Citation: McGrath, L.M., Yu, D., Marshall, C., Davis, L.K., Thiruvahindrapuram, B., Li, B., Cappi, C., Gerber, G., Wolf, A., Schroeder, F., Osiecki, L., O'Dushlaine, C., Kirby, A., Illman, C., Haddad, S., Gallagher, P., Fagerness, J. (…) Cook, E., Pauls, D. L., Wang, K., Scharf, J. (2014). Copy number variation in obsessive-compulsive disorder and Tourette syndrome: A cross-disorder study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 53(8), 910-919.
http://www.jaacap.com/article/S0890-8567%2814%2900404-3/abstract

Copy number variation in obsessive-compulsive disorder and Tourette syndrome: A cross-disorder study

Lauren M. McGrath, PhD, Dongmei Yu, MS, Christian Marshall, PhD, Lea K. Davis, PhD, Bhooma Thiruvahindrapuram, MSc, Bingbin Li, HBSc, Carolina Cappi, MSc, Gloria Gerber, BA, Aaron Wolf, BA, Frederick A. Schroeder, PhD, Lisa Osiecki, BA, Colm O'Dushlaine, PhD, Andrew Kirby, BA, Cornelia Illmann, PhD, Stephen Haddad, MS, Patience Gallagher, BS, Jesen A. Fagerness, JD, Cathy L. Barr, PhD, Laura Bellodi, MD, Fortu Benarroch, MD, O. Joseph Bienvenu, MD, PhD, Donald W. Black, MD, Michael H. Bloch, MD, MS, Ruth D. Bruun, MD, Cathy L. Budman, MD, Beatriz Camarena, MSc, Danielle C. Cath, MD, Maria C. Cavallini, MD, Denise A. Chavira, PhD, Sylvain Chouinard, MD, Vladimir Coric, MD, Bernadette Cullen, MB, BCh. BAO, Richard Delorme, MD, PhD, Damiaan Denys, MD, PhD, Eske M. Derks, PhD, Yves Dion, MD, Maria C. Rosário, MD, PhD, Valsama Eapen, MD, PhD, Karin Egberts, MD, Patrick Evans, PhD, Peter Falkai, MD, Thomas Fernandez, MD, Eduardo Fournier, MS, Helena Garrido, MA, Daniel Geller, MD, Simon L. Girard, PhD, Hans J. Grabe, MD, Marco A. Grados, MD, MPH, Benjamin D. Greenberg, MD, PhD, Varda Gross-Tsur, MD, Edna Grünblatt, PhD, Gary A. Heiman, PhD, Sian M.J. Hemmings, PhD, Luis D. Herrera, MD, MPH, Ana G. Hounie, MD, PhD, Joseph Jankovic, MD, James L Kennedy, MD, Robert A. King, MD, Roger Kurlan, MD, Nuria Lanzagorta, PsyD, Marion Leboyer, MD, PhD, James F. Leckman, MD, Leonhard Lennertz, MSc, Christine Lochner, PhD, Thomas L. Lowe, MD, Gholson J. Lyon, MD, PhD, Fabio Macciardi, MD, PhD, Wolfgang Maier, MD, James T. McCracken, MD, William McMahon, MD, Jubel Morgan, RN, Dennis L. Murphy, MD, Allan L Naarden, MD, Benjamin M Neale, PhD, Andrew J. Pakstis, PhD, Michele T. Pato, MD, Carlos N. Pato, MD, PhD, John Piacentini, PhD, Christopher Pittenger, MD, PhD, Yehuda Pollak, PhD, Tobias Renner, MD, Victor I. Reus, MD, Margaret A. Richter, MD, Mary M. Robertson, MD, PhD, Roxana Romero, MA, David Rosenberg, MD, Guy A. Rouleau, MD, Stephan Ruhrmann, MD, Aline S. Sampaio, MD, PhD, Jack Samuels, PhD, Paul Sandor, MD, Brooke Sheppard, ScM, Harvey S. Singer, MD, Jan H. Smit, PhD, Dan J. Stein, MD, PhD, Jay A. Tischfield, PhD, Homero Vallada, MD, PhD, Jeremy Veenstra-VanderWeele, MD, Susanne Walitza, MD, John Walkup, MD, Ying Wang, MSc, Jens R. Wendland, MD, Herman G. Westenberg, PhD[∞], Yin Yao Shugart, PhD, Euripedes C. Miguel, MD, PhD, Humberto Nicolini, MD, PhD, Ben A. Oostra, PhD, Rainald Moessner, MD, Michael Wagner, PhD, Andres Ruiz-Linares, MD, PhD, Peter Heutink, PhD, Gerald Nestadt, MD, Nelson Freimer, MD, Tracey Petryshen, PhD, Danielle Posthuma, PhD, Michael A. Jenike, MD, Nancy J. Cox, PhD, Gregory L. Hanna, MD, Helena Brentani, MD, PhD, Stephen W. Scherer, PhD, Paul D. Arnold, MD, PhD, S. Evelyn Stewart, MD, Carol A. Mathews, MD, James A. Knowles, MD, PhD, Edwin H. Cook, MD, David L. Pauls, PhD, Kai Wang, PhD, Jeremiah M. Scharf, MD, PhD[*]

Affiliations: Drs. McGrath, Illmann, Fagerness, Geller, Jenike, Neale, Petryshen, Pauls, Scharf, Schroeder, Mses.Yu, Gerber, Osiecki, Gallagher, and Messrs. Wolf, Haddad, and Kirby are with Massachusetts General Hospital. Dr. McGrath is also with American University. Drs. O'Dushlaine, McGrath, Neale, Petryshen, Scharf, and Ms. Yu are also with the Harvard-MIT Broad Institute. Drs. Marshall, Scherer, Barr, Richter, Arnold, Sandor, Ms. Thiruvahindrapuram and Mr. Li are with the University of Toronto and the Hospital for Sick Children. Dr. Kennedy is with the Centre for Addiction and Mental Health. Drs. Davis, Evans, and Cox are with the University of Chicago. Ms. Cappi, Drs. Vallada, Miguel, and Brentani are with the University of São Paulo Medical School. Dr. Bellodi is with the Università Vita-Salute San Raffaele. Dr. Benarroch is with Hadassah-Hebrew University Medical Center. Drs. Bienvenu, Cullen, Grados, Samuels, Singer, Nestadt, and Ms. Wang are with Johns Hopkins University School of Medicine. Dr. Black is with the University of Iowa College of Medicine. Drs. Bloch, Coric, Fernandez, Leckman, Pakstis, Pittenger, and King are with the Yale University School of Medicine. Drs. Bruun and Budman are with

North Shore-Long Island Jewish Medical Center. Ms. Camarena is with the Instituto Nacional de Psiquiatría Ramon de la Fuente Muñiz. Dr. Cath is with Utrecht University. Dr. Cavallini is with the Ospedale San Raffaele. Dr. Chavira is with UCLA and UCSD. Dr. Chouinard, Dr. Dion, and Mr. Girard are with the University of Montreal. Drs. Delorme and Leboyer are with the Robert Debre University Hospital and the French National Science Foundation. Dr. Denys is with the Netherlands Institute for Neuroscience. Drs. Denys and Derks are with the Academic Medical Center, Amsterdam. Drs. Rosário, Hounie, and Sampaio are with the Federal University of São Paulo. Dr. Eapen is with the University of New South Wales. Drs. Egberts and Renner are with the University of Würzburg. Dr. Falkai is with the University of Munich. Mr. Fournier, Dr. Herrera, and Ms. Romero and Garrido are with the Hospital Nacional de Niños. Dr. Grabe is with the University Medicine Greifswald. Dr. Greenberg is with Brown Medical School. Drs. Gross-Tsur and Pollak are with Shaare Zedek Medical Center. Drs. Grünblatt and Walitza are with the University of Zurich. Drs. Heiman and Tischfield are with Rutgers University. Dr. Jankovic is with Baylor College of Medicine. Dr. Kurlan is with the Atlantic Neuroscience Institute. Drs. Lanzagorta and Nicolini are with the Carracci Medical Group. Mr. Lennertz, and Drs. Maier, Moessner, and Wagner are with the University of Bonn. Drs. Lochner and Hemmings are with the University of Stellenbosch. Drs. Lowe, Reus, Mathews and Ms. Sheppard are with the UCSF School of Medicine. Dr. Lyon is with the Cold Spring Harbor Laboratory. Dr. Macciardi is with University of California Irvine. Drs. McCracken, Piacentini and Freimer are with the UCLA School of Medicine. Dr. McMahon and Ms. Morgan are with the University of Utah. Drs. Murphy, Wendland and Shugart are with the NIMH Intramural Research Program. Dr. Naarden is with Medical City Dallas Hospital. Drs. M. Pato, C. Pato, Knowles, and Wang are with the Zilkha Neurogenetic Institute. Dr. Richter is with Sunnybrook Health Sciences Centre. Drs. Robertson and Ruiz-Linares are with University College London. Dr. Rosenberg is with Wayne State University and Detroit Medical Center. Dr. Rouleau is with the Montreal Neurological Institute. Dr. Ruhrmann is with the University of Cologne. Dr. Stein is with the University of Cape Town. Dr. Veenstra-VanderWeele is with Vanderbilt University. Dr. Walitza is with the University of Würzburg. Dr. Walkup is with Weill Cornell Medical Center. Dr. Oostra is with Erasmus Medical Center. Dr. Heutink is with the German Center for Neurodegenerative Diseases. Drs. Posthuma and Smit are with VU University Amsterdam. Dr. Hanna is with the University of Michigan. Dr. Stewart is with the University of British Columbia. Dr. Cook is with the University of Illinois at Chicago. Dr. Westenberg was with the University Medical Center Utrecht.

[∝]Deceased

[*]Corresponding author:
Jeremiah Scharf, M.D., Ph.D.
Psychiatric & Neurodevelopmental Genetics Unit
Massachusetts General Hospital
185 Cambridge Street
Boston, MA 02114
TEL: 617-643-3139
FAX: 617-726-0830

Running Title: Copy number variation in OCD and TS

Keywords: Tourette syndrome, obsessive-compulsive disorder, copy number variation, genetics, 16p13.11

Word count:
Abstract: 225
Acknowledgments: 115
Manuscript: 6000
Figures: 1; Tables: 3; Supplementary Material: 1 Word document, 5 tables

Acknowledgments: We thank the families, patients, and control volunteers who participated in this research. This work was supported by grants from the David Judah Fund, the Tourette Syndrome Association, the International OCD Foundation, NIH grant U01NS40024 (DLP/JMS/TSA International Consortium for Genetics), R01NS16648, R01MH079489, and MH073250 (DLP), K23MH085057 (JMS), T32MH16259 (LMM), NS037484 and P30NS062691 (NBF), K20MH01065 and R01MH58376 (GLH), R01MH092293 (GAH/RAK/JAT), American Recovery and Re-investment Act (ARRA) awards NS40024-07S1 and NS16648-29S1 (DLP). Additional acknowledgements are provided as supplementary material. None of the funding agencies for this project had any influence on a) the design or conduct of the study; b) management, analysis or interpretation of the data; c) preparation, review or approval of the manuscript.

ABSTRACT

Objective: Obsessive-compulsive disorder (OCD) and Tourette syndrome (TS) are heritable, neurodevelopmental disorders with a partially shared genetic etiology. This study represents the first genome-wide investigation of large (>500kb), rare (<1%) copy number variants (CNVs) in OCD and the largest genome-wide CNV analysis in TS to date.

Methods: The primary analyses utilized a cross-disorder design in 2699 cases (1613 ascertained for OCD, 1086 ascertained for TS) and 1789 controls. Parental data facilitated a *de novo* analysis in 348 OCD trios.

Results: Although no global CNV burden was detected in the cross-disorder analysis or in secondary, disease-specific analyses, there was a 3.3-fold increased burden of large deletions previously associated with other neurodevelopmental disorders (p=.09). Half of these neurodevelopmental deletions were located in a single locus, 16p13.11 (5 case deletions: 0 control deletions, p=0.08 in current study, p=0.025 compared to published controls). Three 16p13.11 deletions were confirmed *de novo*, providing further support to the etiological significance of this region. The overall OCD *de novo* rate was 1.4%, which is intermediate between published rates in controls (0.7%) and in autism or schizophrenia (2-4%).

Conclusions: Several converging lines of evidence implicate 16p13.11 deletions in OCD, with weaker evidence for a role in TS. The trend toward increased overall neurodevelopmental CNV burden in TS and OCD suggests that deletions previously associated with other neurodevelopmental disorders may also contribute to these phenotypes.

INTRODUCTION

Obsessive-compulsive disorder (OCD) and Tourette syndrome (TS) are neurodevelopmental disorders with significant phenotypic and genetic overlap.[1,2] One promising avenue for identifying cross-disorder genetic risk factors in neurodevelopmental disorders is the study of genomic copy number variants (CNVs), segments of DNA ranging from 1 kilobase to several megabases that show deletions or duplications compared to a reference.[3] The association of large, rare CNVs with neurodevelopmental disorders including autism spectrum disorders (ASD), schizophrenia, and intellectual disability (ID) has been one of the most important recent advances in psychiatric genomics.[4] CNVs predisposing to these disorders overlap substantially, highlighting the cross-disorder effects of this class of genetic variation.[5,6] Given this robust literature, an important, unanswered question is whether large, rare CNVs are also relevant for the genetic architectures of OCD and TS.

Both OCD and TS are highly heritable, and have long been suspected to share genetic liability, though specific gene variants have been difficult to identify.[7-9] Both disorders frequently co-occur in individuals,[10] and there is evidence for shared OCD/TS genetic risk from family studies,[9,11] with genetic correlation estimates ranging from 41-90%.[2,12]. In OCD, locus-specific CNV analyses have been reported,[13,14] but no prior genome-wide CNV analysis has been performed. In TS, the three previous genome-wide surveys of CNVs have been limited by small sample sizes (<500 cases), and results differ with regard to whether there is an increased CNV burden in TS compared to controls.[15-17] No specific CNV region has received strong statistical support across studies, although exonic *NRXN1* deletions have been identified in two studies.[15,17]

Given the evidence for shared genetic underpinnings of OCD and TS and cross-disorder effects of specific neurodevelopmental CNVs, along with the need for large samples when investigating rare events, we chose a cross-disorder design that combined OCD and TS samples into a single case group, with follow-up analyses examining the individual disorders. This study is the first genome-wide CNV analysis in OCD and the largest to date in TS, and addressed three key questions. First, is there an increased burden of large, rare CNVs in OCD/TS? Second, are the recurrent and/or *de novo* CNVs implicated in other neurodevelopmental disorders also etiologically relevant for OCD/TS? Third, is there evidence of association between any specific genomic region and OCD/TS?

METHOD

Participants

Individuals with OCD or TS were recruited for a multi-center collaborative GWAS (described in[18] and[19]). Participants ages 18 and older provided written, voluntary informed consent for participation in genetic studies. Individuals under age 18 provided assent; written parental consent was also obtained. The study was approved by the Ethics Committees of all participating sites. Recruitment sites varied in screening and exclusions related to other neurodevelopmental disorders; see Tables S9, S10, S11 for available clinical information regarding ID, ASD, ADHD, and seizures. OCD and TS samples were collected independently, but were genotyped jointly to facilitate cross-disorder analyses. All cases were genotyped on the Illumina Human610-Quadv1_B platform.

OCD. The initial OCD sample consisted of 1565 cases and 437 parent-child trios (n=406 independent families, 31 affected siblings) recruited from 22 sites in the US, Canada, Europe, Latin America, and South Africa, predominantly through OCD specialty clinics. In total, 1971 independent OCD cases (including trio probands) were eligible for analysis. 1613 OCD cases (82%) survived quality control (QC), and were included in the final analyses. Mean age of OCD symptom onset was 13.8 years (SD=9.1). 327 cases and 21 affected siblings had parents available for *de novo* analysis (N=348 total trios). TS or chronic tics (CT) were assessed in 57% of OCD probands using DSM-IV-TR criteria. Of those assessed, TS was present in 10% of OCD cases, and an additional 5% had CT.

TS. The initial TS sample consisted of 1235 individuals recruited from 19 sites in the US, Canada, Europe, and Israel. Participants with DSM-IV-TR-diagnosed TS were recruited primarily from TS specialty clinics or from the Tourette Syndrome Association (TSA). 1086 (88%) individuals passed sample-level QC. Mean age of tic onset was 6.3 years (SD=3.5). OCD as defined by DSM-IV-TR criteria was assessed in 88% of cases; OCD was present in 46% of assessed TS individuals.

Controls. Ancestry-matched controls (N=720) were collected in parallel with their respective cases for the French-Canadian (n=269), German (n=224), South African (n=188), and Dutch (n=39) samples. These controls were screened for TS and OCD and genotyped with cases on the Illumina Human610-Quadv1_B[20,21], (referred to here as ‘Hap610 controls’).

1279 additional European-ancestry controls were obtained through the Database of Genotypes and Phenotypes (dbGaP) from the Studies of Addiction: Genetics and Environment (SAGE) cohort.[20] SAGE controls were excluded for lifetime substance dependence but were not screened for other psychopathology. The SAGE controls were genotyped on the Illumina Human-Hap1Mv1_C (referred to here as 'Hap1M controls').

CNV Calling and Quality Control (QC)

Data from the Hap610 (cases and controls) and Hap1M (controls) platforms were processed and cleaned separately using standard procedures (Supplementary Methods). CNV calls were generated with PennCNV (version 2010-05-01)[21] and iPattern[22,23] using hg18 genomic coordinates. Analyses were limited to autosomal events. Trio analyses utilized the trio functions in PennCNV to improve calling accuracy and to estimate the likelihood of a *de novo* event.[21]

Both sample and CNV-specific QC was conducted by examining distributions of QC metrics informed by comparable published CNV analyses, e.g.,[22,24,25]. Since distributions were similar for calls from PennCNV and iPattern, the same QC thresholds were used for both algorithms to maximize comparability (Supplementary Methods). QC-filtered PennCNV and iPattern callsets were merged at the sample level using CNVision (http://futo.cs.yale.edu/mw/index.php/CNVision). Only calls with >50% overlap based on the union of the CNV region were included in the analysis. Overlap percentages were higher for the Hap1M (86%-98%) compared to the Hap610 chip (59%–82%) (Table S1). Due to the presence of a batch effect within the Hap610 samples, analyses were restricted to large (>500kb) events, the size at which batch effects were no longer observed (Figure S1, Table S2). All CNVs were also filtered for rare events (<1% frequency in the Database of Genomic Variants).

Ancestry matching

The case-control sample was predominantly composed of individuals self-reporting European (EU) ancestry (n=4410), but did include a small number of individuals from Brazil, Mexico, and Costa Rica (n=78) to maximize the power to detect rare events. However, a sensitivity analysis restricting to genetically-defined EU ancestry (n=4276) via multidimensional scaling (Figures S2, S3) confirmed that results were not biased by population stratification.

Statistical Analysis

CNV burden, region-specific analyses, and permutations were performed in PLINK using the rare CNV functions.[26] The primary case-control analyses grouped OCD and TS cases vs. Hap610 and Hap1M controls to maximize sample sizes. No additional covariates were included, though follow-up analyses were stratified by EU ancestry. To evaluate whether OCD and TS patients harbor large, pathogenic CNVs that have been repeatedly implicated in other neurodevelopmental disorders, we assembled a curated list of CNVs drawn from the ASD, schizophrenia, and ID literature, including 47 regions of interest (all >500kb)[22,24,27,28] (Table S3).

Quantitative PCR (qPCR) Validation

Validation of neurodevelopmental or putative *de novo* CNVs was performed with SYBR green qPCR. Two qPCR primers per CNV were designed against NCBI build hg18 sequence to obtain converging evidence for the called event. If one primer pair failed or gave ambiguous results, an additional primer pair was run to resolve the discrepancy (Supplementary Methods).

RESULTS

After filtering and QC, the final sample consisted of 2699 cases (1086 ascertained for TS and 1613 ascertained for OCD), and 1789 controls.

Overall CNV Burden Analysis

There was no significant difference in burden of large, rare CNVs between OCD/TS cases and controls for CNV rate (average number of CNVs per person), CNV proportion (proportion of samples carrying ≥1 CNV), gene rate (the average number of genes spanned by a CNV), or by restricting to CNVs containing exons (Tables 1, S4, S5, S6). Similarly, no increased CNV burden was identified in secondary, disorder-specific analyses (OCD vs. controls or TS vs. controls) (Tables S7, S8).

Neurodevelopmental CNV Burden Analysis

Given that various neurodevelopmental disorders have previously been associated with large, rare, recurrent CNVs in specific regions of the genome[4], we examined 47 known, pathogenic neurodevelopmental loci for an excess of large, rare CNVs in OCD/TS cases compared to controls (Tables 2, 3, S3, S9). We found a 3.3-fold, trend-level increase in large deletions overlapping these loci for TS/OCD cases (p=.09; Table 2). In contrast, there was no enrichment of duplication events (case/control ratio 1.16, p=0.46) and no difference in overall CNV size within these regions (p=.31).

In disorder-specific analyses, the neurodevelopmental deletion burden was larger in OCD (case/control ratio=4.44, p=0.04, one-sided) than in TS (case/control ratio=1.65, p=0.49, one-sided) (Tables S7, S8). The most frequently observed neurodevelopmental CNVs were located at 16p13.11, 22q11, and *PARK2* (Figure 1, Tables 3, S4, S5).

Laboratory Validation

We confirmed 10 of 11 neurodevelopmental deletion events (91%) with qPCR; the one unconfirmed deletion was near the 9q34 telomere and was excluded from the neurodevelopmental burden analysis (Supplementary Methods). We also confirmed 12 of 14 duplication events (86%) with 1 or 2 sets of primers. The two remaining duplications could not be confirmed but had qPCR results trending towards duplication. Of note, the LRR and BAF plots (Figure S6) strongly supported all CNV events with the exception of the unconfirmed 9q34 deletion.

Region specific analyses

Half of the 10 large neurodevelopmental deletions were in the same genomic region on 16p13.11 (case:control ratio=5:0, one-sided Fisher's exact p=0.08) (Figure 1, Table 3). Using previously published estimates to more accurately calculate the control rate for 16p13.11 deletions (3/8329)[28], we found a statistically significant excess in the OCD/TS cases (one-sided Fisher's exact p=0.025). Of note, the rate of 16p13.11 deletions in this sample (0.19%) was comparable to published rates from large samples of children with neurodevelopmental disorders (N's ~15,000) referred for genetic testing (0.11%-0.14%).[28,29] Interestingly, the clinical phenotype of the five 16p13.11 deletions did not respect traditional diagnostic boundaries: three subjects had OCD without tics, one

had TS without OCD, and one had OCD+CT. None of the cases had ASD, ID, or a seizure disorder (Table S10). All 16p13.11 deletions were validated with qPCR.

16p13.11 was also the top recurrent region in genome-wide, region-specific analyses combining deletions and duplications (case:control ratio=7:1, one-sided empirical p=0.13; genome-wide permutation corrected p=0.86). This region-specific effect was driven by deletions; no excess of 16p13.11 duplications was found in OCD/TS cases compared either to sample controls (case:control ratio=2:1, one-sided Fisher's exact p=0.65) or to published controls (10/8329)[28] (one-sided Fisher's exact p=0.83).

We also examined genome-wide, region-specific associations in each disorder separately, combining deletions and duplications. In OCD, 16p13.11 again emerged as the locus with the most notable case:control excess (6:1; one-sided empirical p=0.046, genome-wide permutation-corrected p=0.35). In TS, the 3p26.3 region had the largest case:control excess (7:2 ratio of duplications 50kb upstream of *CNTN6*; one-sided empirical p=0.018, genome-wide permutation-corrected p=0.15); exonic *CNTN6* regions had a more equivocal case-control ratio (8:6; one-sided empirical p=0.11, genome-wide permutation-corrected p=0.99) (Figure S7).

***De novo* analyses**

The OCD parent-proband trios (total trios=348) were examined for the presence of large (>500kb), rare, *de novo* CNVs. We detected five high confidence, *de novo* CNVs at 4q24, 7p21.1-7p21.2, 16p13.11, 17q12, and 22q11.21, resulting in a *de novo* rate of 1.44% (Figure S8). Three of these CNVs were in known pathogenic neurodevelopmental loci: 16p13.11, 17q12, 22q11.21 (Table S3). All 5 events were validated *in silico* and by qPCR (Table S11).

Given that 16p13.11 contained both a *de novo* CNV and the largest case/control difference across the genome, we undertook further investigation of the *de novo* status of other 16p13.11 case CNVs in our sample. Parental DNA was available for one of the five 16p13.11 deletions (TS only), which we confirmed as *de novo* using qPCR. We also re-examined trios removed during QC for evidence of large 16p13.11 events and found one additional deletion (OCD only) with a statistically significant *in silico* probability of being *de novo*, $p=5.68 \times 10^{-14}$ that we subsequently validated using qPCR (Figure S8). This increased the total number of 16p13.11 deletions from five to six, three of which were *de novo* (2 OCD only, 1 TS only) (Table S10). The *de novo* status of two 16p13.11 deletions could not be determined because parental DNA was not available; one deletion was inherited.

DISCUSSION

In this genome-wide analysis of large, rare CNVs in OCD and TS, although there was no global increase in CNV burden, we did find suggestive evidence for an increased burden of known, pathogenic neurodevelopmental deletions in OCD/TS cases compared to controls. The 3.3-fold increased risk associated with this finding only reached trend-level significance, potentially due to the conservative bias toward the null introduced by having the majority of controls genotyped on a more sensitive, higher density genotyping array than cases.

Deletions at 16p13.11, which contributed disproportionately to the neurodevelopmental burden, have been implicated in a wide range of disorders, including ID/developmental delay[28,29], seizures[30,31], and, less strongly, ASD.[32] The confirmation of three *de novo* events among our six OCD/TS 16p13.11 case deletions, as well as the absence of co-morbid ID, seizures, or ASD in the cases assessed, suggests that these events may be pathogenic in our sample and that the phenotypic spectrum of 16p13.11 deletions should be expanded to include OCD and TS. Importantly, the phenotypic profiles indicate that 16p13.11 deletions are primarily associated with OCD (4 OCD only, 1 OCD+CT, 1 TS only). The presence of a TS case without OCD raises the possibility of a pleiotropic effect of this locus, though this hypothesis remains preliminary as it is based only on a single case. It is also likely that additional genetic and environmental factors shape the ultimate phenotypic outcome of these CNV events, including patterns of comorbidity.

Three of the five large *de novo* CNVs reported in this study were located in regions previously associated with other neurodevelopmental disorders. The *de novo* events at 16p13.11 and 22q11 also had additional supporting case events in the same locus, whereas the deletion at 17q12 was a singleton event in an OCD case. The final two *de novo* events were singleton deletions at novel loci: 4q24 and 7p21.1-7p21.2. The clinical significance of both events remains unclear, although pathogenic CNVs have been documented in both regions (www.iscaconsortium.org),[33] including a case report of a patient with Saethre-Chotzen syndrome and co-occurring TS and OCD.[34]

The overall *de novo* rate in the OCD trio sample was 1.44% for large CNVs (> 500kb), which is intermediate between estimates in healthy controls (0.7%)[27] and estimates in ASD (1.8% multiplex, 3.9% simplex)[27] and schizophrenia (2-3%).[35-37] Additional studies with larger samples and more sensitive CNV calling will be needed to refine this estimate.

Previous CNV studies have implicated *NRXN1* deletions in TS.[15,17] We detected one 600kb *NRXN1* deletion in an OCD case (TS status unknown) (chr2:50185814-50799877, hg18) that was called by iPattern and qPCR validated, although it did not pass initial QC because <50% of the region was called by PennCNV. We also observed three 22q11 case duplications, all with OCD (two OCD only, one OCD+CT), one *de novo* deletion (OCD only), and 1 control duplication (Figure S4). The *de novo* 22q11 deletion was smaller (~700kb) than the canonical 1.5-3Mb 22q11 deletion associated with velocardiofacial and DiGeorge Syndrome (OMIM 192430, 188400), whereas the duplications ranged in size from 700kb-2MB. Although the 3:1 22q11 CNV duplication excess in our sample is not significant, it is notable that three other 22q11 duplications, including one *de novo* event, have previously been reported in TS cases[16,17,38,39] and thus this region warrants further study.

The results of this study should be interpreted in the context of some limitations. First, the majority of controls were genotyped on a higher resolution array (Hap1M) than cases (Hap610), resulting in a conservative bias toward the null hypothesis due to better CNV detection in controls than cases. This effect is evident when comparing the higher CNV rates in Hap610 controls to Hap1M controls (Table 1). However, a comparison of Hap610 cases to Hap610 controls did not reveal overall burden differences within the limits of this restricted sample size (Table S12).

Second, we were unable to call CNVs smaller than 500kb due to genotyping batch effects. Although previous research has shown that >500kb events are most likely to be pathogenic,[22,25] we may have missed smaller pathogenic CNVs in this sample. Third, albeit the largest for OCD and TS to date, our sample is still small compared to large-scale investigations of CNVs in other disorders.[4,28,29] For this reason and because the number of OCD/TS cases with rare, pathogenic CNVs appears to be small, we recommend caution in interpreting these results pending further studies in larger OCD/TS samples that can refine the global and neurodevelopmental CNV burden estimates.

Fourth, missing data on TS/CT and OCD comorbidity in some individuals prevented us from dividing cases into mutually exclusive subgroups (TS only, OCD only, OCD+TS) for analysis. Instead, we identified CNVs in the combined OCD/TS sample and then reviewed the diagnostic profile of each case with a CNV. Moreover, TS and OCD participants were not universally screened for other neurodevelopmental disorders, though we documented this information when available (Tables S9, S10, S11). Without comprehensive screening, we cannot exclude two possibilities regarding cases with neurodevelopmental CNVs: (1) the primary TS/OCD diagnosis was misclassified (i.e., stereotypies or restricted interests/repetitive behaviors in the context of ASD were misdiagnosed as TS or

OCD, respectively) and (2) cases with complex comorbidities were more likely to harbor neurodevelopmental CNVs. All assessments were completed by internationally-recognized expert clinicians, reducing the likelihood of misclassification. However, some individuals with neurodevelopmental CNVs may have had subtle or unassessed ASD, ADHD, cognitive impairment, or psychotic symptoms.

SUMMARY

These results suggest that deletions associated with other neurodevelopmental disorders may also contribute to OCD and TS. Converging lines of evidence specifically implicate 16p13.11 deletions, with stronger evidence for OCD than TS. While it is premature to make clinical recommendations based on these observations, we note that tic and obsessive-compulsive symptoms often occur in the context of other neurodevelopmental disorders, such as ASD and ID, where practice parameters do recommend chromosomal microarray testing.[40] Future studies should help to refine clinical guidelines as to whether CNV testing might be indicated for children with TS and/or OCD in general or be restricted to those with multiple co-occurring neurodevelopmental disorders.

REFERENCES

1. Grados MA. The genetics of obsessive-compulsive disorder and Tourette's syndrome: what are the common factors? *Curr Psychiatry Rep.* 2009;11:162-166.
2. Davis LK, Yu D, Keenan CL, et al. Partitioning the heritability of Tourette syndrome and obsessive compulsive disorder reveals differences in genetic architecture. *PLoS genetics.* Oct 2013;9(10):e1003864.
3. Cook Jr. EH, Scherer SW, Cook EH. Copy-number variations associated with neuropsychiatric conditions. *Nature.* 2008;455:919-923.
4. Malhotra D, Sebat J. CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell.* 2012;148:1223-1241.
5. Thapar A, Cooper M. Copy number variation: what is it and what has it told us about child psychiatric disorders? *Journal of the American Academy of Child and Adolescent Psychiatry.* 2013;52:772-774.
6. Pescosolido MF, Gamsiz ED, Nagpal S, Morrow EM. Distribution of disease-associated copy number variants across distinct disorders of cognitive development. *Journal of the American Academy of Child and Adolescent Psychiatry.* 2013;52:414-430.e414.
7. van Grootheest DS, Cath DC, Beekman AT, Boomsma DI. Twin studies on obsessive-compulsive disorder: a review. *Twin Res Hum Genet.* 2005;8:450-458.
8. Bolton D, Rijsdijk F, O'Connor TG, Perrin S, Eley TC. Obsessive-compulsive disorder, tics and anxiety in 6-year-old twins. *Psychological medicine.* 2007;37:39-48.
9. O'Rourke JA, Scharf JM, Yu D, Pauls DL. The genetics of Tourette syndrome: a review. *Journal of psychosomatic research.* 2009;67:533-545.
10. Scharf JM, Miller LL, Mathews CA, Ben-Shlomo Y. Prevalence of Tourette syndrome and chronic tics in the population-based Avon longitudinal study of parents and children cohort. *Journal of the American Academy of Child and Adolescent Psychiatry.* 2012;51:192-201 e195.
11. Pauls DL. The genetics of obsessive-compulsive disorder: a review. *Dialogues in clinical neuroscience.* 2010;12(2):149-163.

12. Mathews CA, Grados MA. Familiality of Tourette syndrome, obsessive-compulsive disorder, and attention-deficit/hyperactivity disorder: heritability analysis in a large sib-pair sample. *Journal of the American Academy of Child and Adolescent Psychiatry.* 2011;50:46-54.

13. Delorme R, Moreno-De-Luca D, Gennetier A, et al. Search for copy number variants in chromosomes 15q11-q13 and 22q11.2 in obsessive compulsive disorder. *BMC Med Genet.* 2010;11:100.

14. Walitza S, Bove DS, Romanos M, et al. Pilot study on HTR2A promoter polymorphism, -1438G/A (rs6311) and a nearby copy number variation showed association with onset and severity in early onset obsessive-compulsive disorder. *Journal of Neural Transmission.* 2012;119:507-515.

15. Sundaram SK, Huq AM, Wilson BJ, Chugani HT. Tourette syndrome is associated with recurrent exonic copy number variants. *Neurology.* 2010;74:1583-1590.

16. Fernandez TV, Sanders SJ, Yurkiewicz IR, et al. Rare copy number variants in tourette syndrome disrupt genes in histaminergic pathways and overlap with autism. *Biological psychiatry.* 2012;71:392-402.

17. Nag A, Bochukova EG, Kremeyer B, et al. CNV analysis in Tourette syndrome implicates large genomic rearrangements in COL8A1 and NRXN1. *PLoS One.* 2013.

18. Scharf JM, Yu D, Mathews CA, et al. Genome-wide association study of Tourette's syndrome. *Molecular psychiatry.* Jun 2013;18(6):721-728.

19. Stewart SE, Yu D, Scharf JM, et al. Genome-wide association study of obsessive-compulsive disorder. *Molecular psychiatry.* Jul 2013;18(7):788-798.

20. Bierut LJ, Agrawal A, Bucholz KK, et al. A genome-wide association study of alcohol dependence. *Proceedings of the National Academy of Sciences of the United States of America.* Mar 16 2010;107(11):5082-5087.

21. Wang K, Li M, Hadley D, et al. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome research.* 2007;17:1665-1674.

22. Pinto D, Pagnamenta AT, Klei L, et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature.* 2010;466:368-372.

23. Pinto D, Darvishi K, Shi X, et al. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. *Nature biotechnology.* 2011;29:512-520.

24. Williams NM, Zaharieva I, Martin A, et al. Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: a genome-wide analysis. *Lancet.* 2010;376:1401-1408.

25. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature.* 2008;455:237-241.

26. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics.* 2007;81:559-575.

27. Sanders SJ, Ercan-Sencicek AG, Hus V, et al. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron.* 2011;70:863-885.

28. Cooper GM, Coe BP, Girirajan S, et al. A copy number variation morbidity map of developmental delay. *Nature Genetics.* 2011;43:838-846.

29. Kaminsky EB, Kaul V, Paschall J, et al. An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. *Genet Med.* 2011;13:777-784.

30. de Kovel CG, Trucks H, Helbig I, et al. Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. *Brain.* 2010;133:23-32.

31. Heinzen EL, Radtke RA, Urban TJ, et al. Rare deletions at 16p13.11 predispose to a diverse spectrum of sporadic epilepsy syndromes. *American journal of human genetics.* 2010;86:707-718.

32. Ullmann R, Turner G, Kirchhoff M, et al. Array CGH identifies reciprocal 16p13.1 duplications and deletions that predispose to autism and/or mental retardation. *Human Mutation.* 2007;28:674-682.

33. Firth HV, Richards SM, Bevan AP, et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *American journal of human genetics.* 2009;84:524-533.

34. Kerbeshian J, Burd LJ, Martsolf JT. Psychiatric comorbidity in a girl with Saethre-Chotzen syndrome. *Psychiatric Genetics.* 2008;18:258-259.

35. Kirov G, Pocklington AJ, Holmans P, et al. De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Molecular psychiatry.* 2012;17:142-153.

36. Malhotra D, McCarthy S, Michaelson JJ, et al. High frequencies of de novo CNVs in bipolar disorder and schizophrenia. *Neuron.* 2011;72:951-963.

37. Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M. Strong association of de novo copy number mutations with sporadic schizophrenia. *Nature Genetics.* 2008;40:880-885.

38. Robertson MM, Shelley BP, Dalwai S, Brewer C, Critchley HD. A patient with both Gilles de la Tourette's syndrome and chromosome 22q11 deletion syndrome: clue to the genetics of Gilles de la Tourette's syndrome? *Journal of Psychosomatic Research.* 2006;61:365-368.

39. Clarke RA, Fang ZM, Diwan AD, Gilbert DL. Tourette syndrome and klippel-feil anomaly in a child with chromosome 22q11 duplication. *Case Rep Med.* 2009;2009:361518.

40. Miller DT, Adam MP, Aradhya S, et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *American journal of human genetics.* 2010;86:749-764.

Table 1. Global burden analysis of large (>500kb), rare (<1%) CNVs in cases ascertained for OCD or TS compared to controls.

Description		Combined OCD/TS N=2699	Hap610 Controls N=561	Hap1M Controls N=1228	Case/ Control Ratio	P-value[a]
All CNVs	#CNVs	186	39	111		
> 500kb	Rate	0.069	0.070	0.090	0.82	0.97
	Proportion	0.067	0.068	0.086	0.84	0.95
	Gene Rate	7.23	5.16	6.25	1.21	0.19
Deletions	#CNVs	60	7	31		
> 500kb	Rate	0.022	0.012	0.025	1.05	0.45
	Proportion	0.022	0.012	0.025	1.03	0.48
	Gene Rate	6.29	4.00	4.39	1.46	0.14
Duplications	#CNVs	126	32	80		
>500kb	Rate	0.047	0.057	0.065	0.75	0.99
	Proportion	0.046	0.055	0.063	0.75	0.99
	Gene Rate	7.63	5.42	6.75	1.20	0.26

[a]one-sided, empirical p-value

Rate: average # CNVs per person, Proportion: proportion of samples with ≥1 CNV, Gene Rate: average # genes spanned by CNVs.

Table 2. Neurodevelopmental burden analysis of large, rare CNVs in OCD or TS cases compared to controls.

Description		Combined OCD/TS N=2699	Hap610 Controls N=561	Hap1M Controls N=1228	Case/ Control Ratio	P-value[a]
All CNVs	#CNVs	24	2	8		
> 500kb	Rate/Proportion[b]	0.0089	0.0036	0.0065	1.59	0.14
Deletions	#CNVs	10	0	2		
> 500kb	Rate/Proportion[b]	0.0037	0	0.0016	3.31	0.09
Duplications	#CNVs	14	2	6		
>500kb	Rate/Proportion[b]	0.0052	0.0036	0.0049	1.16	0.46

[a]one-sided, empirical p-value

[b]rate and proportion are identical because no samples had >1 large, neurodevelopmental CNV

Table 3. Large, rare CNVs in cases and controls overlapping previously identified neurodevelopmental loci.

		Deletions			Duplications		
Chr	Region	Cases (OCDonly/TSonly/OCD+TS/CT) N=2699	Controls N=1789	p-value[1]	Cases (OCDonly/TSonly/OCD+TS/CT) N=2699	Controls N=1789	p-value[1]
2	2p15-16.1	0	0	--	0	1	1.00
2	2q11.2	1 (0/0/1)	0	0.60	0	0	--
3	CNTN4	1 (0/0/1)	1	0.84	1 (0/0/1)	0	0.60
3	3q29	0	0	--	0	1	1.00
6	PARK2	1 (1/0/0)	0	0.60	3 (2/0/1)	0	0.22
7	7q11.23	0	0	--	1 (0/0/1)	0	0.60
12	12q14	0	0	--	0	1	1.00
15	15q11-q13	0	0	--	2 (0/0/2)	0	0.36
15	15q24	0	1	1.00	0	0	--
16	16p13.11	5 (3/1/1)	0	0.08	2 (2/0/0)	1	0.65
16	16p11.2	0	0	--	1 (1/0/0)	2	0.94
17	NF1	0	0	--	1 (0/1/0)	0	0.60
17	17q12	1 (1/0/0)	0	0.60	0	1	1.00
22	22q11.21	1 (1/0/0)	0	0.60	3 (2/0/1)	1	0.48

[1]Fisher's exact 1-sided p-value

FIGURE LEGENDS

Figure 1. 16p13.11 region with 8 case and 1 control CNVs. Red, deletions; Blue, duplications. Gray denotes the *de novo* deletion that did not survive strict QC filters but was validated *in silico* and by qPCR.