Identification of rare variants in KCTD13 at the schizophrenia risk locus 16p11.2
Franziska Degenhardt, Barbara Heinemann, Jana Strohmaier, Marvin A. Pfohl, Ina Giegling, Andrea Hofmann, Kerstin U. Ludwig, Stephanie H. Witt, Michael Ludwig, Andreas J. Forstner, Margot Albus, Sibylle G. Schwab, Margitta Borrmann-Hassenbach, Leonard Lennertz, Michael Wagner, Per Hoffmann, Dan Rujescu, Wolfgang Maier, Sven Cichon, Marcella Rietschel, and Markus M. Nóthen

Duplications in 16p11.2 are a risk factor for schizophrenia (SCZ). Using genetically modified zebrafish, Golzio and colleagues identified KCTD13 within 16p11.2 as a major driver of the neuropsychiatric phenotype observed in humans. The aims of the present study were to explore the role of KCTD13 in the development of SCZ and to provide a more complete picture of the allelic architecture at this risk locus. The exons of KCTD13 were sequenced in 576 patients. The mutations c.6G＞T and c.598G＞A were identified in one patient each. Both mutations were predicted to be functionally relevant and were absent from the 1000 Genomes Project data and the Exome Variant Server. The mutation c.6G＞T was predicted to abolish a potential transcription factor-binding site for specificity protein 1. Altered specificity protein 1 expression has been reported in SCZ patients compared with controls. Further studies in large cohorts are warranted to determine the relevance of the two identified mutations.

Introduction
Schizophrenia (SCZ) is a common and often severely disabling neuropsychiatric disorder. Patients show a variety of symptoms, including hallucinations, delusions, disorganized speech, affective flattening, and avolition (American Psychiatric Association, 1994). The estimated heritability of SCZ ranges between 60 and 80% (Sullivan et al., 2003; Wray and Gottesman, 2012). Previous research has implicated a small number of rare copy number variants (CNVs) from specific chromosomal regions in the risk of SCZ (Malhotra and Sebat, 2012; Sullivan et al., 2012; Rees et al., 2014).

One of these chromosomal regions is 16p11.2. Duplications and deletions in 16p11.2 are particularly interesting as they are implicated in mirrored phenotypes. Deletions in this region are associated with macrocephaly and obesity (Bochukova et al., 2010; Shinawi et al., 2010; Walters et al., 2010; Jacquemont et al., 2011), whereas duplications in 16p11.2 are associated with microcephaly and low BMI (Shinawi et al., 2010; Jacquemont et al., 2011). Furthermore, deletions and duplications in this chromosomal region are established risk factors for developmental delay, intellectual disability, autism spectrum disorder, and epilepsy (Kumar et al., 2008; Weiss et al., 2008; Shinawi et al., 2010; Xiang et al., 2010; Sullivan et al., 2012). However, only duplications in 16p11.2 increase the susceptibility to SCZ (odds ratio ~10; McCarthy et al., 2009; Kirov et al., 2012; Sullivan et al., 2012; Rees et al., 2014; Szatkiewicz et al., 2014).

The CNVs reported to date in 16p11.2 are flanked by segmental duplications and are typically ~600 kb in size, spanning more than 25 genes (Weiss et al., 2008). Several of these are interesting candidate genes for neuropsychiatric phenotypes. Golzio et al. (2012) published the results of an analysis in genetically modified zebrafish. The authors generated evidence that KCTD13 is a major driver of the mirrored neuroanatomical phenotypes of the zebrafish.
CNVs in 16p11.2 (Malhotra and Sebat, 2012). In zebrafish, overexpression of the KCTD13 human transcript caused microcephaly, which resembled the microcephaly phenotype associated with the 16p11.2 duplication. In contrast, inhibition of KCTD13 expression caused a macrocephalic phenotype, which has been associated previously with the 16p11.2 deletion (Golzio et al., 2012; Malhotra and Sebat, 2012). In addition, KCTD13 is located in one of the 108 genome-wide significant loci reported in the largest SCZ genome-wide association study worldwide (36,989 patients and 113,075 controls; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014).

The aims of the present study were to explore the role of KCTD13 in the development of SCZ and to provide a more complete picture of the allelic architecture at the 16p11.2 risk locus. The identification of rarer variants in this gene might provide genetic evidence for the role of KCTD13 in susceptibility to SCZ. Furthermore, rarer variants with higher penetrance might be more suitable for functional follow-up studies than common variants with small effects.

Methods
The study was approved by the respective ethics committees and all participants provided written informed consent before inclusion. All study procedures were carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). All participants were of German descent according to self-reported ancestry.

Sample description
In total, 576 patients were included. The patients were recruited from consecutive admissions to psychiatric inpatient units in Germany. A lifetime ‘best estimate’ diagnosis (Leckman et al., 1982) of SCZ according to the Diagnostic and Statistical Manual of Mental Disorders, 4th ed., criteria (American Psychiatric Association, 1994) was assigned on the basis of the medical records, family history, and the Structured Clinical Interview (Spitzer et al., 1992) and/or the OPCRIT (McGuffin et al., 1991). Each individual was of German descent according to self-reported ancestry.

Sanger sequencing
Primer design was based on the NCBI37/hg19 reference sequence (Ensembl transcript ID: ENST00000568000). All six coding exons and their flanking sequences (±30 bp of each exon analyzed) were amplified. Exons 3 and 4 were grouped together in one amplicon. Sanger sequencing was performed in part at Beckman Coulter Genomics (Takeley, UK) and in part at the Institute of Human Genetics in Bonn. The variants identified were confirmed at the Institute of Human Genetics in Bonn by sequencing the complementary strand of a second, independent amplicon. For the verification step, the 3130xl Genetic Analyzer (Applied Biosystems, Foster City, California, USA) was used. The nucleotide sequences obtained were analyzed using SeqMan II (DNASTAR, Madison, Wisconsin, USA). Primer sequences are obtainable upon request.

Analysis of sequence variants
To predict the effect of an amino-acid change on protein function, scores from the following three programs were used: MutationTaster (Schwarz et al., 2010; http://mutationtaster.org/); PolyPhen-2, version 2.2.2 (Adzhubei et al., 2010; http://genetics.bwh.harvard.edu/pph2/); and SIFT (Ng and Henikoff, 2001; http://sift.jcvi.org/). To obtain information on transcription-binding sites that might be altered by the identified mutations, a search was performed of the TRANSFAC public database (Wingender et al., 1996; http://www.gene-regulation.com/pub/databases.html).

To maximize the number of patients included in the sequencing step, no controls were sequenced and publically available datasets were used to calculate the allele frequency of the identified variants. The allele frequency of the identified variants was checked in the 1000 Genomes Project data (Abecasis et al., 2010; Total European Ancestry EUR; http://evs.gs.washington.edu/EVS/), and the Exome Variant Server (European American population; Exome Variant Server, NHLBI GO Exome Sequencing Project, Seattle, Washington, USA (http://evs.gs.washington.edu/EVS/) (November 2015)).

Results
High-quality sequencing data were obtained from (i) 552 patients for exon 1; (ii) 554 patients for exon 2; (iii) 563 patients for exons 3, 4, and 5; and (iv) 571 patients for exon 6. Two variants were identified and verified in one patient each: (i) c.6G>T in exon 1 and (ii) c.598G>A in exon 5. These variants were present in neither the 1000 Genomes Project data nor the Exome Variant Server. No additional variants were identified in our sample.

The mutation c.6G>T in exon 1 is a synonymous substitution, which was in-silico predicted to be disease causing (MutationTaster). According to TRANSFAC (Wingender et al., 1996), the alteration in the DNA sequence abolishes a potential transcription factor-binding site for Sp1. The non-synonymous substitution c.598G>A in exon 5 p.Asp200Asn was predicted to be functionally relevant by Polyphen-2 (probably damaging); SIFT (damaging); and MutationTaster (disease causing).

No parental DNA was available to test whether the mutation c.6G>T in exon 1 was inherited or de novo. The mutation c.598G>A in exon 5 had not been inherited from the respective patient’s mother. However, the patient’s brother was shown to carry the same mutation. No paternal DNA was available for testing. The brother of the patient was diagnosed with recurrent major depression, agoraphobia, and an unspecified eating disorder. No information on the head size of the mutation carriers was available.
Discussion
The synonymous mutation c.6G>T was in-silico predicted to abolish a potential transcription factor-binding site for specificity protein 1 (SP1). The zinc finger transcription factor SP1 is located on chromosome 12q13.13 and regulates the expression of a number of genes by binding to GC-rich sequences (Suske, 1999). Several studies have reported altered SP1 expressions in patients with SCZ compared with controls (Ben-Shachar and Karry, 2007; Fusté et al., 2013; Pinacho et al., 2013, 2014). In a small sample of first-episode psychosis patients, Fusté et al. (2013) found reduced SP1 protein levels in mononuclear cells from peripheral blood (Fusté et al., 2013). Ben-Shachar and Karry (2007) carried out postmortem expression analyses in various human brain regions. The authors identified significantly decreased SP1 messenger RNA levels in the prefrontal cortex and in the striatum, with increased levels in the ventral parieto-occipital cortex and in lymphocytes (Ben-Shachar and Karry, 2007). Pinacho et al. (2014) reported significantly increased SP1 messenger RNA expression levels in the postmortem hippocampus of patients with chronic SCZ (Pinacho et al., 2014).

The mutation c.6G>T is not reported in either the 1000 Genomes Project data or the Exome Variant Server (total of ∼4,700 individuals). In the European (non-Finnish) sample ascertained by the Exome Aggregation Consortium (ExAC) (Cambridge, Massachusetts, USA) (http://exac.broadinstitute.org), the c.6G>T mutation was detected in 43 of 7838 European individuals (allele frequency = 0.0003). Notably, ExAC includes data from the 1000 Genomes Project, the Exome Variant Server [NHLBI GO Exome Sequencing Project, and sequencing studies in patients with psychiatric disorders. The mutation was not detected in the Schizophrenia Exome Sequencing Genebook (Purcell et al., 2014; http://atgu.mgh.harvard.edu/~spurr-cell/genebook/genebook.cgi?user=guest&cmd=search-gene&box=KCTD13)], which contains the exome sequencing data of 2536 SCZ patients and 2543 controls.

The nonsynonymous substitution c.598G>A in exon 5 (p.Asp200Asn) is not reported in the 1000 Genomes Project data or the Exome Variant Server. In the European (non-Finnish) ExAC sample, the mutation c.598G>A was identified in 3 of 33,154 individuals (allele frequency = 0.00005). In the Swedish SCZ exome-sequencing study, this mutation was identified in one patient and in one control (Purcell et al., 2014; http://atgu.mgh.harvard.edu/~spurr-cell/genebook/genebook.cgi?user=guest&cmd=search-gene&box=KCTD13). At the protein level, the G>A substitution causes an exchange of the charged acidic amino acid aspartic acid to the polar uncharged amino acid asparagine. Moreover, Asp200 is strictly conserved at its corresponding position in KCTD13 as far down as Danio rerio (data from Swiss-Prot).

To date, exome sequencing data from eight studies analyzing de-novo mutations in more than 850 patients with SCZ have been published (Girard et al., 2011; Xu et al., 2011, 2012; Gulsuner et al., 2013; Fromer et al., 2014; Guipponi et al., 2014; McCarthy et al., 2014; Kranz et al., 2015). None of these studies reported a mutation in KCTD13.

The present study has three main limitations. First, the sequencing of KCTD13 was restricted only to patients. During the project design phase, we opted to sequence KCTD13 in as many patients as possible, rather than reducing the number of patients to cover the cost of sequencing controls. Second, we could not determine the phenotype of the mutation carriers identified in ExAC. This hampers the interpretation of the allele frequency reported for the two mutations identified in the present analyses. In particular, information on the variant carriers’ head size and their mental wellbeing would have improved the data interpretation. Third, we focused our analysis on exonic variants and therefore cannot rule out the presence of phenotypically relevant mutations in regulatory regions. Furthermore, KCTD13 is an interesting candidate gene on the basis of a study in genetically modified zebrafish (Golzio et al., 2012). Therefore, we cannot rule out that genetic variants in a/several other genes located in 16p11.2 contribute toward the neuropsychiatric phenotype observed among human CNV carriers. Future studies sequencing all genes located in 16p11.2 are warranted to obtain more information on their relevance to disease pathogenesis.

The lack of an association between single base pair mutations in KCTD13 and SCZ, both in the present study and in the previously published exome sequencing data, may indicate that rare point mutations in this gene do not contribute toward the genetic architecture of SCZ, or alternatively, that mutations in this gene are extremely rare. Our study generated no strong evidence for the involvement of damaging mutations in KCTD13 in the development of SCZ. Therefore, the relevance of the identified rare mutations in KCTD13 remains unclear. Further studies in large, independent cohorts are now warranted.

Acknowledgements
The authors thank all patients for participating in this study. They also thank the NHLBI GO Exome Sequencing Project and its ongoing studies, which produced and provided exome variant calls for comparison: the Lung GO Sequencing Project (HL-102923), the WHI Sequencing Project (HL-102924), the Broad GO Sequencing Project (HL-102925), the Seattle GO Sequencing Project (HL-102926), and the Heart GO Sequencing Project (HL-103010). And they would also like to thank the Exome Aggregation Consortium and the groups that provided exome variant data for comparison. A full list of contributing groups can be found at http://exac.broadinstitute.org/about. Furthermore, they thank Rainald Mössner for the provision of DNA samples.

This study was funded by the German Federal Ministry of Education and Research (BMBF) through: (i) the Integrated...


---

**References**


