Molecular Psychiatry (2013) 18, 788 – 798

www.nature.com/mp

ORIGINAL ARTICLE

Genome-wide association study of obsessive-compulsive disorder

SE Stewart[1,2,3,70], D Yu[1,2,70], JM Scharf[1,2,4,5,7,70], BM Neale[1,6,7,70], JA Fagerness[1,2,70], CA Mathews[8,70], PD Arnold[9,10,70], PD Evans[11], ER Gamazon[11], L Osiecki[1,2], L McGrath[1,2], S Haddad[1,2], J Crane[1,2], D Hezel[1,2], C Illman[1,2], C Mayerfeld[1,2], A Konkashbaev[11], C Liu[12], A Pluzhnikov[11], A Tikhomirov[11], CK Edlund[13,14], SL Rauch[15], R Moessner[16], P Falkai[17], W Maier[16], S Ruhrmann[18], H-J Grabe[19], L Lennertz[16], M Wagner[16], L Bellodi[20], MC Cavallini[21], MA Richter[10,22], EH Cook Jr[23], JL Kennedy[10,24], D Rosenberg[25,26], DJ Stein[27], SMJ Hemmings[28], C Lochner[28], A Azzam[8], DA Chavira[29], E Fournier[30], H Garrido[30], B Sheppard[8], P Umaña[30], DL Murphy[31], JR Wendland[31,32], J Veenstra-VanderWeele[33], D Denys[34], R Blom[34], D Deforce[35], F Van Nieuwerburgh[35], HGM Westenberg[34], S Walitza[36], K Egberts[37], T Renner[37], EC Miguel[38], C Cappi[38], AG Hounie[38], M Conceição do Rosário[38], AS Sampaio[38,39], H Vallada[38], H Nicolini[40,41], N Lanzagorta[41], B Camarena[42], R Delorme[43,44,45], M Leboyer[43,44,45], CN Pato[46], MT Pato[46], E Voyiaziakis[46], P Heutink[47], DC Cath[48,49], D Posthuma[47], JH Smit[49], J Samuels[50], OJ Bienvenu[50], B Cullen[50], AJ Fyer[51,52], MA Grados[50], BD Greenberg[53], JT McCracken[54], MA Riddle[50], Y Wang[50], V Coric[55], JF Leckman[56], M Bloch[55], C Pittenger[55], V Eapen[57], DW Black[58], RA Ophoff[59,60], E Strengman[59,60], D Cusi[61,62], M Turiel[63], F Frau[61,64], F Macciardi[65], JR Gibbs[66], MR Cookson[66], A Singleton[66] for the North American Brain Expression Consortium, J Hardy[67], for the UK Brain Expression Database, AT Crenshaw[68], MA Parkin[68], DB Mirel[68], DV Conti[13,14], S Purcell[1,2,5,68,71], G Nestadt[50,71], GL Hanna[69,71], MA Jenike[1,2,71], JA Knowles[46,71], N Cox[11,71] and DL Pauls[1,2,71]

[1]Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetics Research, Harvard Medical School, Boston, MA, USA; [2]Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA; [3]British Columbia Mental Health and Addictions Research Institute, University of British Columbia, Vancouver, BC, Canada; [4]Department of Neurology, Massachusetts General Hospital, Boston, MA, USA; [5]Department of Neurology, Brigham and Women's Hospital, Boston, MA, USA; [6]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA; [7]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA; [8]Department of Psychiatry, University of California San Francisco, San Francisco, CA, USA; [9]Program in Genetics and Genome Biology, The Hospital for Sick Children, Toronto, ON, Canada; [10]Department of Psychiatry, University of Toronto, Toronto, ON, Canada; [11]Section of Genetic Medicine, Department of Medicine, University of Chicago, Chicago, IL, USA; [12]Department of Psychiatry, University of Chicago, Chicago, IL, USA; [13]Department of Preventative Medicine, Division of Biostatistics, Keck School of Medicine, Zilkha Neurogenetic Institute, University of Southern California, Los Angeles, CA, USA; [14]Epigenome Center, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; [15]Partners Psychiatry and McLean Hospital, Boston, MA, USA; [16]Department of Psychiatry and Psychotherapy, University of Bonn, Bonn, Germany; [17]Department of Psychiatry and Psychotherapy, University of Göttingen, Göttingen, Germany; [18]Department of Psychiatry and Psychotherapy, University of Cologne, Cologne, Germany; [19]Department of Psychiatry and Psychotherapy, Helios-Hospital Stralsund, University Medicine Greifswald, Greifswald, Germany; [20]Psychiatry Universita Vita-Salute San Raffaele, Milano, Italy; [21]Fondazione Centro San Raffaele del Monte Tabor, Milano, Italy; [22]Department of Psychiatry, Sunnybrook Health Sciences Centre, Toronto, ON, Canada; [23]Institute for Juvenile Research, Department of Psychiatry, University of Illinois at Chicago, Chicago, IL, USA; [24]Centre for Addiction and Mental Health, Toronto, ON, Canada; [25]Child Psychiatry and Psychology, Wayne State University, Detroit, MI, USA; [26]Children's Hospital of Michigan, Detroit, MI, USA; [27]University of Cape Town, Cape Town, South Africa; [28]University of Stellenbosch, Stellenbosch, South Africa; [29]Department of Psychiatry, University of California San Diego, La Jolla, CA, USA; [30]Hospital Nacional de Niños, San José, Costa Rica; [31]Laboratory of Clinical Science, NIMH Intramural Research Program, Bethesda, MD, USA; [32]CNS Clinical Biomarker Group, Pharma Research and Early Development, F. Hoffmann-La Roche Ltd., Basel, Switzerland; [33]Departments of Psychiatry, Pediatrics, and Pharmacology, Kennedy Center for Research on Human Development, and Brain Institute, Vanderbilt University, Nashville, TN, USA; [34]Department of Psychiatry, Academic Medical Center and Netherlands Institute for Neuroscience, an Institute of the Royal Netherlands Academy of Arts and Sciences (NIN-KNAW), Amsterdam, The Netherlands; [35]Laboratory of Pharmaceutical Biotechnology, Ghent University, Ghent, Belgium; [36]Department of Child and Adolescent Psychiatry, University of Zurich, Zurich, Switzerland; [37]Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University of Würzburg, Würzburg, Germany; [38]Department of Psychiatry, Faculdade de Medicina da Universidade de Sao Paulo, Sao Paulo, Brazil; [39]University Health Care Services - SMURB, Federal University of Bahia, Salvador, Brazil; [40]Centre for Genomic Sciences, University of Mexico City, Mexico City, Mexico; [41]Carracci Medical Group, Mexico City, Mexico; [42]Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz, Depto. de Genética Psiquiátrica, Mexico City, Mexico; [43]AP-HP, Robert Debré Hospital, Department of Child and Adolescent Psychiatry, INSERM U955, Paris, France; [44]Institut Mondor de Recherche Biomédicale, Psychiatric Genetics, Créteil, France; [45]Foundation Fondamental, French National Science Foundation, Créteil, France; [46]Department of Psychiatry and The Behavioral Sciences, Zilkha Neurogenetic Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; [47]Section of Medical Genomics, Department of Clinical Genetics, VU University Medical Center Amsterdam, Amsterdam, The Netherlands; [48]Department of Psychiatry, VU University Medical Center and Department of Clinical and Health Psychology, Utrecht University, Utrecht, The Netherlands; [49]Department of Psychiatry, EMGO Institute, VU University Medical Center, Utrecht, The Netherlands; [50]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA; [51]Department of Psychiatry, College of Physicians and Surgeons at Columbia University, New York City, NY, USA; [52]New York State Psychiatric Institute, New York City, NY, USA; [53]Department of Psychiatry and Human Behavior, Brown Medical School, Butler Hospital, Providence, RI, USA; [54]Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, School of Medicine, Los Angeles, CA, USA; [55]Child Study Centre and Department of Psychiatry, Yale University, New Haven, CT, USA; [56]Child Study Centre, Pediatrics and Psychology, Yale University, New Haven, CT, USA; [57]Infant, Child and Adolescent Psychiatry, University of New South Wales, Academic Unit of Child Psychiatry South West Sydney (AUCS), Sydney, NSW, Australia; [58]University of Iowa, Roy J. and Lucille A. Carver College of Medicine, Iowa City, IA, USA; [59]UCLA Center for Neurobehavioral Genetics, Los Angeles, CA, USA; [60]University Medical Center Utrecht, Utrecht, The Netherlands; [61]Department of Medicine, Surgery and Dentistry, Graduate School of Nephrology, University of Milano, Milano, Italy; [62]Division of Nephrology, San Paolo Hospital, Milano, Italy; [63]Department of Health Technologies, University of Milano, Milano, Italy; [64]Filarete Foundation, Milano, Italy; [65]Department of Psychiatry and Human Behavior, School of Medicine, University of California Irvine (UCI), Irvine, CA, USA; [66]Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA; [67]Department of Molecular Neuroscience, University College of London, Institute of Neurology, London, UK; [68]The Broad Institute of Harvard-MIT, Cambridge, MA, USA; [69]Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA. Correspondence: Dr SE Stewart, Department of Psychiatry, University of British Columbia, A3-118, 938 West 28th Avenue, OCD Clinic, BCMHARI at CFRI, Vancouver, BC V5Z 4H4, Canada or Dr DL Pauls, Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA 02114, USA.
E-mail: sevelynstewart@gmail.com or dpauls@pngu.mgh.harvard.edu
[70]These authors contributed equally to this work.
[71]These are co-senior authors.
Received 4 January 2012; revised 3 May 2012; accepted 7 May 2012; published online 14 August 2012

Obsessive-compulsive disorder (OCD) is a common, debilitating neuropsychiatric illness with complex genetic etiology. The International OCD Foundation Genetics Collaborative (IOCDF-GC) is a multi-national collaboration established to discover the genetic variation predisposing to OCD. A set of individuals affected with DSM-IV OCD, a subset of their parents, and unselected controls, were genotyped with several different Illumina SNP microarrays. After extensive data cleaning, 1465 cases, 5557 ancestry-matched controls and 400 complete trios remained, with a common set of 469 410 autosomal and 9657 X-chromosome single nucleotide polymorphisms (SNPs). Ancestry-stratified case–control association analyses were conducted for three genetically-defined subpopulations and combined in two meta-analyses, with and without the trio-based analysis. In the case–control analysis, the lowest two *P*-values were located within *DLGAP1* ($P = 2.49 \times 10^{-6}$ and $P = 3.44 \times 10^{-6}$), a member of the neuronal postsynaptic density complex. In the trio analysis, rs6131295, near *BTBD3*, exceeded the genome-wide significance threshold with a *P*-value $= 3.84 \times 10^{-8}$. However, when trios were meta-analyzed with the case–control samples, the *P*-value for this variant was 3.62×10^{-5}, losing genome-wide significance. Although no SNPs were identified to be associated with OCD at a genome-wide significant level in the combined trio–case–control sample, a significant enrichment of methylation QTLs ($P < 0.001$) and frontal lobe expression quantitative trait loci (eQTLs) ($P = 0.001$) was observed within the top-ranked SNPs ($P < 0.01$) from the trio–case–control analysis, suggesting these top signals may have a broad role in gene expression in the brain, and possibly in the etiology of OCD.

Molecular Psychiatry (2013) **18,** 788–798; doi:10.1038/mp.2012.85; published online 14 August 2012

Keywords: DLGAP; genetic; genomic; GWAS; neurodevelopmental disorder; obsessive-compulsive disorder

INTRODUCTION

Obsessive-compulsive disorder (OCD) is a neuropsychiatric disorder characterized by obsessions and/or compulsions that are distressing, time consuming or significantly impairing. It is the fourth most common psychiatric illness[1] with a lifetime prevalence of 1–3%.[2,3] OCD was identified as the anxiety disorder with the highest proportion (50.6%) of serious cases by the National Comorbidity Study Replication[4] and as a leading global cause of non-fatal illness burden by the World Health Organization (WHO) in 2006.[5]

Genetic studies have demonstrated that both biological and environmental factors are important in the etiology of OCD. A multitude of OCD family studies published since the 1930s provide strong evidence for an approximate four- to tenfold OCD risk increase among first-degree relatives of OCD-affected children and adults, respectively, as compared with relatives of controls.[6–14] A review of twin studies concluded that obsessive-compulsive (OC) symptoms are heritable, with greater genetic influences in child-onset (45–65%), than in adult-onset OCD cases (27–47%).[15] This finding has been supported by subsequent twin studies.[16–18] Linkage study results have been somewhat encouraging,[19] identifying peaks on chromosomes 3q,[20] 9p,[21] 10p,[22,23] 15q[20,24] and 19q[19] for OCD and on chromosome 14 for compulsive hoarding.[25] Unfortunately, none of these peaks exceeded the threshold for genome-wide significance, and only the 9p peak has reached suggestive significance in more than one sample.[19–21]

In addition, more than 80 positional and functional candidate gene studies of OCD have been reported, predominantly for variants within genes in the serotonin, dopamine and glutamate[26,27] pathways and within those involved in immune and white matter pathways.[28] *SLC1A1,* which encodes a neuronal glutamate transporter and which is located within the linkage peak on chromosome 9p, is the only candidate gene observed to be associated in multiple independent samples, although the specific-associated variant has varied.[29–32]

Excessive grooming and anxiety-like behaviors have been observed in mice lacking expression of *SAPAP3*, a postsynaptic scaffolding protein located at excitatory synapses. This finding, coupled with high *SAPAP3* expression levels in the striatum, identify its human ortholog (*DLGAP3)* as an appealing candidate gene in OCD.[33] Human studies have provided some support for a possible role of *DLGAP3* in OCD-related disorders, suggesting increased rare non-synonymous variant frequencies in OCD/trichotillomania subjects[34] and association of common *DLGAP3* variants with pathological grooming in a family-based study,[35] albeit with some inconsistencies.[36]

In recent years, the genome-wide association study (GWAS) approach has led to the identification of many genetic associations for common complex traits.[37] This model-free approach to gene discovery has led to a greater pathophysiological understanding of many disorders, although only small proportions of the total genetic variance have so far been explained, and many of the identified variants have not brought new biological understanding.[38] To address the latter concern, functional support for GWAS findings has been sought by determining their effects on gene expression (expression quantitative trait loci- eQTLs) and methylation level (methylation quantitative trait loci-mQTLs).[38] Top single nucleotide polymorphisms (SNPs) have also been examined for potential enrichment of eQTLs and mQTLs, compared with expected rates. Moreover, examination for overrepresentation of micro-RNA (miRNA)-binding sites has also been adopted as an informative approach,[39] given the role of miRNA in regulating gene expression. In addition, pathway analyses have been conducted to determine whether specific gene pathways are enriched among the strongest associated variants.[40]

The International OCD Foundation Genetic Collaborative (IOCDF-GC), consisting of more than 20 research groups, has performed a GWAS to search for common SNPs predisposing to OCD. We present our findings from an analysis of the genetic association between OCD and a genome-wide set of common SNPs among case–control and trio samples and their combined trio–case–control results. We also present analyses of top GWAS findings with respect to their biological function in OCD-related and other brain regions.

MATERIALS AND METHODS

Subjects and genotyping

Our initial sample consisted of 1817 DSM-IV[41] OCD cases, 504 controls and 663 complete trios, genotyped using the Illumina Human610-Quadv1_B SNP array (San Diego, CA, USA). This work was approved by the relevant institutional review boards at all participating sites, and all participants provided written informed consent. The majority of the control subjects genotyped as a part of this project were not screened for the absence of OCD. We also used data from 5654 unscreened controls, previously genotyped on two different Illumina SNP arrays (Supplementary Table S1).

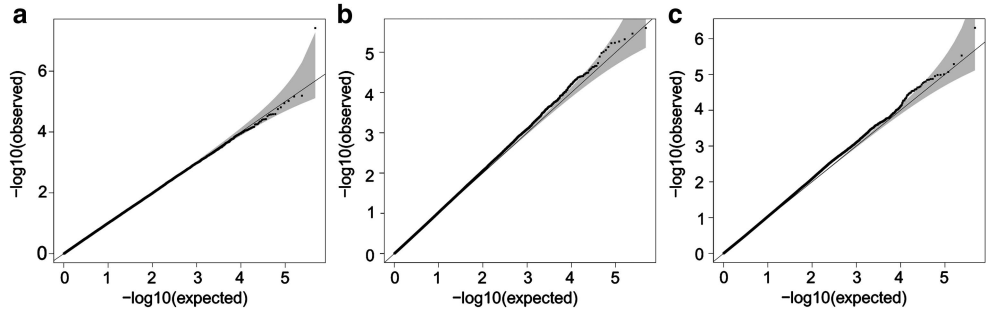

Figure 1. Quantile–quantile (QQ) plots of observed versus expected –log (*P*) statistics for: (**a**) trio samples ($\lambda = 1.015$), (**b**) case–control samples ($\lambda = 1.002$) and (**c**) combined trio–case–control samples ($\lambda = 1.011$). The 95% confidence interval of expected values is indicated in gray. Corresponding genomic control lambda values are indicated within each plot.

Quality control

The data for this study underwent quality control and data cleaning with a concurrent GWAS of Tourette Syndrome (Scharf *et al.*[42]) using PLINK,[43] to exclude samples and SNPs for each array type (Supplementary Figure S1).

Statistical analyses

To control for Type I error due to residual population stratification, case and control samples were separated into subpopulations of European (EU), South African Afrikaner (SA) and Ashkenazi Jewish (AJ) ancestry, using Multi-Dimensional Scaling (MDS) analysis (Supplementary Figures S2–4). Population stratification outliers and those lacking genomically matched controls or cases were excluded, as were samples with excessive low-level relatedness to many others within each subpopulation. Separate association analyses were conducted for each of the case–control subsamples (EU, SA and AJ) and for the trio samples. For the former, logistic regression was employed using an additive test model (1 degree of freedom), with diagnostic status as the dependent variable and each single SNP as the predictor, including specific ancestry-informative MDS axes as covariates (EU = 4 factors, SA = 2 factors and AJ = 1 factor). For the latter, the transmission disequilibrium test was used.

Two meta-analyses were conducted using METAL[44] by combining the three case–control sub-populations, and by combining the three case–control subgroups and the trio group, weighting by the number of cases or trios (Supplementary Materials). Each SNP was tested separately, defining a genome-wide significance threshold at $P < 5 \times 10^{-8}$, based on a 5% Type I error rate.[37] Using the PLINK retrieval interface,[43] SNP annotations were created using the TAMAL database[45] based chiefly on UCSC genome browser files,[46] HapMap[47] and dbSNP.[48] Further annotation was conducted using SCAN[49] and SPOT,[50] and top SNPs ($P < 0.001$) were also manually annotated using the UCSC genome browser.[51] For analysis of sex chromosome SNPs, males and females were assessed separately for each subgroup, with adjustment by MDS factors as described above, and subsequent combination via meta-analysis, using the number of cases or trios as a weighting factor. A sign test was conducted to examine for increased consistent directionality of effect for the most strongly associated SNPs between the case–control and trio samples. Analyses of potential enrichment of SNPs from: (a) 22 previously identified candidate genes, (b) pre-defined gene pathways and (c) target gene intervals containing miRNA-binding sites,[52] among the top hits from the trio, case–control or trio–case–control GWAS results were performed using INRICH.[40]

eQTL and mQTL annotation and enrichment tests

Functional support for the SNPs with the strongest evidence of association in the trio–case–control meta-analysis was sought by determining effects of the most significantly associated SNPs ($P < 0.001$) on both gene expression (eQTLs) and on methylation level (mQTLs). This was done with eQTLs from frontal lobes,[53] parietal lobes,[53] lymphoblastoid cell lines[54] and the cerebellum,[54] and with mQTLs[54] from cerebellum, using previously collected data.[54] To test whether the SNPs with the strongest observed associations were enriched for eQTLS or mQTLs, the linkage disequilibrium (LD)-independent SNPs from the trio–case–control analyses with $P < 0.001$ and with $P < 0.01$ were compared with 1000 random sets of the same size, conditioning on allele frequency, to yield an empirical distribution. An enrichment *P*-value was then calculated as the proportion of randomized sets in which the eQTL (or mQTL) count matched or exceeded the actual observed count in the list of top SNP associations, as previously described[53] (see Supplementary Materials).

Imputation of SNPs

Imputation of SNPs was conducted proximal to any SNPs with genome-wide significance from the trio, case–control or trio–case–control samples. This was completed using the 1000 Genomes Project via IMPUTE2,[55] and haplotypes from the 1092 individuals in a 1000 Genomes Data Release[56] as a reference dataset. Post-imputation quality control and allelic dosage analysis were conducted in PLINK (see Supplementary Materials).

RESULTS

MDS analyses identified three distinct genetic subpopulations within the case–control sample, which corresponded to: European (EU), South African (SA) and Ashkenazi Jewish (AJ) ancestries (Supplementary Materials). After quality control, a total of 1465 cases (1279 EU, 93 SA and 93 AJ), 5557 controls (5139 EU, 260 AJ and 158 SA) and 400 complete trios (299 EU) remained and each had genotypes for a common set of 469 410 autosomal and 9657 X-chromosome SNPs (Supplementary Table S1). Quantile–quantile plots of the observed versus expected log(*P*) values under the null hypothesis were used to calculate genomic control lambda values for the trio ($\lambda = 1.015$), case–control ($\lambda = 1.002$) and trio–case–control samples ($\lambda = 1.011$) (Figure 1). Quantile–quantile plots for EU ($\lambda = 1.009$), SA ($\lambda = 0.969$) and AJ ($\lambda = 0.982$) case–control subpopulations were also constructed (Supplementary Figure S7). There was no evidence for significant residual stratification effects in any of the comparisons.

Trio sample results

An overview of the *P*-values for the trio analysis plotted against genomic location is illustrated in Figure 2a. Of the top four OCD-associated SNPs in the trio sample with *P*-values $< 1 \times 10^{-5}$, one SNP, rs6131295 (11 996,267 bp (hg19) on 20p12.1-2), exceeded the threshold for genome-wide significance of $P < 5 \times 10^{-8}$ with a $P = 3.84 \times 10^{-8}$.[57] This SNP is located ~90 kb 3′ to *BTBD3* (Figure 3). None of the other 442 SNPs with *P*-values < 0.001 were in LD ($r^2 > 0.2$) with this SNP (Supplementary Table S2).

Case–control sample results

In the case–control sample, no SNPs exceeded the genome-wide threshold for significance (Table 1, Figure 2). Nine OCD-associated

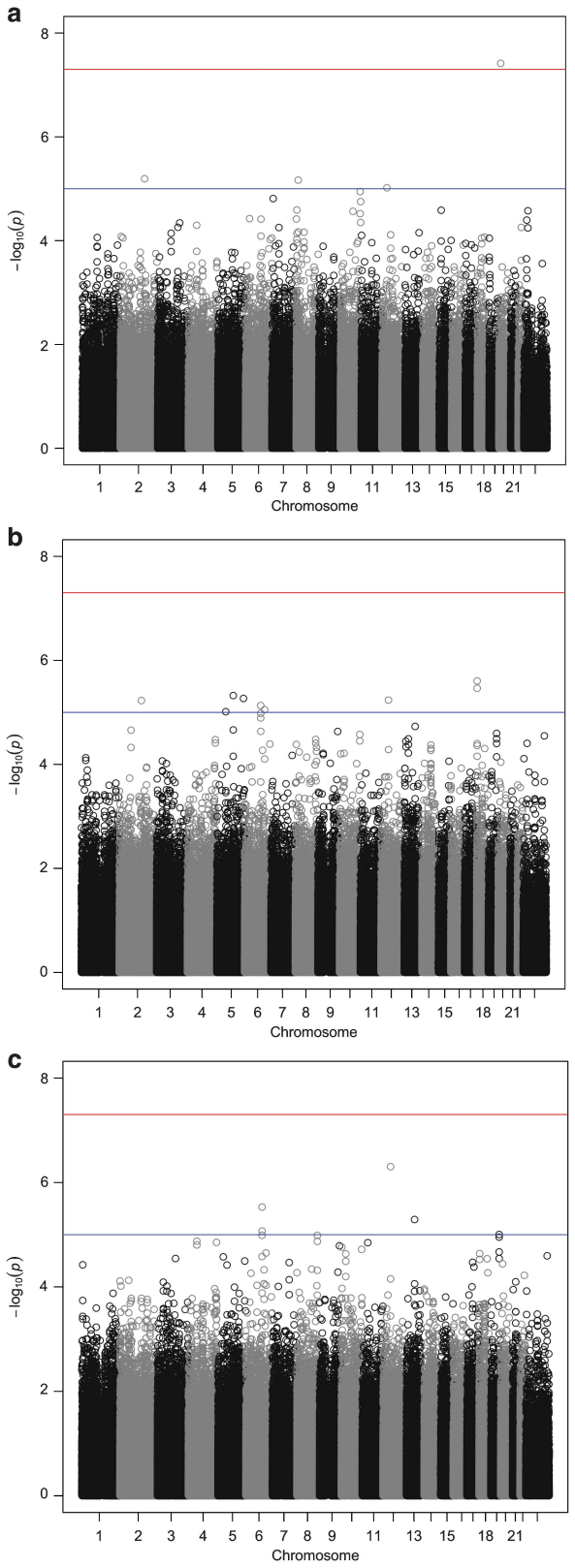

Figure 2. Manhattan plots of all genotyped single-nucleotide polymorphisms (SNPs) for (**a**) trio samples; (**b**) case–control samples; and (**c**) combined trio–case–control samples. Red and blue lines indicate significance thresholds of 5×10^{-8} and 1×10^{-5}, respectively.

SNPs had *P*-values $< 1 \times 10^{-5}$ (Table 1). The lowest two *P*-values were for SNPs rs11081062 ($P = 2.49 \times 10^{-6}$) and rs11663827 ($P = 3.44 \times 10^{-6}$), located at chromosome 18 within an intron of *DLGAP1* (Figure 3). *DLGAP1* (also known as *SAPAP1*) encodes the discs, large (Drosophila) homolog-associated protein a member of the neuronal postsynaptic density complex. The third lowest *P*-value was for the SNP rs26728 ($P = 4.75 \times 10^{-6}$), located within an intron of *EFNA5*, encoding Ephrin-A5 (Supplementary Figure S12). Ephrins are important for development of the neocortex through regulation of axonal inhibition or repulsion,[58] and *EFNA5* was also among top hits in an Alzheimer's disease GWAS.[59] The fourth lowest *P*-value $= 5.40 \times 10^{-6}$, was for rs4868342, lying within an intron of *HMP19*, encoding the brain-specific HMP19 protein (Supplementary Figure S12), which is expressed in the Golgi complex.[60] The fifth lowest *P*-value $= 5.81 \times 10^{-6}$, was for rs297941, which is located approximately 21 kb 5′ to the gene encoding *FAIM2* (also known as *LFG)* and about 25 kb from a cluster of genes encoding a group of aquaporins (*AQP5*, *AQP6*, *AQP2*) and lies within a putative coding region of mRNA BC034605, isolated from testis (Supplementary Figure S12).

Trio–case–control meta-analysis results

None of the SNPs exceeded the genome-wide threshold for significance, although several of the top hits were also identified among top hits in either the trio analysis or in the case–control analysis (Supplementary Figure S12). Using the sign test with 3616 LD-pruned SNPs with $P < 0.01$, there was evidence for increased consistent directionality ($1907/3616 = 0.52$; $P = 5.25 \times 10^{-4}$ for one-sided binomial test) between the trios and the combined case–controls. The top 38 OCD-associated SNPs in this meta-analysis, with *P*-values $< 5 \times 10^{-5}$, are presented in Table 1. For example, the top signal ($P = 4.99 \times 10^{-7}$), rs297941 near *FAIM2,* (LFG), was also the fifth-ranked SNP in the case–control analysis. *FAIM2* is highly expressed in the central nervous system and has a role in Fas-mediated cell death.[61] When rs6131295 (the SNP with significant genome-wide association in the trio sample) was meta-analyzed along with the case–control sample, the combined *P*-value significance decreased to 3.62×10^{-5}.

Examination of prior OCD linkage regions and candidate genes

There was no evidence found for genome-wide significant association with OCD in either previously identified putative linkage regions (Supplementary Table S3) or in 22 previously identified candidate genes when examining the trio, case–control and trio–case–control groups. The Q–Q plot of candidate gene SNPs for the case–control group showed little inflation ($\lambda = 1.085$, Supplementary Figure S8), suggesting no evidence for over-representation within these genes. While the Q–Q plot of the combined trio–case–control sample indicated small inflation ($\lambda = 1.168$, Supplementary Figure S8), the follow-up enrichment test demonstrated no overrepresentation of top hits ($P < 0.001$ and $P < 0.01$) within previously identified candidate genes ($P = 0.15$ and $P = 0.10$, respectively). For the 22 OCD candidate genes examined, the lowest SNP *P*-values are reported in Supplementary Table S4. The strongest finding was observed for *ADARB2*,[22] with a *P*-value $= 1.6 \times 10^{-4}$, which did not survive correction for multiple testing of candidate gene SNPs (corrected $P = 0.53$).

eQTL and mQTL annotation and enrichment analyses

Support for the SNPs with the strongest evidence of association in the combined trio–case–control sample was sought by determining functional effects of the most significantly associated autosomal SNPs. These top SNPs were annotated with eQTL) data from frontal, parietal and cerebellar brain regions (Table 1), along

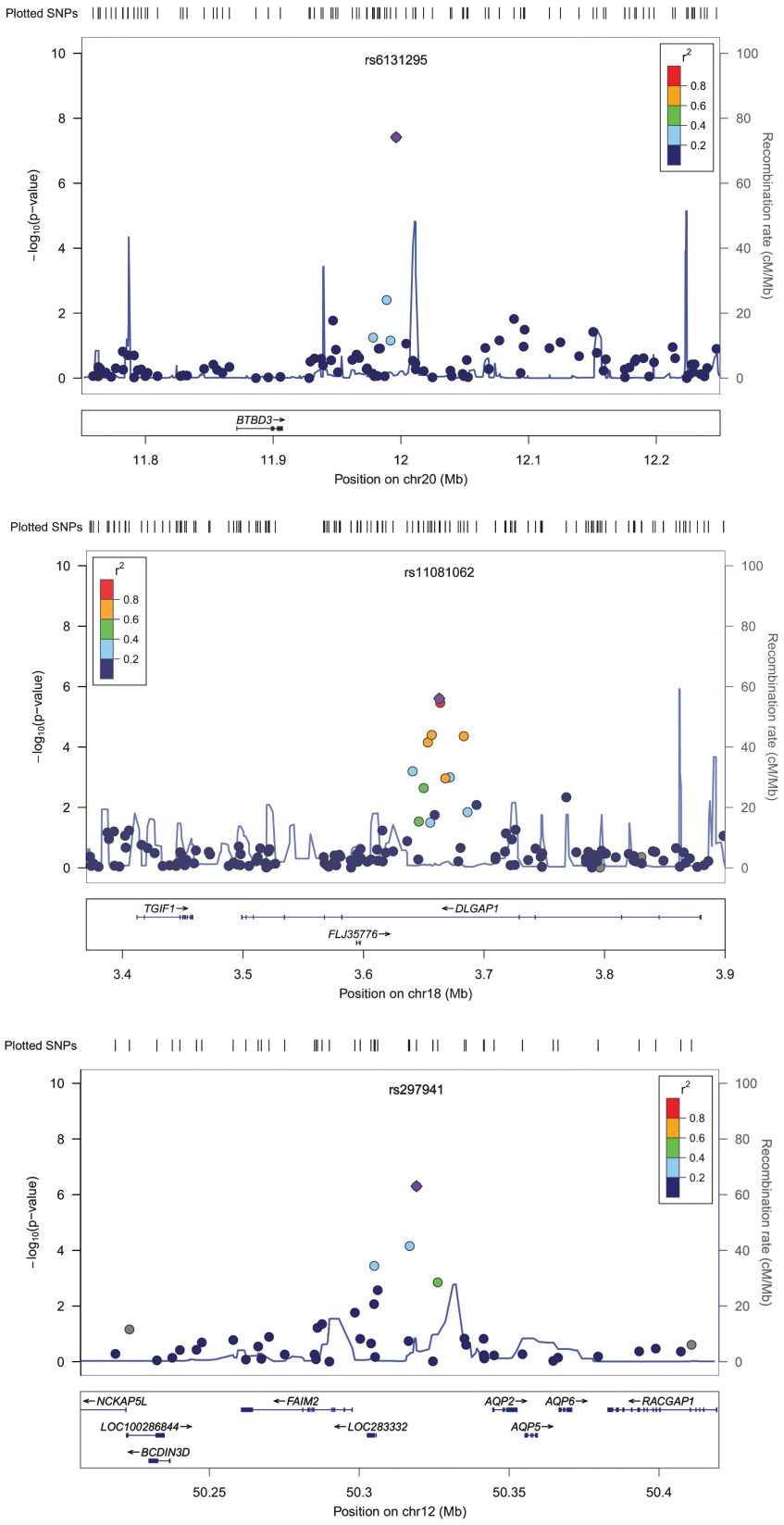

Figure 3. Locus Plots for single-nucleotide polymorphisms (SNPs) rs6131295 (near *BTBD3*), rs11081062 (within *DLGAP1*) and rs297941 (near *FAIM2*). Regional association plots of the best supported SNPs from the (top) trio, (middle) case–control and (bottom) trio–case–control analyses. Locations and observed -log (*P*-values) for genotyped SNPs are shown with circles. Linkage disequilibrium (LD), in r^2, to the lowest *P*-value SNP in each plot is indicated using shading (dark blue, low LD, red-high LD). Light blue lines indicate the estimated recombination rate from HapMap release 22.

Table 1. Strongest associated GWAS variants in trio, case-control and combined trio–case–control samples

No.	Location		Trio		Case–Control	Trio–Case–Control		Left gene (kb) Right gene (kb)		QTL
rs#	Chr	base pair (HG19)	OR	P	P	P	Direction	Intragenic (location)		
rs6131295	20	11 996 267	0.51	3.84E−08	0.0738	3.63E−05	−−−+	*BTBD3* (89)	*SPTLC3* (993)	F:*BTBD3*; P:*JSM1*; m:*NOS1*
rs10165908	2	158 315 629	2.05	6.43E−06	0.735	0.0169	+++−	*CYTIP* (15)	*ACVR1C* (68)	
rs6531002	8	12 722 703	1.63	6.84E−06	0.477	0.00665	−++−	*LONRF1* (110)	*KIAA1456* (80)	C:*C8orf12*; m:*NOMO1*,*GNAL*
rs11611761	12	33 025 612	0.56	9.58E−06	0.593	0.115	−−+−	*PKP2* (intronic)		m:*TRIM31*
rs11081062	18	3 662 879	1	1	2.49E−06	2.92E−05	+0++	*DLGAP1* (intronic)		C:*DLGAP1*
rs11663827	18	3 663 631	1.03	0.799	3.44E−06	2.31E−05	++++	*DLGAP1* (intronic)		C:*TYMS*, *DLGAP1*
rs26728	5	106 946 056	0.96	0.718	4.75E−06	0.000101	−++−	*EFNA5* (intronic)		
rs4868342	5	173 504 522	1.03	0.783	5.40E−06	3.20E−05	++++	*HMP19* (intronic)		C:*CPEB4*; F:*CPEB4*
rs297941	12	50 319 086	0.81	0.0294	5.81E−06	4.99E−07	−−−−	*FAIM2* (21)	*AQP2* (25)	F:*BCDIN3D*
rs11898020	2	144 282 078	1.1	0.44	5.93E−06	0.000256	+−+−	*ARHGAP15* (intronic)		m:*CXCL9*
rs2205748	6	104 462 555	1.11	0.301	7.38E−06	8.52E−06	++++	*GRIK2* (1944)	*HACE1* (713)	
rs182320	6	130 073 291	1.07	0.516	8.87E−06	2.25E−05	++++	*ARHGAP18* (42)	*C6orf191* (79)	P:*TAAR5*
rs1838733	5	58 533 392	0.96	0.669	9.71E−06	3.82E−05	−−−−	*PDE4D* (intronic)		C:*ELOVL*; P:*PLK2*; m:*TTYH1*
rs297941	12	50 319 086	0.81	0.0294	5.81E−06	4.99E−07	−−−−	*FAIM2* (21)	*AQP2* (25)	C:*C12orf62*, *LASS5*, *TUBA1A*; F:*CDIN3D*
rs9499708	6	104 445 367	0.83	0.0818	1.28E−05	2.96E−06	−−+−	*GRIK2* (1927)	*HACE1* (731)	
rs9652236	13	72 688 774	1.4	0.0101	0.0001445	5.14E−06	++++	*DACH1* (247)	*MZT1* (594)	
rs2205748	6	104 462 555	1.11	0.301	7.38E−06	8.52E−06	++++	*GRIK2* (1944)	*HACE1* (713)	
rs485186	19	49 207 206	1.16	0.138	2.54E−05	9.94E−06	++++	*FUT2* (coding-synon T (ACA)−> T (ACG))		C:*SNORD33*; F:*SNAR-A1*; m:*HAS1*, *RASIP1*, *IZUMO1*, *FGF21*, *FUT1*, *M-RIP*
rs6919215	6	104 475 419	1.12	0.277	1.04E−05	1.03E−05	++++	*GRIK2* (1957)	*HACE1* (701)	m:*C9orf24*, *SYN1*
rs7459733	8	131 936 877	0.84	0.114	3.32E−05	1.03E−05	−−−−	*ADCY8* (intronic)		P:*ADCY8*; m:*MS4A12*
rs602662	19	49 206 985	1.17	0.126	3.21E−05	1.11E−05	++++	*FUT2* (missense G (GGT) −> S (AGT))		F:*PRMT1*, *RASIP1*, *SNAR-A1*, *FCGRT*, *RPS11*, *NOSIP*; m:*HAS1*, *RASIP1*, *IZUMO1*, *FGF21*, *FUT1*, *M-RIP*
rs759082	4	55 491 904	1.24	0.0309	0.0001535	1.33E−05	++++	*PDGFRA* (327)	*KIT* (32)	
rs7461923	8	131 936 279	0.85	0.127	3.91E−05	1.35E−05	−−−−	*ADCY8* (intronic)		P: *ADCY8*
rs17070275	4	181 813 489	0.86	0.151	3.34E−05	1.41E−05	−−−−	*LOC285501* (2902)	*LINC00290* (171)	P: *HELT*; m:*PIWIL2*
rs7124427	11	36 089 700	1.29	0.0348	0.0001482	1.42E−05	++−+	*LDLRAD3* (intronic)		
rs7675203	4	55 486 182	1.26	0.0215	0.0002427	1.57E−05	+++-	*PDGFRA* (322)	*KIT* (38)	
rs1874777	9	129 601 002	1.17	0.227	2.33E−05	1.64E−05	++++	*ZBTB43* (4)	*ZBTB34* (22)	C:*LRSAM1*, *RPL12*, *STXBP1*, *NA*, *ZBTB34*; F:*FAM129B*
rs11252374	10	4 186 228	2.3	0.00137	0.001496	1.73E−05	++++	*AK055803* (55)	*LOC100216001* (506)	
rs3824760	10	131 861 906	2	0.197	2.66E−05	1.91E−05	++++	*EBF3* (100)	*LOC38772* (3236)	
rs504963	19	49 208 865	1.17	0.115	7.16E−05	2.15E−05	++++	*FUT2* (UTR-3)		F:*SNAR-A1*, *RASIP1*, *RPS11*, *PRMT1*, *FCGRT*, *NOSIP*;C:*PRMT1*, *PRMT1*, *GRWD1*, *SNORD33*; m:*FGF21*, *FUT1*, *HAS1*, *IZUMO1*, *RASIP1*, *M-RIP*
rs182320	6	130 073 291	1.07	0.516	8.87E−06	2.25E−05	++++	*ARHGAP18* (42)	*C6orf191* (79)	P: *TAAR5*
rs11663827	18	3 663 631	1.03	0.799	3.44E−06	2.31E−05	++++	*DLGAP1* (intronic)		C: *TYMS*, *DLGAP1*
rs749631	10	29 402 289	0.78	0.0395	0.0002177	2.33E−05	−−−−	*LOC100507605* (317)	*LYZL1* (176)	P: *ARX*; C: *BAMBI*
rs5908139	X	141 175 215	1.21	0.2805	2.83E−05	2.54E−05	++-+	*MAGEC1* (178)	*MAGEC2* (115)	
rs6919443	6	104 493 098	1.11	0.328	2.32E−05	2.63E−05	++++	*GRIK2* (1975)	*HACE1* (683)	P: *OPRD1*
rs6897719	5	34 795 981	1.28	0.0668	0.0001558	2.66E−05	++++	*RAI14* (intronic)		C: *LIFR*; m:*DCAMKL1*, *FYB*, *KENAE*
rs1392261	3	115 144 685	0.79	0.0244	0.000393	2.86E−05	−−−+	*ZBTB20* (279)	*GAP43* (198)	C: *QTRTD1*
rs676388	19	49 211 969	1.13	0.242	3.97E−05	2.87E−05	++++	*FUT2* (3)	*FLJ36070* (4)	C: *PRMT1*, *GRWD1*; F:*SNAR-A1*, *NUCB1*, *RASIP1*, *RPS11*, *PRMT1*, *FCGRT*, *NOSIP*, *NA3365*; m:*FGF21*, *FUT1*, *IZUMO1*, *RASIP1*, *TSKS*
rs1005419	18	54 308 996	1.16	0.207	4.94E−05	2.90E−05	++++	*TXNL1* (intronic)		F: *NARS*
rs11081062	18	3 662 879	1	1	2.49E−06	2.98E−05	0+++	*DLGAP1* (intronic)		C: *DLGAP1*
rs4868342	5	173 504 522	1.03	0.783	5.40E−06	3.20E−05	++++	*HMP19* (intronic)		C: *CPEB4*; F: *CPEB4*
rs2793345	10	29 409 192	0.79	0.0466	0.000263	3.24E−05	−−−−	*LOC100507605* (324)	*LYZL1* (169)	P: *ARX*; C:*BAMBI*
rs2857254	17	39 620 034	1.21	0.0946	0.0001426	3.39E−05	++++	*KRT32* (intronic)		F: *DHX58*
rs12705610	7	109 161 891	1.33	0.006065	0.001191	3.42E−05	+++−	*DNAJB9* (947)	*LRRN3* (1569)	P:*C7orf60*; C:*DLD*, *LOC401397*
rs6131295	20	11 996 267	0.51	3.84E−08	0.07383	3.62E−05	−−−+	*BTBD3* (89)	*SPTLC3* (993)	F: *BTBD3*; P:*ISM1*; m:*NOS1*
rs4908608	1	7 292 772	1.21	0.0655	0.0002276	3.79E−05	+++−	*CAMTA1* (intronic)		C: *UBE4B*; m:*TNFRSF25*
rs1838733	5	58 533 392	0.96	0.67	9.71E−06	3.82E−05	−−−−	*PDE4D* (intronic)		P:*PLK2*; C:*ELOVL7*; m:*TTYH1*
rs2586494	17	48 273 155	1.57	0.002075	0.002568	4.17E−05	++++	*COL1A1* (intronic)		F: *LUC7L3*; C:*SPAG9*, *NXPH3*
rs2515144	8	95 289 326	1.23	0.0792	0.0002315	4.58E−05	+++−	*GEM* (15)	*RAD54B* (95)	C: *DPY19L4*, *POP1*; m:*SAFB2*, *KIAA0802*, *EFCAB3*
rs636252	6	117 157 774	1.12	0.2905	5.37E−05	4.72E−05	++++	*GPRC6A* (8)	*RFX6* (41)	F: *TSPYL1*; m:*C6orf78*; *BCAP31*
rs12334868	8	131 954 652	0.76	0.0284	0.0005812	4.82E−05	−−−−	*ADCY8* (intronic)		

Abbreviations: OR, odds ratio; QTL, quantitative trait loci.
Single nucleotide polymorphisms (SNP) listed by rs number include those with association *P*-values $< 10^{-5}$ for the trio and case–control samples, and those with *P*-values $< 5 \times 10^{-5}$ for the combined trio–case–control sample association results. The chromosome (Chr) and base pair location for each SNP are listed in columns to the right of the SNP column. SNPs are listed separately for the analyses of trios (top section of table, with box around results), case–control samples including combined EU, SA and AJ MDS-defined ancestry subgroups (middle section and box), and for combined trio–case–control samples (lower section and box). SNPs with $P < 10^{-3}$ for any of the following are available in online Supplementary Table S2: EU, AJ and SA case–control subgroups individually and combined, trios and combined case–control–trios). OR indicates the odds ratio for the tested allele in the trio sample. Direction indicates whether the direction of association between OCD and the A1 allele is either positive (+) or negative (−) for individual subgroups within the combined (EU, AJ, SA, trios) samples. The left gene and right gene columns lists the closest genes in the SNP region, either located within the gene (no distance given) or as right and left flanking genes (distance in kilobases). For SNPs located within genes, other functional elements in the region are as noted. QTL columns list genes whose expression (eQTL) or methylation levels (m) are associated (*P*-value) with the specified SNP in that row, specifically as identified previously in EU-ancestry frontal (F), parietal (P) or cerebellar (C) tissue. mQTL and F eQTL data were unavailable for X-chromosome SNPs.

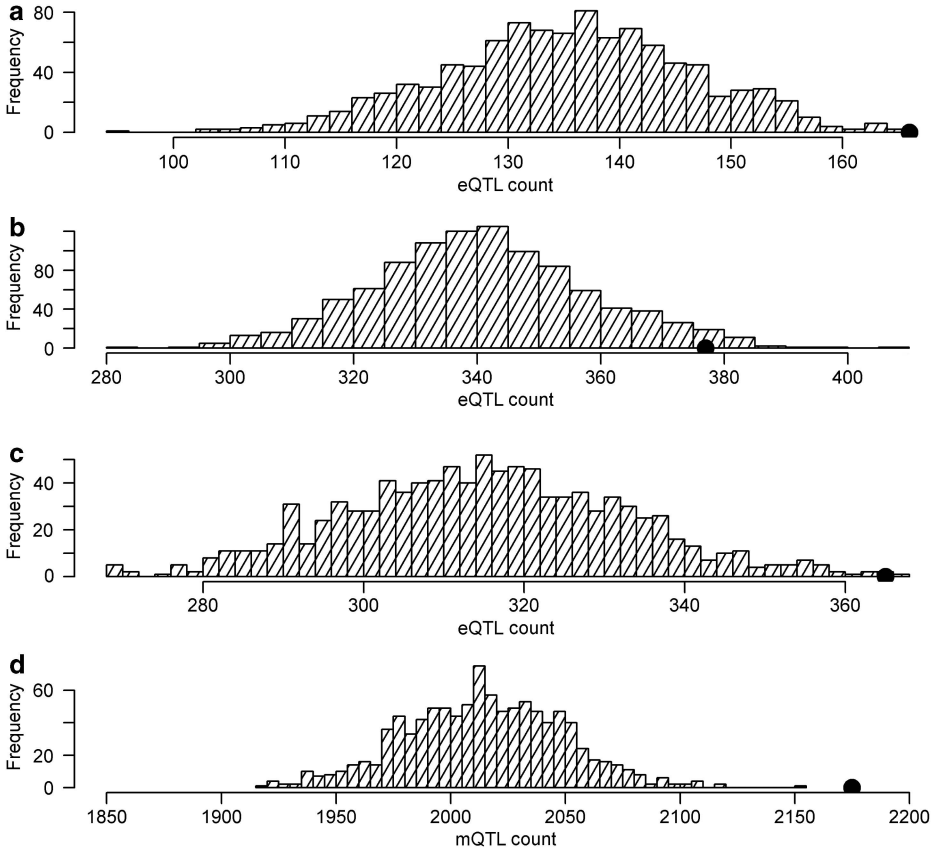

Figure 4. Enrichment analyses for quantitative trait loci (QTLs) among genome-wide association study (GWAS) variants with $P<0.01$. Enrichment of (**a**) frontal lobe expression QTLs ($P=0.001$), (**b**) cerebellum expression QTLs ($P=0.033$), (**c**) parietal lobe expression QTLs ($P=0.003$) and (**d**) methylation QTLs ($P<0.001$) among GWAS single-nucleotide polymorphisms (SNPs) with $P<0.01$ ($N=5321$). Distribution of the count of QTLs in 1000 simulations are displayed, each matching the minor-allele-frequency distribution of the obsessive-compulsive disorder (OCD)-associated SNPs. The black dot identifies the observed expression QTL or methylation QTL count in the OCD susceptibility-associated SNPs.

with lymphoblastoid cell lines (Supplementary Table S2) and methylation levels (mQTLs) in cerebellum (Table 1).

SNPs with association *P*-values <0.01 ($n=3521$) were then examined for enrichment of eQTLs and mQTLs. Significant enrichment was observed for frontal eQTLs ($P=0.001$) as well as for cerebellar eQTLs ($P=0.033$) and parietal eQTLs ($P=0.003$) (Figure 4a-c). Furthermore, enrichment of cerebellar mQTLs was observed ($P<0.001$) with an enrichment *P*-value of $P<0.001$ (Figure 4d), suggesting that these SNPs are more likely to influence the methylation state than expected by chance. No significant enrichment for either genic ($P=0.54$) or missense variants (0.34) was observed. A similar analysis examining only the top SNPs with association *P*-values <0.001 ($n=415$) demonstrated no significant enrichment for mQTLs or for eQTLs ($P>0.05$).

miRNA and pathway analyses

After correction for multiple hypothesis testing, there was no evidence for enrichment of specific miRNA-binding sites among the LD blocks containing top SNPs compared with the genes matched by size and marker density (see Supplementary Table S5). The strongest enrichment was found in 49 high-confidence (TargetScan probability >0.9) predicted-miRNA-219-5p/508/508-3p/4782-3p targets, two of which have at least one SNP with $P<0.001$ (empirical $P=0.011$, corrected $P=0.060$) in the case–control GWAS result. A similar level of enrichment was also found in 89 high-confidence predicted-miR-130ac/301ab/301b/301b-3p/454/721/4295/3666 targets, two of which have at least one SNP with $P<0.001$ in the trio transmission disequilibrium test result. In the pathway analyses, no results achieved significance at the corrected *P*-value (lowest-corrected $P=0.55$) (see Supplementary Table S6).

DISCUSSION

We report results from the first GWAS to search for common DNA sequence variation predisposing individuals to OCD. After removing low-performing SNP assays and DNA samples, we analyzed 400 trios, 1465 cases and 5557 controls for 469 410 autosomal and 9657 X-chromosome SNPs. The trio and case–control subsamples were analyzed individually, and then these results were combined in both case–control and trio–case–control meta-analyses. One SNP, rs6131295, located on chromosome 20p12.1-p12.2, approximately ~90 kb from the *BTBD3* gene, achieved genome-wide significance in the trio analysis ($P=3.84\times10^{-8}$), but not in the combined trio–case–control meta-analysis, suggesting that further examination will be required using independent samples. *BTBD3* is a member of a large family of transcription factors, which includes *BTBD9,* a gene that has been associated with Tourette Syndrome, a disorder frequently co-morbid with OCD.[62] *BTBD3* participates in multiple cellular functions including transcriptional regulation, cytoskeleton dynamics, ion channel assembly and gating, protein ubiquitination and degradation[63] and has also been associated with primary open-angle glaucoma.[64] *BTBD3* is expressed in the brain, with the highest observed levels in childhood and

adolescence (http://www.brainspan.org, Release 3),[63] when OCD frequently emerges.[65] rs6131295 is a cis-eQTL for *BTBD3* in the frontal cortex ($P = 0.028$), a region that has repeatedly been implicated in OCD. This SNP is also a parietal cis-eQTL for *ISM1* (20p12; $P = 0.0036$) and a lymphoblastoid cell line trans-eQTL for *DHRS11* (17q11.2; $P = 0.0001$).

Interestingly, the brain-wide expression pattern of *DHRS11* and *ISM1* are highly correlated with the expression of several of the other genes found among the top hits of both the case–control and the trio–case–control meta-analyses (http://www.brainspan.org, Release 3) (Supplementary Figure S12).[66] Furthermore, many of these genes have been implicated in glutamate signaling. Specifically, *ISM1 (C20orf82)* is correlated with expression of presynaptically located *ADCY8* (0.61, rank 11 of 22 328 transcripts), the gene with the seventh strongest OCD-association in the trio–case–control meta-analysis, which has also been associated with bipolar disorder[67] and with fear memory.[68] *ISM1* is also correlated with brain-wide expression of numerous glutamate-related genes including *GRIK4* (0.565, rank 66), *DLGAP3* (0.576, rank 44), *GRIK1* (0.595, rank 22), *SHANK3* (0.598, rank 21) as well as *ADARB2* (0.600, rank 19), which contains the SNP with the best *P*-value in this study among previously reported candidate genes (Supplementary Table S4), and lies within a childhood-onset OCD linkage peak.[22] Similarly, the expression of *DHRS11 (MGC4172)* is strongly correlated (0.847, rank 25 of 22 328 transcripts) with that of *FAIM2*, which is located in the same LD block as the best SNP (rs297941) in the trio–case–control, and fifth best in the case–control meta-analyses. *FAIM2* has been associated with neuroprotection following transient brain ischemia.[68] The rat homologue of *FAIM2*, neural membrane protein 35 (*NMP35)*, is expressed at the postsynaptic membrane in a subset of synapses and in dendrites, and co-localizes with the glutamate receptor *GluR2*.[61] Thus, there is a potential relationship between rs6131295 (trio analysis) and *FAIM2* and *ADCY8* (tagged by the SNPs ranked numbers 1 and 7 in the trio–case–control analysis).

The top two SNPs associated in the case–control meta-analysis (both with $P < 3 \times 10^{-5}$ in the trio–case–control meta-analysis) are located in *DLGAP1*, another gene that influences glutamate signaling. *DLGAP1* encodes a Shank-associated protein and has been associated with schizophrenia and with a smoking-cessation phenotype[69] and *DLGAP1* deletions have also been observed (two in schizophrenia cases versus one in controls).[70] Another member of this gene family, *DLGAP3,* has been implicated in compulsive-like behavior in a mouse model (*SAPAP3*). Specifically, knockout mice for the striatum-expressed *SAPAP3* gene (which codes for a postsynaptic protein at cortico-striatal glutamatergic excitatory synapses) developed repetitive grooming behaviors and anxiety that were reversed with an SSRI and with gene replacement.[24]

Several of the top associations in the combined trio–case–control meta-analysis are in or near genes that have been implicated in other studies of psychiatric disorders, including ADCY8,[59,71,72] ARHGAP18[47] and JMJD2C [62] in bipolar disorder, schizophrenia and autism spectrum disorders, respectively. Enrichment for eQTLs was observed among the top associated GWAS SNPs ($N = 5321$; $P < 0.01$), with empirical *P*-values of 0.001 for frontal cortex, 0.003 for parietal tissue and 0.033 for cerebellum. Marked enrichment was also observed for methylation QTLs ($P < 0.001$). This is consistent with the finding by Nicolae *et al*.,[53] who reported that disease-associated SNPs from GWAS were significantly more likely to be eQTL, than other random sets of SNPs with similar minor-allele-frequencies.

It remains unclear whether the finding at rs6131295, which exceeded genome-wide significance with $P = 3.84 \times 10^{-8}$ in the trio sample, is a false positive or not. Certainly the decrease in significance of the *P*-value to 3.62×10^{-5} when the trio data is meta-analyzed with the much larger case–control sample data suggests so. On the other hand, our attempts to determine whether this finding was spurious did not find any evidence of such, as detailed here: (1) The intensity plot for this SNP has three tight, separated, clusters (Supplementary Figure S10a); (2) There were no missing genotypes in the trio sample and there were no Mendelian errors; (3) Two nearby directly genotyped SNPs with low r^2 values (0.2–0.4) had *P*-values within the 10^{-2} range, demonstrating very low statistical significance (Supplementary Figure S10b); and (4) Imputation of the trio sample provided additional results that are not inconsistent with a true positive finding. Of the 40 regional SNPs examined, those with large r^2 values (>0.90) and similar minor allele frequencies to rs6131295 had strong *P*-values in the range of 10^{-6} and 10^{-7} (Supplementary Table S7 and Supplementary Figure S11). Moreover, the surrounding SNPs in low r^2 with rs6131295 all have an opposite direction of risk effect, which may partially explain why they have much less significant *P*-values. Although these imputed data and the above noted facts cannot prove that rs6131295 is a true positive, they do not support the hypothesis that it is a false positive. Replication with additional samples will be required to provide further clarification.

In summary, although no SNPs were identified to be associated with OCD at a genome-wide significant level in the combined trio–case–control sample, a highly significant enrichment of methylation QTLs ($P < 0.001$) and frontal lobe eQTLs ($P = 0.001$) was observed within the top-ranked SNPs ($P < 0.01$). This suggests that these top signals may have a broad role in gene expression in the brain, and possibly in the etiology of OCD. In the trio sample, we observed a genome-wide significant result for rs6131295, which is located near *BTBD3*, and is an eQTL for *BTBD3*, *DHRS11* and *ISM1*. The expression of these latter two genes are highly correlated with other top hits, many of which are related to glutamatergic neurotransmission and signaling. So, although no genome-wide significant associations were found in the entire sample, the convergence of results from both the trio and combined trio–case–control analyses suggest the possibility that *BTBD3*, *FAIM2* and *ADCY8* are involved in the pathogenesis of OCD. In the case–control sample, the two most significant *P*-values were located within *DLGAP1,* a member of the same gene family as *DLGAP3,* which is also expressed in the neuronal postsynaptic density complex and which has been implicated in a mouse model of OCD,[33] making these results intriguing. Future exploration and attempts to replicate these findings with additional independent samples is warranted.

CONFLICT OF INTEREST

SE Stewart has received funding from the International OCD Foundation (IOCDF) and is a member of the IOCDF Scientific Advisory Committee. PD Arnold reports funding sources including the CIHR, NIH, Ontario Research Foundation, Ontario Brain Institute, DNA Genotek, and the McLaughlin Centre. AB Singleton served as an unpaid consultant to Teva Pharmaceuticals. HJ Grabe has received funds from the German Research Foundation; Federal Ministry of Education and Research Germany; speakers honoraria from Bristol-Myers Squibb, Eli Lilly, Novartis, Eisai, Boehringer Ingelheim; speaker and travel funds from Janssen-Cilag, Eli Lilly, Novartis, AstraZeneca, Lundbeck and SALUS-Institute for Trend-Research and Therapy Evaluation in Mental Health. R Moessner has been supported by the German Research Foundation (DFG) (grants Wa 731/6 and 731/4), and by the German Federal Ministry for Education and Research (BMBF grant 01GV0907). M Wagner has been supported by the German Research Foundation (DFG) (grants Wa 731/6 and 731/4), and by the German Federal Ministry for Education and Research (BMBF grant 01GV0907). MA Richter has received honoraria from Lundbeck and she is recipient of grant funding from Eli Lilly Canada, Ontario Mental Health Foundation and the Obsessive-Compulsive Foundation. DJ Stein has received research grants and/or consultancy honoraria from Abbott, Astrazeneca, Eli-Lilly, GlaxoSmithKline, Jazz Pharmaceuticals, Johnson & Johnson, Lundbeck, Orion, Pfizer, Pharmacia, Roche, Servier, Solvay, Sumitomo, Takeda and Tikvah. JR Wendland is now a full-time employee of F Hoffmann-La Roche J Veenstra-VanderWeele receives research funding from Seaside Therapeutics, Roche Pharmaceuticals and Novartis. GL Hanna, MT Pato and CN Pato receive NIH funding. K Egberts, T Renner and S Walitza received sample collection funding by DFG WA168/1-1. SL Rauch has received research funding from Cyberonics and Medtronic. JA Knowles is a recipient of grant funding from NIH and from NARSAD; he sits on the Scientific Advisory Committee for Next-Generation Sequencing of Life Technologies and is a technical advisor to SoftGenetics; he is on the Scientific Advisory Committee for

Next-Generation Sequencing of Life Technologies and is a technical advisor to SoftGenetics. JF Leckman has been funded by the NIH, the TSA, Talecris Biotherapeutics, Klingenstein Third Generation Foundation, John Wiley and Sons, McGraw Hill and Oxford University Press. V Coric works for Bristol Myers-Squibb. DW Black has received NIH funding and support from AstraZeneca and Psysadon in addition to royalties from American Psychiatric Publishing and Oxford University Press. The remaining authors declare no conflicts of interest.

ACKNOWLEDGEMENTS

The authors would like to express their utmost gratitude to the OCD-affected families who participated in this research. In addition, they would like to thank the International OCD Foundation (IOCDF) for their role in establishing the IOCDF Genetics Collaborative, as well as other individuals who played roles in assisting this study, including Rhonda Ellwyn, Katherine Beattie, Colm O'Dushlaine, Doug Ruderfer, Priya Moorjani and V. Guttenthaler. This work was supported primarily by a grant from the Judah Foundation (a private, non-industry related foundation established by a family affected by OCD), NIH grants MH079489 and MH073250 to DLP, American Recovery and Re-investment Act (ARRA) awards NS40024-07S1 and NS16648-29S1 to DLP, by an American Academy of Child and Adolescent Psychiatry (AACAP) Early Investigator Research Grant, an Anxiety Disorders Association of America (ADAA) Junior Investigator Research Grant, the University of British Columbia and a Michael Smith Foundation Clinical Research Scholar Award to SES, and grants from the Tourette Syndrome Association (DLP and JMS), the American Academy of Neurology Foundation (JMS) and NIH grant MH085057 to JMS. The Broad Institute Center for Genotyping and Analysis was supported by grant U54 RR020278 from the National Center for Research Resources. Funding support for the Study of Addiction: Genetics and Environment (SAGE) was provided through the NIH Genes, Environment and Health Initiative [GEI] (U01 HG004422). SAGE is one of the genome-wide association studies funded as part of the Gene Environment Association Studies (GENEVA) under GEI. Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01 HG004446). Assistance with data cleaning was provided by the National Center for Biotechnology Information. Support for collection of data sets and samples was provided by the Collaborative Study on the Genetics of Alcoholism (COGA; U10 AA008401), the Collaborative Genetic Study of Nicotine Dependence (COGEND; P01 CA089392) and the Family Study of Cocaine Dependence (FSCD; R01 DA013423). Funding support for related genotyping, which was performed at the Johns Hopkins University Center for Inherited Disease Research, was provided by the NIH GEI (U01HG004438), the National Institute on Alcohol Abuse and Alcoholism, the National Institute on Drug Abuse, and the NIH contract 'High throughput genotyping for studying the genetic contributions to human disease' (HHSN268200782096C). The data sets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id = phs000092.v1.p1 through dbGaP accession number phs000092.v1.p. Frontal lobe eQTL data was provided by the North American Brain Expression Consortium and the UK Human Brain Expression Database. Funding support for generation of the eQTL data was provided by the UK Medical Research Council and the Intramural Research Program of the National Institute on Aging, National Institutes of Health, Department of Health and Human Services project Z01 AG000932-02. The North American Brain Expression Consortium comprises: Sampath Arepalli, Mark R Cookson, Allissa Dillman, Luigi Ferrucci, J Raphael Gibbs, Dena G Hernandez, Robert Johnson, Dan L Longo, Michael A Nalls, Richard O'Brien, Andrew Singleton, Bryan Traynor, Juan Troncoso, Marcel van der Brug, H Ronald Zielke and Alan Zonderman. The UK Human Brain Expression Database membership comprises: John A Hardy, Mina Ryten, Colin Smith, Daniah Trabzuni, Robert Walker and Mike Weale. None of the funding sources supporting this work had any influence or played any role in: (a) the design or conduct of the study; (b) management, analysis or interpretation of the data; or c) preparation, review or approval of the manuscript.

AUTHOR CONTRIBUTIONS

Manuscript preparation: SE Stewart, JA Knowles, D Yu, JM Scharf, CA Mathews, PD Arnold, E Gamazon, PD Evans, GL Hanna, NJ Cox and DL Pauls. Study design: SE Stewart, JM Scharf, D Yu, JA Knowles, PD Arnold, CA Mathews, BM Neale, JA Fagerness, EH Cook, S Purcell, NJ Cox, G Nestadt and DL Pauls. Data analysis: D Yu, BM Neale, S Purcell, JM Scharf, PD Evans, ER Gamazon, A Tikhomirov, A Pluzhnikov, A Konkashbaev, LK Davis, D Posthuma, E Eskin, C Sabatti, CK Edlund, DV Conti, JA Knowles, NJ Cox. Project management: SE Stewart, JM Scharf, JA Fagerness, MA Jenike and DL Pauls. Sample management and processing: JA Fagerness, S Haddad, JM Scharf, J Crane, C Mayerfeld and DL Pauls. Genotyping: AT Crenshaw, MA Parkin and DB Mirel. Phenotype management: SE Stewart, L Osiecki, D Hezel, C Illmann, JM Scharf and DL Pauls. *Case sample collection* (ordered by numbers of submitted samples): University of Bonn, Germany: M Wagner, R Moessner (Site PI), P Falkai, W Maier, S Ruhrmann, H-J Grabe, L Lennertz. Italy: L Bellodi, MC Cavallini. Toronto, Canada/Wayne State collaborative: PD Arnold, MA Richter, EH Cook, Jr, JL Kennedy, D Rosenberg. University of Cape Town, South Africa: DJ Stein (Site PI), SMJ Hemmings, C Lochner. UCSF/ Costa Rica collaborative: CA Mathews (Site PI), A Azzam, DA Chavira, E Fournier, H Garrido, B Sheppard, P Umana. National Institute of Mental Health: DL Murphy, JR Wendland. Michigan: GL Hanna (Site PI), J Veenstra-VanderWeele. AMC, Netherlands: D Denys (Site PI), R Blom, D Deforce, F Van Nieuwerburgh, HGM Westenberg. Wurzburg Germany: S Walitza (Site PI), K Egberts, T Renner. Massachusetts General Hospital, Boston: DL Pauls (Site PI), C Illmann, SE Stewart, JM Scharf, SL Rauch. Brazil: EC Miguel (Site PI), C Cappi, AG Hounie, MC do Rosario, AS Sampaio, H Vallada. Mexico: H Nicolini (Site PI), N Lanzagorta, B Camarena. Paris, France: M Leboyer (Site PI), R Delorme. University of Southern California: MT Pato (Site PI), CN Pato, JA Knowles, E Voyiaziakis. VUMC, Netherlands: DC Cath (Site PI), P Heutink, D Posthuma, JH Smit. OCGS, Johns Hopkins collaborative: G Nestadt (Site PI), J Samuels, OJ Bienvenu, B Cullen, AJ Fyer, MA Grados, BD Greenberg, JT McCracken, MA Riddle, Y Wang. Yale University: JF Leckman (Site PI), M Bloch, C Pittenger, V Coric. United Arab Emirates: V Eapen. Iowa: DW Black. *Control Sample Collection*: University Medical Center, Utrecht: RA Ophoff, E Strengman. University of Bonn: R Moessner (Site PI), M Wagner, P Falkai, W Maier, S Ruhrmann, H-J Grabe, L Lennertz. *Data Collection*: Italian Control data: F Macciardi, D Cusi, M Turiel, F Frau. eQTL and mQTL data: C Liu. MR Cookson, JR Gibbs and A Singleton for the North American Brain Expression Consortium; J Hardy for the UK Human Brain Expression Database.

North American Brain Expression Consortium

S Arepalli[1], MR Cookson[1], A Dillman[1], L Ferrucci[2], JR Gibbs[1,3], DG Hernandez[1,3], R Johnson[4], DL Longo[5], MA Nalls[1], R O'Brien[6], A Singleton[1], B Traynor[1], J Troncoso[6], M van der Brug[1,7], HR Zielke[4], A Zonderman[8];

UK Human Brain Expression Database

JA Hardy[3], M Ryten[3], C Smith[9], D Trabzuni[3], R Walker[9], Mike Weale[10]

[1]Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA; [2]Clinical Research Branch, National Institute on Aging, Baltimore, MD, USA; [3]Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK; [4]NICHD Brain and Tissue Bank for Developmental Disorders, University of Maryland Medical School, Baltimore, MD, USA; [5]Lymphocyte Cell Biology Unit, Laboratory of Immunology, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA; [6]Brain Resource Center, Johns Hopkins University, Baltimore, MD, USA; [7]ITGR Biomarker Discovery Group, Genentech, South San Francisco, CA, USA; [8]Research Resources Branch, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA; [9]Department of Pathology, The University of Edinburgh, Edinburgh, UK and [10]King's College London, Department of Medical & Molecular Genetics, UK.

REFERENCES

1 Kaplan HI, Sadock BJ. Study guide and self-examination review for Kaplan and Sadock's synopsis of psychiatry. 6th edn. Williams & Wilkins: Baltimore, 1998, pp 541.
2 Karno M, Golding JM, Sorenson SB, Burnam MA. The epidemiology of obsessive-compulsive disorder in five US communities. *Arch Gen Psychiatry* 1988; **45**: 1094–1099.
3 Ruscio AM, Stein DJ, Chiu WT, Kessler RC. The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Mol Psychiatry* 2010; **15**: 53–63.
4 Blazer DG, Kessler RC, McGonagle KA, Swartz MS. The prevalence and distribution of major depression in a national community sample: the National Comorbidity Survey. *Am J Psychiatry* 1994; **151**: 979–986.
5 Ayuso-Mateos J. Global Burden of obsessive-compulsive disorder in the year 2000. *World Health Organization:* Geneva, 2006.
6 Luxenburger H. Hereditat und Familientypus der Zwangsneurotiker. *Arch Psychiatry* 1930; **91**: 590–594.
7 Lewis A. Problems of obsessional illness. *Proc R Soc Med* 1935; **29**: 325–336.
8 Brown FW. Heredity in the Psychoneuroses (Summary). *Proc R Soc Med* 1942; **35**: 785–790.
9 Rudin E. On the problem of compulsive disease with special reference to its hereditary relations]. *Arch Psychiatr Nervenkr Z Gesamte Neurol Psychiatr* 1953; **191**: 14–54.
10 Kringlen E. Obsessional neurotics: a long-term follow-up. *Br J Psychiatry* 1965; **111**: 709–722.
11 Rosenberg CM. Familial aspects of obsessional neurosis. *Br J Psychiatry* 1967; **113**: 405–413.
12 Carey G, Goldsmith H, Tellegen A, Gottesman I. Genetics and personality inventories: The limits of replication with twin data. *Behav Genet* 1978; **8**: 299–313.
13 Insel TR, Hoover C, Murphy DL. Parents of patients with obsessive-compulsive disorder. *Psychol Med* 1983; **13**: 807–811.
14 Rasmussen SA, Tsuang MT. Clinical characteristics and family history in DSM-III obsessive-compulsive disorder. *Am J Psychiatry* 1986; **143**: 317–322.

15 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.
16 Iervolino AC, Rijsdijk FV, Cherkas L, Fullana MA, Mataix-Cols D. A multivariate twin study of obsessive-compulsive symptom dimensions. *Arch Gen Psychiatry* 2011; **68**: 637–644.
17 Moore J, Smith GW, Shevlin M, O'Neill FA. Alternative factor models and heritability of the Short Leyton Obsessional Inventory-Children's Version. *J Abnorm Child Psychol* 2010; **38**: 921–934.
18 Bolton D, Rijsdijk F, Eley TC, O'Connor TG, Briskman J, Perrin S. Normative childhood repetitive routines and obsessive compulsive symptomatology in 6-year-old twins. *J Child Psychol Psychiatry* 2009; **50**: 1139–1146.
19 Hanna GL, Veenstra-VanderWeele J, Cox NJ, Boehnke M, Himle JA, Curtis GC *et al.* Genome-wide linkage analysis of families with obsessive-compulsive disorder ascertained through pediatric probands. *Am J Med Genet* 2002; **114**: 541–552.
20 Shugart YY, Samuels J, Willour VL, Grados MA, Greenberg BD, Knowles JA *et al.* Genomewide linkage scan for obsessive-compulsive disorder: evidence for susceptibility loci on chromosomes 3q, 7p, 1q, 15q, and 6q. *Mol Psychiatry* 2006; **11**: 763–770.
21 Willour VL, Yao Shugart Y, Samuels J, Grados M, Cullen B, Bienvenu OJ *et al.* Replication study supports evidence for linkage to 9p24 in obsessive-compulsive disorder. *Am J Hum Genet* 2004; **75**: 508–513.
22 Hanna GL, Veenstra-Vanderweele J, Cox NJ, Van Etten M, Fischer DJ, Himle JA *et al.* Evidence for a susceptibility locus on chromosome 10p15 in early-onset obsessive-compulsive disorder. *Biol Psychiatry* 2007; **62**: 856–862.
23 Liang KY, Wang Y, Shugart YY, Grados M, Fyer AJ, Rauch S *et al.* Evidence for potential relationship between SLC1A1 and a putative genetic linkage region on chromosome 14q to obsessive-compulsive disorder with compulsive hoarding. *Am J Med Genet B Neuropsychiatr Genet* 2008; **147B**: 1000–1002.
24 Ross J, Badner J, Garrido H, Sheppard B, Chavira DA, Grados M *et al.* Genomewide linkage analysis in Costa Rican families implicates chromosome 15q14 as a candidate region for OCD. *Hum Genet* 2011; **6**: 795–805.
25 Samuels J, Shugart YY, Grados MA, Willour VL, Bienvenu OJ, Greenberg BD *et al.* Significant linkage to compulsive hoarding on chromosome 14 in families with obsessive-compulsive disorder: results from the OCD Collaborative Genetics Study. *Am J Psychiatry* 2007; **164**: 493–499.
26 Pittenger C, Bloch MH, Williams K. Glutamate abnormalities in obsessive compulsive disorder: neurobiology, pathophysiology, and treatment. *Pharmacol Ther* 2011; **132**: 314–332.
27 Wu K, Hanna GL, Rosenberg DR, Arnold PD. The role of glutamate signaling in the pathogenesis and treatment of obsessive-compulsive disorder. *Pharmacol Biochem Behav* 2011; **100**: 726–735.
28 Stewart SE, Platko J, Fagerness J, Birns J, Jenike E, Smoller JW *et al.* A genetic family-based association study of OLIG2 in obsessive-compulsive disorder. *Arch Gen Psychiatry* 2007; **64**: 209–214.
29 Stewart SE, Fagerness JA, Platko J, Smoller JW, Scharf JM, Illmann C *et al.* Association of the SLC1A1 glutamate transporter gene and obsessive-compulsive disorder. *Am J Med Genet B Neuropsychiatr Genet* 2007; **144B**: 1027–1033.
30 Arnold PD, Sicard T, Burroughs E, Richter MA, Kennedy JL. Glutamate transporter gene SLC1A1 associated with obsessive-compulsive disorder. *Arch Gen Psychiatry* 2006; **63**: 769–776.
31 Dickel DE, Veenstra-VanderWeele J, Cox NJ, Wu X, Fischer DJ, Van Etten-Lee M *et al.* Association testing of the positional and functional candidate gene SLC1A1/EAAC1 in early-onset obsessive-compulsive disorder. *Arch Gen Psychiatry* 2006; **63**: 778–785.
32 Wendland JR, Moya PR, Timpano KR, Anavitarte AP, Kruse MR, Wheaton MG *et al.* A haplotype containing quantitative trait loci for SLC1A1 gene expression and its association with obsessive-compulsive disorder. *Arch Gen Psychiatry* 2009; **66**: 408–416.
33 Welch JM, Lu J, Rodriguiz RM, Trotta NC, Peca J, Ding JD *et al.* Cortico-striatal synaptic defects and OCD-like behaviours in Sapap3-mutant mice. *Nature* 2007; **448**: 894–900.
34 Züchner S, Wendland JR, Ashley-Koch AE, Collins AL, Tran-Viet KN, Quinn K *et al.* Multiple rare SAPAP3 missense variants in trichotillomania and OCD. *Mol Psychiatry* 2009; **14**: 6–9.
35 Bienvenu OJ, Wang Y, Shugart YY, Welch JM, Grados MA, Fyer AJ *et al.* Sapap3 and pathological grooming in humans: Results from the OCD collaborative genetics study. *Am J Med Genet B Neuropsychiatr Genet* 2009; **150B**: 710–720.
36 Boardman L, van der Merwe L, Lochner C, Kinnear CJ, Seedat S, Stein DJ *et al.* Investigating SAPAP3 variants in the etiology of obsessive-compulsive disorder and trichotillomania in the South African white population. *Compr Psychiatry* 2011; **52**: 181–187.
37 Cichon S, Craddock N, Daly M, Faraone SV, Gejman PV, Kelsoe J *et al.* Genomewide association studies: history, rationale, and prospects for psychiatric disorders. *Am J Psychiatry* 2009; **166**: 540–556.
38 Gamazon ER, Huang RS, Cox NJ, Dolan ME. Chemotherapeutic drug susceptibility associated SNPs are enriched in expression quantitative trait loci. *Proc Natl Acad Sci USA* 2010; **107**: 9287–9292.
39 Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA *et al.* Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 2011; **43**: 969–976.
40 Lee PH, O'Dushlaine C, Thomas B, Purcell SM. INRICH: interval-based enrichment analysis for genome wide association studies. *Bioinformatics* 2012; **28**: 1797–1799.
41 American Psychiatric Association.. *Diagnostic criteria from DSM-IV-TR*. American Psychiatric Association: Washington, D.C., 2000. xii, pp 370.
42 Scharf J. Genome-wide association study of Tourette Syndrome. *Mol Psychiatry* 2012.
43 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
44 Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010; **26**: 2190–2191.
45 Hemminger BM, Saelim B, Sullivan PF. TAMAL: an integrated approach to choosing SNPs for genetic studies of human complex traits. *Bioinformatics* 2006; **22**: 626–627.
46 Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H *et al.* The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res* 2006; **34**(Database issue): D590–D598.
47 The International HapMap C. A haplotype map of the human genome. *Nature* 2005; **437**: 1299–1320.
48 Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V *et al.* Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 2007; **35**(Database issue): D5–12.
49 Gamazon ER, Zhang W, Konkashbaev A, Duan S, Kistner EO, Nicolae DL *et al.* SCAN: SNP and copy number annotation. *Bioinformatics* 2010; **26**: 259–262.
50 Saccone SF, Saccone NL, Swan GE, Madden PA, Goate AM, Rice JP *et al.* Systematic biological prioritization after a genome-wide association study: an application to nicotine dependence. *Bioinformatics* 2008; **24**: 1805–1811.
51 Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM *et al.* The human genome browser at UCSC. *Genome Res* 2002; **12**: 996–1006.
52 Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 2005; **120**: 15–20.
53 Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet* 2010; **6**: e1000888.
54 Gibbs JR, van der Brug MP, Hernandez DG, Traynor BJ, Nalls MA, Lai SL *et al.* Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet* 2010; **6**: e1000952.
55 Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 2009; **5**: e1000529.
56 Consortium GP. A map of human genome variation from population-scale sequencing. *Nature* 2010; **467**: 1061–1073.
57 Dudbridge F, Gusnanto A. Estimation of significance thresholds for genomewide association scans. *Genet Epidemiol* 2008; **32**: 227–234.
58 Torii M, Hashimoto-Torii K, Levitt P, Rakic P. Integration of neuronal clones in the radial cortical columns by EphA and ephrin-A signalling. *Nature* 2009; **461**: 524–528.
59 Potkin SG, Turner JA, Fallon JA, Lakatos A, Keator DB, Guffanti G *et al.* Gene discovery through imaging genetics: identification of two novel genes associated with schizophrenia. *Mol Psychiatry* 2009; **14**: 416–428.
60 Sabéran-Djoneidi D, Marey-Semper I, Picart R, Studler JM, Tougard C, Glowinski J *et al.* A 19-kDa protein belonging to a new family is expressed in the Golgi apparatus of neural cells. *J Biol Chem* 1995; **270**: 1888–1893.
61 Schweitzer B, Suter U, Taylor V. Neural membrane protein 35/Lifeguard is localized at postsynaptic sites and in dendrites. *Brain Res Mol Brain Res* 2002; **107**: 47–56.
62 Rivière JB, Xiong L, Levchenko A, St-Onge J, Gaspar C, Dion Y *et al.* Association of intronic variants of the BTBD9 gene with Tourette syndrome. *Arch Neurol* 2009; **66**: 1267–1272.
63 Perez-Torrado R, Yamada D, Defossez PA. Born to bind: the BTB protein-protein interaction domain. *Bioessays* 2006; **28**: 1194–1202.
64 Sud A, Del Bono EA, Haines JL, Wiggs JL. Fine mapping of the GLC1 K juvenile primary open-angle glaucoma locus and exclusion of candidate genes. *Mol Vis* 2008; **14**: 1319–1326.
65 Stewart SE, Rosario MC, Brown TA, Carter AS, Leckman JF, Sukhodolsky D *et al.* Principal components analysis of obsessive-compulsive disorder symptoms in children and adolescents. *Biol Psychiatry* 2007; **61**: 285–291.
66 NIMH Transcriptional Atlas of Human Brain Development. Accessed on 11 November 2011. www.developinghumanbrain.org.

67 Zhang P, Xiang N, Chen Y, Sliwerska E, McInnis MG, Burmeister M *et al*. Family-based association analysis to finemap bipolar linkage peak on chromosome 8q24 using 2500 genotyped SNPs and 15 000 imputed SNPs. *Bipolar Disord* 2010; **12**: 786–792.
68 Wieczorek L, Maas Jr JW, Muglia LM, Vogt SK, Muglia LJ. Temporal and regional regulation of gene expression by calcium-stimulated adenylyl cyclase activity during fear memory. *PLoS One* 2010; **5**: e13385.
69 Rose JE, Behm FM, Drgon T, Johnson C, Uhl GR. Personalized smoking cessation: interactions between nicotine dose, dependence and quit-success genotype score. *Mol Med* 2010; **16**: 247–253.
70 Kirov G, Pocklington AJ, Holmans P, Ivanov D, Ikeda M, Ruderfer D *et al*. De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Mol Psychiatry* 2012; **17**: 142–153.
71 de Mooij-van Malsen AJ, van Lith HA, Oppelaar H, Hendriks J, de Wit M, Kostrzewa E *et al*. Interspecies trait genetics reveals association of Adcy8 with mouse avoidance behavior and a human mood disorder. *Biol Psychiatry* 2009; **66**: 1123–1130.
72 Kantojärvi K, Onkamo P, Vanhala R, Alen R, Hedman M, Sajantila A *et al*. Analysis of 9p24 and 11p12-13 regions in autism spectrum disorders: rs1340513 in the JMJD2C gene is associated with ASDs in Finnish sample. *Psychiatr Genet* 2010; **20**: 102–108.

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