

UCSF

UC San Francisco Previously Published Works

Title

Genome-Wide Linkage Analysis of Obsessive-Compulsive Disorder Implicates Chromosome 1p36

Permalink

<https://escholarship.org/uc/item/8fh9q06m>

Journal

Biological Psychiatry, 72(8)

ISSN

0006-3223

Authors

Mathews, Carol A
Badner, Judith A
Andresen, J Michael
[et al.](#)

Publication Date

2012-10-01

DOI

10.1016/j.biopsych.2012.03.037

Peer reviewed



Published in final edited form as:

Biol Psychiatry. 2012 October 15; 72(8): 629–636. doi:10.1016/j.biopsych.2012.03.037.

Genomewide Linkage Analysis of Obsessive Compulsive Disorder Implicates Chromosome 1p36

Carol A. Mathews, M.D.¹, Judith A. Badner, M.D., Ph.D.², J. Michael Andresen, Ph.D.³, Brooke Sheppard, B.S.¹, Joseph A. Himle, Ph.D.^{4,5}, Jon E. Grant, M.D.⁶, Kyle A Williams, M.D.⁶, Denise A. Chavira, Ph.D.⁷, Amin Azzam, M.D., M.A.¹, Maxine Schwartz, B.S.¹, Victor I. Reus, M.D.¹, Suck Won Kim, M.D.⁶, Edwin H. Cook, M.D.⁸, and Gregory L. Hanna, M.D.⁴

¹Department of Psychiatry, University of California, San Francisco, San Francisco, California

²Department of Psychiatry, University of Chicago, Chicago, Illinois

³Massachusetts Institute of Technology, Cambridge, Massachusetts

⁴Department of Psychiatry, University of Michigan Medical School, Ann Arbor, Michigan

⁵School of Social Work, University of Michigan, Ann Arbor, Michigan

⁶Department of Psychiatry, University of Minnesota, Minneapolis, Minnesota

⁷Department of Psychiatry, University of California, San Diego, La Jolla, California

⁸Department of Psychiatry, University of Illinois at Chicago, Chicago, Illinois

Abstract

Background—Obsessive compulsive disorder (OCD) has a complex etiology involving both genetic and environmental factors. However, the genetic causes of OCD are largely unknown, despite the identification of several promising candidate genes and linkage regions.

Methods—Our objective was to conduct genetic linkage studies of the type of OCD thought to have the strongest genetic etiology (i.e., childhood-onset OCD), in 33 Caucasian families with 2 childhood-onset OCD-affected individuals from the United States (US) (N=245 individuals with genotype data). Parametric and non-parametric genome-wide linkage analyses were conducted with Morgan and Merlin in these families using a selected panel of single nucleotide repeat polymorphisms (SNPs) from the Illumina 610-Quad Bead Chip. The initial analyses were followed by fine-mapping analyses in genomic regions with initial heterogeneity LOD (HLOD) scores of ≥ 2.0 .

Results—We identified five areas of interest (HLOD score ≥ 2) on chromosomes 1p36, 2p14, 5q13, 6p25, and 10p13. The strongest result was on chromosome 1p36.33-p36.32 (HLOD=3.77, suggestive evidence for linkage after fine-mapping). At this location, several of the families showed haplotypes co-segregating with OCD.

© 2012 Society of Biological Psychiatry. Published by Elsevier Inc. All rights reserved.

^{*}Correspondence and reprint requests to: Carol A Mathews, M.D., Department of Psychiatry, University of California San Francisco, 401 Parnassus Avenue, San Francisco CA 94143-0984; cmathews@lppi.ucsf.edu.; phone: 415-476-7702; fax: 415-476-7389.

FINANCIAL DISCLOSURES: The authors report no biomedical financial interests or potential conflicts of interest.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conclusions—The results of this study represent the strongest linkage finding for OCD in a primary analysis to date, and suggest that chromosome 1p36, and possibly several other genomic regions, may harbor susceptibility loci for OCD. Multiple brain-expressed genes lie under the primary linkage peak (approximately 4 mb in size). Follow-up studies, including replication in additional samples and targeted sequencing of the areas of interest, are needed to confirm these findings and to identify specific OCD risk variants.

Keywords

linkage; genomewide; pedigree; obsessive-compulsive; genetics; multigenerational

INTRODUCTION

Obsessive compulsive disorder (OCD) [MIM 164230] is a common neuropsychiatric disorder consisting of repeated, distressing ego-dystonic thoughts (obsessions) and behaviors (compulsions) with a world-wide prevalence of 1–3% (1–4). Family, segregation, and twin studies clearly demonstrate that OCD is familial, with estimated heritabilities for obsessive-compulsive (OC) symptoms of 27% to 47% for adults and 45% to 65% for children (5–18). Genetic epidemiological studies also indicate that OCD with symptom onset before age 18 has a stronger genetic contribution than OCD with later onset, with a doubling of the OCD risk for first-degree family members in families of probands with childhood-onset compared to adult-onset of symptoms (11–18).

Although several promising genes and genomic regions of interest have been identified, clear susceptibility genes for OCD have not yet been demonstrated. Most studies have focused on candidate gene approaches, although four primary genome-wide linkage studies have also been conducted (15, 19–28). These studies have identified eleven genomic regions with LOD scores of ≥ 1.4 on chromosomes 1q, 3q, 6p, 6q, 7p, 9p, 10p, 11p, 14q, 15q, and 19q, most with a broad definition of the OCD phenotype (15, 19–28). The strongest linkage finding in a primary analysis of OCD reported to date was on chromosome 15q14 in three Costa Rican families (LOD score=3.13); this region was also previously identified in a Caucasian sample (25, 28). Other than 15q14 and the 9p region identified by Hanna et al, and subsequently examined as a targeted replication in a separate sample (26, 27), no linkage region has been identified in more than one study. Genome-wide association studies (GWAS) for OCD and Tourette Syndrome (TS), a related disorder, have recently been completed. However, in the context of genetic and environmental heterogeneity, multiple approaches are appropriate, and linkage studies continue to play an important role. While GWAS are useful for the identification of common variants with relatively small effect sizes, linkage studies of multiplex families are particularly useful for the identification of rare variants with larger effect sizes that are increasingly believed to underlie a substantial proportion of the risk for complex disorders (29). Individual *variants* identified via linkage approaches may be family specific, each accounting for a small proportion of the overall variance; however, the *genes* implicated by these variants will be of interest in multiple samples and populations, likely accounting for a much larger proportion of the overall OCD risk, in addition to providing insights about the biology of OCD.

The aim of this study was to search for genomic regions that potentially harbor OCD susceptibility genes using genome-wide linkage approaches in Caucasian families with childhood-onset OCD.

MATERIALS AND METHODS

Sample Collection

The sample consisted of 33 families (245 individuals with genotype data) ascertained for ongoing genetic studies of OCD in the US (Table 1). We included families for whom phenotype data were available for 2 OCD-affected individuals (broad or narrow phenotype) and who had 1 affected (narrow phenotype) and 1 unaffected individual with genotype data. Families ranged in size from 4 individuals in two generations to 58 individuals in four generations (examples in Figure 1). Families were ascertained via probands with DSM-IV OCD whose symptoms began before age 18 and who did not have a pervasive developmental disorder, bipolar disorder, schizophrenia, or a primary psychotic disorder. All families were Caucasian ethnicity of European (primarily Northern European) descent. Families were ascertained and collected at the University of California, San Diego and subsequently the University of California, San Francisco (CAM), the University of Michigan (GLH), and the University of Minnesota (SWK). Genome-wide linkage analyses using a less dense set of microsatellite markers in 17 of the 18 Michigan families have been previously reported(24, 26). The study was approved by the Institutional Review Boards of the participating sites. After complete discussion of the study with the participants, written informed consent or assent was obtained; parental permission was also obtained for participants under age 18.

Clinical assessments at UCSF/UCSD and Minnesota were conducted by psychiatrists or PhD-level psychologists specializing in OCD and trained in the research instruments. Clinical assessments at Michigan were conducted by interviewers with at least a master's degree plus clinical training who were trained to 90% diagnostic agreement with the assessment instruments. The primary assessment instruments for all sites included the adult and child versions of the Yale-Brown Obsessive Compulsive Scale (Y-BOCS and CY-BOCS, respectively) (UCSF/UCSD and Minnesota) or the Schedule for Tourette and Other Behavioral Syndromes (STOBS), which includes a modified version of the Y-BOCS (Michigan), complemented with the Diagnostic Interview for Genetics Studies (DIGS) (UCSF/UCSD) or the Structured Clinical Interview for DSM-IV Axis I diagnoses (SCID) (Michigan and Minnesota) for adults, and the Schedule for Affective Disorders and Schizophrenia for School-Age Children (KSADS) (all sites) (30–35) (Additional details in Supplement). Because phenotypic data were collected independently and at different time periods by the sites, clinical data collection could not be standardized prospectively. Instead, phenotypes were standardized at the diagnostic level, using a common phenotype matrix across all sites that included two OCD phenotypes, an unaffected phenotype, and an unknown phenotype (see below). Concordance between sites was achieved via a best estimate (BE) consensus approach(36). This approach, which uses all available sources of information (e.g., medical records, clinical interviews, self-report questionnaires, and family history interviews), requires 100% concordance on all elements of the diagnostic criteria, and reduces the phenotypic heterogeneity that may arise from the use of different assessment instruments (details in Supplement).

Two OCD diagnoses were assigned: narrow and broad OCD. A narrow OCD diagnosis was given if the individual met all DSM-IV criteria for OCD. The broad OCD diagnosis encompassed both DSM-IV OCD and subclinical OCD, which was considered present if the individual had clear obsessions and/or compulsions, but did not quite meet the impairment or distress criteria (e.g., OC symptoms taking less than an hour and causing mild rather than moderate to severe distress and/or impairment). The broad definition was designed to capture a robust phenotype that is likely to be etiologically related to OCD, but was not severe enough to meet strict DSM-IV criteria. Participants were considered unknown for both phenotypes if there was a history of thoughts or behaviors suggestive of OC symptoms

that met most, but not all criteria for subclinical OCD, or if they were under age 40 and did not have OC symptoms. Individuals with subclinical OCD who were coded as affected for the broad analyses were coded as unknown in the narrow analyses. Participants with no history of any OC symptoms who were 40 years old at the time of the interview were classified as unaffected. The mean age of onset of OC symptoms was 8.7, and the mean lifetime worst-ever Y-BOCS/CY-BOCS severity score was 24.0 for the broad phenotype and 24.9 for the narrow phenotype (Table 2). 15% had a co-occurring chronic tic disorder (Tourette Syndrome, chronic motor or vocal tic disorder), and 4% had a co-occurring eating disorder.

Genotyping

DNA extraction was performed from blood or immortalized lymphoblastoid cell lines according to standard procedures. A small number of individuals from the UCSF sample were genotyped using the Illumina Linkage Panel IVb at the UCSF Genome Core Facility (UCSF GCF). The rest were genotyped using the Illumina Human 610-Quad BeadChip at the Broad Institute (Massachusetts General Hospital). Data were analyzed for quality control and Mendel errors using GenomeStudio software (Illumina). 540,123 SNPs were retained for analysis for the samples genotyped with the 610-Quad BeadChip. Those samples that were genotyped on the Illumina Linkage Panel 4b had 2,157 markers that overlapped with the Human 610-Quad BeadChip; genotypes for the additional markers were coded as missing.

Statistical analysis

Pedigree relationships were confirmed prior to analysis using PREST and PLINK(37, 38). In two families, pedigree structures were altered to incorporate non-paternities that were identified through these assessments. Parametric and nonparametric linkage analyses were conducted using Morgan version 3.0 and Merlin (details in Supplement)(39, 40). We chose to use both Merlin and Morgan because of the size and complexity of the pedigrees combined with the number of genetic markers available. While Morgan can analyze very large pedigrees, the number of markers that can be analyzed are limited, and must be in linkage equilibrium. In contrast, Merlin controls for the effects of linkage disequilibrium between markers, utilizing all available genotype information, but cannot use all individuals due to the size and complexity of the largest families. PedShrink was used to trim the pedigrees as needed for the Merlin analysis, with priority on trimming uninformative individuals(41).

We used a model-based (dominant and recessive) approach because simulation studies show that formulating a genetic model that approximates the true inheritance may have more power than nonparametric analyses, in part because parametric models can utilize information about unaffected individuals, which is not the case for nonparametric analyses(42). We also conducted non-parametric analyses (details in Supplement) because, for loci with high frequencies and low penetrances, non-parametric models may be more powerful. Because we had three different analytic approaches, each with different strengths, we were able to compare results across approaches, identifying and prioritizing those that were consistent across analyses as the most likely to represent true linkage regions.

The linkage parameters (see Supplement) were chosen to model a relatively rare locus with a large effect size and to reduce the risk of false positives due to phenocopies, given the high degree of bilineality that is seen in OCD families, including ours. We note that power to detect linkage is not sensitive to misspecification of penetrance or allele frequency, but instead is most sensitive to degree of dominance(43). Heterogeneity LOD (HLOD) scores

were calculated by allowing the proportion of linked families to vary and estimating the proportion that gave the highest LOD scores for a given region.

Fine-mapping

Fine-mapping using additional SNP markers from the Illumina 610-Quad Bead Chip was conducted on chromosomal regions where the HLOD scores were ≥ 2.0 using the model and phenotype that showed the strongest evidence of linkage. All SNPs from the 610-Quad marker panel that were under the linkage peak of interest and had a MAF of >0.25 were identified and used for the Merlin analyses; for the Morgan analyses, this marker set was then pruned for linkage disequilibrium so that only SNPs with a pairwise $r^2 < 0.1$ were included.

Haplotype analysis

Haplotypes of SNPs pruned for linkage disequilibrium were generated for all linked pedigrees in the genomic region with the highest HLOD score using the “haplotype analysis” command in Simwalk2snp and visualized using HaploPainter(44). Haplotypes that were inherited identical by descent and co-segregated with the OCD phenotype were assessed within each family.

Estimation of genome-wide significance values

Estimates of genome-wide significance values, incorporating both the markers used in the original linkage analyses and those used for fine-mapping, were calculated using 1) simulations and 2) the autoregressive method described by Bacanau(45). Permutations were performed using gene-dropping simulations, as implemented in Morgan and Merlin. For both Morgan and Merlin, 1000 replicates were simulated and each replicate was analyzed with both parametric models and both affection statuses. The significance for each LOD score was assessed by: 1) counting the number of replicates (nr) in which the maximum LOD score exceeded the observed lod score; and 2) calculating the p-value as $(nr + 1)/1001$. The threshold for genome-wide significant linkage was taken to be the 49th highest LOD score of the 1000 replicates. Criteria for significant genome-wide linkage (occurring in 5% of genome scans by chance) was determined to be LOD of 2.8–2.9 for a single analysis in Morgan (3.1–3.3 in Merlin), and 3.3 considering all four parametric analyses (3.8 in Merlin). It is likely that the thresholds are higher for Merlin because many more markers were analyzed. In comparison, Lander and Kruglyak suggested that a LOD score of 3.3 for a parametric linkage analysis of an “infinitely dense” map be considered the threshold for a genome-wide significant result(46). We also used the autoregressive method to generate genome-wide significance thresholds from the data and made a (conservative) Bonferroni correction for the number of genome scans that were done (both parametric and non-parametric). The range of LOD score thresholds for suggestive linkage using this approach was 3.1 to 3.5, and the range of LOD score thresholds for significant linkage was 4.1 to 4.8 (Table S1 in the Supplement).

RESULTS

Genome-wide linkage analysis

We identified eleven chromosomal regions with HLOD scores ≥ 1.5 , a threshold commonly used to identify linkage regions of interest, and five with a HLOD score ≥ 2 (Table 3). Figures S1–S3 in the Supplement shows the results of the genome-wide parametric and non-parametric analyses. The region with the highest HLOD score was on chromosome 1p36, with a maximum HLOD score of 2.96 using Merlin under the dominant model and broad phenotype (LOD score without correction for heterogeneity = -3.88) and a maximum

HLOD score of 2.88 with the dominant model and narrow phenotype (LOD score without correction for heterogeneity = -2.74). The maximum HLOD score in this region with Morgan was 2.66 under a dominant model using the narrow phenotype, and the maximum LOD score in this region using the nonparametric approach was 0.87 at marker rs2377041 with Morgan and 0.93 at markers rs6676961 to rs6677984 with Morgan. Note that the difference between the parametric and nonparametric LOD scores is most likely due to the added information provided by the unaffected individuals in the parametric analyses.

Seventeen of the 33 families showed LOD scores >0 in this region. Only one of the 11 identified linkage regions has been previously reported as potentially harboring OCD susceptibility genes; the linkage region on chromosome 6p25, at ~ 3 Mb, had a maximum HLOD score of 2.56, and is near the 6p25 region identified by Hanna et al at ~ 5 Mb(26). When the six pedigrees that were linked in both the previous and the current studies were excluded, the evidence for linkage in this region remained, although somewhat diminished (Table 4).

Individual family LOD scores

Individual family LOD scores for the genomic regions with HLODs ≥ 2 are shown in Table S2 in the Supplement. The strongest genome-wide individual family LOD scores for the three largest families (shown in Figure 1) were 3.4 for family 1 on chromosome 1p36 (broad phenotype, dominant inheritance), 2.0 on chromosome 18q22.1 for family 2 (narrow phenotype, dominant inheritance), and 1.4 on chromosome 1q31 for family 3 (narrow phenotype, dominant inheritance).

Fine mapping

Fine-mapping analyses were conducted on chromosomes 1p36, 2p14, 5q13, 6p25, and 10p13 (Figure 2). For chromosome 1p36, which had similar high HLOD scores for both the broad and narrow phenotypes in Merlin, we conducted fine-mapping analyses for the narrow phenotype in both Morgan and Merlin and for the broad phenotype in Merlin only, as Morgan gave a HLOD score of only 1.35 for the broad phenotype. For chromosome 2, the HLOD score decreased with the inclusion of additional markers, and for chromosomes 5 and 10, the HLOD score increased in one of the two analyses only. For all other genomic regions, the HLOD scores increased with the inclusion of additional markers for both analyses (Table 4). As in the genome-wide analysis, the highest overall HLOD score was obtained at chromosome 1p36.33 to 1p36.32, with a maximum HLOD score of 3.77 at marker rs897615 using Merlin and 3.08 at marker rs884080 using Morgan (dominant model, narrow phenotype in both) (Figure 2). The confidence interval for this linkage peak using $\text{HLOD} > 2.0$ as the cutoff was bounded by SNPs rs884080 to rs7518255 for the Morgan analysis and by SNPs rs4475691 to rs1874266 for the Merlin analysis.

Examination of haplotypes

We examined the haplotypes generated by Simwalk2snp in all families linked to chromosome 1p36 using the LD-pruned SNP set. We identified haplotypes that co-segregated with OCD, encompassing the region with the highest genome-wide HLOD scores, in the majority of the linked families. In the largest family, which had a LOD score of 2.9 under the narrow phenotype (LOD = 3.4 under the broad phenotype), 11 of the 14 individuals with the narrow OCD phenotype carried a common haplotype inherited from the founder, along with all four individuals with the broad OCD phenotype and one obligate carrier. In the next largest family, which had a LOD score of 0.8 under the narrow phenotype, five of the seven biologically related individuals with the narrow OCD diagnosis (including the founder) carry a shared haplotype, along with three of the four biologically related individuals with the broad OCD phenotype, and two obligate carriers (Figure S4 in

the Supplement). While there was haplotype sharing within each family, we did not identify a haplotype that was shared between families.

DISCUSSION

In this study we report the results of a genome-wide linkage analysis in multiply-affected pedigrees with childhood-onset OCD. We identified several genomic regions of potential interest for OCD on chromosomes 1p36, 2p14, 5q13, 6p25, and 10p13. The linkage region on chromosome 1p36.33-1p36.32, which meets genome-wide criteria for suggestive linkage after fine-mapping based on our calculated significance thresholds ($HLOD=3.77$) and spans 4 Mb, is the strongest linkage finding for OCD reported to date, and the most interesting region that we identified. The majority of the linkage signal on 1p36 comes from our largest family, however, 16 other families also contributed to the LOD score, and had haplotypes that co-segregated with either the narrow or the broad OCD phenotype, suggesting that the finding is not specific to a single family.

Although this is the first reported linkage for OCD on chromosome 1p36, several other neuropsychiatric disorders have been linked to the 1p36 region or nearby including major depressive disorder (MDD), eating disorders (ED), childhood-onset mood disorders, and childhood epilepsy(47–50). A whole genome linkage scan of recurrent MDD identified a maximum LOD score of 3.03 in females (there was no evidence of linkage in males) at 1p36.23 to 1p36.22 (7.6 Mb to 12.3 Mb), a region that adjoins our linkage region(47). Major depression co-occurs with OCD in about 50% of cases, including in our families, and a recent family study has suggested that childhood-onset OCD with MDD may represent a distinct etiological syndrome(51).

Similarly, a whole-genome linkage scan for ED identified a linkage region on chromosome 1p33 to 1p36 (fine mapping peak multipoint NPL score of 3.45, restricting subtype of anorexia nervosa), just centromeric to our linkage region (1p36.33 to 1p36.32, or 0 to 4 Mb) (49, 50). As with MDD, there is evidence that OCD and ED show substantial phenotypic and etiological overlap(49, 52–55). A recent comprehensive review of epidemiological, longitudinal, and family studies suggests a much higher rate of co-occurrence of ED and OCD than expected by chance, as well as a clear etiological relationship between these disorders. The data suggest that the two most likely models are 1) OCD and ED are alternate expressions (or different phases) of the same underlying etiological risk factors, or 2) OCD, which often has an earlier age of onset than ED, is a risk factor for the development of an ED(56). Rates of ED were very low in our families (~4%), and were not concentrated in the families that were linked to chromosome 1, suggesting that the second model is not likely for our sample. However, both models are consistent with shared genetic risk factors, and therefore candidate genomic regions that are identified in both OCD and ED are of increased interest.

The 1p36 region has also been implicated in a deletion syndrome (1p36 syndrome) with a complex phenotype characterized by intellectual disability and multiple system anomalies(57). Behavioral disorders, including self-biting, temper tantrums, reduced social interactions, stereotypies or other repetitive movements, and hyperphagia have been reported in approximately half of the deletion 1p36 cases(57). OCD symptoms have not specifically been reported, but given the phenomenological similarities between stereotypies and compulsive behaviors, and the role of the striatum in both phenotypes, the overlap of this known deletion syndrome with our primary linkage region strengthens the hypothesis that it may harbor susceptibility genes for OCD and perhaps for other related disorders of childhood, as do the linkage findings for MDD and ED.

STRENGTHS AND LIMITATIONS

The major strengths of our study are the size of the sample and the informativeness of the families for linkage analyses. We believe that the use of both parametric and nonparametric analyses is also a strength—this approach was chosen to maximize the information available from complex pedigrees that are not easily accommodated by a single phenotype or model in a disorder with an unknown mode of inheritance. Nevertheless, we recognize that multiple analyses can lead to an inflation of the LOD scores. We have corrected for this by calculating the relevant genome-wide significance thresholds using the actual data, and by examining the evidence for linkage across the multiple analyses, as well as by examining haplotypes and segregation patterns in our families. The results of the simulations indicate that our LOD scores are not artificially inflated. We believe that, for chromosome 1 at least, the convergence of results across analyses strengthens rather than weakens the evidence for linkage.

The primary limitation of this study relates to heterogeneity. For example, there may be phenotypic heterogeneity among families ascertained from the different sites due to the nonuniformity of clinical assessments used (although all sites used a version of the Y-BOCS or CY-BOCS and an additional semi-structured, well-validated instrument for clinical assessment). We have addressed this by conducting BE diagnoses and requiring 100% concordance on all of the diagnostic criteria for phenotypic assignment of both OCD (narrow phenotype) and subclinical OCD (broad phenotype). We believe that this rigorous approach minimizes the problem of potential phenotypic heterogeneity.

There is also genetic heterogeneity in the study population, perhaps due to subtle ethnic variation (e.g., northern vs southern European descent). However, the use of an HLOD approach, which identifies and incorporates heterogeneity in the linkage analysis, addresses this concern. Genetic heterogeneity is also evident in the observation that, while we identified haplotypes on chromosome 1p36 that co-segregate with OCD within families, we did not identify a consistent haplotype that co-segregated with OCD across families. This is not surprising, given that the families in our sample are from outbred Caucasian populations rather than from a genetic isolate, and does not necessarily reduce our confidence in the results. It does highlight the need for further investigation of this region and the other genomic regions of interest identified by our study, however, as discussed below.

CONCLUSION

In conclusion, this work identifies a new region of interest for OCD on chromosome 1p36. Despite meeting suggestive rather than significant linkage criteria based on our simulations, this is the strongest linkage result reported to date for this complex disorder. This region has previously been associated with several neuropsychiatric phenotypes that are related to OCD, including eating and mood disorders, and the 1p36 deletion syndrome. As with all genetic investigations, follow-up studies are needed to validate and extend these findings, including replication of the linkage findings, and sequencing of the region to identify possible functional variants within particular gene(s).

Although for a time supplanted in favor of case control or trio-based approaches such as GWAS, the advent of high-throughput sequencing technology has caused a renewed interest in linkage studies for complex traits. Such studies are needed to help determine which of the many rare, potentially deleterious, variants identified through sequencing co-segregate with disease. Genome-wide linkage studies such as this one can help to identify and prioritize genomic regions of interest, as well as helping to identify the most informative individuals within linked pedigrees for either targeted or complete genome sequencing.

In addition, although the common disease/common variant approach has driven the interest in GWAS, rare variants have also been shown to be important in the etiology of common disease (e.g., Crohn's disease)(58). Variants identified through family-based approaches, such as linkage and translocation studies, while individually affecting only a small proportion of families or individuals, have led to the identification of biologically relevant genes, gene clusters, or gene networks that are involved in the pathogenesis of Alzheimer disease (presenilin genes) and schizophrenia (DISC1 and 2)(59, 60). As with these complex traits, in order to more fully understand the biological underpinnings of OCD, multiple approaches, including (but likely not limited to) linkage, GWAS, whole-genome sequencing, and animal model studies, will ultimately be required.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by grants to CAM from the National Center for Research Resources (K23 RR015533), the National Alliance for Research on Schizophrenia and Affective Disorders, the Obsessive Compulsive Foundation, and the Althea Foundation, by grants to GLH from the National Institute of Mental Health (K20 MH 01065 and R01 MH 58376) and the Obsessive Compulsive Foundation, and to DAC from the National Institute of Mental Health (K01 MH072952).

References

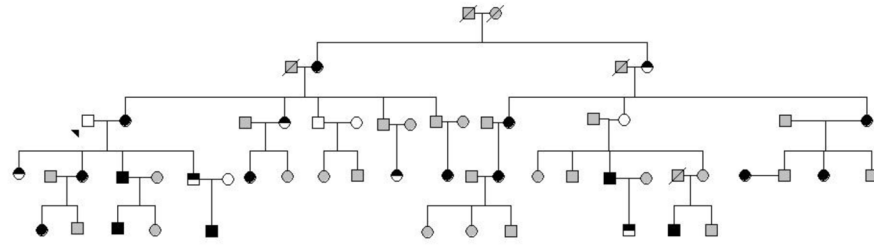
1. Ruscio AM, Stein DJ, Chiu WT, Kessler RC. The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Mol Psychiatry*. 2008 Aug 26.
2. APA. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV TR). 4. Washington, DC: American Psychiatric Association; 2000. Text Revision
3. Fontenelle LF, Mendlowicz MV, Versiani M. The descriptive epidemiology of obsessive-compulsive disorder. *Prog Neuropsychopharmacol Biol Psychiatry*. 2006 May; 30(3):327–37. [PubMed: 16412548]
4. Horwath E, Weissman MM. The epidemiology and cross-national presentation of obsessive-compulsive disorder. *Psychiatr Clin North Am*. 2000 Sep; 23(3):493–507. [PubMed: 10986723]
5. Eley TC, Bolton D, O'Connor TG, Perrin S, Smith P, Plomin R. A twin study of anxiety-related behaviours in pre-school children. *J Child Psychol Psychiatry*. 2003 Oct; 44(7):945–60. [PubMed: 14531577]
6. Van Grootheest DS, Cath DC, Beekman AT, Boomsma DI. Genetic and environmental influences on obsessive-compulsive symptoms in adults: a population-based twin-family study. *Psychol Med*. 2007 Nov; 37(11):1635–44. [PubMed: 17592667]
7. van Grootheest, DS.; Cath, DC.; Beekman, ATF.; Boomsma, DI., editors. International Congress on Psychiatric Genetics. Boston, MA: 2005. Genetic and environmental influences on OC symptoms in adults: a population based twin-family study.
8. Hudziak JJ, Van Beijsterveldt CE, Althoff RR, Stanger C, Rettew DC, Nelson EC, et al. Genetic and environmental contributions to the Child Behavior Checklist Obsessive-Compulsive Scale: a cross-cultural twin study. *Arch Gen Psychiatry*. 2004 Jun; 61(6):608–16. [PubMed: 15184240]
9. van Grootheest DS, Cath DC, Beekman AT, Boomsma DI. Twin studies on obsessive-compulsive disorder: a review. *Twin Res Hum Genet*. 2005 Oct; 8(5):450–8. [PubMed: 16212834]
10. van Grootheest DS, Bartels M, Cath DC, Beekman AT, Hudziak JJ, Boomsma DI. Genetic and environmental contributions underlying stability in childhood obsessive-compulsive behavior. *Biol Psychiatry*. 2007 Feb 1; 61(3):308–15. [PubMed: 16950209]
11. Pauls DL, Alsobrook JP 2nd, Goodman W, Rasmussen S, Leckman JF. A family study of obsessive-compulsive disorder. *Am J Psychiatry*. 1995 Jan; 152(1):76–84. [PubMed: 7802125]

12. Nestadt G, Lan T, Samuels J, Riddle M, Bienvenu OJ 3rd, Liang KY, et al. Complex segregation analysis provides compelling evidence for a major gene underlying obsessive-compulsive disorder and for heterogeneity by sex. *Am J Hum Genet.* 2000 Dec; 67(6):1611–6. [PubMed: 11058433]
13. Nestadt G, Samuels J, Riddle M, Bienvenu OJ 3rd, Liang KY, LaBuda M, et al. A family study of obsessive-compulsive disorder. *Arch Gen Psychiatry.* 2000 Apr; 57(4):358–63. [PubMed: 10768697]
14. Geller D, Biederman J, Jones J, Park K, Schwartz S, Shapiro S, et al. Is juvenile obsessive-compulsive disorder a developmental subtype of the disorder? A review of the pediatric literature. *J Am Acad Child Adolesc Psychiatry.* 1998 Apr; 37(4):420–7. [PubMed: 9549963]
15. Pauls DL. The genetics of obsessive compulsive disorder: a review of the evidence. *Am J Med Genet C Semin Med Genet.* 2008 May 15; 148C(2):133–9. [PubMed: 18412099]
16. Hanna GL, Fingerlin TE, Himle JA, Boehnke M. Complex Segregation Analysis of Obsessive-Compulsive Disorder in Families with Pediatric Proband. *Hum Hered.* 2005 Jul 27; 60(1):1–9. [PubMed: 16088199]
17. Hanna GL, Fischer DJ, Chadha KR, Himle JA, Van Etten M. Familial and sporadic subtypes of early-onset Obsessive-Compulsive disorder. *Biol Psychiatry.* 2005 Apr 15; 57(8):895–900. [PubMed: 15820710]
18. Hanna GL, Himle JA, Curtis GC, Gillespie BW. A family study of obsessive-compulsive disorder with pediatric probands. *Am J Med Genet B Neuropsychiatr Genet.* 2005 Apr 5; 134(1):13–9. [PubMed: 15635694]
19. Dickel DE, Veenstra-VanderWeele J, Cox NJ, Wu X, Fischer DJ, Van Etten-Lee M, et al. Association testing of the positional and functional candidate gene SLC1A1/EAAC1 in early-onset obsessive-compulsive disorder. *Arch Gen Psychiatry.* 2006 Jul; 63(7):778–85. [PubMed: 16818867]
20. Arnold PD, Sicard T, Burroughs E, Richter MA, Kennedy JL. Glutamate transporter gene SLC1A1 associated with obsessive-compulsive disorder. *Arch Gen Psychiatry.* 2006 Jul; 63(7):769–76. [PubMed: 16818866]
21. Bienvenu OJ, Wang Y, Shugart YY, Welch JM, Grados MA, Fyer AJ, et al. Sapap3 and pathological grooming in humans: Results from the OCD collaborative genetics study. *Am J Med Genet B Neuropsychiatr Genet.* 2008 Dec 2.
22. Wang Y, Samuels JF, Chang YC, Grados MA, Greenberg BD, Knowles JA, et al. Gender differences in genetic linkage and association on 11p15 in obsessive-compulsive disorder families. *Am J Med Genet B Neuropsychiatr Genet.* 2009 Jan 5; 150B(1):33–40. [PubMed: 18425788]
23. Samuels J, Shugart YY, Grados MA, Willour VL, Bienvenu OJ, Greenberg BD, et al. Significant linkage to compulsive hoarding on chromosome 14 in families with obsessive-compulsive disorder: results from the OCD Collaborative Genetics Study. *Am J Psychiatry.* 2007 Mar; 164(3):493–9. [PubMed: 17329475]
24. Hanna GL, Veenstra-Vanderweele J, Cox NJ, Van Etten M, Fischer DJ, Himle JA, et al. Evidence for a susceptibility locus on chromosome 10p15 in early-onset obsessive-compulsive disorder. *Biol Psychiatry.* 2007 Oct 15; 62(8):856–62. [PubMed: 17544380]
25. Shugart YY, Samuels J, Willour VL, Grados MA, Greenberg BD, Knowles JA, et al. Genomewide linkage scan for obsessive-compulsive disorder: evidence for susceptibility loci on chromosomes 3q, 7p, 1q, 15q, and 6q. *Mol Psychiatry.* 2006 Aug; 11(8):763–70. [PubMed: 16755275]
26. Hanna GL, Veenstra-VanderWeele J, Cox NJ, Boehnke M, Himle JA, Curtis GC, et al. Genome-wide linkage analysis of families with obsessive-compulsive disorder ascertained through pediatric probands. *Am J Med Genet.* 2002 Jul 8; 114(5):541–52. [PubMed: 12116192]
27. Willour VL, Yao Shugart Y, Samuels J, Grados M, Cullen B, Bienvenu OJ 3rd, et al. Replication study supports evidence for linkage to 9p24 in obsessive-compulsive disorder. *Am J Hum Genet.* 2004 Sep; 75(3):508–13. [PubMed: 15272418]
28. Ross J, Badner J, Garrido H, Sheppard B, Chavira DA, Grados M, et al. Genomewide linkage analysis in Costa Rican families implicates chromosome 15q14 as a candidate region for OCD. *Hum Genet.* 2011 Jun 21.
29. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet.* 2010 Jun; 11(6):415–25. [PubMed: 20479773]

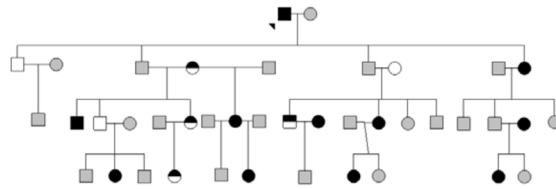
30. Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, et al. The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. *Arch Gen Psychiatry*. 1989 Nov; 46(11):1006–11. [PubMed: 2684084]
31. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, et al. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry*. 1997 Jul; 36(7): 980–8. [PubMed: 9204677]
32. Nurnberger JI Jr, Blehar MC, Kaufmann CA, York-Cooler C, Simpson SG, Harkavy-Friedman J, et al. Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. *Arch Gen Psychiatry*. 1994 Nov; 51(11):849–59. discussion 63-4. [PubMed: 7944874]
33. Scahill L, Riddle MA, McSwiggin-Hardin M, Ort SI, King RA, Goodman WK, et al. Children's Yale-Brown Obsessive Compulsive Scale: reliability and validity. *J Am Acad Child Adolesc Psychiatry*. 1997 Jun; 36(6):844–52. [PubMed: 9183141]
34. Pauls, DLHC. Schedule for Tourette and other behavioral syndromes. New Haven, CT: Child Study Center, Yale University; 1991.
35. Orvaschel, H. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Epidemiologic Version-5 (K-SADS-E-5). Fort Lauderdale, FL: Nova Southeastern University, Center for Psychological Studies; 1995.
36. Leckman JF, Sholomskas D, Thompson WD, Belanger A, Weissman MM. Best estimate of lifetime psychiatric diagnosis: a methodological study. *Arch Gen Psychiatry*. 1982 Aug; 39(8): 879–83. [PubMed: 7103676]
37. 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 Sep; 81(3):559–75. [PubMed: 17701901]
38. McPeck MS, Sun L. Statistical tests for detection of misspecified relationships by use of genome-screen data. *Am J Hum Genet*. 2000 Mar; 66(3):1076–94. [PubMed: 10712219]
39. Wijsman EM, Rothstein JH, Thompson EA. Multipoint linkage analysis with many multiallelic or dense diallelic markers: Markov chain-Monte Carlo provides practical approaches for genome scans on general pedigrees. *Am J Hum Genet*. 2006 Nov; 79(5):846–58. [PubMed: 17033961]
40. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet*. 2002 Jan; 30(1):97–101. [PubMed: 11731797]
41. Schaid, DJ. pedigree.shrink. 2009. Available from: http://mayoresearch.mayo.edu/schaid_lab/software.cfm
42. Greenberg DA, Abreu P, Hodge SE. The power to detect linkage in complex disease by means of simple LOD-score analyses. *Am J Hum Genet*. 1998 Sep; 63(3):870–9. [PubMed: 9718328]
43. Clerget-Darpoux F, Bonaiti-Pellie C, Hochez J. Effects of misspecifying genetic parameters in lod score analysis. *Biometrics*. 1986 Jun; 42(2):393–9. [PubMed: 3741977]
44. Thiele H, Nurnberg P. HaploPainter: a tool for drawing pedigrees with complex haplotypes. *Bioinformatics*. 2005 Apr 15; 21(8):1730–2. [PubMed: 15377505]
45. Bacanu SA. Robust estimation of critical values for genome scans to detect linkage. *Genet Epidemiol*. 2005 Jan; 28(1):24–32. [PubMed: 15372617]
46. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet*. 1995 Nov; 11(3):241–7. [PubMed: 7581446]
47. McGuffin P, Knight J, Breen G, Brewster S, Boyd PR, Craddock N, et al. Whole genome linkage scan of recurrent depressive disorder from the depression network study. *Hum Mol Genet*. 2005 Nov 15; 14(22):3337–45. [PubMed: 16203746]
48. Feng Y, Kapornai K, Kiss E, Tamas Z, Mayer L, Baji I, et al. Association of the GABRD gene and childhood-onset mood disorders. *Genes Brain Behav*. 2010 Aug; 9(6):668–72. [PubMed: 20561060]
49. Grice DE, Halmi KA, Fichter MM, Strober M, Woodside DB, Treasure JT, et al. Evidence for a susceptibility gene for anorexia nervosa on chromosome 1. *Am J Hum Genet*. 2002 Mar; 70(3): 787–92. [PubMed: 11799475]

50. Bergen AW, van den Bree MB, Yeager M, Welch R, Ganjei JK, Haque K, et al. Candidate genes for anorexia nervosa in the 1p33–36 linkage region: serotonin 1D and delta opioid receptor loci exhibit significant association to anorexia nervosa. *Mol Psychiatry*. 2003 Apr; 8(4):397–406. [PubMed: 12740597]
51. Hanna GL, Himle JA, Hanna BS, Gold KJ, Gillespie BW. Major depressive disorder in a family study of obsessive-compulsive disorder with pediatric probands. *Depress Anxiety*. 2011 Jun; 28(6):501–8. [PubMed: 21538726]
52. Cavallini MC, Bertelli S, Chiapparino D, Riboldi S, Bellodi L. Complex segregation analysis of obsessive-compulsive disorder in 141 families of eating disorder probands, with and without obsessive-compulsive disorder. *Am J Med Genet*. 2000 Jun 12; 96(3):384–91. [PubMed: 10898919]
53. Bellodi L, Cavallini MC, Bertelli S, Chiapparino D, Riboldi C, Smeraldi E. Morbidity risk for obsessive-compulsive spectrum disorders in first-degree relatives of patients with eating disorders. *Am J Psychiatry*. 2001 Apr; 158(4):563–9. [PubMed: 11282689]
54. Murphy R, Nutzinger DO, Paul T, Leflow B. Conditional-associative learning in eating disorders: a comparison with OCD. *J Clin Exp Neuropsychol*. 2004 Apr; 26(2):190–9. [PubMed: 15202539]
55. Boghi A, Sterpone S, Sales S, D'Agata F, Bradac GB, Zullo G, et al. In vivo evidence of global and focal brain alterations in anorexia nervosa. *Psychiatry Res*. 2011 Jun 30; 192(3):154–9. [PubMed: 21546219]
56. Altman SE, Shankman SA. What is the association between obsessive-compulsive disorder and eating disorders? *Clin Psychol Rev*. 2009 Nov; 29(7):638–46.
57. Battaglia A, Hoyme HE, Dallapiccola B, Zackai E, Hudgins L, McDonald-McGinn D, et al. Further delineation of deletion 1p36 syndrome in 60 patients: a recognizable phenotype and common cause of developmental delay and mental retardation. *Pediatrics*. 2008 Feb; 121(2):404–10. [PubMed: 18245432]
58. Pritchard JK, Cox NJ. The allelic architecture of human disease genes: common disease-common variant...or not? *Hum Mol Genet*. 2002 Oct 1; 11(20):2417–23. [PubMed: 12351577]
59. St George-Hyslop PH, Petit A. Molecular biology and genetics of Alzheimer's disease. *C R Biol*. 2005 Feb; 328(2):119–30. [PubMed: 15770998]
60. Brandon N, Millar K, Korth C, Sive H, Sing KK, Sawa A. Understanding the role of DISC1 in psychiatric disease and during normal development. *J Neurosci*. 2009; 29(41):12768–75. [PubMed: 19828788]

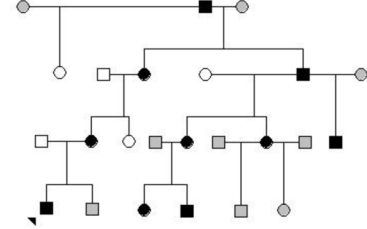
Family 1



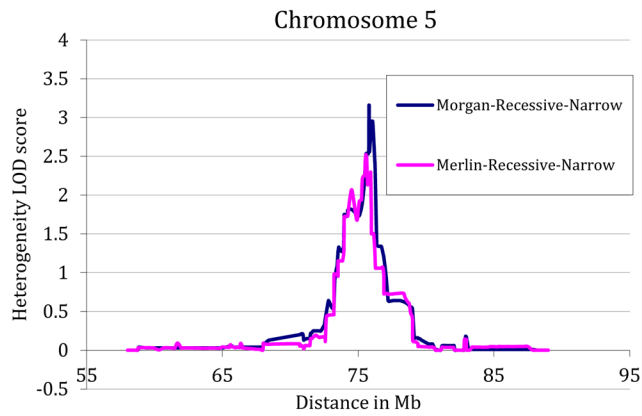
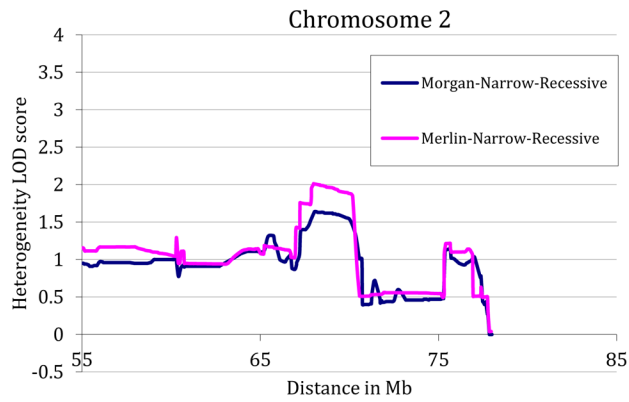
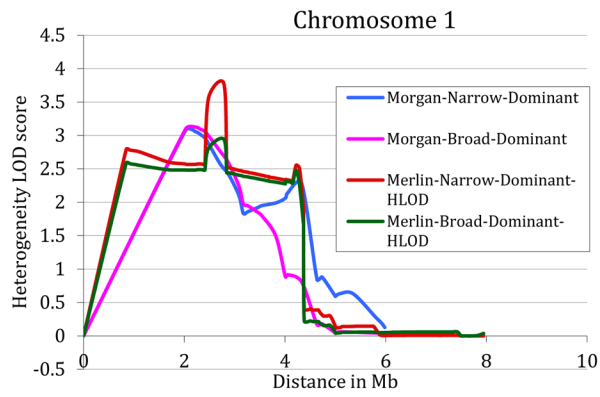
Family 2



Family 3

**Figure 1.**

Examples of large, multigenerational pedigrees used in the genome-wide linkage analysis of OCD. Filled black symbols indicate narrow OCD phenotype, half filled black symbols indicate broad OCD phenotype, grey symbols indicate unknown phenotype, and unfilled symbols indicate unaffected phenotype.



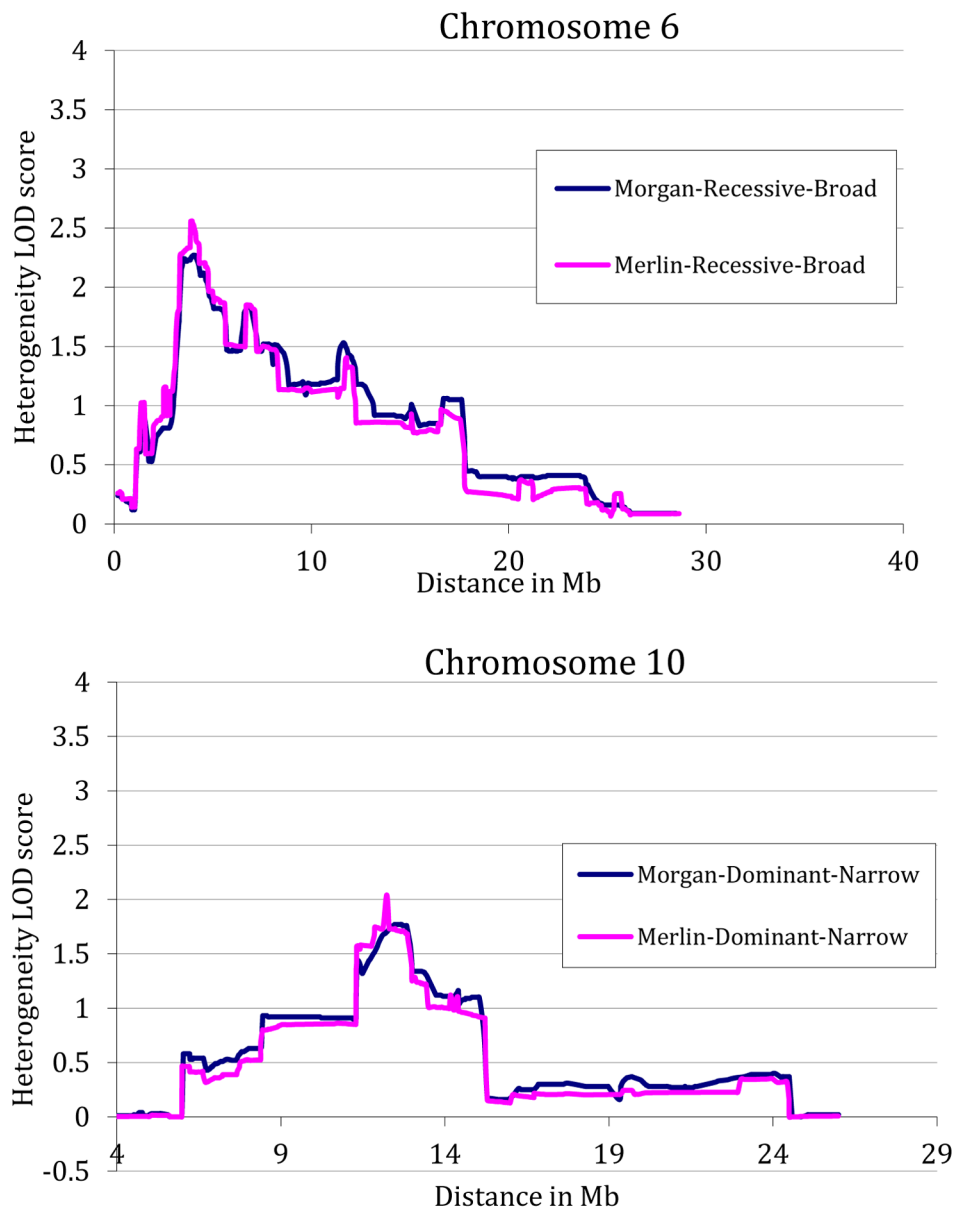


Figure 2. Fine-mapping analyses for genomic regions with HLOD scores ≥ 2.0 . Y axis indicates HLOD score; X axis indicates genomic position. Graphs are presented for phenotype, model, and analysis type (Merlin or Morgan) with the highest HLOD score.

Table 1

Characteristics of individuals in 33 early-onset OCD families included in the linkage analysis. All families had 2 OCD-affected members (narrow or broad definition) with available clinical data, with genotype data available on 1 OCD-affected individual and 1 unaffected individual.

	All sites	UCSF/UCSD	Minnesota	Michigan
Total # families	33	14	1	18
Total # individuals genotyped	245	105	39	101
	Narrow OCD			
Mean # affected individuals per pedigree (median, range)	4.1 (3, 1–17)	3.8 (3, 1–11)	17	3.6 (3, 2–10)
Mean # genotyped affected individuals per pedigree (median, range)	3.8 (3, 1–15)	3.4 (3, 1–11)	15	3.4 (3, 2–10)
	Broad OCD			
Mean # affected individuals per pedigree (median, range)	4.7 (3, 2–23)	4.4 (3.5, 2–15)	23	3.9 (3, 2–10)
Mean # genotyped affected individuals per pedigree (median, range)	4.2 (3, 1–19)	4.0 (3, 2–15)	19	3.6 (3, 2–10)

Clinical characteristics of individuals with either narrow or broad OCD (N=158) in the 33 early-onset OCD families included in the linkage analysis.

Table 2

	Mean age of onset of OC symptoms (SD, range)	Mean YBOCS/C-YBOCS severity score (SD, range) Broad phenotype	Mean YBOCS/C-YBOCS severity score (SD, range) Narrow phenotype	Number (%) of individuals with a tic disorder	Number (%) of individuals with an eating disorder
Total	8.7 (4.3, 2–27)	24.0 (9.4, 0–40)	24.9 (8.6, 2–40)	24 (15%)	7 (4%)
UCSF/UCSD	7.4 (3.1, 2–15)	21.2 (9.3, 0–40)	22.1 (8.6, 2–40)	4 (6%)	4 (6%)
Minnesota	8.5 (2.8, 3–12)	15.2 (10.2, 0–36)	15.3 (8.7, 4–36)	6 (26%)	0 (0%)
Michigan	9.6 (5.0, 3–27)	28.6 (6.4, 8–40)	29.6 (5.0, 18–40)	14 (19%)	3 (4%)

Mean YBOCS/C-YBOCS total severity score is lifetime worst ever. Tic disorder=Tourette Syndrome, chronic motor tic disorder, or chronic vocal tic disorder. Eating disorder=anorexia nervosa or bulimia nervosa.

Table 3

Chromosomal regions with LOD scores > 1.5.

Chrom	SNP	LOD score	HLOD score	Analysis	Model	Phenotype	cM	alpha
1	rs884080	-2.74	2.66	Morgan	Dominant	Narrow	4	0.4
1	rs6676961	-3.18	2.88	Merlin	Dominant	Narrow	11	0.435
1	rs6676961	-3.88	2.96	Merlin	Dominant	Broad	11	0.408
2	rs10197684	1.78		Merlin	Nonparametric	Narrow	62	
2	rs7560138	-0.69	2.41	Merlin	Recessive	Narrow	90	0.545
2	rs7560138	-2.12	1.95	Merlin	Recessive	Broad	90	0.465
2	rs1986601	-4.72	1.78	Morgan	Recessive	Narrow	91	0.43
5	rs1700680	1.04	2.64	Merlin	Recessive	Narrow	90	0.647
5	rs1700680	1.77		Merlin	Nonparametric	Narrow	90	
5	rs1700680	-1.00	1.56	Merlin	Recessive	Broad	90	0.492
5	rs1700680	1.74		Merlin	Nonparametric	Broad	90	
6	rs226938	-4.05	1.68	Merlin	Recessive	Narrow	11	0.397
6	rs226938	-10.92	1.52	Merlin	Dominant	Broad	11	0.385
6	rs4959264	-2.64	2.16	Merlin	Recessive	Broad	12	0.488
6	rs2025267	-3.55	1.64	Morgan	Recessive	Broad	12	0.41
10	rs2096396	-7.19	1.54	Merlin	Dominant	Narrow	29	0.334
10	rs1111060	-4.35	2.00	Morgan	Dominant	Narrow	30	0.43
14	rs10139518	1.84		Merlin	Nonparametric	Broad	105	
16	rs26762	-2.49	1.997	Merlin	Recessive	Broad	48	0.411
20	rs976192	-0.29	1.56	Morgan	Recessive	Broad	6	0.48

cM=centiMorgans. Results presented are for the SNP with maximum HLOD score in the region. LOD score= LOD score when all families are included. HLOD= heterogeneity LOD score, excluding unlinked families.

Table 4

Finemapping results for chromosomal regions with initial LOD scores > 2.0.

Chr	Model	Phenotype	Merlin				Morgan				
			Original HLOD score	HLOD score	Individual model genome-wide p-value	Overall parametric genome-wide p-value	SNP	HLOD score	Individual model genome-wide p-value	Overall parametric genome-wide p-value	SNP
1	Dom	Narrow	2.88	3.77	0.015	0.050	rs897615	3.08	0.032	0.087	rs884080
2	Rec	Narrow	2.41	2.01	0.499	0.845	rs1346870	1.64	0.56	0.904	rs1503236
5	Rec	Narrow	2.64	2.52	0.197	0.506	rs2972831	3.16	0.031	0.075	rs7727095
6	Rec	Broad	2.16	2.56	0.189	0.477	rs4959902	2.27	0.17	0.446	rs1414102
6*	Rec	Broad	1.01	1.34	0.948	1.00	rs13196874	1.12	0.93	1.0	rs2147374
10	Rec	Narrow	2.00	2.04	0.438	0.830	rs12260965	1.77	0.42	0.812	rs10906164

cM=centiMorgans. Results presented are for the SNP with maximum LOD score in the region. With the exception of chromosome 10, all of the original maximum LOD scores were obtained using Merlin. Chr=chromosome Dom=dominant; Rec=recessive. 6*=linkage results on chromosome 6 excluding the six Michigan families who were previously linked to this region in an earlier analysis. Significance thresholds for Morgan analyses as determined by gene-dropping simulations = 2.9 for narrow phenotype (both dominant and recessive), 2.8 for broad phenotype (dominant and recessive), and 3.3 controlling for all four parametric analyses. Significance thresholds for Merlin =3.3 for narrow phenotype (dominant and recessive), 3.1 for dominant inheritance and broad phenotype, 3.2 for recessive inheritance and broad phenotype, and 3.8 controlling for all four parametric analyses. Genome-wide p-values were determined using the simulation approach described in the methods.