Genome-wide association study of Tourette Syndrome

Citation

Published Version

Permanent link

Terms of Use

Share Your Story

Genome-wide association study of Tourette Syndrome

Jeremiah M. Scharf, MD PhD[1,2,3,4,5,#,*], **Dongmei Yu, MS**[1,2,#], **Carol A. Mathews, MD**[6,#], **Benjamin M. Neale, PhD**[1,2,5,7,8,#], **S. Evelyn Stewart, MD**[1,2,9,#], **Jesen A Fagerness, JD**[1,2,#], **Patrick Evans, PhD**[10,11], **Eric Gamazon, MS**[10,11], **Christopher K. Edlund, MS**[12,13], **Susan Service, MS**[14], **Anna Tikhomirov, PhD**[10,11], **Lisa Osiecki, BA**[1,2], **Cornelia Illmann, PhD**[1,2], **Anna Pluzhnikov, PhD**[10,11], **Anuar Konkashbaev, MS**[10,11], **Lea K Davis, PhD**[10,11], **Buhm Han, PhD**[15], **Jacquelyn Crane, BA**[1,2], **Priya Moorjani, MS**[7,16], **Andrew T. Crenshaw**[7], **Melissa A. Parkin**[7], **Victor I. Reus, MD**[6], **Thomas L. Lowe, MD**[6], **Martha Rangel-Lugo, PhD**[6], **Sylvain Chouinard, MD**[17], **Yves Dion, MD**[17], **Simon Girard, PhD**[17], **Danielle C Cath, MD**[18,19], **Jan H Smit, MD**[18], **Robert A. King, MD**[20], **Thomas Fernandez, MD PhD**[21], **James F. Leckman, MD**[20], **Kenneth K. Kidd, PhD**[21], **Judith R. Kidd, PhD**[21], **Andrew J. Pakstis, PhD**[21], **Matthew State, MD PhD**[21], **Luis Diego Herrera, MD MPH**[22], **Roxana Romero, MA**[22], **Eduardo Fournier, MS**[22], **Paul Sandor, MD**[23,24], **Cathy L Barr, PhD**[23,25], **Nam Phan, MSc**[23], **Varda Gross-Tsur,**

[*]**Corresponding Authors:** Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA 02114. Tel: (617) 726-0793; Fax: (617) 726-0830; dpauls@pngu.mgh.harvard.edu Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA 02114. Tel: (617) 643-3139; Fax: (617) 726-0830; jscharf@partners.org.
[#]These authors contributed equally to this work.

MD[26], **Fortu Benarroch, MD**[27], **Yehuda Pollak, PhD**[26], **Cathy L. Budman, MD**[28,29], **Ruth D. Bruun, MD**[28,30], **Gerald Erenberg, MD**[31], **Allan L Naarden, MD**[32], **Paul C Lee, MD**[6], **Nicholas Weiss, MD**[6], **Barbara Kremeyer, PhD**[33], **Gabriel Bedoya Berrío, MSc**[34], **Desmond Campbell, MSc**[33], **Julio C. Cardona Silgado, MSc**[34], **William Cornejo Ochoa, MD**[34], **Sandra C. Mesa Restrepo, MD**[34], **Heike Muller, MSc**[33], **Ana V. Valencia Duarte, MSc**[34,35], **Gholson J Lyon, MD PhD**[36], **Mark Leppert, PhD**[36], **Jubel Morgan, RN**[36], **Robert Weiss, PhD**[36], **Marco A. Grados, MD MPH**[37], **Kelley Anderson, BA**[37], **Sarah Davarya, MS**[37], **Harvey Singer, MD**[37], **John Walkup, MD**[38], **Joseph Jankovic, MD**[39], **Jay A. Tischfield, PhD**[40,41], **Gary A. Heiman, PhD**[40,41], **Donald L. Gilbert, MD**[42], **Pieter J. Hoekstra, MD PhD**[43], **Mary M. Robertson, MD, DSc (Med), FRCPsych**[33,44], **Roger Kurlan, MD**[45], **Chunyu Liu, PhD**[46], **J. Raphael Gibbs**[47] for the North American Brain Expression Consortium,, **Andrew Singleton, PhD**[47] for the UK Human Brain Expression Database,, **John Hardy**[33], **Eric Strengman, MSc**[14,48], **Roel Ophoff, PhD**[14,48], **Michael Wagner, PhD**[49], **Rainald Moessner, MD**[49], **Daniel B. Mirel**[7], **Danielle Posthuma, PhD**[50,51,52], **Chiara Sabatti, PhD**[53], **Eleazar Eskin, PhD**[15,54], **David V. Conti, PhD**[12], **James A. Knowles, PhD**[55], **Andres Ruiz-Linares, MD**[33], **Guy A. Rouleau, MD**[17], **Shaun Purcell, PhD**[1,2,7,8], **Peter Heutink, PhD**[51], **Ben A. Oostra, PhD**[56], **William McMahon, MD**[36], **Nelson Freimer, MD**[14], **Nancy J. Cox, PhD**[10,11], and **David L. Pauls, PhD**[1,2,*]

[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 [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, Quebec, 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, Texas, USA [33]University College London, London, UK [34]Universidad de Antioquia, Medellín, Colombia [35]Universidad Pontificia Bolivariana, Medellín, Colombia [36]University of Utah, Salt Lake City, Utah, USA [37]Johns Hopkins University School of Medicine, Baltimore, MD, USA [38]Division of Child and Adolescent Psychiatry, Department of Psychiatry, Weill Cornell Medical Center, New

York, NY, USA [39]Parkinson's Disease Center and Movement Disorders Clinic, Department of Neurology, Baylor College of Medicine, Houston, TX, USA [40]Department of Genetics, Rutgers University, Piscataway, NJ, US [41]Human Genetics Institute of New Jersey, Rutgers University, Piscataway, NJ, US [42]Cincinnati Children's Hospital Medical Center and the University of Cincinnati, Cincinnati, OH, USA [43]Department of Psychiatry, University Medical Center, University of Groningen, Groningen, The Netherlands [44]St George's Hospital and Medical School, London, UK [45]Atlantic Neuroscience Institute, Overlook Hospital, Summit, NJ, USA [46]Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL, USA [47]Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA [48]University Medical Center, Utrecht, The Netherlands [49]Department of Psychiatry and Psychotherapy, University of Bonn, Bonn, Germany [50]Department of Functional Genomics, Center for Neurogenomics and Cognitive Research, Neuroscience Campus Amsterdam, VU University Amsterdam, De Boelelaan Amsterdam, The Netherlands [51]Section of Medical Genomics, Department of Clinical Genetics, VU Medical Centre, De Boelelaan, Amsterdam, The Netherlands, [52]Department of Child and Adolescent Psychiatry/Psychology, Erasmus University Medical Center-Sophia Children's Hospital, Wytemaweg, Rotterdam, The Netherlands. [53]Department of Health Research and Policy, Stanford University, Stanford, CA USA [54]Department of Human Genetics, University of California Los Angeles, Los Angeles, CA, USA [55]Department of Psychiatry and the Behavioral Sciences, Zilkha Neurogenetic Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA [56]Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands

Abstract

Tourette 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 ($p < 5 \times 10^{-8}$); the top signal was found in rs7868992 on chromosome 9q32 within *COL27A1* ($p = 1.85 \times 10^{-6}$). 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, 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.

Keywords

Tourette Syndrome; tics; genetics; GWAS; neurodevelopmental disorder

INTRODUCTION

Tourette Syndrome (TS) is a chronic, childhood-onset neuropsychiatric disorder characterized by multiple motor tics and at least one phonic tic that persist for greater than one year.[1-2] TS has a population prevalence of ~0.3-0.8%, and, like many neurodevelopmental disorders, occurs more frequently in boys, with male:female ratios ranging between 3:1-4:1.[3-4] It is frequently accompanied by a wide range of additional

psychiatric co-morbidities, in particular obsessive-compulsive disorder (OCD) and attention-deficit hyperactivity disorder (ADHD).[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 5-15-fold increased risk of TS compared to 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 multi-generational 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 GWAS in a large cohort of samples of general European ancestry, as well as two European-derived population isolates, Ashkenazi Jews from the US 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

1998 TS cases were recruited from 20 sites in the US, Canada, UK, Netherlands, Israel, Costa Rica and Colombia and divided into four strata based on self-reported ancestry: 1) 1252 European ancestry, non-isolate cases from North America and Europe (EU); 2) 210 Ashkenazi Jewish cases from the US and Israel (AJ); 3) 302 French Canadian cases (FC); 4) Cases from two closely-related population isolates from the Central Valley of Costa Rica (CVCR) (n=137) and Antioquia, Colombia (ANT) (n=97) (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 (ID), 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 ADHD using a standardized and validated semi-structured interview that has high validity and reliability for TS ($\kappa=1.00$) and OCD ($\kappa=0.97$).[10] Subjects from the other 3 sites were assessed only for a lifetime diagnosis of definite TS.

Controls

5403 European ancestry controls were derived primarily from cohorts of previously genotyped, unselected population controls (Supplementary Methods, 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 HumanHap1Mv1_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.

297 FC and 380 ANT ancestry-matched controls were collected in parallel with their respective cases (Supplementary Methods). ANT controls were used for analysis of both ANT and CVCR cases given their shared ancestry.[21-22] All participants 18 years of age and older gave informed consent. Individuals under 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, CVCR), as well as 298 EU and all FC and ANT controls, was conducted on the Illumina Human610-Quadv1_B SNP array (Illumina, San Diego, CA, USA) 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, San Diego, CA). 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[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, Figure S1). In general, two thresholds were used for 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 $p<1\times10^{-3}$ in any analysis are annotated in Tables S2-S4.

Sample and SNP QC were initially performed within each platform separately (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 due to increased rates of missing SNP data relative to other sites (Figure S2). Platform-specific SNP QC included removing monomorphic SNPs, CNV-targeted SNP probes, SNPs with genotyping rate <98%, and strand-ambiguous SNPs with significant allele frequency differences or aberrant 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 610Quad and 550v1. In addition, concordance was examined in HapMap duplicates from the Illumina database genotyped on 2 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 an MAF difference >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 (Stewart et al., accompanying manuscript). Any SNP not present on the three major common platforms (550v1, 610-Quad, 1M) was removed, leaving 496 877 SNPs for population-specific QC.

Multi-dimensional scaling (MDS) 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) (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 (Figures S4-S6). Within each of the MDS-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, 4414 controls; AJ: 242 cases, 354 controls; FC: 265 cases, 196 controls) (Table 1; 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 HWE $p<10^{-10}$ in controls (flagged for HWE $p<10^{-5}$) and two additional cross-platform QC steps to remove SNPs with differential missingness between cases and controls across the 5 Illumina datasets (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[23] employing logistic regression under an additive model with significant subpopulation-specific MDS dimensions included as covariates to control for residual population stratification. Strata were then combined in a case-weighted meta-analysis in METAL[25] 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, 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.[26-27] Similarly, methylation QTLs (mQTLs), which represent SNPs that are associated with variation in genome-wide patterns of methylation, were derived from adult cerebellum.[28] 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 to 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[29] and haplotypes from all 1,092 individuals in the 1000 Genomes June 2011 Data Release[30] as a reference dataset (Supplementary Methods). Post-imputation QC and allelic dosage analysis were conducted in each subpopulation separately in PLINK followed by case-weighted meta-analysis in METAL.

RESULTS

Quality control analyses in individual ancestral subpopulations

After QC filtering, 1285 cases and 4964 controls remained across the three European ancestry strata (EU, AJ, 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, λ=1.011; AJ, λ=0.993; FC, λ=0.971; Figure S8a-c). The Latin-American population isolate stratum (CVCR/ANT) showed a small inflation of the median test statistic (λ=1.044), indicative of some residual stratification between CVCR and ANT samples (Figure S6). However, no SNPs in this subpopulation-specific analysis had extreme p-values outside the expected null distribution (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 $p < 5.0 \times 10^{-8}$ (Figure 1). The top 5 LD-independent loci are annotated in Table 2; full annotation of all SNPs with $p < 1 \times 10^{-3}$ are provided in Table S2. The SNP with the strongest signal, rs7868992, lies on chromosome 9q32 within an intron of *COL27A1* ($p = 1.85 \times 10^{-6}$; Figure S9). The other four top independent GWAS signals include rs6539267, an intronic SNP within *POLR3B* on chromosome 12q23 ($p = 7.41 \times 10^{-6}$; Figure S10); rs13063502, a SNP that lies in a 1.7 Mb intergenic region on chromosome 3q13 ($p = 8.96 \times 10^{-6}$; Figure S11); rs7336083, located on chromosome 13q31 within a 1.9 Mb intergenic region between *SLITRK6* and *SLITRK1*[12] ($p = 9.49 \times 10^{-6}$; Figure S12); and rs769111, an intergenic SNP on chromosome 7p21 between *THSD7A* and *TMEM106B* ($p = 1.20 \times 10^{-5}$; Figure S13). No effect-size heterogeneity was present between the three European-derived subpopulations for SNPs rs7868992, rs6539267 and rs7336083 (Figures S9-13). rs13063592 and rs769111 demonstrated moderate heterogeneity ($I^2 = 45.4\%$ and 64.2%, respectively), though the direction of effect was consistent across the EU, AJ and FC populations.

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 $p = 2.94 \times 10^{-8}$; Table S5, 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 non-significant 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; Tables S2,S3).

Analysis of imputed data

Imputation was performed using 1000 Genomes Project data[30] 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 (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 (Figure S9c).

Enrichment analyses of expression and methylation quantitative trait loci

Since many of the top signals in the primary meta-analysis (p<0.001) appeared to lie within or adjacent to known brain-expressed genes (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 expression QTL (eQTL) information derived previously from lymphoblast cell lines (LCLs), adult cerebellum, and frontal cortex as well as methylation QTL (mQTL) information from adult cerebellum (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) (Figure 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), though missense SNPs demonstrated a borderline enrichment (p=0.098).

Examination of previously reported TS candidate genes

As an additional exploratory analysis, we examined the associations of SNPs within 50kb of 24 previously reported candidate TS genes (Tables S6-S7). We found no excess of lower p-values among the 2135 SNPs within these genes compared to those expected under the null, suggesting that these candidate genes are not enriched for common SNPs associated with TS (Figure S16). One signal in the primary European ancestry meta-analysis had a nominal $p < 1 \times 10^{-3}$ (rs10277969 within *CNTNAP2*, $p=7.8 \times 10^{-4}$), but this locus did not survive a Bonferroni correction for gene size (266 LD-independent SNPs within *CNTNAP2*, corrected p=.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 $p \; 5 \times 10^{-8}$.[31-32] This observation is not surprising, given that GWA studies for other highly heritable neuropsychiatric diseases (e.g., autism, bipolar disorder and schizophrenia) have required sample sizes of 5000-10000 cases to identify definitive common risk alleles with modest effect sizes (odds ratios <1.3).[33] 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 to eQTLs in cerebellum and LCLs is anatomically consistent with the hypothesis that TS is caused by abnormalities in fronto-striatal circuitry.[34] 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 (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 $p=3.61 \times 10^{-7}$, most likely due to 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 (Figure S9d) and the 100% concordance rate in all cross-platform comparisons of this SNP in HapMap 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 alpha chain gene. COL27A1 is a fibrillar collagen primarily expressed in cartilage, though it is expressed in the cerebellum during many stages of human development.[35-36] While non-fibrillar collagens have been implicated in various neurodevelopmental processes (e.g. axon guidance and synaptogenesis), the function of *COL27A1* in the developing nervous system is unknown.[37]

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 non-coding RNAs including tRNAs, small rRNAs and microRNAs.[38] Recessive mutations in *POLR3B* cause hypomyelinating leukodystrophy with a severe neurological phenotype (developmental delay, spasticity, dysarthria and ataxia), though no reported tics.[39-40] Both the secondary meta-analysis and imputed data provide additional support for this locus, and expand the region of LD to ~300kb, including adjacent genes *CKAP4*, *TCP11L2* and *RFX4* (Table S5, Figure S10).

The other 3 top loci in the primary analysis are located within large intergenic regions. rs13063502 on 3q13.1 lies between the non-coding cDNA *FLJ25363* and *PVRL3*, which resides 1.5 Mb telomeric to rs13063502 and is expressed primarily in placenta and testis.[41] rs769111 on 7p21.3 is situated between *THSD7A*, a gene expressed almost exclusively in developing endothelial cells[42], and *TMEM106B*, a gene recently associated with fronto-temporal dementia with TDP-43 inclusions (FTD-TDP), whose primary function in the brain remains to be elucidated.[43] Lastly, rs7336083 lies in a 1.9 Mb intergenic region between *SLITRK1* and *SLITRK6* on chromosome 13q31. While *SLITRK1* is an *a priori* candidate TS susceptibility gene based on previous identification of both rare functional variants[12] and common haplotypes[44] 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 (Table S9), including two SNPs recently reported to be associated with TS in a separate European-ancestry sample[45] (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 (Figures S11-S12, 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).[46] The minimal inflation of the median test statistic in the primary meta-analysis (λ_{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 Central Valley of Costa Rica (CVCR) and control samples from Antioquia, Colombia (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, N. Freimer, personal communication). Though 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. While 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[47], 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, 45, 48-50] 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.[51-53] Finally, the current data will facilitate examination of the proposed genetic relationships between TS and its common co-occurring conditions, OCD and ADHD[8], as well as those from additional psychiatric disorders[33], with the goal of identifying the biological pathways shared by these common neurodevelopmental conditions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

AUTHOR CONTRIBUTIONS

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

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

Data analysis: D Yu, BM Neale, JM Scharf, P Evans, E Gamazon, CK Edlund, S 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, NJ Cox

Project management: JM Scharf, JA Fagerness, DL Pauls

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

Genotyping: AT Crenshaw, MA Parkin, DB Mirel

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

Case Sample Collection: European Ancestry Samples (ordered by numbers of submitted samples): **Yale University**: RA King (Site PI), T 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:** W 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:** N Freimer (Site PI)

Antioquia Colombian Samples: **University of College London:** A Ruiz-Linares (Site PI), B Kremeyer, D 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: R 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], 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, Maryland, 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; [10]Department of Medical & Molecular Genetics, King's College London, UK

REFERENCES

1. APA. Diagnostic and Statistical Manual of Mental Disorders. 4th edition, text revision (DSM-IV-TR). American Psychiatric Association; Washington, DC: 2000.

2. Jankovic J, Kurlan R. Tourette syndrome: evolving concepts. Mov Disord. 2011; 26(6):1149–1156. [PubMed: 21484868]

3. Scharf, JM.; Pauls, DL. Genetics of Tic Disorders. In: Rimoin, DL.; Connor, JM.; Pyeritz, RE.; Korf, BR., editors. Emery and Rimoin's Principles and Practices of Medical Genetics. 5th edn. Churchill Livingstone/Elsevier; Philadelphia: 2007. p. 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(5):461–472. [PubMed: 18940377]

5. Freeman RD, Fast DK, Burd L, Kerbeshian J, Robertson MM, Sandor P. An international perspective on Tourette syndrome: selected findings from 3,500 individuals in 22 countries. Dev Med Child Neurol. 2000; 42(7):436–447. [PubMed: 10972415]

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(1):52–59. [PubMed: 11202129]

7. Leckman JF, Bloch MH, King RA, Scahill L. Phenomenology of tics and natural history of tic disorders. Adv Neurol. 2006; 99:1–16. [PubMed: 16536348]

8. O'Rourke JA, Scharf JM, Yu D, Pauls DL. The genetics of Tourette syndrome: a review. J Psychosom Res. 2009; 67(6):533–545. [PubMed: 19913658]

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 multigenerational families. Am J Hum Genet. 2007; 80(2):265–272. [PubMed: 17304708]

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):1–9. [PubMed: 12809671]

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(5746):317–320. [PubMed: 16224024]

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(20):1901–1908. [PubMed: 20445167]

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. [PubMed: 18413575]

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(5):447–450. [PubMed: 20351724]

16. TSCSG. Definitions and classification of tic disorders. Arch Neurol. 1993; 50(10):1013–1016. [PubMed: 8215958]

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(3):208–213. [PubMed: 12875049]

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(1):24–35. [PubMed: 17158188]

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(1-2):14–22. [PubMed: 18243582]

20. Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D, et al. Common variants conferring risk of schizophrenia. Nature. 2009; 460(7256):744–747. [PubMed: 19571808]

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(5-6):534–541. [PubMed: 12601469]

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(5):556–560. [PubMed: 16582909]

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(3):559–575. [PubMed: 17701901]

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(8):904–909. [PubMed: 16862161]

25. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010; 26(17):2190–2191. [PubMed: 20616382]

26. Gamazon ER, Nicolae DL, Cox NJ. A study of CNVs as trait-associated polymorphisms and as expression quantitative trait loci. PLoS Genet. 2011; 7(2):e1001292. [PubMed: 21304891]

27. 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(5):e1000952. [PubMed: 20485568]

28. 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(3):411–419. [PubMed: 20215007]

29. 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(6):e1000529. [PubMed: 19543373]

30. 1000 Genomes Project. A map of human genome variation from population-scale sequencing. Nature. 2010; 467(7319):1061–1073. [PubMed: 20981092]

31. 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(4):381–385. [PubMed: 18348202]

32. Dudbridge F, Gusnanto A. Estimation of significance thresholds for genomewide association scans. Genet Epidemiol. 2008; 32(3):227–234. [PubMed: 18300295]

33. Sullivan PF. The psychiatric GWAS consortium: big science comes to psychiatry. Neuron. 2010; 68(2):182–186. [PubMed: 20955924]

34. Mink JW. Neurobiology of basal ganglia and Tourette syndrome: basal ganglia circuits and thalamocortical outputs. Adv Neurol. 2006; 99:89–98. [PubMed: 16536354]

35. [Accessed May 26th 2011] NIMH Transcriptional Atlas of Human Brain Development. 2011. http://developinghumanbrain.org

36. Pace JM, Corrado M, Missero C, Byers PH. Identification, characterization and expression analysis of a new fibrillar collagen gene, COL27A1. Matrix Biol. 2003; 22(1):3–14. [PubMed: 12714037]

37. Fox MA. Novel roles for collagens in wiring the vertebrate nervous system. Curr Opin Cell Biol. 2008; 20(5):508–513. [PubMed: 18573651]

38. Dieci G, Fiorino G, Castelnuovo M, Teichmann M, Pagano A. The expanding RNA polymerase III transcriptome. Trends Genet. 2007; 23(12):614–622. [PubMed: 17977614]

39. 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(5):644–651. [PubMed: 22036171]

40. 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(5):652–655. [PubMed: 22036172]

41. 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(2):347–355. [PubMed: 11024295]

42. 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(3):685–694. [PubMed: 20020485]

43. 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(3):234–239. [PubMed: 20154673]

44. 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(4):483–486. [PubMed: 18698576]

45. 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. 2011

46. Sebastiani P, Solovieff N, Puca A, Hartley SW, Melista E, Andersen S, et al. Genetic signatures of exceptional longevity in humans. Science. Jul 1.2010 2010.

47. 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(3):e1000037. [PubMed: 18369456]

48. Sundaram SK, Huq AM, Wilson BJ, Chugani HT. Tourette syndrome is associated with recurrent exonic copy number variants. Neurology. 2010; 74(20):1583–1590. [PubMed: 20427753]

49. 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(5):901–904. [PubMed: 21520241]

50. 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(5):392–402. [PubMed: 22169095]

51. International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature. 2008; 455(7210):237–241. [PubMed: 18668038]

52. 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(3):247–250. [PubMed: 22344220]

53. Sullivan P. Don't give up on GWAS. Mol Psychiatry. 2012; 17(1):2–3. [PubMed: 21826059]
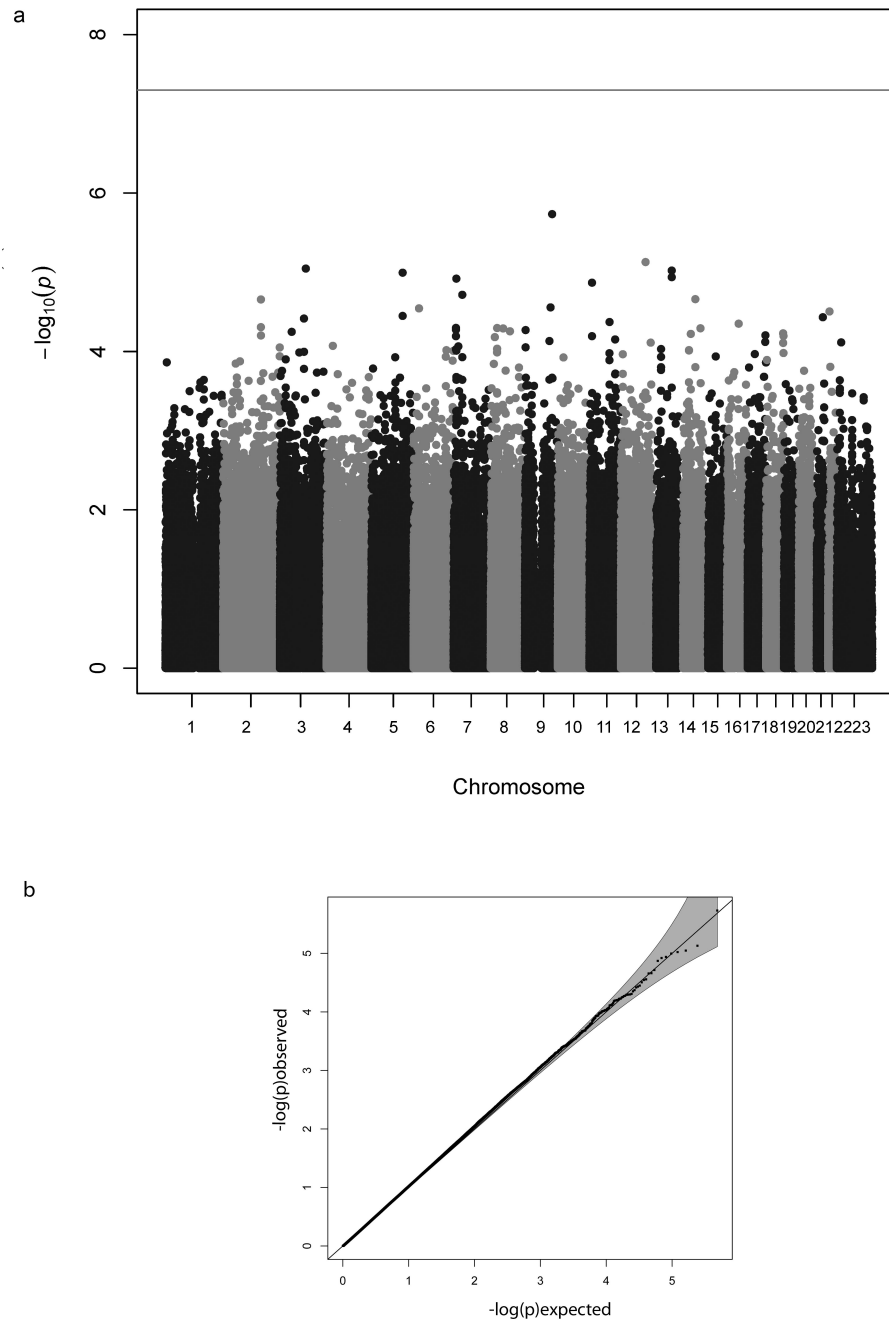
Figure 1. Results of the primary meta-analysis from the three European ancestry TS populations
a) Manhattan plot of all genotyped SNPs for 1285 TS cases and 4964 controls from the EU, AJ and FC populations. Grey line indicates the genome-wide significance threshold of 5 $\times 10^{-8}$. b) Quantile-quantile 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.
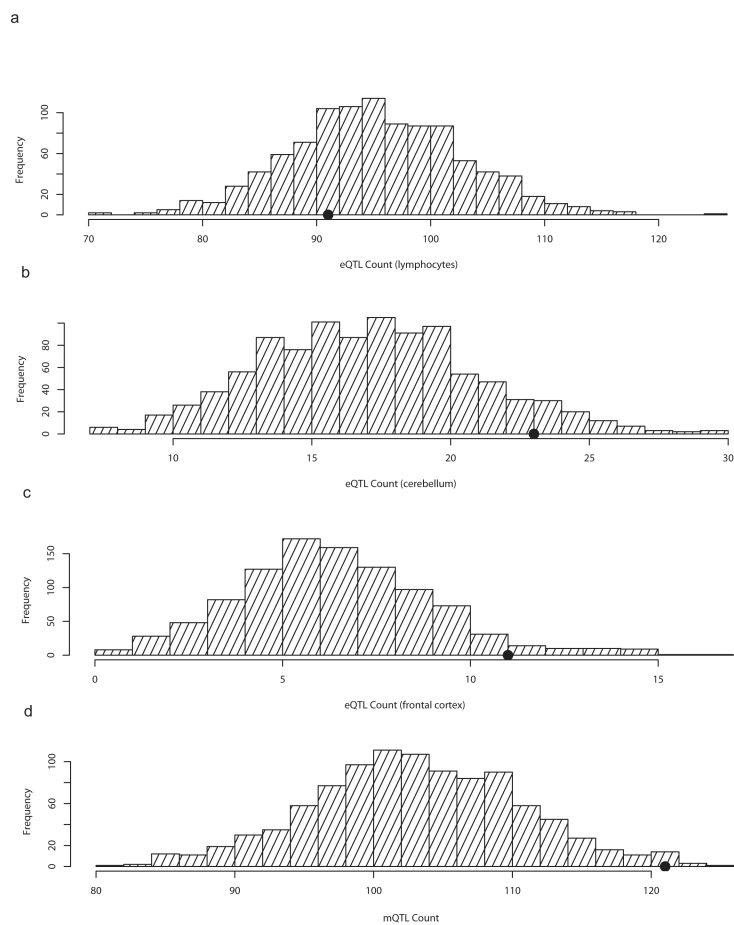
Figure 2. Enrichment analysis of functional SNPs within the top signals of the primary TS meta-analysis

Filled circles indicate the observed count of expression quantitative trait loci (eQTLs) or methylation QTLs (mQTLs) among the top loci (p<1×10^{-3}) in the primary European-derived meta-analysis following 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. d) Cerebellar mQTLs, p=0.011.

Table 1

Characteristics of the final TS GWAS samples

	Cases	Controls
N	1496	5249
Gender (% male)	79%	39%
Age at assessment, y (mean, s.d)[1]	16.6 ± 11.5	
Age of tic onset, y (mean, s.d.)[2]	6.0 ± 2.8	
OCD (%)[3]	42%	
ADHD (%)[4]	61%	

[1] Based on 1247 cases with available data

[2] Based on 1110 cases

[3] Based on 1223 cases

[4] Based on 1048 cases.

Table 2

Top 5 LD-independent signals in the primary European-derived TS meta-analysis.

CHR	SNP	BP	A1/A2	Primary European Meta-analysis			# SNPs in LD[1]	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	-	LOC 387939	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

CHR, chromosome; BP, hg19 position; A1, reference allele; A2, alternative allele; MAF, minor allele frequency; OR, odds ratio;

[1] # SNPs in LD, number of additional SNPs in linkage disequilibrium (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 $<1 \times 10^{-3}$ are provided in Supplementary Table S2.