reduction in the MAP2 (microtubule-associated protein 2) expression through the ubiquitin proteasome pathway. Reduced expression of *MAP2* may be a useful marker for diagnosis of schizophrenia and bipolar disorder *in vivo*[35,36] and *in vitro*.[37,38] Mice with a disrupted *NTAN1* gene show less locomotion in an open field and impairment of several spatial memory tasks.[39,40]

NDE1 and *NDEL* are highly homologous genes involved in brain development, neuronal proliferation, migration and synapse formation. They encode proteins that biologically interact with DISC1 and LISI proteins, with *NDE1* appearing to be interchangeable with its homolog *NDEL*, except that *NDE1* is expressed earlier in development. *nde1*-null mice are viable and display microcephaly with thinning cortical layer and reduced numbers of neurons. Interestingly, two out of three reported autism cases with duplication had increased head circumference.[12] The *nde1*-null mice display defects in neuronal proliferation and neuronal migration. The NDE1 protein directly interacts with the DISC1 protein at the C-terminal end that is distal to a truncating mutation that is reported to segregate with schizophrenia and other forms of major mental illness in a large Scottish *DISC1* translocation family.[41,42] As truncated DISC1 is known to alter NDE1/NDEL function, the duplications we report here may have a similar biological effect as the truncating mutation associated with schizophrenia in the Scottish family.

All duplications and deletions in our study involve interval II that harbors the *NDE1* gene and this makes dysregulation of NDE1 expression the most parsimonious explanation for the increased risk of the phenotypes reported by Ullmann *et al*.[12] and by the authors of this paper. On the other hand, the strongest association is with duplications that also involve interval I. It is possible that combined changes in expression of *NTAN1* and *NDE1* increase susceptibility over changes in expression of *NDE1* alone. No clear-cut findings emerged from our examination of the 16p13.1 region for allelic association for schizophrenia. The findings are discussed along with association studies following conditioning for *DISC1* Ser704Cys in Supplementary Information, part 3.

Further work is required before the clinical implications of our findings become clear. On the one hand, the data strongly suggest that recurrent duplications at 16p13.1 locus increase the risk of schizophrenia. They also strengthen the hypothesis that there are shared genetic risk factors between schizophrenia, MR and autism. However, the odds ratios, even for the Dup I+II, are substantially less than the increased risks we have observed for three out of four recurrent deletions on chromosomes 1, 15 and 22.[6] Whether the smaller odds ratio we observe for duplications is a feature of the 16p13.1 locus itself, or it is part of a broader rule that recurrent duplications are generally less penetrant than recurrent

deletions, remains to be determined. The 16p13.1 duplications we observe are rare, at a rate of about 3 or 4 per 1000 schizophrenia cases, and, estimating from the control population in the present study, about 0.08% in the general population. This makes it difficult to obtain precise measurements of overall risk or risk for individual neuropsychiatric disorders. Analysis of CNV data from sets of cases and controls considerably larger than the sets we report in this paper, which itself to date is one of the largest assembled, will be required. These and many other questions will need to be answered before the exciting findings that arise from CNV analysis can be used in clinical practice for diagnostics, disease classification or genetic testing.

Acknowledgments

We thank the participating subjects and their relatives, and staff at the recruitment centres. We thank David Goldstein for permission to use the genotype data from the Scottish samples typed at Duke University. This work was sponsored by EU grant LSHM-CT-2006–037761 (Project SGENE). Genotyping of the Dutch samples was sponsored by NIMH funding, R01 MH078075 (to RAO).

Conflict of interest

The authors declare no conflict of interest.

References

1 Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T *et al.* Strong association of *de novo* copy number mutations with autism. *Science* 2007; **316**: 445–449.
2 Zhao X, Leotta A, Kustanovich V, Lajonchere C, Geschwind DH, Law K *et al.* A unified genetic theory for sporadic and inherited autism. *Proc Natl Acad Sci USA* 2007; **104**: 12831–12836.
3 Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R *et al.* Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 2008; **358**: 737–739.
4 Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM *et al.* Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 2008; **320**: 539–543.
5 Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M. Strong association of *de novo* copy number mutations with sporadic schizophrenia. *Nat Genet* 2008; **40**: 880–885.
6 Stefansson H, Rujescu D, Cichon S, Pietiläinen OPH, Ingason A, Fossdal R *et al.* Large recurrent deletions associated with schizophrenia. *Nature* 2008; **455**: 232–236.
7 The International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* 2008; **445**: 237–241.
8 Mefford H, Sharp A, Baker C, Itsara Z, Jiang K, Buyysse S *et al.* Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N Eng J Med* 2008; **359**: 1685–1699.
9 Cantor RM, Geschwind DH. Schizophrenia: genome, interrupted. *Neuron* 2008; **58**: 165–167.
10 O'Donovan MC, Kirov G, Owen MJ. Phenotypic variations on the theme of CNVs. *Nat Genet* 2008; **12**: 1392–1393.
11 St Clair D. Copy number variation and schizophrenia. *Schizophr Bull* 2009; **35**: 9–12.
12 Ullmann R, Turner G, Kirchhoff M, Chen W, Tonge B, Rosenberg C *et al.* Array CGH Identifies reciprocal 16p13.1 duplications and deletions that predispose to autism and /or mental retardation. *Hum Mutat* 2007; **28**: 674–682.
13 Hannes FD, Sharp AJ, Mefford HC, de Ravel T, Ruivenkamp CA, Breuning MH *et al.* Recurrent reciprocal deletions and duplications of 16p13.11: the deletion is a risk factor for MR/MCA while the duplication may be a rare benign variant. *J Med Genet* 2008; **46**: 223–232.
14 McInnes LA, Escamilla MA, Service SK, Reus VI, Leon P, Silva S *et al.* A complete genome screen for genes predisposing to severe bipolar disorder in two Costa Rican pedigrees. *Proc Natl Acad Sci USA* 1996; **493**: 13060–13065.
15 Ewald H, Flint T, Kruse TA, Mors O. A genome-wide scan shows significant linkage between bipolar disorder and chromosome 12q24.3 and suggestive linkage to chromosomes 1p22–21, 4p16, 6q14–22, 10q26 and 16p13.3. *Mol Psychiatry* 2002; **7**: 734–744.
16 Ekholm JM, Kieseppä T, Hiekkalinna T, Partonen T, Paunio T, Perola M *et al.* Evidence of susceptibility loci 4q32 and 16p12 for bipolar disorder. *Hum Mol Genet* 2003; **12**: 1907–1915.
17 Kassem L, Lopez V, Hedeker D, Steele J, Zandi P, Bipolar Disorder Consortium NIMH Genetics Initiative *et al.* Familiarity of polarity at illness onset in bipolar affective disorder. *Am J Psychiatry* 2006; **163**: 1099–1104.
18 Jones I, Hamshere M, Nangle JM, Bennett P, Green E, Heron J *et al.* Bipolar affective puerperal psychosis. Genome wide significant evidence for 16p13 linkage. *Am J Psychiatry* 2007; **164**: 248–258.
19 Ekelund J, Hennah W, Hiekkalinna T, Parker A, Meyer J, Lonnqvist J *et al.* Replication of 1q42 linkage in Finnish schizophrenia pedigrees. *Mol Psychiatry* 2004; **9**: 1037–1041.
20 Hennah W, Tomppo L, Hiekkalinna T, Palo OM, Kilpinen H, Ekelund J *et al.* Families with the risk allele of DISC1 reveal a link between schizophrenia and another component of the same molecular pathway, NDE1. *Hum Mol Genet* 2007; **16**: 453–462.
21 Numata S, Ueno S, Iga J, Nakataki M, Ohmori T, Tanahashi T *et al.* No association between NDE1 gene and schizophrenia in the Japanese population. *Schizophr Res* 2008; **99**: 367–369.
22 Burdick KE, Kamiya A, Hodgkinson CA, Lencz T, Derosse P, Ishizuka K *et al.* Elucidating the relationship between DISC1, NDEL1, and NDE1 and the risk for schizophrenia: evidence of epistasis and competitive binding. *Hum Mol Genet* 2008; **17**: 2462–2473.
23 Need AC, Ge D, Weale ME, Feng S, Heinzen EL, Shianna K *et al.* A Genome wide investigation of SNPs and CNVs in schizophrenia. *PloS Genet* 2009; **5**: e1000421.
24 Colella S, Yau C, Taylor JM, Mirza G, Butler H, Clouston P *et al.* QuantiSNP: an Objective Bayes Hidden–Markov Model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Res* 2007; **35**: 2013–2025.
25 Shaw CJ, Lupski JR. Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum Mol Genet* 2004; **13**: R57–R64.
26 Tuzun E, Sharp AJ, Bailey JA, Kaul R, Morrison VA, Pertz LM *et al.* Fine-scale structural variation of the human genome. *Nat Genet* 2005; **37**: 727–732.
27 Zhang J, Feuk L, Duggan G, Khaja R, Scherer SW. Development of bioinformatics resources for display and analysis of copy number and other structural variants in the human genome Cytogen. *Genome Res* 2006; **115**: 205–214.
28 Sharp AJ, Hansen S, Selzer RR, Cheng Z, Regan R, Hurst JA *et al.* Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat Genet* 2006; **38**: 1038–1042.
29 Kriek M, White SJ, Bouma MC, Dauwerse HG, Hansson KB, Nijhuis JV *et al.* Genomic imbalances in mental retardation. *J Med Genet* 2004; **41**: 249–255.
30 Finelli P, Natacci F, Bonati MT, Gottardi G, Engelen JJ, de Die-Smulders CE *et al.* FISH characterisation of an identical (16)(p11.2p12.2) tandem duplication in two unrelated patients with autistic behaviour. *J Med Genet* 2004; **41**: e90.
31 Amminger GP, Pape S, Rock D, Roberts SA, Ott SL, Squires-Wheeler E *et al.* Relationship between childhood behavioural disturbance and later schizophrenia in the New York High-Risk Project. *Am J Psychiatry* 1999; **156**: 525–530.

32 Keshavan MS, Sujata M, Mehra A, Montrose DM, Sweeney JA. Psychosis proneness and ADHD in young relatives of schizophrenia patients. *Schizophr Res* 2003; **59**: 85–92.
33 Oner O, Munir K. Attentional and neurocognitive characteristics of high-risk offspring of parents with schizophrenia compared with DSM-IV attention deficit hyperactivity disorder children. *Schizophr Res* 2005; **76**: 293–299.
34 Kirov G, Groseva D, Norton N, Ivanov D, Mantriparga K, Holmans P *et al.* Support for involvement of large copy number variants in the pathogenesis of schizophrenia. *Hum Mol Genet* 2009; **18**: 1497–1503.
35 Whitaker-Azmitia PM, Borella A, Raio N. Serotonin depletion in the adult rat causes loss of the dendritic marker MAP-2. A new animal model of Schizophrenia? *Neuropsychopharmacology* 1995; **12**: 269–272.
36 Mazer C, Muneyyirci J, Taheny K, Raio N, Borella A, Whitaker-Azmitia P. Serotonin depletion during synaptogenesis leads to decreased synaptic density and learning deficits in the adult rat: a possible model of neurodevelopmental disorders with cognitive deficits. *Brain Res* 1997; **760**: 68–73.
37 Marx CE, Jarskog LF, Lauder JM, Lieberman JA, Gilmore JH. Cytokine effects on cortical neuron MAP-2 immunoreactivity: implications for schizophrenia. *Biol Psychiatry* 2001; **50**: 743–749.
38 Bouras C, Kovari E, Hof PR, Riederer BM, Giannakopoulos P. Anterior cingulate cortex pathology in schizophrenia and bipolar disorder. *Acta Neuropathol* 2001; **102**: 373–379.
39 Kwon YT, Balogh SA, Davydov IV, Kashina AS, Yoon JK, Xie Y *et al.* Altered activity, social behaviour, and spatial memory in mice lacking the NTAN1p amidase and the asparagine branch of the N-end rule pathway. *Mol Cell Biol* 2000; **20**: 4135–4148.
40 Balogh SA, Kwon YT, Denenberg VH. Varying inter-trial intervals reveals temporally defined memory deficits and enhancements in NTAN1-defecient mice. *Learn Mem* 2000; **7**: 279–286.
41 Kamiya A, Kubo K, Tomoda T, Takaki M, Youn R, Ozeki Y *et al.* A schizophrenia associated mutation of DISC1 perturbs cerebral cortex development. *Nat Cell Biol* 2005; **12**: 1167–1178.
42 St Clair D, Blackwood D, Muir W, Carothers A, Walker M, Spowart G *et al.* Association within a family of a balanced autosomal translocation with major mental illness. *Lancet* 1990; **336**: 13–16.
43 Vrijenhoek T, Buizer-Voskamp JE, van der Stelt I, Strengman E, Sabatti C, Geurts van KA, *et al.*, Genetic Risk and Outcome in Psychosis (GROUP) Consortium. Recurrent CNVs disrupt three candidate genes in schizophrenia patients. *Am J Hum Genet* 2008; **83**: 504–510.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)

Appendix

GROUP investigators include: René S Kahn and Wiepke Cahn, Department of Psychiatry, Rudolf Magnus Institute of Neuroscience, University Medical Centre Utrecht, Utrecht, The Netherlands; Don Linszen and Lieuwe de Haan, Department of Psychiatry, Academic Medical Centre University of Amsterdam, Amsterdam, The Netherlands; Jim van Os, Lydia Krabbendam and Inez Myin-Germeys, Maastricht University Medical Centre, EURON South Limburg Mental Health Research and Teaching Network, Maastricht, The Netherlands; Durk Wiersma and Richard Bruggeman, Department of Psychiatry, University Medical Centre Groningen, University of Groningen, The Netherlands.