Tourette Syndrome GWAS Replication Workshop
Meeting Summary

The Tourette Syndrome GWAS Replication Workshop was held at the Hyatt Regency Bethesda in Bethesda, MD on March 10-11, 2009 to discuss the preliminary results of the Tourette Syndrome genome-wide association study (GWAS) and to discuss potential efforts and new international collaborations to replicate and extend the findings of this initial study. The meeting, sponsored by the National Institute of Neurological Disorders and Stroke (NINDS), the National Institute of Mental Health (NIMH) and the Office of Rare Diseases (ORD) at the NIH brought together a panel of international Tourette Syndrome (TS) experts from more than a dozen countries. Meeting participants included experts from the Americas, Japan, China, Israel, Austria, Brazil, Hungary, Netherlands, Germany, Canada, the United Kingdom, in addition to the United States. This two-day workshop began with a welcome and introductions by Dr. Story Landis, Director of the NINDS, and Judit Ungar, President of the Tourette Syndrome Association.

Background and Introduction

Dr. David Pauls at the Harvard Medical School provided the history of the Tourette Syndrome Association International Consortium for Genetics (TSAICG), formed approximately twenty years ago, and reviewed the progress to date on the genetics of Tourette Syndrome. Tourette Syndrome (TS) is a neuropsychiatric disorder characterized by multiple motor and phonic tics that wax and wane over a lifetime. The prevalence of TS is estimated to be between 10 to 100 per 10,000. Males are more likely to have TS with a ratio as high as 4 to 1. The heritability of TS has been well established. Early segregation analysis suggested the existence of a single major gene and an autosomal dominant mode of inheritance was proposed. However, reports of bilineal transmission and a lack of consensus from the many linkage studies undertaken to identify major genes suggests that the mode of inheritance is likely more complex. The lack of consensus in results from genome-wide linkage studies for TS is characteristic for psychiatric disorders and points to a more complex disease model than a single major gene. There have been recent encouraging linkage findings on chromosome 2p, but there is a need to use an alternative approach both to fine-map chromosome 2p and to identify additional TS susceptibility loci.

Funding and Contributing Sites

The TSAICG efforts are currently supported by a cooperative grant mechanism (U01 NS40024) from the NINDS with co-funding from the Tourette Syndrome Association (TSA). For the present GWAS, a private donor provided funding for the genotyping aspects of this study along with a parallel GWAS for obsessive compulsive disorder (OCD). The OCD GWAS was also supported by a supplement from the NIMH. Contributors to the TS GWAS project included 11 sites from the TSAICG in the US, Canada and Europe (Johns Hopkins, MGH, Rochester, UCLA/UCSF, Utah, Yale, as well as University of Toronto, University de Montreal, University College London, Erasmus University, Rotterdam and the Free University of Amsterdam). In addition, two additional US sites contributed samples (Baylor and Rutgers/NJ Center for TS Sharing Repository) as well as sites contributing samples from four population isolates with TS: Ashkenazi Jewish (UCSF, UCLA, North Shore-LIJ, UT Southwestern, Cleveland Clinic and
Shaare Zedek, Jerusalem), French Canadian (Montreal), Central Valley Costa Rica (UCSF, UCLA, Hospital Nacional de Ninos) and Antioquia region of Colombia (UCL, Universidad de Antioquia).

**Study Design**

The TS GWAS was designed as a case-control study, although parental DNA samples are available for the majority of the cases for replication and follow-up genotyping. Study inclusion criteria included: Tourette Syndrome Classification Study Group (TSCSG) Research Criteria for TS diagnosis (DSM-IV-TR criteria PLUS tics observed by an experienced clinician) and DNA available from blood or cell lines. Exclusion criteria included: Mental Retardation (VIQ <80), tardive Tourettism/neuroleptic-induced tics, or other known genetic, metabolic or acquired tics. Phenotypic data collected include: TS diagnosis, demographic variables, age of symptom onset, comorbid OCD and ADHD status, and YGTSS Worst Ever Severity Score. In addition, for ~1200 of the cases, item-level symptom data are available. In total, 983 cases with European-derived ancestry were submitted, as well as 215 Ashkenazi, 309 French Canadian, 140 Central Valley of Costa Rica (CVCR) and 102 Colombian cases (total cases = 1749). Controls were submitted for the French Canadian, CVCR and Colombian samples and were genotyped simultaneously with the cases. For European-ancestry derived cases and Ashkenazi cases, previously genotyped controls were used (3300 European controls on the Illumina 550K platform from the Illumina iControl database and 430 Ashkenazi controls on the Illumina 317K platform contributed by Peter Gregerson). Genotyping was performed at the Harvard-MIT Broad Institute, with the Illumina 610-Quad BeadChip platform, containing approximately 550,000 single nucleotide polymorphisms (SNPs) that capture 89% of the 2.6 million SNPs in European HapMap Phase II at r² ≥ 0.8 as well as 60,000 additional markers to assess common and rare copy-number variants (CNVs) including “unSNPable” regions. Including the combined OCD and TS efforts, there are a total ~8000 samples in the GWAS.

**Preliminary Data Analysis**

Genotypes from approximately 90% of the samples (i.e., a 90% data freeze as described below) were received from the Harvard-MIT Broad Institute about three weeks prior to the meeting. Dr. Jeremiah Scharf at MGH presented the initial preliminary results from the analysis of the TS genotypic data. The data analysis was primarily performed by the teams of Nelson Freimer at UCLA, Nancy Cox at University of Chicago, and David Pauls at Harvard Medical School.

Initially, 2,429 samples for the TS GWAS, including 1,749 cases and 680 controls, were submitted to the Broad for genotyping. 237 samples (138 cases, 99 controls) were removed immediately because of chip failure or for call rates <97%. The remaining 1,611 case genotypes were available for subsequent analysis. Of note, approximately 10% of the samples from the entire TS and OCD GWAS were either not yet genotyped (400) at the time of the meeting or will have to be re-genotyped (800-900) because call rates were not sufficient.

The genotyping data underwent extensive quality control steps to clean the data and to identify subtle population differences among samples. The quality control included the removal of problematic SNPs (>5% missingness, low minor allele frequency, lack of Hardy-Weinberg Equilibrium, and differential missingness by case-control status, site or batch) and problematic individuals (sex incompatibility, unintended duplicates or relatives, outlying rates of
genomewide heterozygosity, or population outliers). Quality control was performed within each population separately. After this extensive quality control, there were 1453 TS cases and 4027 controls remaining for analyses.

Analyses were performed in each population separately, employing logistic regression and incorporating 2-4 principal components from the pairwise IBD estimation as covariates to control for population stratification. For the European-derived ancestry sample, the residual genomic inflation factor (lambda) was 1.11 after correction; for each of the isolates, lambda was much lower (1.015 for Ashkenazi, 1.019 for French Canadian, 1.013 for Antioqua Colombians, and 1.044 for the combined CVCR and Colombian samples).

Although the results are still very preliminary, some exciting observations were presented at the meeting that warrant further investigation and follow-up. First, in the non-isolate, European-derived sample, 7 loci on chromosomes 2,3,7,9,11,22 and X showed strong associations with TS, several of which met the threshold for genomewide significance. Furthermore, in the Ashkenazi sample, three SNPs in strong LD showed a strong association to chromosome 8p. In addition to these signals, there were a plethora of signals, many (~50) of which are shared across all the different populations, that are in the subthreshold range of significance. All of the loci need to be explored further in the complete sample to separate the true signals from the background of false positives. It is also necessary for these findings to be replicated in a second independent sample.

**Generation of a Replication Sample**

The primary goal of the meeting was to foster collaboration among a wider range of TS clinical and genetic research groups and to identify additional TS DNA samples that could serve as a source of replication of the initial GWAS data. The second day of the meeting was dedicated to discussion of each site's existing samples as well as their estimates of prospective samples that could be collected within the next 15-18 months if funds were available. Among the 17 sites not currently part of the TSAICG, it was estimated that an additional 1,571 extant TS DNA samples currently exist as well as ~400 additional samples with chronic tics. The eight current sites from the TSAICG (funded by NINDS U01 NS40024) estimated that they currently have 478 cases on hand, for a total of 2,049 TS cases which, if funds were available, could be available immediately for replication and extension efforts. Of these 2,049 cases, the vast majority were of European ancestry (1,559); other ancestries included 265 Han Chinese, 78 French Canadian, 118 Brazilian and 29 Arabic Middle Eastern samples.

In terms of a possible replication initiative, during the course of this meeting, this group of international sites identified another 2,000 extant samples and agreed in principle to contribute them to a follow-up GWAS. Furthermore, it was estimated that if additional funds were available for prospective sample collection, the consortium could collect an additional 1,500-2,000 new samples within the next 2 years (in addition to the 1,200 additional samples currently slated to be collected by the current members of the TSAICG under NINDS U01 NS40024). Thus, the entire group was extremely excited to realize that a replication sample of 2,000-6,000 more TS cases might be available pending funding, which should have enough power to replicate the preliminary findings and extend the findings to additional TS susceptibility genes. The entire group agreed in principle to join in a replication initiative, pending review and acceptance of a memorandum of understanding (MOU).
The remainder of the meeting was spent discussing preliminary logistics of a replication effort, including: phenotypic data available at each site; a method for coordinating, collecting and confirming phenotypic data; discussion of plans for replication (targeted genotypic of select loci vs. whole genome genotyping on the entire replication sample); potential obstacles to each site such as IRB/ethics committee requirements; issues of translation of phenotypic data; and funding needed per sample to perform new collections. The group also discussed strategies for applying for funding to initiate this effort including applying for ARRA (Stimulus) supplement funding by each TS investigator with an active R01 grant as well as writing a separate application to CIDR to fund the genotyping. The meeting adjourned with a directive to a subgroup to draft a memorandum of understanding (MOU) within the upcoming two weeks to be distributed to all attendees. This subcommittee consisted of: Harald Aschauer, Ana Hounie, Gholson Lyon, Carol Mathews, Yoshiko Nomura, David Pauls, Perry Paschou, and Jeremiah Scharf.

The meeting organizers and participants expressed the view that the meeting was a tremendous success, which bodes well for future genetics research for TS.