Molecular Psychiatry (2013) 18, 721–728
© 2013 Macmillan Publishers Limited All rights reserved 1359-4184/13
www.nature.com/mp

ORIGINAL ARTICLE

Genome-wide association study of Tourette's syndrome

JM Scharf[1,2,3,4,5,61], D Yu[1,2,61], CA Mathews[6,61], BM Neale[1,2,5,7,8,61], SE Stewart[1,2,9,61], JA Fagerness[1,2,61], P Evans[10,11], E Gamazon[10,11], CK Edlund[12,13], SK Service[14], A Tikhomirov[10,11], L Osiecki[1,2], C Illmann[1,2], A Pluzhnikov[10,11], A Konkashbaev[10,11], LK Davis[10,11], B Han[15], J Crane[1,2], P Moorjani[7,16], AT Crenshaw[7], MA Parkin[7], VI Reus[6], TL Lowe[6], M Rangel-Lugo[6], S Chouinard[17], Y Dion[17], S Girard[17], DC Cath[18,19], JH Smit[18], RA King[20], TV Fernandez[20], JF Leckman[20], KK Kidd[21], JR Kidd[21], AJ Pakstis[21], MW State[21], LD Herrera[22], R Romero[22], E Fournier[22], P Sandor[23,24], CL Barr[23,25], N Phan[23], V Gross-Tsur[26], F Benarroch[27], Y Pollak[26], CL Budman[28,29], RD Bruun[28,30], G Erenberg[31], AL Naarden[32], PC Lee[6], N Weiss[6], B Kremeyer[33,34], GB Berrío[35], DD Campbell[33,36], JC Cardona Silgado[35], WC Ochoa[35], SC Mesa Restrepo[35], H Muller[33], AV Valencia Duarte[35,37], GJ Lyon[38,39], M Leppert[38], J Morgan[38], R Weiss[38], MA Grados[40], K Anderson[40], S Davarya[40], H Singer[40], J Walkup[41], J Jankovic[42], JA Tischfield[43,44], GA Heiman[43,44], DL Gilbert[45], PJ Hoekstra[46], MM Robertson[33,47], R Kurlan[48], C Liu[49], JR Gibbs[50], A Singleton[50] for the North American Brain Expression Consortium, J Hardy[33] for the UK Human Brain Expression Database, E Strengman[14,51], RA Ophoff[14,51], M Wagner[52], R Moessner[52], DB Mirel[7], D Posthuma[53,54,55], C Sabatti[56], E Eskin[15,57], DV Conti[12], JA Knowles[58], A Ruiz-Linares[33], GA Rouleau[17], S Purcell[1,2,5,7,8,59], P Heutink[54], BA Oostra[60], WM McMahon[38], NB Freimer[14], NJ Cox[10,11] and DL Pauls[1,2]

Tourette's syndrome (TS) is a developmental disorder that has one of the highest familial recurrence rates among neuropsychiatric diseases with complex inheritance. However, the identification of definitive TS susceptibility genes remains elusive. Here, we report the first genome-wide association study (GWAS) of TS in 1285 cases and 4964 ancestry-matched controls of European ancestry, including two European-derived population isolates, Ashkenazi Jews from North America and Israel and French Canadians from Quebec, Canada. In a primary meta-analysis of GWAS data from these European ancestry samples, no markers achieved a genome-wide threshold of significance ($P<5\times10^{-8}$); the top signal was found in rs7868992 on chromosome 9q32 within *COL27A1* ($P=1.85\times10^{-6}$). A secondary analysis including an additional 211 cases and 285 controls from two closely related Latin American population isolates from the Central Valley of Costa Rica and Antioquia,

[1]Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetics Research, Boston, MA, USA; [2]Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA; [3]Movement Disorders Unit, Department of Neurology, Massachusetts General Hospital, Boston, MA, USA; [4]Division of Cognitive and Behavioral Neurology, Brigham and Women's Hospital, Boston, MA, USA; [5]Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA, USA; [6]Department of Psychiatry, University of California San Francisco, San Francisco, CA, USA; [7]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA; [8]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA; [9]British Columbia Mental Health and Addictions Research Institute, University of British Columbia, Vancouver, BC, Canada; [10]Department of Human Genetics, University of Chicago, Chicago, IL, USA; [11]Department of Medicine, University of Chicago, Chicago, IL, USA; [12]Department of Preventative Medicine, Division of Biostatistics, University of Southern California, Los Angeles, CA, USA; [13]USC Epigenome Center, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; [14]Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, CA, USA; [15]Department of Computer Science, University of California Los Angeles, Los Angeles, CA, USA; [16]Department of Genetics, Harvard University, Cambridge, MA, USA; [17]University of Montreal, Montreal, QC, Canada; [18]Department of Psychiatry, VU University Medical Center, Amsterdam, The Netherlands; [19]Department of Clinical & Health Psychology, Utrecht University, Utrecht, The Netherlands; [20]Yale Child Study Center, Yale University School of Medicine, New Haven, CT, USA; [21]Department of Genetics, Yale University School of Medicine, New Haven, CT, USA; [22]Hospital Nacional de Niños, San Jose, Costa Rica; [23]The Toronto Western Research Institute, University Health Network, Toronto, ON, Canada; [24]Division of Child Psychiatry, Department of Psychiatry, University of Toronto, Toronto, ON, Canada; [25]The Hospital for Sick Children, Toronto, ON, Canada; [26]Neuropediatric Unit, Shaare Zedek Medical Center, Jerusalem, Israel; [27]Herman Dana Division of Child and Adolescent Psychiatry, Hadassah-Hebrew University Medical Center, Jerusalem, Israel; [28]North Shore-Long Island Jewish Medical Center, Manhasset, NY, USA; [29]Hofstra University School of Medicine, Hempstead, NY, USA; [30]New York University Medical Center, New York, NY, USA; [31]Cleveland Clinic, Cleveland, OH, USA; [32]Department of Clinical Research, Medical City Dallas Hospital, Dallas, TX, USA; [33]University College London, London, UK; [34]Wellcome Trust Sanger Institute, Cambridge, UK; [35]Universidad de Antioquia, Medellín, Colombia; [36]Department of Psychiatry and Center for Genomic Sciences, The University of Hong Kong, Hong Kong, SAR; [37]Universidad Pontificia Bolivariana, Medellín, Colombia; [38]University of Utah, Salt Lake City, UT, USA; [39]Stanley Institute for Cognitive Genomics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA; [40]Johns Hopkins University School of Medicine, Baltimore, MD, USA; [41]Division of Child and Adolescent Psychiatry, Department of Psychiatry, Weill Cornell Medical Center, New York, NY, USA; [42]Parkinson's Disease Center and Movement Disorders Clinic, Department of Neurology, Baylor College of Medicine, Houston, TX, USA; [43]Department of Genetics, Rutgers University, Piscataway, NJ, USA; [44]Human Genetics Institute of New Jersey, Rutgers University, Piscataway, NJ, USA; [45]Cincinnati Children's Hospital Medical Center and the University of Cincinnati, Cincinnati, OH, USA; [46]Department of Psychiatry, University Medical Center, University of Groningen, Groningen, The Netherlands; [47]St George's Hospital and Medical School, London, UK; [48]Atlantic Neuroscience Institute, Overlook Hospital, Summit, NJ, USA; [49]Department of Psychiatry, University of Illinois at Chicago, Chicago, IL, USA; [50]Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA; [51]Department of Psychiatry, Rudolf Magnus Institute of Neuroscience, University Medical Center, Utrecht, The Netherlands; [52]Department of Psychiatry and Psychotherapy, University of Bonn, Bonn, Germany; [53]Department of Functional Genomics, Center for Neurogenomics and Cognitive Research, Neuroscience Campus Amsterdam, VU University Amsterdam, De Boelelaan, Amsterdam, The Netherlands; [54]Section of Medical Genomics, Department of Clinical Genetics, VU Medical Centre, De Boelelaan, Amsterdam, The Netherlands; [55]Department of Child and Adolescent Psychiatry/Psychology, Erasmus University Medical Center-Sophia Children's Hospital, Wytemaweg, Rotterdam, The Netherlands; [56]Department of Health Research and Policy, Stanford University, Stanford, CA, USA; [57]Department of Human Genetics, University of California Los Angeles, Los Angeles, CA, USA; [58]Department of Psychiatry and the Behavioral Sciences, Zilkha Neurogenetic Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; [59]Division of Psychiatric Genomics, Mount Sinai School of Medicine, New York, NY, USA and [60]Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands. Correspondence: Professor DL Pauls or Dr JM Scharf, Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA 02114, USA.
E-mail: dpauls@pngu.mgh.harvard.edu (DLP) or jscharf@partners.org (JMS)
[61]These authors contributed equally to this work.
Received 21 November 2011; revised 19 April 2012; accepted 24 April 2012; published online 14 August 2012

Colombia also identified rs7868992 as the top signal ($P = 3.6 \times 10^{-7}$ for the combined sample of 1496 cases and 5249 controls following imputation with 1000 Genomes data). This study lays the groundwork for the eventual identification of common TS susceptibility variants in larger cohorts and helps to provide a more complete understanding of the full genetic architecture of this disorder.

Molecular Psychiatry (2013) **18,** 721–728; doi:10.1038/mp.2012.69; published online 14 August 2012

Keywords: genetics; GWAS; neurodevelopmental disorder; tics; Tourette's syndrome

INTRODUCTION

Tourette's syndrome (TS) is a chronic, childhood-onset neuropsychiatric disorder characterized by multiple motor tics and at least one phonic tic that persist for >1 year.[1,2] TS has a population prevalence of $\sim$0.3–0.8% and, like many neurodevelopmental disorders, occurs more frequently in boys, with male/female ratios ranging between 3:1 and 4:1.[3,4] It is frequently accompanied by a wide range of additional psychiatric comorbidities, in particular obsessive–compulsive disorder (OCD) and attention-deficit hyperactivity disorder.[5] TS causes substantial physical and psychosocial morbidity in children and adolescents, and can produce lifelong disability in severe cases.[6,7]

Twin and family studies have repeatedly demonstrated that TS is highly heritable.[8] First-degree relatives of affected individuals have a 5- to 15-fold increased risk of TS compared with that of the general population, representing one of the highest familial recurrence risks among common neuropsychiatric diseases.[3,9] However, despite this strong familiality, identification of TS susceptibility genes has been challenging. Linkage analyses have produced inconsistent results, although a recent study combining multigenerational families with affected sibling pairs has identified at least one major TS locus on chromosome 2p.[10] Multiple candidate genes have also been proposed, although none have been consistently replicated.[8] Mutations in the strongest TS candidate genes (*SLITRK1*, *CNTNAP2* and *HDC*) have been found only in single families or a small number of individuals, suggesting that, if truly causative, they account for only a small proportion of TS cases.[11–15] Thus, additional gene-finding strategies are needed. Here, we report the first TS genome-wide association study (GWAS) in a large cohort of samples of general European ancestry, as well as two European-derived population isolates, Ashkenazi Jews from the United States and Israel (AJ) and French Canadians from Quebec, Canada (FC), and two closely related Latin American population isolates, the Central Valley of Costa Rica (CVCR) and Antioquia, Colombia (ANT).

MATERIALS AND METHODS

Cases

A total of 1998 TS cases were recruited from 20 sites in the United States, Canada, UK, The Netherlands, Israel, Costa Rica and Colombia and divided into four strata based on self-reported ancestry: (1) 1252 European ancestry, nonisolate cases from North America and Europe (EU); (2) 210 Ashkenazi Jewish cases from the United States and Israel (AJ); (3) 302 French Canadian cases (FC); and (4) cases from two closely related population isolates from the Central Valley of Costa Rica (CVCR; $n = 137$) and Antioquia, Colombia (ANT; $n = 97$) (see Supplementary Methods). Inclusion criteria required a TS Classification Study Group (TSCSG) diagnosis of definite TS (a DSM-IV-TR diagnosis of TS plus tics observed by an experienced clinician),[16] and available genomic DNA extracted either from blood or cell lines. Exclusion criteria consisted of a history of intellectual disability, tardive tourettism or other known genetic, metabolic or acquired tic disorders. Subjects from 17 of the 20 sites were assessed for a lifetime diagnosis of TS, OCD and attention-deficit hyperactivity disorder using a standardized and validated semistructured interview that has high validity and reliability for TS ($\kappa = 1.00$) and OCD ($\kappa = 0.97$).[10] Subjects from the other three sites were assessed only for a lifetime diagnosis of definite TS.

Controls

A total of 5403 European ancestry controls were derived primarily from cohorts of previously genotyped, unselected population controls (see Supplementary Methods, Supplementary Table S1). These included 3212 controls from the Illumina Genotype Control Database genotyped on the Illumina HumanHap550v1/v3 platforms (www.Illumina.com, Illumina, San Diego, CA, USA), 1288 controls from the Studies of Addiction: Genetics and Environment (SAGE) cohort[17–19] genotyped on the Illumina Human-Hap1Mv1_C and 653 Dutch ancestry controls genotyped on the Illumina HumanHap550v1.[20] An additional 298 German and Dutch EU controls were genotyped simultaneously with the TS case samples, including 48 duplicates from the Dutch 550v1 control cohort, to facilitate cross-platform and cross-facility comparisons.

A total of 297 FC and 380 ANT ancestry-matched controls were collected in parallel with their respective cases (see Supplementary Methods). ANT controls were used for analysis of both ANT and CVCR cases given their shared ancestry.[21,22] All participants $\geqslant 18$ years of age gave informed consent. Individuals <18 years of age gave assent after a parent signed a consent form on their behalf. The research project was approved by the ethics committees of each participating site.

Genotyping

Genotyping of 908 of the 1252 EU cases and all population-isolate cases (AJ, FC, ANT and CVCR), as well as 298 EU and all FC and ANT controls, was conducted on the Illumina Human610-Quadv1_B SNP array (Illumina) at the Broad Institute of Harvard and MIT (Cambridge, MA, USA) in two batches using standard protocols. Samples were randomized across plates and batches both by originating site and case–control status. Genotype calling was performed using BeadStudio (Illumina). In addition, 432 EU cases were genotyped on the Illumina HumanCNV370-Duo_v1 at the Yale Center for Genome Analysis (New Haven, CT, USA), including 88 duplicate EU samples overlapping with those genotyped on the 610-Quad platform to allow for cross-platform checks of concordance.

Quality control

Quality control (QC) analyses were performed using PLINK v1.07 (ref. 23) and EIGENSTRAT.[24] In addition to standard QC protocols, particular detail focused on cross-platform comparisons of concordance, allele frequency and differential missingness, given the use of control samples genotyped previously on different Illumina platforms (full details and ordered QC pipeline available online, Supplementary Figure S1). In general, two thresholds were used for single-nucleotide polymorphism (SNP) QC: a more stringent threshold at which SNPs were removed, and a second liberalized threshold for which SNPs were flagged and re-examined later for potential QC-related bias. All flagged SNPs with $P < 1 \times 10^{-3}$ in any analysis are annotated in Supplementary Tables S2–S4.

Sample and SNP QC were initially performed within each platform separately (Supplementary Figure S1). Samples were removed for autosomal call rates $<98\%$, discrepancy between phenotypic and genetic sex, and indeterminate genetic sex. In addition, all 151 cases from one site were removed because of increased rates of missing SNP data relative to other sites (Supplementary Figure S2). Platform-specific SNP QC included removing monomorphic SNPs, copy number variation targeted SNP probes,

© 2013 Macmillan Publishers Limited

Table 1. Characteristics of the final TS GWAS samples

	Cases	*Controls*
N	1496	5249
Gender (% male)	79%	39%
Age at assessment, years (mean, s.d.)[a]	16.6 ± 11.5	
Age of tic onset, years (mean, s.d.)[b]	6.0 ± 2.8	
OCD (%)[c]	42%	
ADHD (%)[d]	61%	

Abbreviations: ADHD, attention-deficit hyperactivity disorder; GWAS, genome-wide association study; OCD, obsessive–compulsive disorder; TS, Tourette's syndrome.
[a]Based on 1247 cases with available data.
[b]Based on 1110 cases.
[c]Based on 1223 cases.
[d]Based on 1048 cases.

SNPs with genotyping rate $<98\%$ and strand-ambiguous SNPs with significant allele frequency differences or aberrant linkage disequilibrium (LD) correlations with adjacent SNPs based on the entire HapMap2 reference panel. Concordance was checked between 82 duplicates genotyped both on the 610-Quad (Broad) and 370K (Yale), as well as 41 duplicates genotyped on the 610-Quad and 550v1. In addition, concordance was examined in HapMap duplicates from the Illumina database genotyped on two or more platforms used in this study. No SNPs were identified with significant association between the two 610-Quad genotyping batches.

After merging samples from all platforms, SNPs with minor allele frequency difference of >0.15 between case–case or control–control platforms were flagged, as were SNPs with $>1\%$ Mendelian errors in a parallel sample of 400 OCD trios genotyped simultaneously with the TS cases.[25] Any SNP not present on the three major common platforms (550v1, 610-Quad and 1M) was removed, leaving 496 877 SNPs for population-specific QC.

Multidimensional scaling analysis was used to exclude duplicate and related samples as well as samples of non-European descent (other than the CVCR/ANT samples, which were set aside for subpopulation-specific QC; Supplementary Figure S3). Remaining EU and European-derived isolate samples were separated into three strata (EU, AJ and FC) based on observed genetic ancestry and source population (Supplementary Figures S4–S6). Within each of the multidimensional scaling-defined genetic subpopulations, additional outliers were removed for excess low-level relatedness, abnormal average heterozygosity or inadequate case–control matching. The final European ancestry sample contained 1285 cases and 4964 controls (EU: 778 cases and 4414 controls; AJ: 242 cases and 354 controls; FC: 265 cases and 196 controls; see Table 1 and Supplementary Figure S1). The final CVCR/ANT sample consisted of 211 cases (87 ANT, 124 CVCR) and 285 ANT controls.

Subpopulation-specific SNP QC included removal of SNPs with Hardy–Weinberg equilibrium $P<10^{-10}$ in controls (flagged for Hardy–Weinberg equilibrium $P<10^{-5}$) and two additional cross-platform QC steps to remove SNPs with differential missingness between cases and controls across the five Illumina data sets (Supplementary Figure S7). The final number of SNPs for meta-analyses across all populations was 484 295 SNPs.

Genetic association and meta-analysis

Four ancestry-stratified association analyses were performed using PLINK, version 1.07, (ref. 23) employing logistic regression under an additive model with significant subpopulation-specific multidimensional scaling dimensions included as covariates to control for residual population stratification. Strata were then combined in a case-weighted meta-analysis in METAL,[26] assuming a fixed-effects model. For X-chromosome SNPs, males and females were analyzed separately first and subsequently combined by meta-analysis (Supplementary Methods). For all SNPs, two meta-analyses were conducted: a primary analysis with the European-derived strata only (EU, AJ and FC), and an exploratory, secondary meta-analysis including the CVCR/ANT Latin American samples. Heterogeneity was assessed using Cochran's Q and I^2 statistics.

Enrichment analyses

Expression quantitative trait loci (eQTL) data from lymphoblast cell lines (LCLs), cerebellum and frontal cortex were generated as described previously.[27–29] Similarly, methylation QTLs (mQTLs), which represent SNPs that are associated with variation in genome-wide patterns of methylation, were derived from adult cerebellum.[30] The top distribution of GWAS SNPs from the primary meta-analysis, 412 LD-pruned SNPs with $P<0.001$, were tested for eQTL or mQTL enrichment compared with 1000 randomly drawn, LD-pruned sets of allele frequency-matched SNPs taken from the set of typed SNPs on the Illumina 550K (Supplementary Methods). The number of eQTLs (or mQTLs) in each simulated set yielded an empirical distribution and enrichment *P*-value, calculated as the proportion of randomized sets in which the eQTL/mQTL count matched or exceeded the actual observed count in the list of top SNP associations. A similar analysis was performed to test for enrichment of missense SNPs or SNPs within a gene as defined by dbSNP annotation.

Imputation

Imputation of SNPs from the 1000 Genomes Project was performed using IMPUTE2 (ref. 31) and haplotypes from all 1092 individuals in the 1000 Genomes June 2011 Data Release[32] as a reference data set (Supplementary Methods). Postimputation QC and allelic dosage analysis were conducted in each subpopulation separately in PLINK followed by case-weighted meta-analysis in METAL.

RESULTS

QC analyses in individual ancestral subpopulations

After QC filtering, 1285 cases and 4964 controls remained across the three European ancestry strata (EU, AJ and FC). Examination of quantile-quantile (Q-Q) plots and genomic control λ values of the individual subpopulation-specific analyses revealed no evidence of residual population stratification or systematic technical artifact (EU, $\lambda=1.011$; AJ, $\lambda=0.993$; FC, $\lambda=0.971$; Supplementary Figure S8a–c). The Latin American population isolate stratum (CVCR/ANT) showed a small inflation of the median test statistic ($\lambda=1.044$), indicative of some residual stratification between CVCR and ANT samples (Supplementary Figure S6). However, no SNPs in this subpopulation-specific analysis had extreme *P*-values outside the expected null distribution (Supplementary Figure S8d).

Primary meta-analysis of GWAS data from European-derived subpopulations

In the primary meta-analysis of European-derived samples, no SNP surpassed a genome-wide significant threshold of $P<5.0\times10^{-8}$ (Figure 1). The top five LD-independent loci are annotated in Table 2; full annotation of all SNPs with $P<1\times10^{-3}$ are provided in Supplementary Table S2. The SNP with the strongest signal, rs7868992, lies on chromosome 9q32 within an intron of *COL27A1* ($P=1.85\times10^{-6}$; Supplementary Figure S9). The other four top independent GWAS signals include rs6539267, an intronic SNP within *POLR3B* on chromosome 12q23 ($P=7.41\times10^{-6}$; Supplementary Figure S10); rs13063502, a SNP that lies in a 1.7-Mb intergenic region on chromosome 3q13 ($P=8.96\times10^{-6}$; Supplementary Figure S11); rs7336083, located on chromosome 13q31 within a 1.9-Mb intergenic region between *SLITRK6* and *SLITRK1* (ref. 12) ($P=9.49\times10^{-6}$; Supplementary Figure S12); and rs769111, an intergenic SNP on chromosome 7p21 between *THSD7A* and *TMEM106B* ($P=1.20\times10^{-5}$; Supplementary Figure S13). No effect-size heterogeneity was present between the three European-derived subpopulations for SNPs rs7868992, rs6539267 and rs7336083 (Supplementary Figures S9–13). rs13063592 and rs769111 demonstrated moderate heterogeneity ($I^2=45.4$ and 64.2%, respectively), although the direction of effect was consistent across the EU, AJ and FC populations.

© 2013 Macmillan Publishers Limited

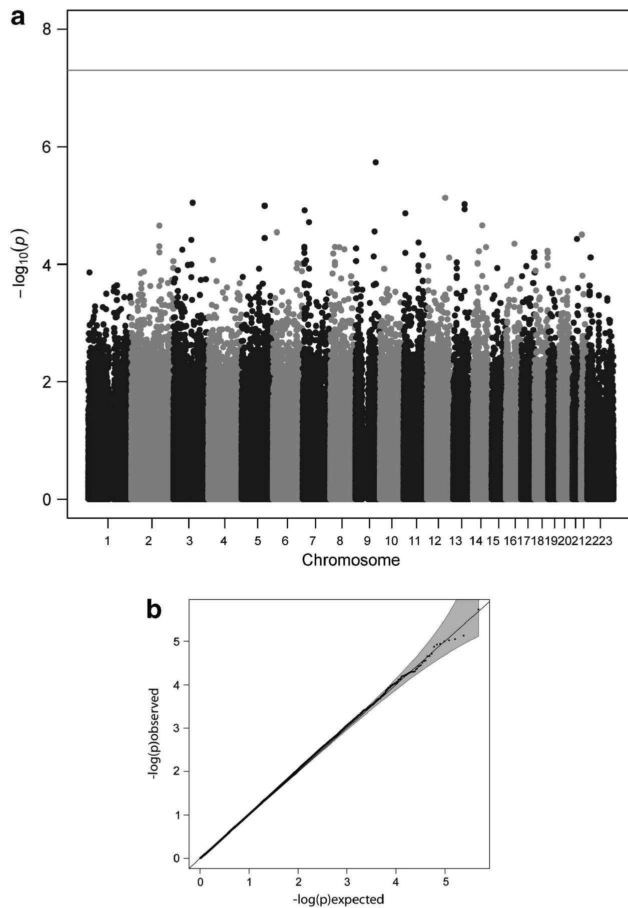

Figure 1. Results of the primary meta-analysis from the three European ancestry Tourette's syndrome (TS) populations. (**a**) Manhattan plot of all genotyped single-nucleotide polymorphisms (SNPs) for 1285 TS cases and 4964 controls from the EU, AJ and FC populations. Grey line indicates the genome-wide significance threshold of 5×10^{-8}. (**b**) Quantile-quantile plot of observed vs expected $-\log$ (*P*) values from the primary meta-analysis. The 95% confidence interval of expected values is indicated in grey. The genomic control λ value is 0.996. AJ, Ashkenazi Jews from the United States and Israel; EU, European ancestry, nonisolate cases from North America and Europe; FC, French Canadians from Quebec, Canada.

Analysis of Latin American TS GWAS data and meta-analysis of all TS samples

In the secondary meta-analysis combining all 1496 TS cases and 5249 controls (European ancestry samples plus 211 cases and 285 controls from the Latin American CVCR/ANT samples), the strongest association was again found in rs7868992 within *COL27A1* on 9q32 (combined $P = 2.94. \times 10^{-8}$; Supplementary Table S5, Supplementary Figures S9c, S14). Examination of an LD-pruned set of top SNPs from the primary meta-analysis (412 SNPs with $P < 1 \times 10^{-3}$) found a slight, but nonsignificant increase in the number of SNPs with the same direction of effect in the CVCR/ANT analysis (223/412, $P = 0.052$, one-sided binomial sign test; Supplementary Tables S2 and S3).

Analysis of imputed data

Imputation was performed using 1000 Genomes Project data[32] to identify additional supportive SNPs within the top signals from each meta-analysis. Q-Q plots of the primary and secondary meta-analyses incorporating imputed data demonstrated minimal inflation of the median test statistic (Supplementary Figure S15). No imputed SNPs in either meta-analysis surpassed the genome-wide significant threshold of $P < 5 \times 10^{-8}$. rs7868992 remained the top SNP overall, although its *P*-value dropped to 3.61×10^{-7} following imputation (Supplementary Figure S9c).

Enrichment analyses of eQTLs and mQTLs

As many of the top signals in the primary meta-analysis ($P < 0.001$) appeared to lie within or adjacent to known brain-expressed genes (Supplementary Table S2), we sought functional evidence to support the observed associations by evaluating the effect of these SNPs on transcriptional expression and DNA methylation levels. We annotated all GWAS SNPs with eQTL information derived previously from LCLs, adult cerebellum and frontal cortex as well as mQTL information from adult cerebellum (Supplementary Table S2). The top LD-independent SNPs (412 SNPs with $P < 0.001$) were subsequently tested for eQTL and mQTL enrichment. These top SNPs from the primary analysis were nominally enriched for eQTLs in frontal cortex (empirical *P*-value $= 0.045$) with a trend toward enrichment in cerebellum ($P = 0.077$), but no enrichment in LCLs ($P = 0.712$; Figures 2a–c). The highest association signals were also nominally enriched for cerebellar mQTLs ($P = 0.011$; Figure 2d). A similar test for SNPs located within gene loci found no enrichment ($P = 0.258$), although missense SNPs demonstrated a borderline enrichment ($P = 0.098$).

Table 2. Top five LD-independent signals in the primary European-derived TS meta-analysis

CHR	*SNP*	*BP*	*A1/A2*	*Primary European meta-analysis*			*No. of SNPs in LD*[a]	*Annotation*				
				MAF	*OR*	*P-value*		*Gene*	*Left gene*	*Right gene*	*eQTL*	*Cerebellar mQTL*
9	rs7868992	116030892	G/A	0.28	1.29	1.85×10^{-6}	1	*COL27A1* (intron)	*KIF12*	*ORM1*	—	*SYTL4, AMBP, HSPC152, OAS2, PWP1, RALBP1*
12	rs6539267	105309684	C/T	0.31	0.79	7.41×10^{-6}	0	*POLR3B* (intron)	*TCP11L2*	*FLJ45508*	—	*TMEM119*
3	rs13063502	110707002	T/C	0.14	1.35	8.96×10^{-6}	0	—	*FLJ25363*	*LOC440973*	—	—
13	rs7336083	84901388	A/G	0.34	0.80	9.49×10^{-6}	2	—	*LOC387939*	*SLITRK6*	*SLITRK6* (cerebellum)	*SORT1, ARFGAP1, CSN3*
7	rs769111	12026331	G/T	0.38	0.81	1.20×10^{-5}	4	—	*THSD7A*	*TMEM106B*	*MEOX2* (cerebellum)	*PLSCR1, PCDHB16*

Abbreviations: A1, reference allele; A2, alternative allele; BP, hg19 position; CHR, chromosome; eQTL, expression quantitative trait loci; LD, linkage disequilibrium; MAF, minor allele frequency; mQTL, methylation quantitative trait loci; OR, odds ratio; SNP, single-nucleotide polymorphism; TS, Tourette's syndrome.
[a]Number of additional SNPs in LD with association *P*-values $< 1 \times 10^{-3}$ in the primary meta-analysis (LD defined as $r^2 > 0.5$). Complete annotation of these SNPs as well as all SNPs with association *P*-values $< 1 \times 10^{-3}$ is provided in Supplementary Table S2.

© 2013 Macmillan Publishers Limited

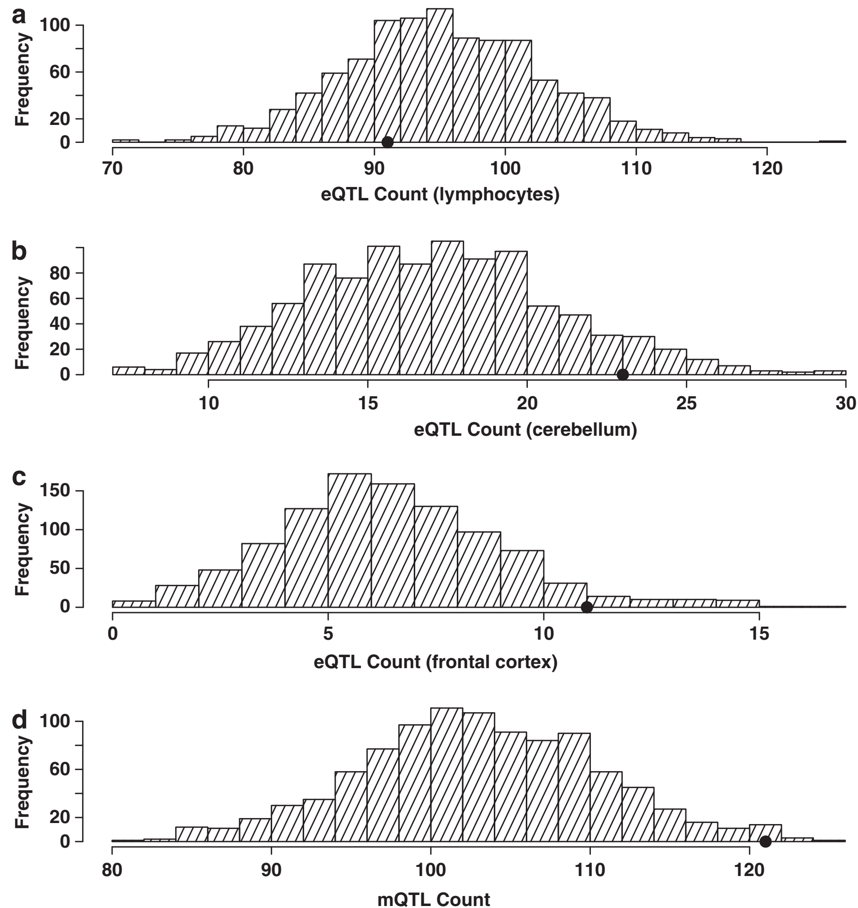

Figure 2. Enrichment analysis of functional single-nucleotide polymorphisms (SNPs) within the top signals of the primary Tourette's syndrome (TS) meta-analysis. Filled circles indicate the observed count of expression quantitative trait loci (eQTLs) or methylation QTLs (mQTLs) among the top loci ($P<1\times10^{-3}$) in the primary European-derived meta-analysis following linkage disequilibrium (LD) pruning. Empirical *P*-values indicate the rank of the observed eQTL (or mQTL) count relative to 1000 random sets of allele frequency-matched SNPs drawn from the entire null distribution of LD-pruned SNPs (hatched boxes). (**a**) Lymphoblast cell line eQTLs, $P=0.712$; (**b**) cerebellar eQTLs, $P=0.077$; (**c**) frontal cortex eQTLs, $P=0.045$; and (**d**) cerebellar mQTLs, $P=0.011$.

Examination of previously reported TS candidate genes

As an additional exploratory analysis, we examined the associations of SNPs within 50 kb of 24 previously reported candidate TS genes (Supplementary Tables S6 and S7). We found no excess of lower *P*-values among the 2135 SNPs within these genes compared with those expected under the null, suggesting that these candidate genes are not enriched for common SNPs associated with TS (Supplementary Figure S16). One signal in the primary European ancestry meta-analysis had a nominal $P<1\times10^{-3}$ (rs10277969 within *CNTNAP2*, $P=7.8\times10^{-4}$), but this locus did not survive a Bonferroni correction for gene size (266 LD-independent SNPs within *CNTNAP2*, corrected $P=0.21$).

DISCUSSION

Although the current sample of 1496 TS cases and 5249 controls is the largest studied to date, no loci in our analysis reached the widely accepted statistical threshold for genome-wide significance of $P\leqslant5\times10^{-8}$ (refs. 33,34). This observation is not surprising, given that GWA studies for other highly heritable neuropsychiatric diseases (for example, autism, bipolar disorder and schizophrenia) have required sample sizes of 5000–10 000 cases to identify definitive common risk alleles with modest effect sizes (odds ratios <1.3) (ref. 35). However, the marginal enrichment of functional brain variants (eQTLs and mQTLs) within the top loci in the primary meta-analysis (Figure 2) suggests that a subset of top signals in our analysis are true associations that may contribute to TS risk through effects on gene expression and methylation. In particular, the trend toward enrichment of frontal cortex eQTLs compared with eQTLs in cerebellum and LCLs is anatomically consistent with the hypothesis that TS is caused by abnormalities in frontostriatal circuitry.[36] Nonetheless, given the nominal significance of these enrichment results, further studies in larger samples are needed before drawing definitive conclusions.

The strongest signal in the primary European ancestry meta-analysis, rs7868992, was also the top locus in the secondary meta-analysis, which incorporated an additional 496 non-European cases and controls from the CVCR and ANT Latin American population isolates (Supplementary Figure S9). In this combined analysis, rs7868992 initially achieved a *P*-value of 2.94×10^{-8}, surpassing the threshold for genome-wide significance. However, following imputation, this signal decreased to $P=3.61\times10^{-7}$, most likely because of the incorporation of imputed data from the 148 European ancestry cases genotyped on the Illumina 370K, which does not directly interrogate rs7868992. Nonetheless, rs7868992 performed robustly on the other Illumina platforms used in this study based on review of the normalized intensity plots (Supplementary Figure S9d) and the 100% concordance rate in all cross-platform comparisons of this SNP in HapMap

© 2013 Macmillan Publishers Limited

duplicates from the Illumina database (Supplementary Materials). Therefore, rs7868992 remains a promising candidate, but cannot be considered a TS susceptibility variant unless it is replicated in an independent sample.

rs7868992 is located within an intron of *COL27A1*, the type XXVII collagen α chain gene. *COL27A1* is a fibrillar collagen primarily expressed in cartilage, although it is expressed in the cerebellum during many stages of human development.[37,38] Although nonfibrillar collagens have been implicated in various neurodevelopmental processes (for example, axon guidance and synaptogenesis), the function of *COL27A1* in the developing nervous system is unknown.[39]

The second top SNP in the primary analysis, rs6539267, is located on chromosome 12q23 within an intron of *POLR3B*. This gene encodes the second largest subunit of RNA polymerase III, which transcribes eukaryotic noncoding RNAs including transfer RNAs, small ribosomal RNAs and microRNAs.[40] Recessive mutations in *POLR3B* cause hypomyelinating leukodystrophy with a severe neurological phenotype (developmental delay, spasticity, dysarthria and ataxia), though no reported tics.[41,42] Both the secondary meta-analysis and imputed data provide additional support for this locus, and expand the region of LD to ~300 kb, including adjacent genes *CKAP4*, *TCP11L2* and *RFX4* (Supplementary Table S5 and Supplementary Figure S10).

The other three top loci in the primary analysis are located within large intergenic regions. rs13063502 on 3q13.1 lies between the noncoding complementary DNA *FLJ25363* and *PVRL3*, which resides 1.5 Mb telomeric to rs13063502 and is expressed primarily in placenta and testis.[43] rs769111 on 7p21.3 is situated between *THSD7A*, a gene expressed almost exclusively in developing endothelial cells,[44] and *TMEM106B*, a gene recently associated with frontotemporal dementia with TDP-43 inclusions, whose primary function in the brain remains to be elucidated.[45] Last, rs7336083 lies in a 1.9-Mb intergenic region between *SLITRK1* and *SLITRK6* on chromosome 13q31. Although *SLITRK1* is an *a priori* candidate TS susceptibility gene based on previous identification of both rare functional variants[12] and common haplotypes[46] in TS patients, functional annotation indicates that rs7336083 is a cerebellar eQTL of *SLITRK6*. Candidate gene analysis of all genotyped SNPs within 50 kb of *SLITRK1* identified no nominally associated SNPs (Supplementary Table S9), including two SNPs recently reported to be associated with TS in a separate European ancestry sample[47] (rs9593835 and rs9546538; $P = 0.52$ and $P = 0.98$, respectively, in this study). Of note, the association signals in rs7336083 and rs13063502 decreased in the secondary meta-analysis (Supplementary Figures S11 and S12 and Supplementary Table S3). It remains to be determined whether these signal reductions are indicative of false positive associations, random signal fluctuations or genetic heterogeneity between the European ancestry samples and the Latin American CVCR/ANT samples used in the secondary analysis.

This study has several potential limitations. The use of shared controls genotyped previously on different Illumina platforms creates the possibility of a systematic technical bias. To address this concern, we employed stringent, iterative individual platform QC procedures, tests of cross-platform concordance using sample duplicates and additional extensive testing for differential missing data between platforms. We also excluded SNPs known to perform differentially across Illumina platforms that can cause spurious results if not recognized (N Cox, personal communication).[48] The minimal inflation of the median test statistic in the primary meta-analysis ($\lambda_{GC} = 0.996$), as well as the nominal enrichment of the top signals for SNPs with known functional significance in brain, argues that these efforts effectively mitigated this potential confound.

Second, there was residual population stratification between the TS cases from the CVCR and control samples from ANT. Although initially thought to have arisen from common founders,[22] recent studies suggest that these populations have slight differences in Native American ancestry (A Ruiz-Linares and NB Freimer, personal communication). Although the resulting λ_{GC} of 1.04 in the CVCR/ANT subpopulation analysis is relatively small and thus is likely not to introduce significant bias in a meta-analysis, we chose to reserve these non-European samples for a secondary analysis to provide supportive evidence to individual candidate susceptibility loci. Although we did not find significant evidence for a consistent direction of effect between the top signals in the primary European ancestry meta-analysis and those in the CVCR/ANT subpopulation analysis, it is important to note that the CVCR/ANT samples are an admixed population with a significant proportion of non-European ancestry,[49] and thus do not represent a true replication sample for the European ancestry meta-analysis.

In summary, this study represents the first GWAS of TS. Despite the lack of genome-wide significant loci, the study provides an important foundation for future replication efforts and lays the groundwork for the eventual identification of definitive common TS susceptibility variants. The data also contribute to the still nascent understanding of the underlying genetic architecture of TS, which is likely to include genetic variation across the allelic frequency spectrum.[13,47,50–52] Our results also parallel those of other common neuropsychiatric disorders, for which increased sample sizes have generated significant findings for both common and rare variants that together provide key insights into previously unknown disease mechanisms.[53–55] Finally, the current data will facilitate examination of the proposed genetic relationships between TS and its common co-occurring conditions, OCD and attention-deficit hyperactivity disorder,[8] as well as those from additional psychiatric disorders,[35] with the goal of identifying the biological pathways shared by these common neurodevelopmental conditions.

CONFLICT OF INTEREST

Drs Pauls, Scharf, Mathews, Cox, Freimer, McMahon, Heutink, Oostra, Grados, King, Rouleau, Sandor and Budman have all received research support from the NIH and the Tourette Syndrome Association (TSA) on behalf of the TSA International Consortium for Genetics (TSAICG). Drs Scharf, Mathews and Grados have received honoraria and travel support from the TSA. Dr Mathews is a member of the TSA Medical Advisory Board. Dr Sandor has received support and consulting fees from Psyadon, Shire, Solway, UCB Pharma, Janssen, Eli Lilly, Pfizer and Prestwick. Dr Budman has been funded by Psyadon Pharmaceuticals, Otsuka Pharmaceuticals and NINDS and is a member of the National TSA Medical Advisory Board, LI TSA and LI CHADD Medical Advisory Boards. Dr Hoekstra has received honoraria for advice through Desitin, Eli Lilly and Shire. Dr Dion has received honoraria from Biovail Pharma, Pfizer and Eli Lilly. Dr 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. Dr Walkup receives research support, travel support for paid and unpaid activities, serves in an unpaid position on the Medical Advisory Board and receives an honorarium for an Educational Meeting from the TSA. He receives royalties from Guilford Press and Oxford Press. He received free medication and placebo from Lilly, Pfizer and Abbott for NIH-funded studies. Dr Jankovic has received research grants from the following: Allergan, Allon Therapeutics, Ceregene, Chelsea Therapeutics, Diana Helis Henry Medical Research Foundation, EMD Serono, Huntington's Disease Society of America, Huntington Study Group, Impax Pharmaceuticals, Ipsen Limited, Lundbeck, Medtronic, Merz Pharmaceuticals, Michael J Fox Foundation for Parkinson Research, National Institutes of Health, National Parkinson Foundation, Neurogen, St Jude Medical, Teva Pharmaceutical Industries, University of Rochester and Parkinson Study Group. He has served as a consultant or advisory committee member for Allergan, Chelsea Therapeutics, EMD Serono, Lundbeck, Merz Pharmaceuticals, Michael J Fox Foundation for Parkinson Research and Teva Pharmaceutical Industries. He also serves on the editorial boards for Elsevier, Medlink: Neurology, Neurology in Clinical Practice, Neurotoxin Institute, Scientiae and UpToDate. Dr Knowles is on the Scientific Advisory Committee for Next-Generation Sequencing of Life Technologies, and is a technical advisor to SoftGenetics. Ms Anderson, Dr Barr, Dr Benarroch, Mr Berrío, Dr Bruun, Mr Campbell, Dr Cath, Dr Chouinard, Dr Conti, Ms Crane, Mr Crenshaw, Ms Davarya, Dr Davis, Ms Duarte, Mr Edlund, Dr Erenberg, Dr Eskin, Dr Evans, Mr Fagerness, Dr Fernandez, Mr Fournier, Mr Gamazon, Mr Gibbs, Dr Gilbert, Dr Girard, Dr Gross-Tsur, Dr Han, Dr Hardy, Dr Heiman, Dr Herrera, Dr Heutink, Dr Illmann, Dr J Kidd, Dr K Kidd,

© 2013 Macmillan Publishers Limited

Mr. Konkashbaev, Dr Kremeyer, Dr Kurlan, Dr Lee, Dr Leppert, Dr Liu, Dr Lowe, Dr Lyon, Dr Mirel, Dr Moessner, Ms Moorjani, Mr Morgan, Mr Muller, Dr Naarden, Dr Neale, Dr Ochoa, Dr Ophoff, Ms Osiecki, Dr Pakstis, Ms Parkin, Mr Phan, Dr Pluzhnikov, Dr Pollak, Dr Posthuma, Dr Purcell, Dr Rangel-Lugo, Dr Restrepo, Dr Reus, Ms Rivas, Dr Robertson, Ms Romero, Dr Ruiz-Linares, Dr Sabatti, Ms Service, Mr Silgado, Dr Singer, Dr Singleton, Dr Smit, Dr State, Dr Stewart, Mr. Strengman, Dr Tikhomirov, Dr Tischfield, Dr Wagner, Dr N Weiss, Dr R Weiss and Ms Yu declare no conflict of interest.

ACKNOWLEDGEMENTS

We are grateful to all the patients with Tourette's syndrome who generously agreed to participate in this study. Furthermore, the members of the Tourette Syndrome Association International Consortium for Genetics are deeply indebted to the Tourette Syndrome Association for their guidance and support. We also thank Libby Bernier and Janelle Alabiso for their assistance in manuscript preparation and Stephan Ripke for help with meta-analysis figures. This work was supported by a grant from the Judah Foundation, NIH Grants NS40024 to DLP and the Tourette Syndrome Association International Consortium for Genetics, NIH Grant NS16648 and a grant from the Tourette Syndrome Association to DLP, NIH Grant NS037484 to NBF, NIH Grant NS043538 to AR-L, American Recovery and Re-investment Act (ARRA) awards NS40024-07S1 and NS16648-29S1 to DLP, NIH Grant MH079489, and an American Academy of Neurology Foundation Grant 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. Support was also provided by the New Jersey Center for Tourette Syndrome & Associated Disorders (through New Jersey Department of Health and Senior Services: 08-1827-FS-N-0) to GAH and JAT and P01MH049351, R01MH061940, K05MH076273 and T32MH018268 to JFL. 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. 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 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. None of the funding agencies for this project (NINDS, NIMH, the Tourette Syndrome Association and the Judah Foundation) had any influence or played any role in (1) the design or conduct of the study; (2) management, analysis or interpretation of the data; and (3) preparation, review or approval of the manuscript.

Author contributions

Manuscript preparation: JM Scharf, D Yu, CA Mathews, BM Neale, SE Stewart, JA Fagerness, E Gamazon, NB Freimer, NJ Cox and DL Pauls.

Study design: JM Scharf, D Yu, CA Mathews, BM Neale, SE Stewart, JA Fagerness, S Purcell, P Heutink, BA Oostra, WM McMahon, NB Freimer, NJ Cox and DL Pauls.

Data analysis: D Yu, BM Neale, JM Scharf, P Evans, E Gamazon, CK Edlund, SK Service, SE Stewart, A Tikhomirov, A Pluzhnikov, A Konkashbaev, LK Davis, B Han, D Posthuma, E Eskin, C Sabatti, DV Conti, JA Knowles, NB Freimer, S Purcell and NJ Cox.

Project management: JM Scharf, JA Fagerness and DL Pauls.

Sample management and processing: JA Fagerness, JM Scharf, J Crane, P Moorjani and DL Pauls.

Genotyping: AT Crenshaw, MA Parkin and DB Mirel.

Phenotype management: CA Mathews, JM Scharf, L Osiecki, C Illmann, SE Stewart, W McMahon and DL Pauls.

Case sample collection:

European Ancestry Samples (ordered by numbers of submitted samples): Yale University: RA King (Site PI), TV Fernandez, KK Kidd, JR Kidd, JF Leckman, AJ Pakstis, MW State; Utrecht University/VU Medical Center: DC Cath (Site PI), JH Smit, P Heutink; University of Toronto: P Sandor (Site PI), CL Barr, N Phan; Massachusetts General Hospital: DL Pauls (Site PI), C Illmann, L Osiecki, JM Scharf; University of Utah: WM McMahon (Site PI), G Lyon, M Leppert, J Morgan, R Weiss; Johns Hopkins School of Medicine: MA Grados (Site PI), K Anderson, S Davarya, H Singer, J Walkup; Baylor College of Medicine: J Jankovic (Site PI); Rutgers University: JA Tischfield (Site PI), GA Heiman, RA King; University of Cincinnati: DL Gilbert (Site PI); University of Groningen: PJ Hoekstra(Site PI); University College London: MM Robertson (Site PI); UCSF: CA Mathews (Site PI), VI Reus, TL Lowe, P Lee, M Rangel-Lugo; University of Rochester School of Medicine: R Kurlan (Site PI).

French Canadian Samples: University of Montreal: GA Rouleau (Site PI), S Chouinard, Y Dion, S Girard.

Ashkenazi Jewish Samples: UCSF: CA Mathews (Site PI), VI Reus, TL Lowe, M Rangel-Lugo. Shaare Zedek Medical Center: V Gross-Tsur (Site PI), Y Pollak; Hadassah Mount Scopus Hospital: F Benarroch; North Shore-Long Island Jewish Medical Center: C Budman (Site PI), R Bruun.

Central Valley Costa Rica (CVCR) Samples: UCSF: CA Mathews (Site PI), VI Reus, TL Lowe, M Rangel-Lugo; N Weiss; Hospital Nacional de Niños: LD Herrera (Site PI), R Romero, E Fournier; UCLA: NB Freimer (Site PI).

Antioquia Colombian Samples: University of College London: A Ruiz-Linares (Site PI), B Kremeyer, DD Campbell, H Muller; Universidad de Antioquia: G Bedoya Berrío, J Cardona Silgado, W Cornejo Ochoa, S Mesa Restrepo, A Valencia Duarte.

Control Sample Collection: University Medical Center, Utrecht: RA Ophoff, E Strengman; University of Bonn: M Wagner, R Moessner.

eQTL and mQTL data: C Liu; 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], Michael A Nalls[1], Richard O'Brien[6], Andrew Singleton[1], Bryan Traynor[1], Juan Troncoso[6], Marcel 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] and 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 APA. *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn, text revision (DSM-IV-TR) edn. American Psychiatric Association: Washington, DC 2000.
2 Jankovic J, Kurlan R. Tourette syndrome: evolving concepts. *Mov Disord* 2011; **26**: 1149–1156.
3 Scharf JM, Pauls DL. Genetics of tic disorders. In: Rimoin DL, Connor JM, Pyeritz RE, Korf BR (eds) *Emery and Rimoin's Principles and Practices of Medical Genetics*, 5th edn Churchill Livingstone/Elsevier: Philadelphia, 2007 pp 2737–2754.
4 Robertson MM. The prevalence and epidemiology of Gilles de la Tourette syndrome. Part 1: the epidemiological and prevalence studies. *J Psychosom Res* 2008; **65**: 461–472.
5 Freeman RD, Fast DK, Burd L, Kerbeshian J, Robertson MM, Sandor P. An international perspective on Tourette syndrome: selected findings from 3500 individuals in 22 countries. *Dev Med Child Neurol* 2000; **42**: 436–447.
6 Elstner K, Selai CE, Trimble MR, Robertson MM. Quality of Life (QOL) of patients with Gilles de la Tourette's syndrome. *Acta Psychiatr Scand* 2001; **103**: 52–59.
7 Leckman JF, Bloch MH, King RA, Scahill L. Phenomenology of tics and natural history of tic disorders. *Adv Neurol* 2006; **99**: 1–16.
8 O'Rourke JA, Scharf JM, Yu D, Pauls DL. The genetics of Tourette syndrome: a review. *J Psychosom Res* 2009; **67**: 533–545.
9 NIMH Genetics Workgroup. *Genetics and Mental Disorders*. National Institute of Mental Health: Rockville, MD, 1998. Report no. 98-4268.
10 TSAICG. Genome scan for Tourette disorder in affected-sibling-pair and multi-generational families. *Am J Hum Genet* 2007; **80**: 265–272.
11 Verkerk AJ, Mathews CA, Joosse M, Eussen BH, Heutink P, Oostra BA. CNTNAP2 is disrupted in a family with Gilles de la Tourette syndrome and obsessive compulsive disorder. *Genomics* 2003; **82**: 1–9.
12 Abelson JF, Kwan KY, O'Roak BJ, Baek DY, Stillman AA, Morgan TM *et al*. Sequence variants in SLITRK1 are associated with Tourette's syndrome. *Science* 2005; **310**: 317–320.

© 2013 Macmillan Publishers Limited

13 Ercan-Sencicek AG, Stillman AA, Ghosh AK, Bilguvar K, O'Roak BJ, Mason CE *et al.* L-histidine decarboxylase and Tourette's syndrome. *N Engl J Med* 2010; **362**: 1901–1908.
14 Scharf JM, Moorjani P, Fagerness J, Platko JV, Illmann C, Galloway B *et al.* Lack of association between SLITRK1var321 and Tourette syndrome in a large family-based sample. *Neurology* 2008; **70**(16 Pt 2): 1495–1496.
15 O'Roak BJ, Morgan TM, Fishman DO, Saus E, Alonso P, Gratacos M *et al.* Additional support for the association of SLITRK1 var321 and Tourette syndrome. *Mol Psychiatry* 2010; **15**: 447–450.
16 TSCSG. Definitions and classification of tic disorders. *Arch Neurol* 1993; **50**: 1013–1016.
17 Bierut LJ, Saccone NL, Rice JP, Goate A, Foroud T, Edenberg H *et al.* Defining alcohol-related phenotypes in humans. The Collaborative Study on the genetics of alcoholism. *Alcohol Res Health* 2002; **26**: 208–213.
18 Bierut LJ, Madden PA, Breslau N, Johnson EO, Hatsukami D, Pomerleau OF *et al.* Novel genes identified in a high-density genome wide association study for nicotine dependence. *Hum Mol Genet* 2007; **16**: 24–35.
19 Bierut LJ, Strickland JR, Thompson JR, Afful SE, Cottler LB. Drug use and dependence in cocaine dependent subjects, community-based individuals, and their siblings. *Drug Alcohol Depend* 2008; **95**: 14–22.
20 Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D *et al.* Common variants conferring risk of schizophrenia. *Nature* 2009; **460**: 744–747.
21 Carvajal-Carmona LG, Ophoff R, Service S, Hartiala J, Molina J, Leon P *et al.* Genetic demography of Antioquia (Colombia) and the Central Valley of Costa Rica. *Hum Genet* 2003; **112**: 534–541.
22 Service S, DeYoung J, Karayiorgou M, Roos JL, Pretorious H, Bedoya G *et al.* Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. *Nat Genet* 2006; **38**: 556–560.
23 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.
24 Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; **38**: 904–909.
25 Stewart SE, Yu D, Scharf JM, Neale BM, Fagerness JA, Mathews CA, Arnold PD *et al.* Genome-wide association study of obsessive-compulsive disorder. *Mol Psychiatry* (in press).
26 Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010; **26**: 2190–2191.
27 Gamazon ER, Nicolae DL, Cox NJ. A study of CNVs as trait-associated polymorphisms and as expression quantitative trait loci. *PLoS Genet* 2011; **7**: e1001292.
28 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.
29 Gamazon ER, Badner JA, Cheng L, Zhang C, Zhang D, Cox NJ *et al.* Enrichment of cis-regulatory gene expression SNPs and methylation quantitative trait loci among bipolar disorder susceptibility variants. *Mol Psychiatry*; advance online publication, 3 January 2012; doi:10.1038/mp.2011.174.
30 Zhang D, Cheng L, Badner JA, Chen C, Chen Q, Luo W *et al.* Genetic control of individual differences in gene-specific methylation in human brain. *Am J Hum Genet* 2010; **86**: 411–419.
31 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.
32 1000 Genomes Project. A map of human genome variation from population-scale sequencing. *Nature* 2010; **467**: 1061–1073.
33 Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol* 2008; **32**: 381–385.
34 Dudbridge F, Gusnanto A. Estimation of significance thresholds for genomewide association scans. *Genet Epidemiol* 2008; **32**: 227–234.
35 Sullivan PF. The psychiatric GWAS consortium: big science comes to psychiatry. *Neuron* 2010; **68**: 182–186.
36 Mink JW. Neurobiology of basal ganglia and Tourette syndrome: basal ganglia circuits and thalamocortical outputs. *Adv Neurol* 2006; **99**: 89–98.
37 NIMH Transcriptional Atlas of Human Brain Development. http://developinghumanbrain.org, 2011, Accessed 26 May 2011..
38 Pace JM, Corrado M, Missero C, Byers PH. Identification, characterization and expression analysis of a new fibrillar collagen gene, COL27A1. *Matrix Biol* 2003; **22**: 3–14.
39 Fox MA. Novel roles for collagens in wiring the vertebrate nervous system. *Curr Opin Cell Biol* 2008; **20**: 508–513.
40 Dieci G, Fiorino G, Castelnuovo M, Teichmann M, Pagano A. The expanding RNA polymerase III transcriptome. *Trends Genet* 2007; **23**: 614–622.
41 Saitsu H, Osaka H, Sasaki M, Takanashi J, Hamada K, Yamashita A *et al.* Mutations in POLR3A and POLR3B encoding RNA Polymerase III subunits cause an autosomal-recessive hypomyelinating leukoencephalopathy. *Am J Hum Genet* 2011; **89**: 644–651.
42 Tetreault M, Choquet K, Orcesi S, Tonduti D, Balottin U, Teichmann M *et al.* Recessive mutations in POLR3B, encoding the second largest subunit of Pol III, cause a rare hypomyelinating leukodystrophy. *Am J Hum Genet* 2011; **89**: 652–655.
43 Reymond N, Borg JP, Lecocq E, Adelaide J, Campadelli-Fiume G, Dubreuil P *et al.* Human nectin3/PRR3: a novel member of the PVR/PRR/nectin family that interacts with afadin. *Gene* 2000; **255**: 347–355.
44 Wang CH, Su PT, Du XY, Kuo MW, Lin CY, Yang CC *et al.* Thrombospondin type I domain containing 7A (THSD7A) mediates endothelial cell migration and tube formation. *J Cell Physiol* 2010; **222**: 685–694.
45 Van Deerlin VM, Sleiman PM, Martinez-Lage M, Chen-Plotkin A, Wang LS, Graff-Radford NR *et al.* Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nat Genet* 2010; **42**: 234–239.
46 Miranda DM, Wigg K, Kabia EM, Feng Y, Sandor P, Barr CL. Association of SLITRK1 to Gilles de la Tourette Syndrome. *Am J Med Genet B Neuropsychiatr Genet* 2009; **150B**: 483–486.
47 Karagiannidis I, Rizzo R, Tarnok Z, Wolanczyk T, Hebebrand J, Nothen MM *et al.* Replication of association between a SLITRK1 haplotype and Tourette Syndrome in a large sample of families. *Mol Psychiatry*; advance online publication, 15 November 2011; doi:10.1038/mp.2011.151.
48 Sebastiani P, Solovieff N, Puca A, Hartley SW, Melista E, Andersen S *et al.* Genetic signatures of exceptional longevity in humans. *Science* 2010 [*Science* advance online publication, 1 July 2010; doi:10.1126/science.1190532; Retraction in: *Science* 2011; **333**: 404].
49 Wang S, Ray N, Rojas W, Parra MV, Bedoya G, Gallo C *et al.* Geographic patterns of genome admixture in Latin American Mestizos. *PLoS Genet* 2008; **4**: e1000037.
50 Sundaram SK, Huq AM, Wilson BJ, Chugani HT. Tourette syndrome is associated with recurrent exonic copy number variants. *Neurology* 2010; **74**: 1583–1590.
51 Sundaram SK, Huq AM, Sun Z, Yu W, Bennett L, Wilson BJ *et al.* Exome sequencing of a pedigree with Tourette syndrome or chronic tic disorder. *Ann Neurol* 2011; **69**: 901–904.
52 Fernandez TV, Sanders SJ, Yurkiewicz IR, Ercan-Sencicek AG, Kim YS, Fishman DO *et al.* Rare copy number variants in Tourette syndrome disrupt genes in histaminergic pathways and overlap with autism. *Biol Psychiatry* 2012; **71**: 392–402.
53 International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* 2008; **455**: 237–241.
54 Lee SH, DeCandia TR, Ripke S, Yang J, Sullivan PF, Goddard ME *et al.* Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nat Genet* 2012; **44**: 247–250.
55 Sullivan P. Don't give up on GWAS. *Mol Psychiatry* 2012; **17**: 2–3.

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

© 2013 Macmillan Publishers Limited