Title

Genomewide linkage analysis in Costa Rican families implicates chromosome 15q14 as a candidate region for OCD

Journal

Human Genetics, 130(6)

ISSN

0340-6717

Authors

Ross, Jessica
Badner, Judith
Garrido, Helena
et al.

Publication Date

2011-12-01

DOI

10.1007/s00439-011-1033-6

Peer reviewed

control critera.

The call rate threshold for samples was ≥90%. We have added a statement to this effect in the Methods.

4. *No rationale provided for selection of individuals for sequencing. For example, why were all individuals in Family 1 sequenced but only the offspring in family 3?*

We selected the individuals within each family that were likely to be the most informative in sequencing studies. For example, the phenotypes of the parents in family 3 are unknown, sequencing data is unlikely to be of additional benefit, as we would be unable to determine whether variants of interest co-segregate with the phenotype. We have added a section in the Methods outlining these choices.

5. *"…several [of the coding variants] were associated with the major haplotype of interest in one of the three families". It would be helpful to know how many variants were associated with the haplotype of interest, without having to go to the Supplemental Table.*

We have added this information in the text of the Results section and have highlighted these variants in grey in Supplemental Table 3.

We hope that these revisions are satisfactory, and submit the revised manuscript for your consideration for publication.

Thank you very much,

Carol A. Mathews, MD

Genomewide Linkage Analysis in Costa Rican Families Implicates Chromosome 15q14 as a Candidate Region for OCD

Jessica Ross, M.D. M.S.

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

Judith Badner, M.D., Ph.D.

Departments of Psychiatry and Human Genetics, University of Chicago, Chicago, Illinois

Helena Garrido, M.A.

Departments of Psychiatry and Pediatrics, Hospital Nacional de Niños, San Josè, Costa Rica

Brooke Sheppard, B.S.

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

Denise Chavira, Ph.D.

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

Marco Grados, M.D., M.P.H.

Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, Baltimore, MD

Jonathan M Woo, M.S.

Genome Core Facility, Institute for Human Genetics, University of California, San Francisco, San Francisco, California

Pamela Doo, B.A.

Genome Core Facility, Institute for Human Genetics, University of California, San Francisco, San Francisco, California

Paula Umaña, B.A.

Departments of Psychiatry and Pediatrics, Hospital Nacional de Niños, San Josè, Costa Rica

Eduardo Fournier, M.A.

Departments of Psychiatry and Pediatrics, Hospital Nacional de Niños, San Josè, Costa Rica

Sarah Shaw Murray, Ph.D

Department of Genetics, Scripps Genomic Medicine, La Jolla, California

Carol A. Mathews, M.D

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

*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; e-mail: cmathews@lppi.ucsf.edu.; phone: 415-476-7702; fax: 415-476-7389

Keywords: Obsessive Compulsive Disorder, Genetic Linkage, Genetic Isolate, Genetic Loci, Humans, Genetic Predisposition to Disease

ABSTRACT

Obsessive compulsive disorder (OCD) has a complex etiology that encompasses both genetic and environmental factors. However, to date, despite the identification of several promising candidate genes and linkage regions, the genetic causes of OCD are largely unknown. The objective of this study was to conduct linkage studies of childhood-onset OCD, which is thought to have the strongest genetic etiology, in several OCD-affected families from the genetically isolated population of the Central Valley of Costa Rica (CVCR). The authors used parametric and non-parametric approaches to conduct genome-wide linkage analyses using 5786 single nucleotide repeat polymorphisms (SNPs) in three CVCR families with multiple childhood-onset OCD-affected individuals. We identified areas of suggestive linkage (LOD score ≥2) on chromosomes 1p21, 15q14, 16q24, and 17p12. The strongest evidence for linkage was on chromosome 15q14 (LOD=3.13), identified using parametric linkage analysis with a recessive model, and overlapping a region identified in a prior linkage study using a Caucasian population. Each CVCR family had a haplotype that co-segregated with OCD across a ~7Mbp interval within this region, which contains 18 identified brain expressed genes, several of which are potentially relevant to OCD. Exonic sequencing of the strongest candidate gene in this region, the ryanodine receptor 3 (RYR3), identified several genetic variants of potential interest, although none cosegregated with OCD in all three families. These findings provide evidence that chromosome 15q14 is linked to OCD in families from the CVCR, and supports previous findings to suggest that this region may contain one or more OCD susceptibility loci.

INTRODUCTION

Although multiple etiologic factors are important in the development of obsessive-compulsive disorder (OCD), there is strong evidence for a genetic contribution, particularly when symptoms begin in childhood(Pauls 2008). Several promising candidate genes have been identified for OCD, including SLC1A1(Dickel et al. 2006; Arnold et al. 2006). However, these genes, if confirmed, will only account for a small proportion of the genetic variance of this complex disorder. Large-scale genome-wide association studies (GWAS) are currently underway for OCD; however, such studies will most likely account for only a proportion of the total disease variation(Yang et al. 2011; Manolio et al. 2009). For example, the estimated proportion of the variance explained by all currently reported genetic loci identified by GWAS for height, a complex trait with a heritability of 80-90%, is approximately 10%; this number increases to 45% with recently developed genomic partitioning approaches, but still does not approach 100%. Similar results are seen with categorical traits such as type 2 diabetes and Crohn's disease, where published susceptibility loci account for 6% and 20% of the total variance, respectively(Manolio et al. 2009). Linkage studies represent a complementary approach to candidate gene studies and GWAS. Where GWAS are useful for the identification of common variants with relatively small effect sizes, linkage studies are 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(Cirulli and Goldstein 2010).

To date, three primary linkage studies for OCD have been published, along with two secondary analyses and one targeted replication; these studies have identified several genomic candidate regions for OCD (Supplemental Table 1)(Wang et al. 2009; Samuels et al. 2007; Hanna et al. 2007; Shugart et al. 2006; Hanna et al. 2002; Willour et al. 2004). None of the

studies using OCD as the primary phenotype identified regions meeting genome-wide significance criteria for linkage, although two re-analyses of a large OCD affected-sib pair sample using alternate phenotypes or sample selection did identify significant linkage regions(Samuels et al. 2007; Wang et al. 2009). The first re-analysis included only families with two or more members affected with compulsive hoarding, and identified a region on chromosome 14q with a LOD score of 3.7, and the second focused on families with male probands only, and identified a region on chromosome 11p with an initial LOD score of 2.92 and a subsequent LOD score of 5.08 following fine mapping with additional markers(Samuels et al. 2007; Wang et al. 2009).

As linkage and segregation studies suggest, the identification of susceptibility genes in OCD is likely confounded by a complex oligogenetic architecture, as well as genetic and environmenal heterogeneity across populations(Cavallini et al. 1999; Geller et al. 2007; Eapen et al. 2006; Hanna et al. 2005; Nestadt et al. 2000). One approach to reducing genetic and environmental heterogeneity in such complex traits is to conduct studies using multiply-affected families in genetically isolated founder populations(Mathews et al. 2004; Hovatta et al. 1999). Studies in these populations benefit from decreased genetic heterogeneity due to decreased migration and, in some cases, increased inbreeding over a defined period of time(Mathews et al. 2004). Additional potential advantages include the possibility for decreased environmental heterogeneity due to the isolation of the population, and for decreased phenotypic heterogeneity due to cultural variations in symptom expression, interpretation, or reporting.

We report on the results of linkage and sequencing studies in three multiply-affected OCD pedigrees from the genetically isolated population of the Central Valley of Costa Rica (CVCR), a population with a well-documented history as a genetic isolate, and one in which founder

haplotypes for several inherited disesases have been demonstrated(Uhrhammer et al. 1995; Leon et al. 1992).

MATERIALS AND METHODS

Families and Subjects

As a part of ongoing genetic studies of OCD in Costa Rica, we identified three families with two or more OCD-affected individuals and CVCR ancestry for whom DNA was available from individuals in at least two generations (Figure 1). Families were originally ascertained via probands with DSM-IV OCD whose obsessive-compulsive symptoms began before age 18, who did not have Tourette Syndrome (TS), a pervasive developmental disorder, bipolar disorder, schizophrenia, or a primary psychotic disorder. The study was approved by the Institutional Review Boards of the University of California, San Francisco, the University of California, San Diego, and the Hospital Nacional de Niños in San José, Costa Rica. 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 were conducted in Spanish by psychiatrists and psychologists specializing in OCD and trained in the research instruments (CAM, HG, and MB). Primary assessment instruments included the Yale Brown Obsessive Compulsive Scale (YBOCS), the Diagnostic Interview for Genetics Studies (DIGS) for adults, and the Kiddie Schedule for Affective Disorders and Schizophrenia (KSADS) for children under 18. Additional clinical assessments included the Leyton Obsessional Inventory, childhood version (LOI-CV), the Structured Clinical Interview for DSM-IV Axis II diagnoses (SCID II), and a semi-structured interview that assessed time course and impact of symptoms, presence and severity of related symptoms (including tics), and developmental, school, social, and family history. Diagnoses

were assigned using a best estimate consensus approach using all available clinical data by two research clinicians (DC, MG) who were blinded to the presumed clinical diagnosis, family history, or relationship within the pedigree(Leckman et al. 1982; Chavira et al. 2007). Two OCD diagnoses were assigned: narrow and broad OCD. A diagnosis of narrow OCD was given if the individual met all of the DSM-IV criteria for OCD. A diagnosis of broad OCD encompassed both DSM-IV OCD and subclinical OCD, which was considered present if the individual had five or more obsessive-compulsive symptoms that were persistent, took less than an hour, and caused some distress and/or impairment (even if mild), or if the symptoms took an hour or more but were associated with minimal distress 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 for OCD. We required the presence of at least five obsessive-compulsive symptoms to ensure that isolated (usually ego-syntonic) symptoms (for example, a cluster of two or three related cleaning behaviors causing little to no distress or impairment), which occur frequently in the population, but may not be etiologically related to OCD, were not included in the phenotype definitions, to avoid the introduction of phenocopies. Consistent with this aim, individuals with one to four obsessive-compulsive symptoms were coded as phenotype unknown. Age of onset of obsessive-compulsive symptoms was assessed in all individuals. The average age of onset was 10.9, and all individuals had childhood-onset of symptoms (range 3-18).

Genotyping

DNA extraction was performed from blood or immortalized lymphoblastoid cell lines according to standard procedures. Genotyping was performed on approximately two-thirds of the sample at the Southern California Genotyping Consortium (SCGC) using the Illumina linkage panel IVa;

the remaining sample was genotyped at the UCSF Genome Core Facility (UCSF GCF) using the Illumina linkage panel IVb. Only markers that were present in both linkage panels were retained (N=5858). Data for each sample were analyzed for quality control and Mendel errors using GenomeStudio Data Analysis software (Illumina). The average call rate for the 5858 SNPs was 0.99. After initial data cleaning on each sample, data from the two sample sets were merged, and additional data cleaning and quality control procedures were performed. Error checking was completed using Illumina GenomeStudio, Progeny and Pedcheck(Lab ; O'Connell and Weeks 1998). Markers that showed Mendelian inconsistency, those that were not in Hardy-Weinberg equilibrium (HWE) (exclusion threshold $p<0.0001$), and those with <95% call rates were removed from the marker set, resulting in a final total of 5786 SNPs (98.8%) retained for analysis. The threshold for exclusion of individual samples was set at genotyping call rates <90%. However, all genotyped individuals passed this threshold and therefore, none were excluded.

Statistical analysis

To maximize the power of this small sample, the primary analysis was a multipoint parametric linkage analysis. We chose to examine a model-based approach (dominant and recessive models) as our primary analysis, in addition to the secondary non-parametric (model-free) analysis, to maximize the power of this small sample, and because examination of the pedigrees suggested that either a dominant or a recessive mode of inheritance was possible (Figure 1). Simulation studies have shown that formulating a genetic model that approximates the true inheritance may have more power than nonparametric analyses(Greenberg et al. 1998). Pedigree relationships were confirmed prior to analysis using PREST and PLINK. Parametric linkage analyses were

conducted using MORGAN version 3.0 and Merlin(Wijsman et al. 2006; Abecasis et al. 2002). MORGAN uses genetic and phenotypic information from all available family members. Prior to performing the linkage analysis, the linkage panel was pruned so that only SNPs with a pairwise $r^2 \leq 0.5$ were analyzed (N=2526), to avoid an increase in false positives that may occur when linkage disequilibrium (LD) between SNPs is not taken into account. In contrast, MERLIN has the capacity to analyze all SNPs, accounting for LD, but is not able to use all individuals due to the size and complexity of the largest family. For the MERLIN analysis, PedShrink was used to trim the pedigrees as needed, with priority on trimming uninformative individuals(Schaid 2009). For the dominant model, we assumed a disease allele frequency of 0.01, a penetrance of 0.01 for the wildtype genotype (dd), a penetrance of 0.6 for the heterozygote (Dd) and homozygote genotypes (DD), where D is the disease-causing allele and d is the wildtype allele, resulting in a phenocopy rate of 45%. For the recessive model we assumed a disease allele frequency of 0.10, a penetrance of 0.01 for the wildtype and heterozygote genotypes, and a penetrance of 0.6 for the homozygote genotype, resulting in a phenocopy rate of 50%. These parameters were chosen to model a relatively rare locus with a large effect size, as might be expected to occur in families with multiple affected individuals collected from a genetically isolated population. We used two phenotypes, narrow and broad OCD, as described above. Individuals with subclinical OCD, who were coded as affected for the broad analyses, were coded as unknown in the narrow analyses. Nonparametric linkage analyses (NPL) were conducted using Simwalk2snp, using the LD-pruned SNP panel(Sobel et al. 2001). The primary parameters examined were the NPL-pairs statistic, which is a measure of allele sharing among affected relative-pairs (roughly equal to the sum of the conditional kinship coefficient for all affected relative-pairs), and the NPL-all statistic, which is a measure of whether a few founder alleles are overly represented in the

affecteds. Critical significance thresholds for suggestive and significant NPL scores were generated for NPL-all (3.38 for significant linkage and 2.35 for suggestive linkage) and NPL-pairs scores (3.92 for significant linkage and 3.07 for suggestive linkage) from the data using the autoregressive method described by Bacanu (Bacanu 2005).

Haplotype analysis and candidate gene selection

Chromosomal regions that had >5 consecutive SNPs with LOD or NPL scores (either NPL-pairs or NPL-all) ≥2.0 were identified as being of interest for further investigation, and were prioritized according to the following criteria: 1) the magnitude of the LOD score; 2) the number of consecutive markers with LOD scores ≥2; 3) identification as a region of interest for OCD in previous linkage analyses, or known to contain candidate genes for OCD; and 4) presence of a haplotype that co-segregated with the OCD phenotype.

Estimated haplotypes were generated for each pedigree using the "haplotype analysis" command in Simwalk2snp. The haplotypes were then visualized using Haplopainter(Thiele and Nurnberg 2005). As a further check, haplotypes were also constructed by hand (blind to the results generated by Simwalk2snp) for the most significant regions of interest and compared to the results of the haplotype analysis for discrepancies. Haplotypes that were inherited identical by descent and co-segregated with the OCD phenotype were assessed within each family. The families were then assessed to identify haplotypes or portions of the haplotypes that were shared across families. After the genomic region of highest interest was determined, candidate genes within the region were identified for further exploration. Initially, all genes and transcripts were identified using the UCSC genome browser, assembly GRCh37 (http://genome.ucsc.edu/). Only those genes that were identified as brain expressed in the UCSC genome browser were examined

further. Candidate genes were then ranked according to 1) whether they had been previously associated with OCD; 2) where in the brain they were expressed, with priority given to those expressed in regions that have been associated with OCD (i.e., striatum, orbitofrontal cortex, thalamus, hippocampus, cingulate); and 3) where they fell under the linkage peak, with priority given to those with the strongest evidence for linkage(Menzies et al. 2008). The top ranking candidate gene was then sequenced as described below.

Sequencing

Sequencing of the top candidate gene was conducted in selected OCD-affected individuals from each family. Individuals for sequencing were chosen from within each family so as to maximize the information and minimize the number of total individuals sequenced. In family 2, three individuals who carried the haplotype in the area of interest (one from each living generation; individuals 1, 13, and 23), and two OCD-affected individuals who did not carry the haplotype (individuals 7 and 16) were sequenced. In family 1, where phenotypes were known for all family members, all individuals were sequenced, and in family 3, the two OCD-affected siblings were sequenced, while the parents, who had unknown phenotypes and were therefore not informative, were not. Note that the mother in family 1 endorsed OCS that did not meet the threshold for subclinical OCD; therefore, she was coded as unknown in the genetic analyses. The exons and 150bp on either side of the exon were sequenced. Sequence data were viewed and analyzed with the Sequencher program (GeneCodes), and sequence variants were identified by comparison with the reference human genome. Sequence variants were confirmed by sequencing the opposite strand using the Sanger sequencing method. All sequence variants that were present in the OCD-affected individuals, both those carrying the haplotype of interest and those who did not carry the haplotype, were investigated further. We examined both non-synonomous and

synonomous coding changes within exonic regions, as well as sequence variants within non-coding regions. Predicted function of identified variants was assessed using SIFT; if SIFT was unable to identify the variant, F-SNP (http://genepi_toolbox.i-med.ac.at/), an integrated functional prediction tool that searches across multiple databases (e.g., Ensembl, RESCUE_ESE, SIFT, etc), was used to assess predicted function(Ng 2001; Ng and Henikoff 2003; Yuan et al. 2006; Lee and Shatkay 2008).

RESULTS:

Parametric linkage analysis

We carried out parametric linkage analysis on the three families shown in Figure 1 for the narrow and broad phenotypes under both autosomal dominant and autosomal recessive models. We identified eleven chromosomal regions with LOD scores ≥1.5, a threshold commonly used to identify areas of interest in linkage studies, and four with LOD scores ≥2 (Table 1). Three of these regions met our criteria for further investigation. The first was on chromosome 15q14, with a maximum LOD score of 3.13 under the recessive model using the broad phenotype, the second was on chromosome 1p21, which had a LOD score of 2.16 under recessive model using the narrow phenotype, and the third was on chromosome 16q24, and had a LOD score of 2.15 under the dominant model using the narrow phenotype. The heterogeneity alpha for the region on chromosome 15 was 1.0, suggesting that all three families contributed to the LOD score in this region. The peak individual family LOD scores in this region were: 0.7813 (family 1), 1.8758 (family 2), and 0.474 (family 3). Figure 2 shows the LOD scores for chromosomes 1, 15, and 16; LOD scores for the other chromosomes are in Supplemental Figure 1. The LOD scores for

family 2, which contributed the most to the overall LOD score on chromosome 15q14, are shown in Supplemental Figure 2.

Examination of haplotypes: We examined the haplotypes generated by Simwalk2snp for the chromosome 15q14, chromosome 1p21, and chromosome 16q24 regions, with a primary focus on chromosome 15q. All three families contributed to the LOD score on chromosome 15q, which showed the strongest evidence of linkage, and had been identified in a previous linkage study of OCD as a region of interest(Shugart et al. 2006). In each of the families, we identified a conserved haplotype that co-segregates with OCD on 15q, encompassing the region with the highest genome-wide LOD scores. In family 1 (see Figure 1 for family structures), all four siblings carried a common haplotype inherited from the OCD-affected father; the three OCD-affected siblings also carried a smaller haplotype in this region inherited from the mother, who endorsed some obsessive compulsive symptoms but did not meet the full best estimate criteria for the broad phenotype and therefore was coded as phenotype unknown. In family 3, the two OCD-affected siblings carried a shared haplotype inherited from the mother and one inherited from the father; both parents' phenotypes are unknown. In family 2, the largest of the families, 9 of 11 individuals with the narrow phenotype, and one individual who is an obligate carrier and has the broad phenotype (subclinical OCD) shared a haplotype, including the three affected siblings in the founding generation (Figure 3). Two individuals with the narrow phenotype (a mother and son pair) did not carry the shared haplotype. While there was haplotype sharing within each family, we did not identify a haplotype that was shared by all three families.

The region of chromosome 15q as defined by the conserved haplotype in family 2 covers approximately 15 Mbp and contains 18 brain-expressed genes (Supplemental Table 2). In

combining the information provided from recombinations that defined the haplotypes in each of the families, we were able to narrow the region of interest on 15q from 15 Mbp to approximately 7 Mbp (bounded by SNPs rs1520942 and rs1003223).

Although all brain-expressed genes in this region are potentially of interest as candidate genes for OCD, the SNPs with the strongest evidence of linkage in all families (6 consecutive SNPs) were located in the intronic region between exons 1 and 2 of the ryanodine 3 receptor (RYR3). According to the Allen Brain Atlas, RYR3 is broadly expressed in human brain, with high levels of expression in the cingulate gyrus, hippocampus, parahippocampal gyrus, occipital and parietal lobes, temporal lobe, globus pallidus, striatum (including the caudate and putamen), dorsal thalamus, pontine tegmentum, and myelencephalon(Allen Brain Atlas Resources 2009). Functional and structural imaging studies have consistently implicated several of these regions in OCD pathophysiology, particularly the cingulate gyrus, the hippocampus, the striatum, and the thalamus(Takeshima et al. 1996; Rauch et al. 2001; Nakashima et al. 1997; Menzies et al. 2008). For these reasons, we chose RYR3 as our top candidate for sequencing. RYR3 is a large gene, with 103 exons, and encoding 4870 amino acids.

Sequencing of RYR3

We identified 184 sequence variants in the exons and surrounding intronic regions that differed between our sample and the reference human genome, 23 (12.5%) of which were novel. The variants identified included 9 non-synonymous coding variants, two of which were novel, and 16 synonymous coding variants, two of which were novel. However, none of the identified coding variants were uniquely associated with OCD in all families, although six were associated with the major haplotype of interest in one of the three families (highlighted in grey in Supplemental Table 3).

We also identified 159 intronic variants; of these, 19 were novel. None of the variants segregated perfectly with the haplotype of interest in all three families, although 48 variants had allele frequencies of <0.5 and were present on the haplotype of interest in at least one of the two larger families (Supplemental Table 3). The intronic variants included several that are predicted by programs such as FastSNP and F-SNP to have potential functional effects (to be involved in splice site regulation, for example) (Supplemental Table 3).

DISCUSSION

Our results provide evidence from linkage and haplotype studies that chromosome 15q14 may harbor one or more susceptibility genes for OCD. This study specifically focused on OCD-affected families from the genetically isolated population of the CVCR, which has the potential advantage of increasing the power to detect linkage even in a relatively small number of families. In fact, even with only three families, the linkage finding on chromosome 15 is the strongest that has yet been reported for a primary analysis of OCD, and is further strengthened by the identification of shared haplotypes within each family that co-segregate with the OCD phenotype. The utility of looking for individual or private mutations in relatively small numbers of families from specific populations such as the CVCR has been previously successful for complex traits such as Alzheimer disease, where linkage studies of a small number of Volga German families identified a linkage on chromosome 1 that eventually led to the identification of the PS-2 mutation, which was then subsequently identified in sporadic Alzheimer disease cases(St George-Hyslop and Petit 2005).

Chromosome 15q14-15 has been previously implicated as a region of interest for OCD in a linkage analysis using affected sib pair families(Shugart et al. 2006). The Shugart et al. study, which used a non-parametric analysis and comparable phenotype definitions to those used in this study, identified a Kong and Cox LOD score of 1.32 under the broad phenotype on chromosome 15q13-15q35. Chromosome 15q14-15 has also been implicated in studies of compulsive behavior in mice, where linkage to compulsive wheel running was found on mouse chromosomes 2, 6, and 7; the linked region of the mouse chromosome 2 is homologous to the human chromosome 15q region identified previously and in our study(Kas et al. 2010). In addition, multiple SNPs in the 15q14 region have recently been found to be associated with both OCD and with Tourette Syndrome, a closely related neuropsychiatric disorder, in GWAS (p values of ~10^{-4}) (J. Scharf and SE Stewart, personal communication, unpublished data). However, to date, no gene in this area has been extensively examined for evidence of a relationship to OCD.

Although there are multiple potential candidate genes of interest for OCD in this region, we chose to examine RYR3, which is of particular interest as a candidate gene for OCD not only because of the localization of the strongest linkage signal in our study to this gene, but also because of its brain expression patterns and its potential relationship to fear conditioning and synaptic plasticity (Rauch et al. 2001; Nakashima et al. 1997; Matsuo et al. 2009; Balschun et al. 1999). We were unable to identify any sequence variants that were both associated with OCD in all three families and were predicted to have a clear functional effect, although we did identify SNPs that were associated with OCD in two of the three families, and were predicted by *in silico* methods to have a potential functional impact on the RYR3 protein.

The ryanodine receptors are calcium induced calcium release (CICR) channels primarily located in smooth and cardiac muscle(Takeshima et al. 1996). RYR3 is expressed in smooth muscle at low levels, but appears to be primarily expressed in brain; as noted previously, it is expressed at high levels in multiple areas implicated in OCD pathology(Rauch et al. 2001; Nakashima et al. 1997). Calcium influx via AMPA-type glutamate receptors has been shown to elicit CICR from ryanodine receptors in rat ganglion cells, suggesting a relationship between glutamate, which has been implicated in OCD pathophysiology, and ryanodine(Morton-Jones et al. 2008). Similarly, RYR3 interacts with BDNF to promote hippocampally-mediated spatial memory learning and memory consolidation(Adasme et al. 2011). Mutant mice lacking RYR3 show impaired spatial learning, impaired hippocampal synaptic plasticity, and impairments in contextual fear conditioning and deficits in social interactions(Matsuo et al. 2009; Balschun et al. 1999; Kouzu et al. 2000). RYR3's effect on fear conditioning appears to be modulated through its ability to activate $Ca2^{+}$/calmodulin-dependent protein kinase II (CaMKII), which is one of the key synaptic molecules involved in long-term fear memory formation, and is highly expressed in the amygdala and hippocampus(Radwanska et al. 2010).

Limitations

There are two main limitations to this study that potentially limit the generalizability. The first is that the sample size is relatively small, and although all three of the families studied contribute to the linkage signal on chromosome 15q14, the majority of the signal comes from family 2, the largest and most complex pedigree. The second limitation is that we used multiple analytic approaches and two phenotypic models, increasing the risk of identifying a false positive. Despite the multiple models tested, however, we believe that the linkage results, which

are consistent across multiple phenotypic and analytic models, are compelling, and as noted, are congruent with findings from previous genetic studies of OCD and OCD-like behaviors. The third limitation is that our sequencing studies did not provide clear and compelling evidence of a causal sequence variant in the exonic regions of the primary candidate gene of interest, RYR3. Given the phenotypic and genetic heterogeneity of OCD, our findings do not entirely rule out RYR3 as a candidate gene, as future studies may identify functionally significant variants in the intronic or regulatory regions. Further, as will likely be the case for many complex neuropsychiatric disorders, genetic heterogeneity, the presence of phenocopies, and incomplete penetrance within families may make the identification of specific causal variants in a given linkage region difficult without complete targeted sequencing efforts followed by functional studies of the identified associated variants.

Summary

In summary, our results, in combination with previously reported OCD linkage studies, recent genome-wide association studies, and mouse studies of the chromosome 15 homologous region, provide multiple lines of evidence to implicate chromosome 15q14 in genetic susceptibility for OCD, and suggest that RYR3, and potentially several other brain-expressed genes under this linkage peak, the nicotinic receptor (CHRNA7) and the muscarinic cholinergic receptor (CHRM5), for example, are worth investigating more fully as potential candidate genes. Follow-up studies are needed to validate and extend these findings, including replication of the linkage findings in additional families, both from the CVCR and other populations, sequencing of the intronic and surrounding regions of the RYR3 gene, examination of the other brain-

expressed genes in this region, and molecular studies aimed at elucidating the functional impact of the identified sequence variants.

ACKNOWLEDGEMENTS: This research was performed at the University of California, San Francisco

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

FINANCIAL DISCLOSURES: The authors report no competing interests.

Parts of this work were presented at the 17th Annual Molecular Psychiatry Conference in Park City, Utah, January 31 to February 2, 2010

References

Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30 (1):97-101. doi:10.1038/ng786 ng786 [pii]

Adasme T, Haeger P, Paula-Lima AC, Espinoza I, Casas-Alarcon MM, Carrasco MA, Hidalgo C (2011) Involvement of ryanodine receptors in neurotrophin-induced hippocampal synaptic plasticity and spatial memory formation. Proc Natl Acad Sci U S A 108 (7):3029-3034. doi:1013580108 [pii] 10.1073/pnas.1013580108

Allen Brain Atlas Resources. (2009) Allen Institute for Brain Science. http://www.brain-map.org/. Accessed 2/25/2011 2011

Arnold PD, Sicard T, Burroughs E, Richter MA, Kennedy JL (2006) Glutamate transporter gene SLC1A1 associated with obsessive-compulsive disorder. Arch Gen Psychiatry 63 (7):769-776

Bacanu SA (2005) Robust estimation of critical values for genome scans to detect linkage. Genet Epidemiol 28 (1):24-32. doi:10.1002/gepi.20030

Balschun D, Wolfer DP, Bertocchini F, Barone V, Conti A, Zuschratter W, Missiaen L, Lipp HP, Frey JU, Sorrentino V (1999) Deletion of the ryanodine receptor type 3 (RyR3) impairs forms of synaptic plasticity and spatial learning. EMBO J 18 (19):5264-5273. doi:10.1093/emboj/18.19.5264

Cavallini MC, Pasquale L, Bellodi L, Smeraldi E (1999) Complex segregation analysis for obsessive compulsive disorder and related disorders. Am J Med Genet 88 (1):38-43

Chavira DA, Garrido H, Bagnarello M, Azzam A, Reus VI, Mathews CA (2007) A comparative study of obsessive-compulsive disorder in Costa Rica and the United States. Depress Anxiety

Cirulli ET, Goldstein DB (2010) Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat Rev Genet 11 (6):415-425. doi:nrg2779 [pii] 10.1038/nrg2779

Dickel DE, Veenstra-VanderWeele J, Cox NJ, Wu X, Fischer DJ, Van Etten-Lee M, Himle JA, Leventhal BL, Cook EH, Jr., Hanna GL (2006) Association testing of the positional and functional candidate gene SLC1A1/EAAC1 in early-onset obsessive-compulsive disorder. Arch Gen Psychiatry 63 (7):778-785

Eapen V, Pauls DL, Robertson MM (2006) The role of clinical phenotypes in understanding the genetics of obsessive-compulsive disorder. J Psychosom Res 61 (3):359-364

Geller D, Petty C, Vivas F, Johnson J, Pauls D, Biederman J (2007) Further evidence for co-segregation between pediatric obsessive compulsive disorder and attention deficit hyperactivity disorder: a familial risk analysis. Biol Psychiatry 61 (12):1388-1394

Greenberg DA, Abreu P, Hodge SE (1998) The power to detect linkage in complex disease by means of simple LOD-score analyses. Am J Hum Genet 63 (3):870-879. doi:S0002-9297(07)61390-1 [pii] 10.1086/301997

Hanna GL, Fingerlin TE, Himle JA, Boehnke M (2005) Complex Segregation Analysis of Obsessive-Compulsive Disorder in Families with Pediatric Probands. Hum Hered 60 (1):1-9

Hanna GL, Veenstra-VanderWeele J, Cox NJ, Boehnke M, Himle JA, Curtis GC, Leventhal BL, Cook EH, Jr. (2002) Genome-wide linkage analysis of families with obsessive-compulsive disorder ascertained through pediatric probands. Am J Med Genet 114 (5):541-552

Hanna GL, Veenstra-Vanderweele J, Cox NJ, Van Etten M, Fischer DJ, Himle JA, Bivens NC, Wu X, Roe CA, Hennessy KA, Dickel DE, Leventhal BL, Cook EH, Jr. (2007) Evidence for a susceptibility locus on chromosome 10p15 in early-onset obsessive-compulsive disorder. Biol Psychiatry 62 (8):856-862

Hovatta I, Varilo T, Suvisaari J, Terwilliger JD, Ollikainen V, Arajarvi R, Juvonen H, Kokko-Sahin ML, Vaisanen L, Mannila H, Lonnqvist J, Peltonen L (1999) A genomewide screen for schizophrenia genes in an isolated Finnish subpopulation, suggesting multiple susceptibility loci. Am J Hum Genet 65 (4):1114-1124. doi:S0002-9297(07)62615-9 [pii]
10.1086/302567

Kas MJ, Gelegen C, van Nieuwerburgh F, Westenberg HG, Deforce D, Denys D (2010) Compulsivity in mouse strains homologous with chromosomes 7p and 15q linked to obsessive-compulsive disorder. Am J Med Genet B Neuropsychiatr Genet 153B (1):252-259. doi:10.1002/ajmg.b.30994

Kouzu Y, Moriya T, Takeshima H, Yoshioka T, Shibata S (2000) Mutant mice lacking ryanodine receptor type 3 exhibit deficits of contextual fear conditioning and activation of calcium/calmodulin-dependent protein kinase II in the hippocampus. Brain Res Mol Brain Res 76 (1):142-150. doi:S0169328X99003447 [pii]

Lab P Progeny Lab. http://www.progenygenetics.com/lab/index.html.

Leckman JF, Sholomskas D, Thompson WD, Belanger A, Weissman MM (1982) Best estimate of lifetime psychiatric diagnosis: a methodological study. Arch Gen Psychiatry 39 (8):879-883

Lee PH, Shatkay H (2008) F-SNP: computationally predicted functional SNPs for disease association studies. Nucleic Acids Res 36 (Database issue):D820-824. doi:gkm904 [pii]
10.1093/nar/gkm904

Leon PE, Raventos H, Lynch E, Morrow J, King MC (1992) The gene for an inherited form of deafness maps to chromosome 5q31. Proc Natl Acad Sci U S A 89 (11):5181-5184

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM (2009) Finding the missing heritability of complex diseases. Nature 461 (7265):747-753. doi:nature08494 [pii]
10.1038/nature08494

Mathews CA, Reus VI, Bejarano J, Escamilla MA, Fournier E, Herrera LD, Lowe TL, McInnes LA, Molina J, Ophoff RA, Raventos H, Sandkuijl LA, Service SK, Spesny M, Leon PE, Freimer NB (2004) Genetic studies of neuropsychiatric disorders in Costa Rica: a model for the use of isolated populations. Psychiatr Genet 14 (1):13-23

Matsuo N, Tanda K, Nakanishi K, Yamasaki N, Toyama K, Takao K, Takeshima H, Miyakawa T (2009) Comprehensive behavioral phenotyping of ryanodine receptor type 3 (RyR3) knockout mice: decreased social contact duration in two social interaction tests. Front Behav Neurosci 3:3. doi:10.3389/neuro.08.003.2009

Menzies L, Chamberlain SR, Laird AR, Thelen SM, Sahakian BJ, Bullmore ET (2008) Integrating evidence from neuroimaging and neuropsychological studies of obsessive-compulsive disorder: the orbitofronto-striatal model revisited. Neurosci Biobehav Rev 32 (3):525-549. doi:S0149-7634(07)00114-5 [pii]
10.1016/j.neubiorev.2007.09.005

Morton-Jones RT, Cannell MB, Housley GD (2008) Ca2+ entry via AMPA-type glutamate receptors triggers Ca2+-induced Ca2+ release from ryanodine receptors in rat spiral ganglion neurons. Cell Calcium 43 (4):356-366. doi:S0143-4160(07)00133-9 [pii]
10.1016/j.ceca.2007.07.003

Nakashima Y, Nishimura S, Maeda A, Barsoumian EL, Hakamata Y, Nakai J, Allen PD, Imoto K, Kita T (1997) Molecular cloning and characterization of a human brain ryanodine receptor. FEBS Lett 417 (1):157-162. doi:S0014-5793(97)01275-1 [pii]

Nestadt G, Lan T, Samuels J, Riddle M, Bienvenu OJ, 3rd, Liang KY, Hoehn-Saric R, Cullen B, Grados M, Beaty TH, Shugart YY (2000) Complex segregation analysis provides compelling evidence for a major gene underlying obsessive-compulsive disorder and for heterogeneity by sex. Am J Hum Genet 67 (6):1611-1616

Ng PC, Henikoff S (2003) SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 31 (13):3812-3814

Ng PC, Henikoff, S. (2001) Predicting deleterious amino acid substitutions. Genome Res 11:12

O'Connell JR, Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 63 (1):259-266. doi:S0002-9297(07)60744-7 [pii]
10.1086/301904

Pauls DL (2008) The genetics of obsessive compulsive disorder: a review of the evidence. Am J Med Genet C Semin Med Genet 148C (2):133-139

Radwanska K, Tudor-Jones AA, Mizuno K, Pereira GS, Lucchesi W, Alfano I, Lach A, Kaczmarek L, Knapp S, Peter Giese K (2010) Differential regulation of CaMKII inhibitor beta protein expression after exposure to a novel context and during contextual fear memory formation. Genes Brain Behav. doi:GBB595 [pii]
10.1111/j.1601-183X.2010.00595.x

Rauch SL, Whalen PJ, Curran T, Shin LM, Coffey BJ, Savage CR, McInerney SC, Baer L, Jenike MA (2001) Probing striato-thalamic function in obsessive-compulsive disorder and Tourette syndrome using neuroimaging methods. Adv Neurol 85:207-224

Samuels J, Shugart YY, Grados MA, Willour VL, Bienvenu OJ, Greenberg BD, Knowles JA, McCracken JT, Rauch SL, Murphy DL, Wang Y, Pinto A, Fyer AJ, Piacentini J, Pauls DL, Cullen B, Rasmussen SA, Hoehn-Saric R, Valle D, Liang KY, Riddle MA, Nestadt G (2007) Significant linkage to compulsive hoarding on chromosome 14 in families with obsessive-compulsive disorder: results from the OCD Collaborative Genetics Study. Am J Psychiatry 164 (3):493-499

Schaid DJ (2009) pedigree.shrink. http://mayoresearch.mayo.edu/schaid_lab/software.cfm.

Shugart YY, Samuels J, Willour VL, Grados MA, Greenberg BD, Knowles JA, McCracken JT, Rauch SL, Murphy DL, Wang Y, Pinto A, Fyer AJ, Piacentini J, Pauls DL, Cullen B, Page J, Rasmussen SA, Bienvenu OJ, Hoehn-Saric R, Valle D, Liang KY, Riddle MA, Nestadt G (2006) Genomewide linkage scan for obsessive-compulsive disorder: evidence for susceptibility loci on chromosomes 3q, 7p, 1q, 15q, and 6q. Mol Psychiatry 11 (8):763-770

Sobel E, Sengul H, Weeks DE (2001) Multipoint estimation of identity-by-descent probabilities at arbitrary positions among marker loci on general pedigrees. Hum Hered 52 (3):121-131. doi:hhe52121 [pii]

St George-Hyslop PH, Petit A (2005) Molecular biology and genetics of Alzheimer's disease. C R Biol 328 (2):119-130

Takeshima H, Ikemoto T, Nishi M, Nishiyama N, Shimuta M, Sugitani Y, Kuno J, Saito I, Saito H, Endo M, Iino M, Noda T (1996) Generation and characterization of mutant mice lacking ryanodine receptor type 3. J Biol Chem 271 (33):19649-19652

Thiele H, Nurnberg P (2005) HaploPainter: a tool for drawing pedigrees with complex haplotypes. Bioinformatics 21 (8):1730-1732. doi:10.1093/bioinformatics/bth488
bth488 [pii]

Uhrhammer N, Lange E, Porras O, Naeim A, Chen X, Sheikhavandi S, Chiplunkar S, Yang L, Dandekar S, Liang T, et al. (1995) Sublocalization of an ataxia-telangiectasia gene distal to D11S384 by ancestral haplotyping in Costa Rican families. Am J Hum Genet 57 (1):103-111

Wang Y, Samuels JF, Chang YC, Grados MA, Greenberg BD, Knowles JA, McCracken JT, Rauch SL, Murphy DL, Rasmussen SA, Cullen B, Hoehn-Saric R, Pinto A, Fyer AJ, Piacentini J, Pauls DL, Bienvenu OJ, Riddle M, Shugart YY, Liang KY, Nestadt G (2009) Gender differences in genetic linkage and association on 11p15 in obsessive-compulsive disorder families. Am J Med Genet B Neuropsychiatr Genet 150B (1):33-40

Wijsman EM, Rothstein JH, Thompson EA (2006) 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 79 (5):846-858. doi:S0002-9297(07)60828-3 [pii]
10.1086/508472

Willour VL, Yao Shugart Y, Samuels J, Grados M, Cullen B, Bienvenu OJ, 3rd, Wang Y, Liang KY, Valle D, Hoehn-Saric R, Riddle M, Nestadt G (2004) Replication study supports evidence for linkage to 9p24 in obsessive-compulsive disorder. Am J Hum Genet 75 (3):508-513

Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N, Cunningham JM, de Andrade M, Feenstra B, Feingold E, Hayes MG, Hill WG, Landi MT, Alonso A, Lettre G, Lin P, Ling H, Lowe W, Mathias RA, Melbye M, Pugh E, Cornelis MC, Weir BS, Goddard ME, Visscher PM (2011) Genome partitioning of genetic variation for complex traits using common SNPs. Nat Genet. doi:ng.823 [pii]
10.1038/ng.823

Yuan HY, Chiou JJ, Tseng WH, Liu CH, Liu CK, Lin YJ, Wang HH, Yao A, Chen YT, Hsu CN (2006) FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. Nucleic Acids Res 34 (Web Server issue):W635-641. doi:34/suppl_2/W635 [pii]
10.1093/nar/gkl236

Table and Figure Legends and Footnotes:

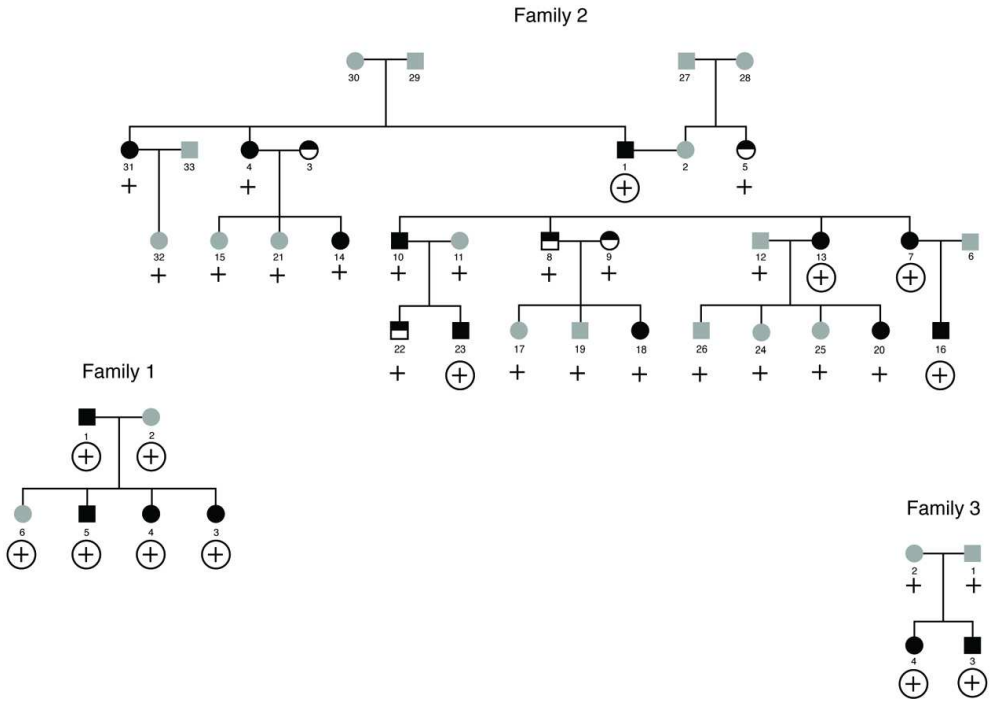
Table 1: Chromosomal regions with LOD scores ≥1.5. * = SNP(s) with highest LOD score. # SNPs = number of consecutive SNPs with LOD scores >1.5. cM=centiMorgans

Figure 1: Pedigrees included in the linkage analysis. Full black symbols indicate DSM-IV OCD. Half black symbols indicate subclinical OCD. Grey symbols= affected or unknown phenotype. **+**= individuals with genotype data. Circles indicate individuals with sequence data.

Figure 2: Plots of genome-wide LOD and NPL scores for chromosomes where LOD scores (MORGAN) or NPL scores ≥2.0. Panel A: Chromosome 1. Panel B: Chromosome 15. Panel C: Chromosome 16. Plots of LOD and NPL scores for all other chromosomes can be found in the supplementary materials. Dominant=autosomal dominant model. Recessive=autosomal recessive model. Narrow=narrow OCD phenotype. Broad=broad OCD phenotype.

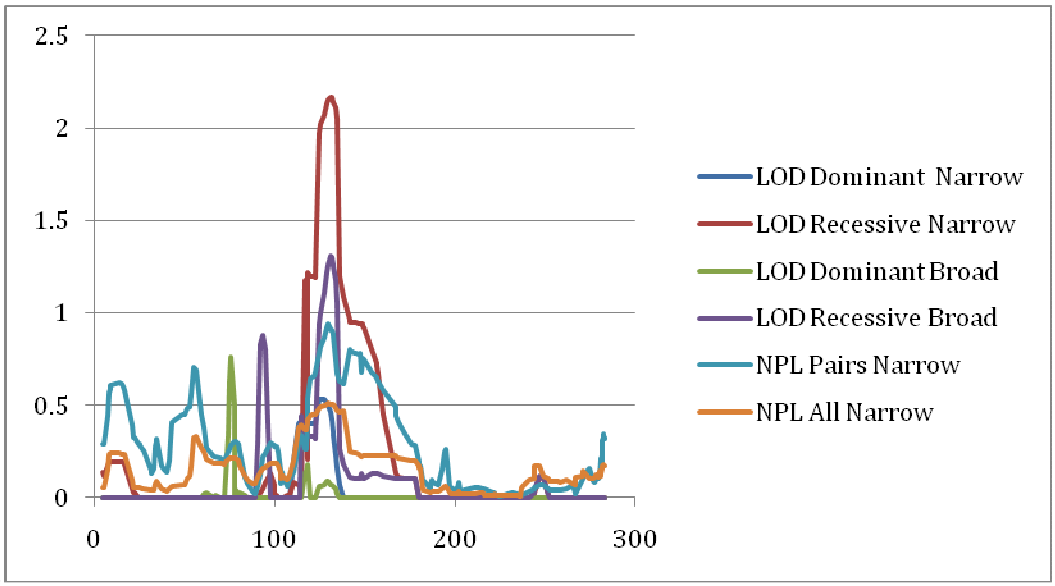
Figure 3: Haplotypes for family 2 in the region of interest on chromosome 15. Black symbols represent the narrow OCD phenotype. Major segregating haplotype is indicated by black line; minor segregating haplotype is indicated by grey line.

Chrom	SNP*	LOD score	# SNPs	Analysis	Model	Phenotype	Location (cM)
1	rs1445225 rs716581 rs977155	2.17	50	MORGAN	Recessive	Narrow	131
2	rs925229	1.81	13	MORGAN	Dominant	Broad	46
3	rs800065	1.73	6	MORGAN	Recessive	Narrow	151
4	rs1981635	1.73	8	MORGAN	Dominant	Broad	24
5	rs998876	1.79	10	MORGAN	Dominant	Broad	199
12	rs4766200	1.83	17	MORGAN	Dominant	Narrow	12
13	rs1886204	1.73	11	MORGAN	Dominant	Broad	22
15	rs965471 rs2059956 rs580839 rs732165	3.13	66	MORGAN	Recessive	Broad	33
16	rs39552 rs922450	2.15	37	MORGAN	Dominant	Narrow	124
17	rs917541	2.04	3	Merlin	Dominant	Narrow	46
22	rs137930 rs6712 rs6520165	1.73	24	MORGAN	Recessive	Broad	81

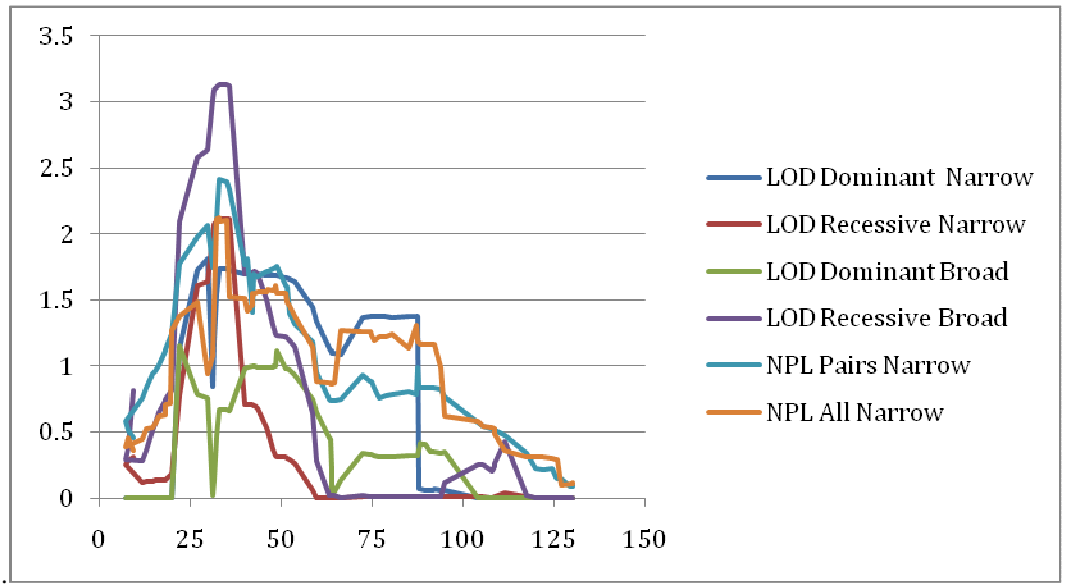


215x149mm (150 x 150 DPI)

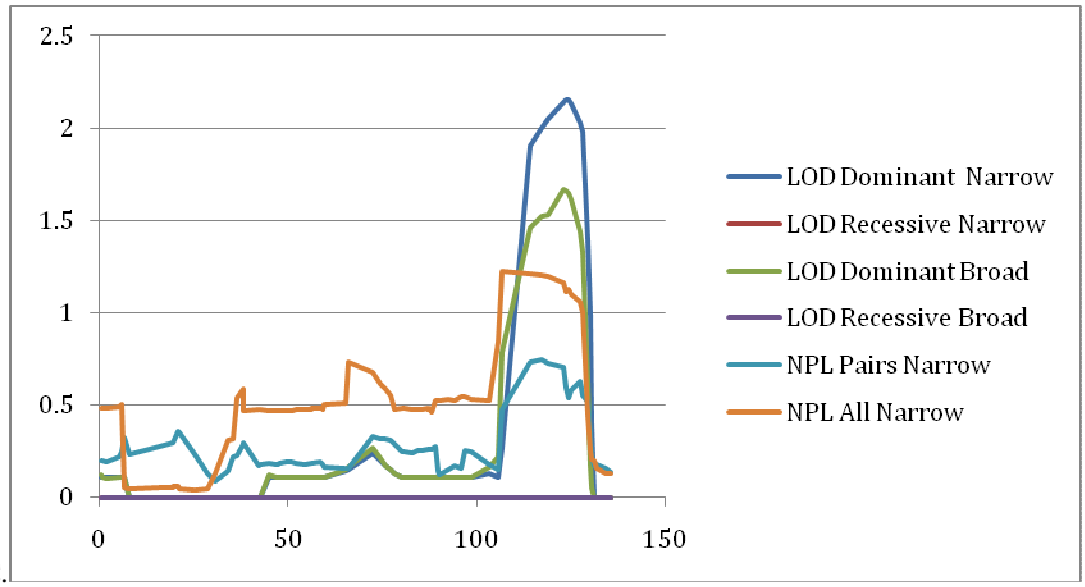
A.


2.5
2
1.5
1
0.5
0
0
100
200
300
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

B.


3.5
3
2.5
2
1.5
1
0.5
0
0
25
50
75
100
125
150
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

C.


2.5
2
1.5
1
0.5
0
0
50
100
150
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

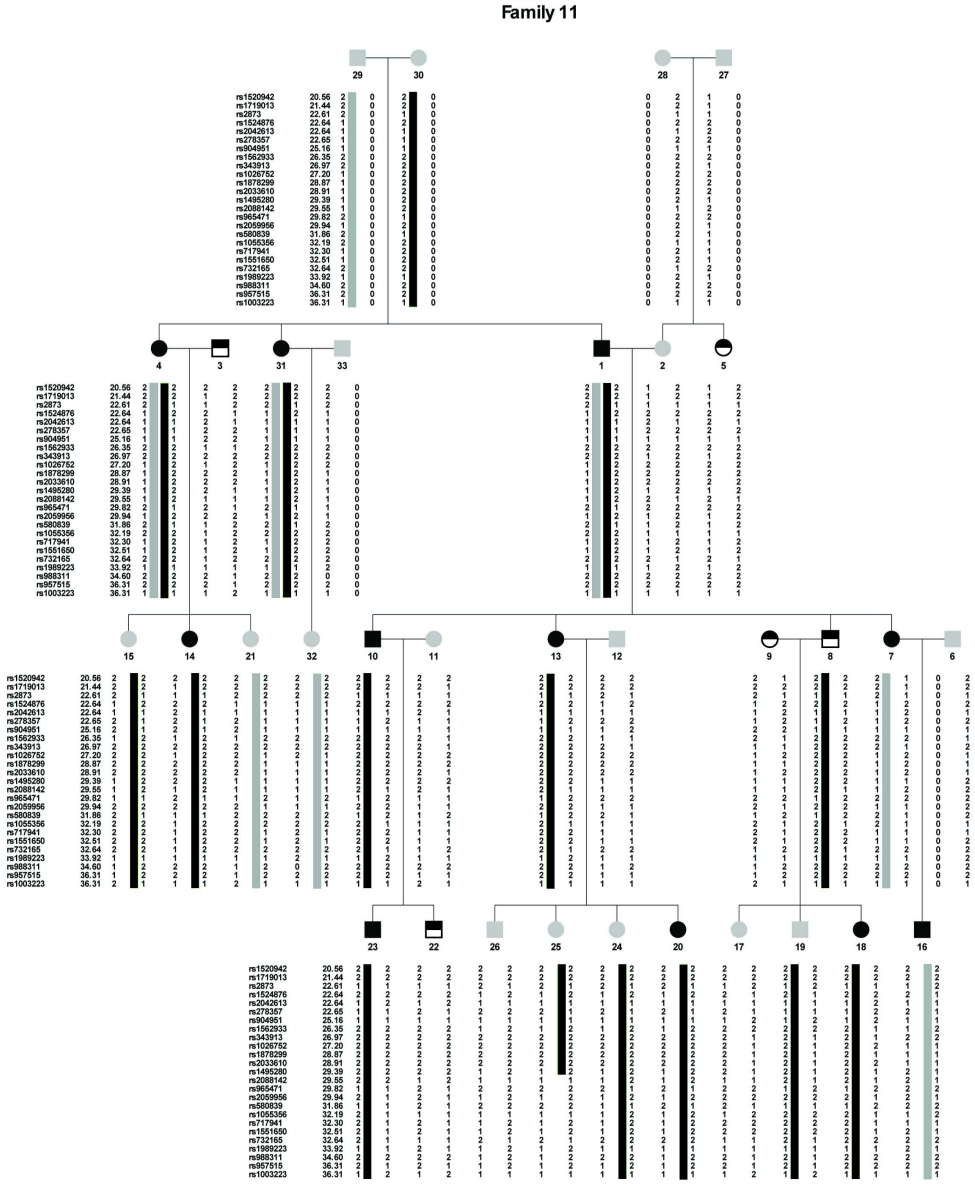


Figure 3
198x239mm (300 x 300 DPI)

Supplemental Data

Chr	Model	Phenotype	Location cM	LOD score	Reference
1q	NPL	OCD Broad	162	1.61	Shugart et al. 2006
3q	NPL	OCD Narrow	209	2.67	Shugart et al. 2006
3q	NPL	OCD Broad	210	2.67	Shugart et al. 2006
6p	NPL	OCD (DSM-III-R)	5	1.40	Hanna et al. 2002
6q	NPL	OCD Broad	187	1.58	Shugart et al. 2006
7p	NPL	OCD Broad	70	1.81	Shugart et al. 2006
9p	Parametric	OCD (DSM-III-R)	9.9	2.25	Hanna et al. 2002
9p	Parametric	OCD Narrow	9.8	2.26	Willour et al. 2004
10p	NPL	Early-Onset OCD	4.4	2.43 (NPL)	Hanna et al. 2007
11p	NPL	OCD Narrow; male probands	5.3	5.08	Wang et al. 2009
14q	NPL	OCD with Hoarding	96	3.7	Samuels et al. 2007*
15q	NPL	OCD Broad	13-35	1.32	Shugart et al. 2006
19q	NPL	OCD (DSM-III-R)	89	1.73	Hanna et al. 2002

Supplemental Table 1: Top linkage results from published OCD linkage studies. NPL=non-parametric linkage. NPL studies reported Kong and Cox LOD scores, except for the Hanna et al study. *=re-analysis of the Shugart et al. sample using only families with two or more members affected with compulsive hoarding.

Abbreviation	Gene name
TJP1	Tight junction protein 1
CHRFAM7A	CHRNA7/FAM7A fusion protein
MTMR15	Myotubularin related protein 15
TRPM1	Transient receptor potential cation channel, subfamily M, member 1
KLF13	Kruppel-like factor 13
OTUD7A	OTU domain containing protein 7A
CHRNA7	Cholinergic receptor, neural nicotinic, alpha polypeptide 7
ARHGAP11A	RHO-GTPase activating protein 11A
SCG5	Secretogranin V
GREM1	Gremlin homolog, cysteinesuperknot family
FMN1	Formin 1
RYR3	Ryanodine receptor 3
CHRM5	Cholinergic receptor, muscarinic 5
LPCAT4	Lysophosphatidylcholineacetyltransferase 4
GOLGA8A	Golgin subfamily A member 8B
NOP10	Nucleolar protein family A member 3
ZNF770	Zinc finger protein 770

Supplemental Table 2: Brain expressed genes in the region of 15q13-15q15

Exon	Position	SNPID	Type	Ref /MutBase	Ref /MutAmino Acid	Ped	Carried on	Mut base CEU	Mut base MEX	Predicted function

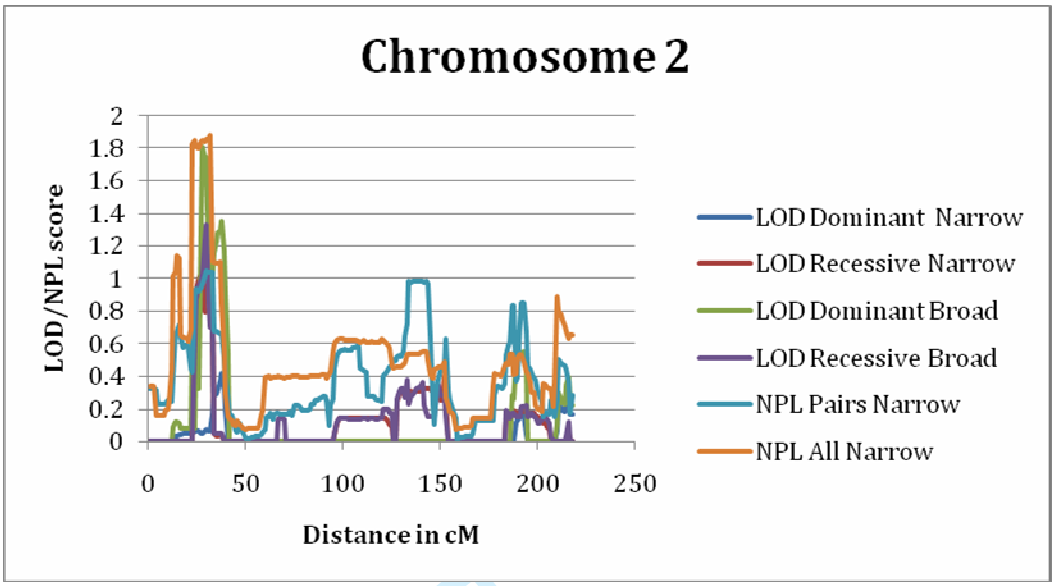
x4	-51	rs41279200	intron	G/A		2	Maj	0.153	N/A	N/A
x7	+122	rs16972012	intron	T/G		2	Maj	0.119	0.026	N/A
x9	-27	rs10519827	intron	G/A		2	Maj	0.148	N/A	N/A
x10	-101	rs2304381	intron	G/A		1,2,3	Min (fam2); Maj (fam 1)	0.181	0.412	N/A
x10	40	rs41279202	syn	A/T	Arg/Arg	2	Maj	0.111	NR	N/A
x13	1	rs674155	syn	C/T	Ser/Ser	1*,2,3	All	0.786	0.784	ESE (1,2,3,4)
x15	-147	rs10431811	intron	G/A		2	Maj	0.509	0.362	N/A
x17	-53	rs12906396	intron	C/G		1*	Maj	0.123	N/A	N/A
x19	27	rs2229116	nonsyn	A/G	Ile/Val	3		0.235	0.362	ESE (1,3)
x20	46	rs2229117	syn	G/C	Leu/Leu	1		0.18	N/A	ESE (1,3,4)
x21	-10	rs35706688	intron	T/C		2	Maj	0.361	N/A	N/A
x21	+150	rs34022625	intron	C/-		2	Maj	N/A	N/A	N/A
x27	-77	rs2291734	intron	G/A		2+,3	Min (fam2)	0.198	0.328	N/A
x27	99	Novel	syn	C/T	Ile/Ile	2+	Min	N/A	N/A	N/A
x30	-130	rs2291736	intron	C/T		1,2	Maj	0.216	0.198	N/A
x33	-11	rs2279662	intron	T/C		1,2	Maj	0.142	0.25	conserved
x34	-72	rs2279663	intron	C/T		1,2	Maj	0.137	0.233	N/A
x35	-127	rs16957279	intron	A/G		1,2	Maj	0.363	0.414	N/A
x35	299	rs4780144	nonsyn	C/T	Arg/Cys	1,2,3	All	0.969	0.922	damaging (orthologues)/ tolerated (homologues) (SNP Effect predicts effect on solvent accessibility); ESE (3)
x38	-179	rs34219516	intron	T/-		1,2	All	N/A	N/A	N/A
x40	-42	rs3736531	intron	A/G		2	Maj	0.159	0.155	N/A
x40	+97	rs1390161	intron	G/A		1,3	All	0.389	N/A	N/A
x42	+39	rs2280419	intron	G/A		2	Maj	0.288	0.353	N/A
x42	+77	rs2280418	intron	G/A		2	Maj	0.165	0.138	N/A
x44	-3	rs2293028	intron	C/T		1,3	Maj	0.279	0.391	Frameshift coding; splice site (5)
x44	98	rs2293027	syn	G/A	Gly/Gly	2#		0.168	0.06	ESE (1,3)
x45	9	rs6495228	nonsyn	G/A	Gly/Glu	1*,2,3	All	0.85	0.879	ESE (1,2,3,4)

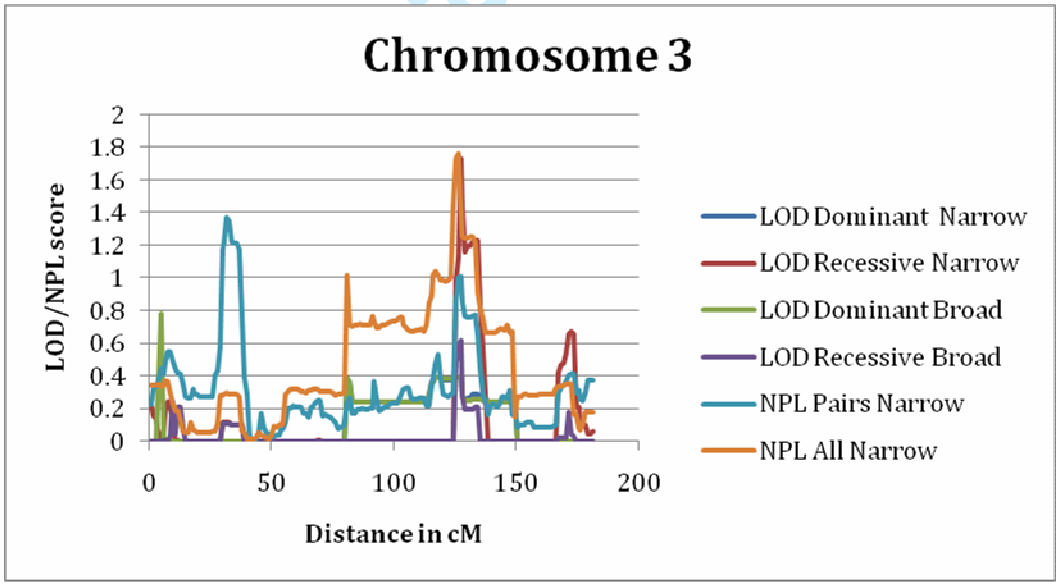
x47	-144	rs28460434	intron	T/G		2	All	0.097	N/A	N/A
x49	-181	rs2288607	intron	G/A		2#,3	All	0.273	N/A	N/A
x49	+125	rs12909708	intron	T/C		2	All	0.095	0.017	N/A
x50	+74	rs4780171	intron	T/C		1,3	Maj	0.496	0.707	N/A
x51	+71	rs12911737	intron	G/A		2	All	0.032	N/A	N/A
x54	+35	rs1565936	intron	G/A		2	All	0.102	0.034	N/A
x62	+150	rs1565936	intron	C/G		2	All	0.102	0.034	N/A
x62	+173	rs2278315	intron	G/A		2	All	0.062	N/A	N/A
x63	-78	rs1429754	intron	T/C		1,2,3	All	N/A	N/A	N/A
x69	-22	rs35048413	intron	G/T		1,2,3	All (fam 1,3) Min (fam 2)	0.389	N/A	N/A
x69	+39	rs2115747	intron	C/T		1,2,3	All (fam 1,3) Min (fam 2)	0.416	0.25	N/A
x70	+52	rs2879436	intron	G/A		1,2,3	All (fam 1,3) Min (fam 2)	0.254	NA	conserved
x78	54	rs2288613	syn	C/T	Phe/Phe	3		0.556	N/A	ESE (1,2,3)
x79	66	rs2288614	syn	A/G	Ala/Ala	1,2#,3		0.644	N/A	ESE (1,2,3,4)
x81	+8	rs12914825	intron	C/T		1,2	Maj	0.075	0.009	Frameshift coding; splice site (5)
x84	+41/+42	rs33969322/ rs4041449	intron	-/AG		1,2,3	All	N/A	N/A	N/A
x87	+183/184	rs6495270	intron	C/G		1,2,3	All	N/A	N/A	N/A
x89	-109	Novel	intron	G/C		1	Maj	N/A	N/A	N/A
x89	866	Novel	nonsyn	G/C	Val/Leu	1		N/A	N/A	tolerated
x89	+121	rs73372076	intron	G/C		1	Maj	N/A	N/A	N/A
x90	-159	rs11639042	intron	C/T		1	Maj	0.032	N/A	N/A
x90	-154	rs55726623	intron	A/G		1	Maj	N/A	N/A	N/A
x90	-151	rs4780184	intron	G/A		1*,2+,3	Min	0.361	N/A	N/A
x94	-13	rs41279232	intron	C/T		1	Maj	0.042	N/A	N/A
x95	+23	rs56258848	intron	A/G		1	Maj	N/A	N/A	N/A
x95	+84	rs79847071	intron	G/C		1	Maj	0.028	N/A	N/A
x100	-91	Novel	intron	C/A		1	Maj	N/A	N/A	N/A
x100	-61	rs80248888	intron	G/A		1	Maj	0.042	N/A	N/A
x100	126	rs79305633	syn	C/T	Phe/Phe	1		0.042	N/A	N/A

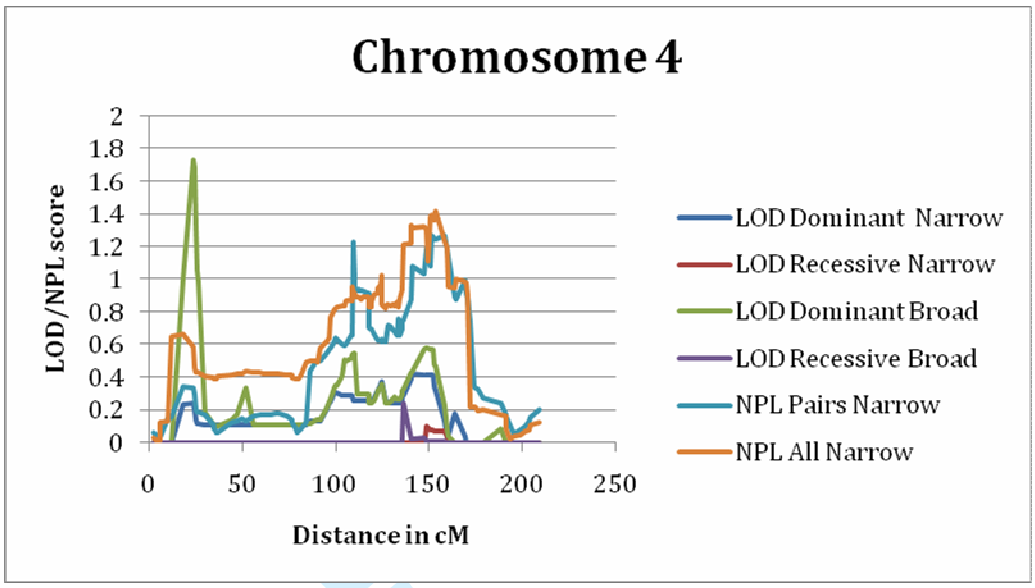
x102	-148	rs71891566	intron	**/-		1	Maj	N/A	N/A	N/A
x104	512-515	rs5811800/ rs66707641	intron	ACTC/-		1,2,3	All	N/A	N/A	N/A

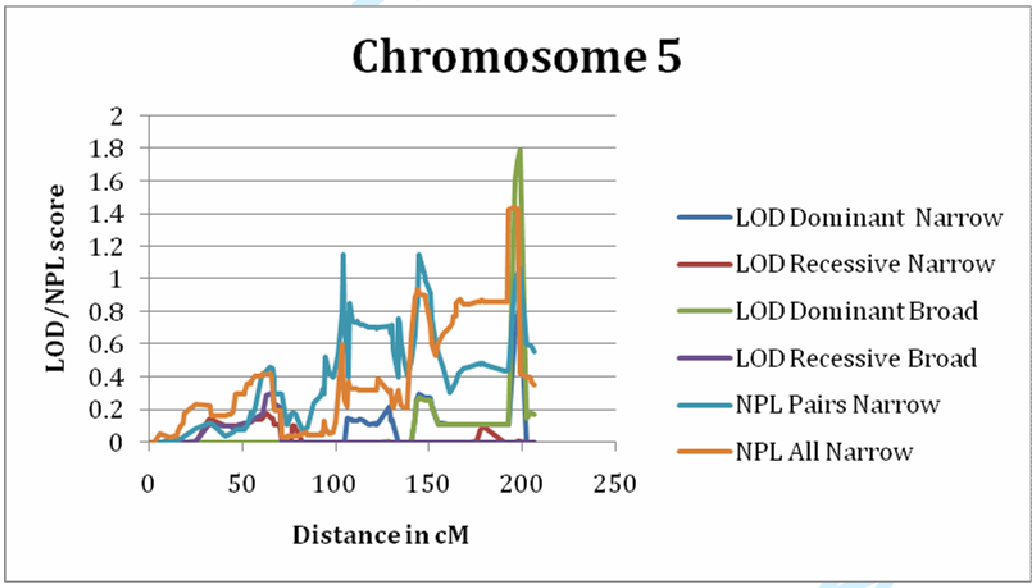
Supplemental Table 3. Sequence variants in the RYR3 gene carried by OCD- affected individuals in at least one of the three CVCR OCD pedigrees. Grey rows indicate identified coding variants that are carried on the haplotype of interest in at least one family. Maj=carried on the major haplotype for that family. Min=carried on the minor haplotype for that family. All=carried by all OCD-affected individuals in that family.Unless otherwise indicated, all variants occur only in family members who carry the major haplotype co-segregating with OCD within that family. For family 1, the major haplotype is considered to be the maternal haplotype. *indicates that the variant is carried by an unaffected sibling. $^{+}$indicates that the variant is carried by the founder and the OCD-affected individuals who do not carry the major co-segregating haplotype. #carried by all OCD-affected individuals in family 2. N/A=information not available. CEU= allele frequency in the CEPH HapMap sample. MEX= allele frequency in the Mexican HapMap sample. ESE= exonic splice enhancer 1=ESE Finder, 2= RESCUE_ESE, 3=ESRS Search, 4= PESX, 5=Ensembl

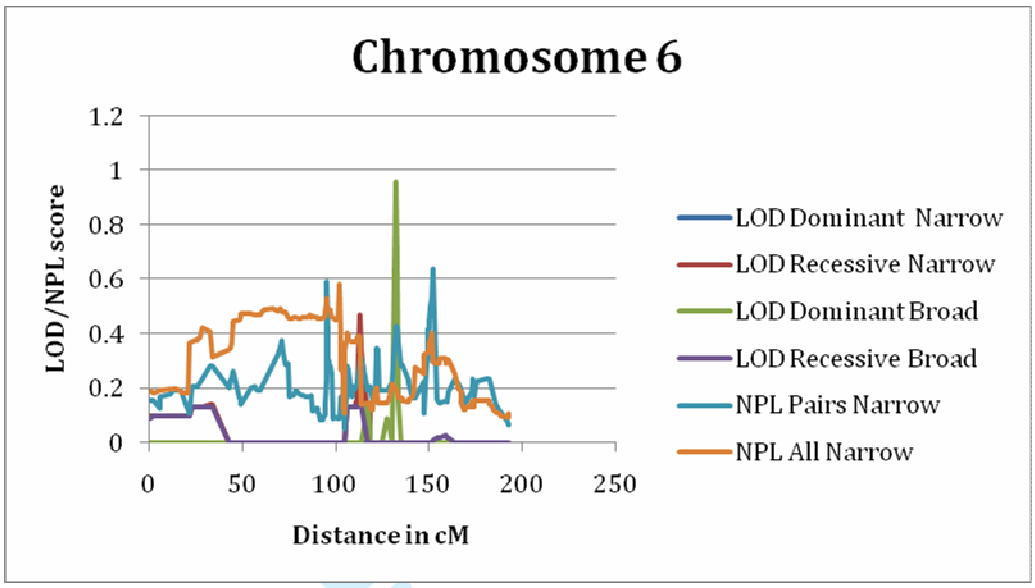
**TGTATGGCACTACTGAGTGAATGACTAATAGCCAA

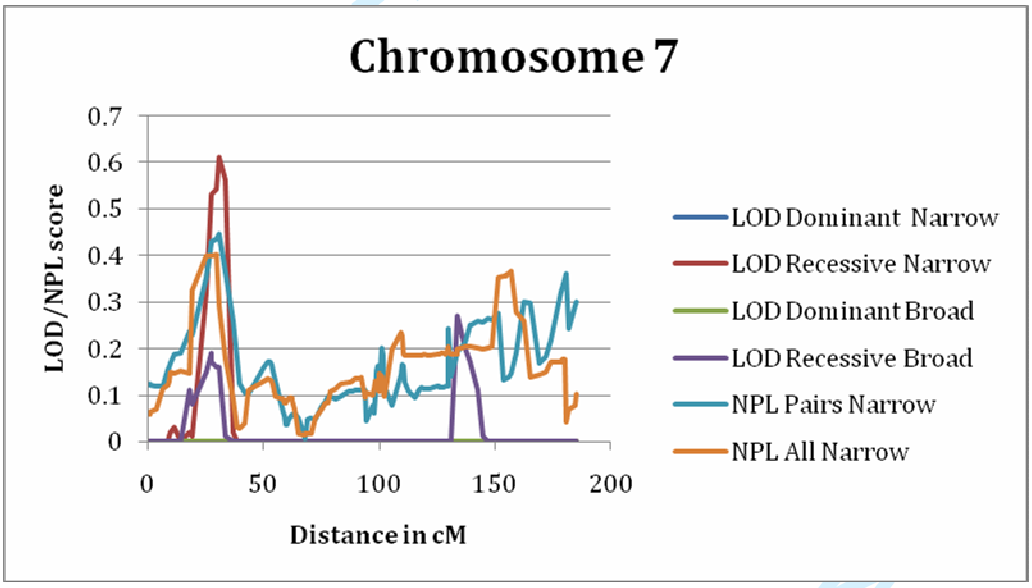

Chromosome 2
LOD/NPL score
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
200
250
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

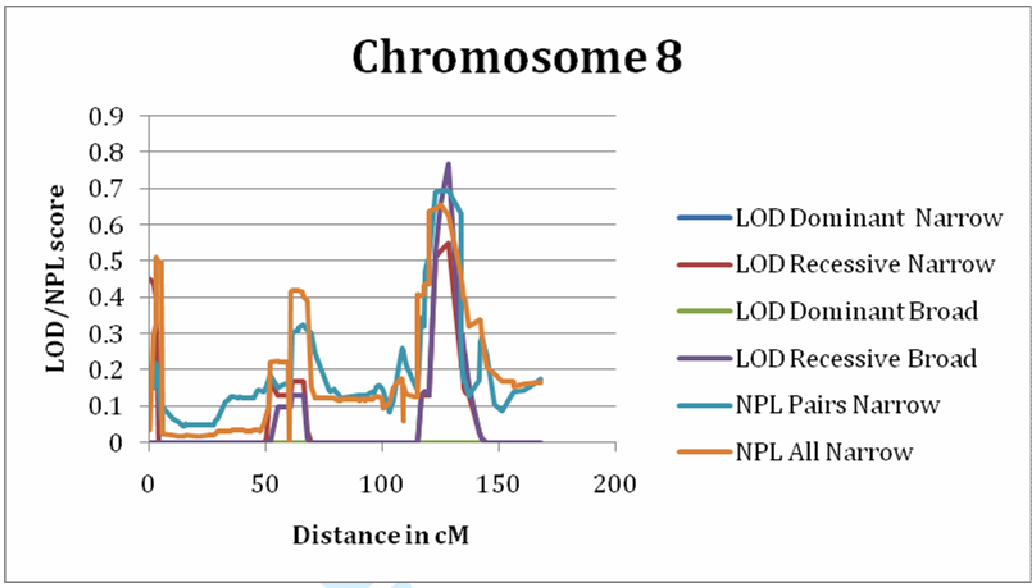

Chromosome 3
LOD/NPL score
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
200
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

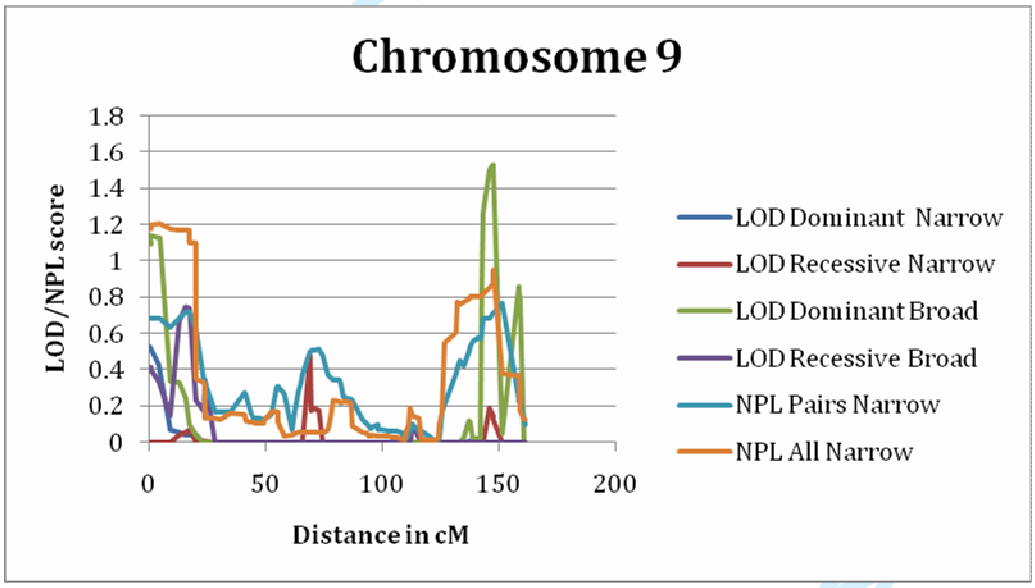

Chromosome 4
LOD/NPL score
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
200
250
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

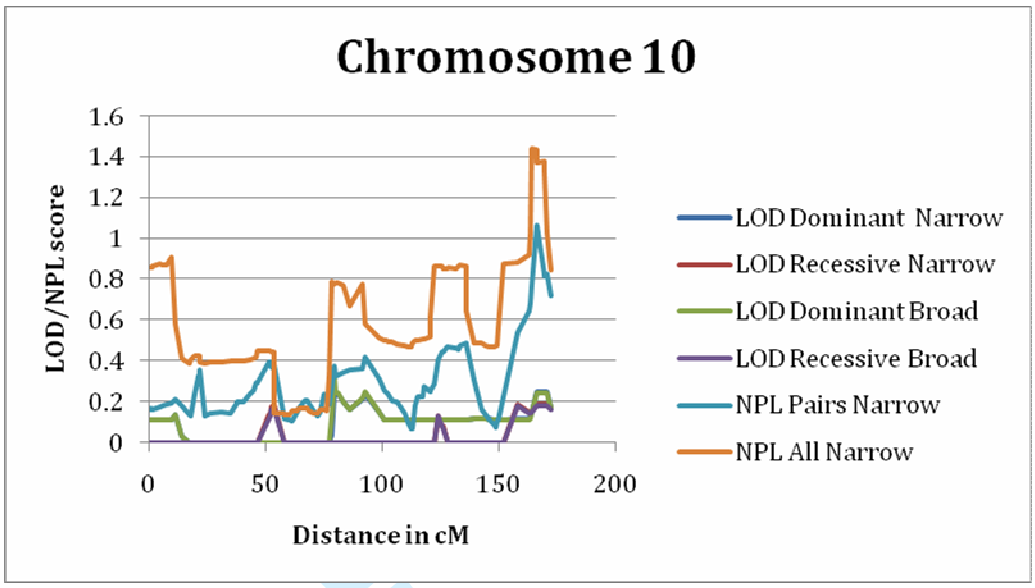

Chromosome 5
LOD/NPL score
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
200
250
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

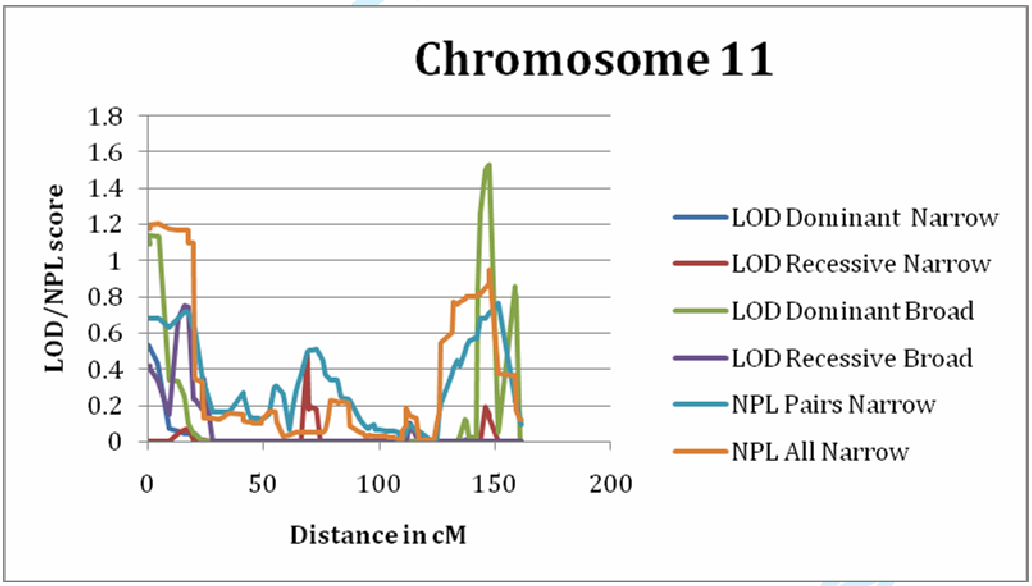

Chromosome 6
LOD/NPL score
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
200
250
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

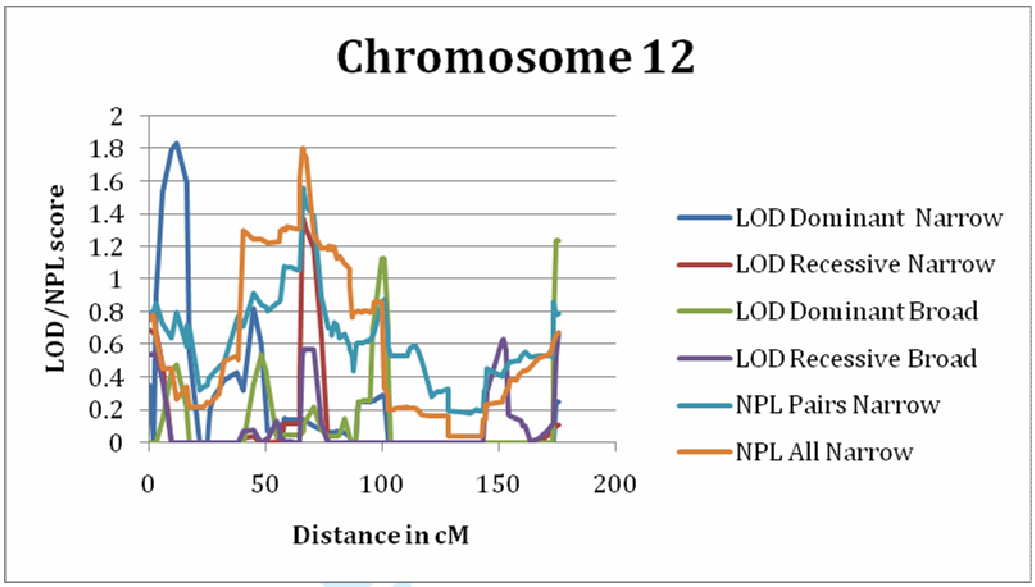

Chromosome 7
LOD/NPL score
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
50
100
150
200
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

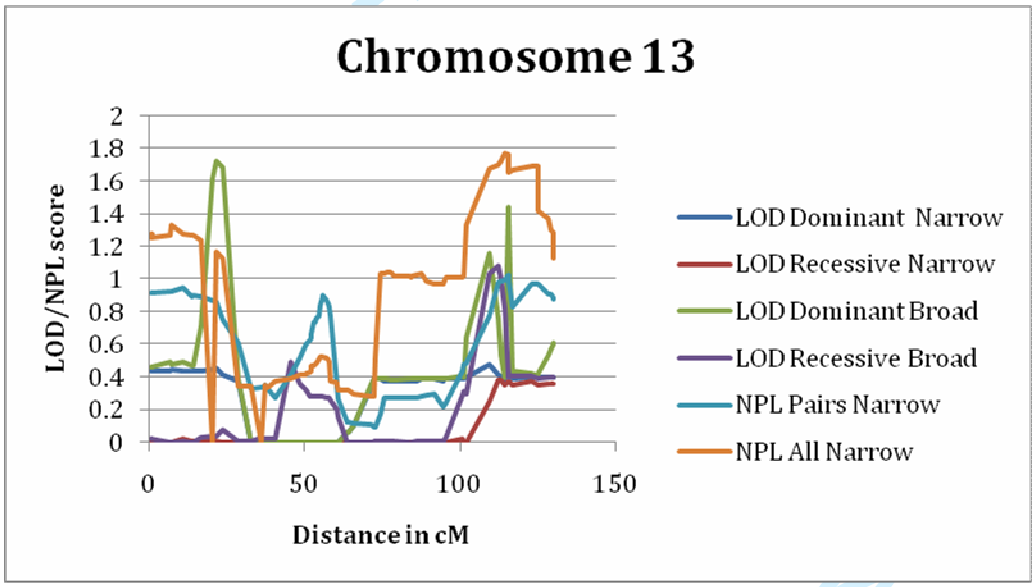

Chromosome 8
LOD/NPL score
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
50
100
150
200
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

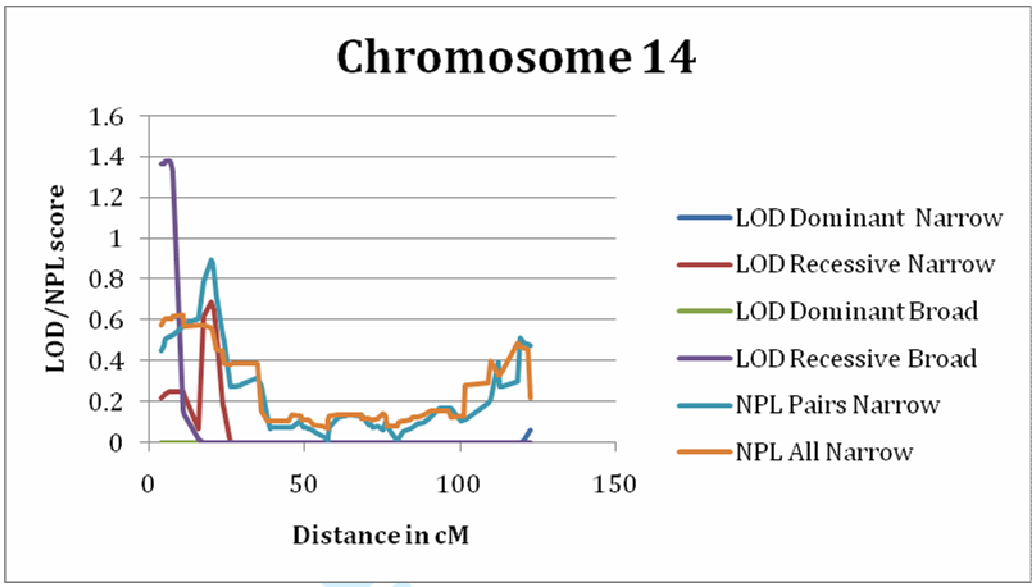

Chromosome 9
LOD/NPL score
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
200
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

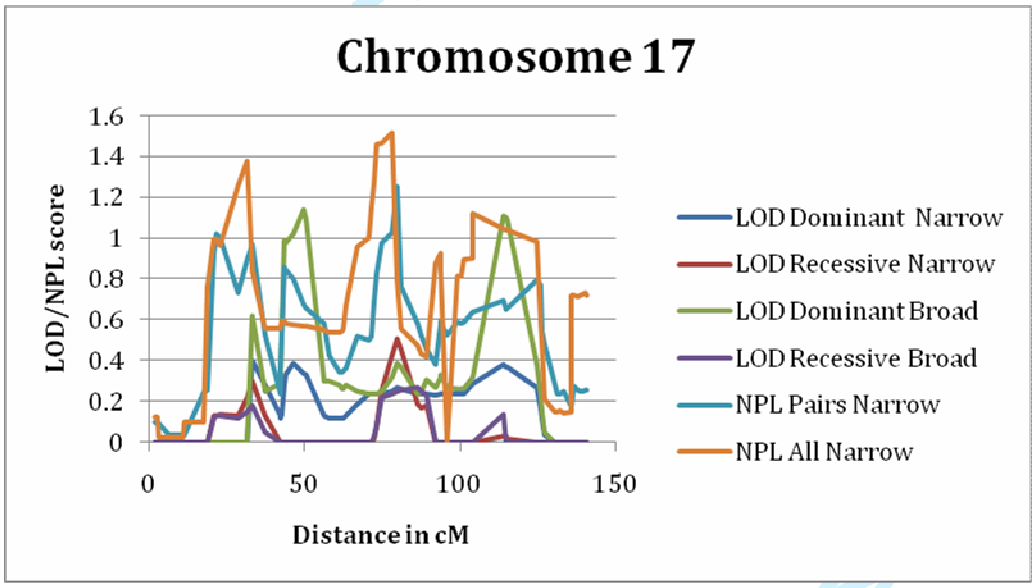

Chromosome 10
LOD/NPL score
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
200
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

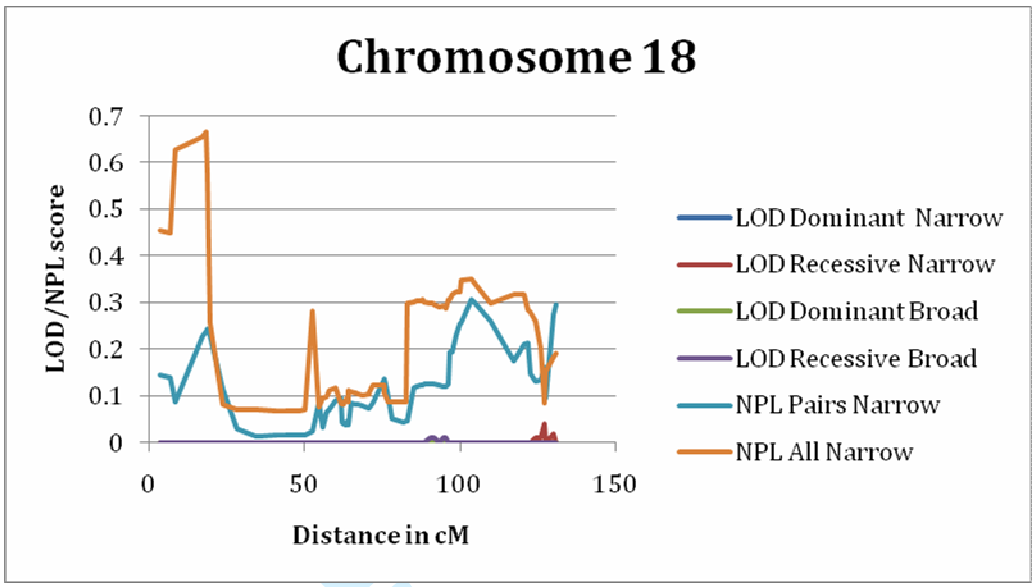

Chromosome 11
LOD/NPL score
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
200
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

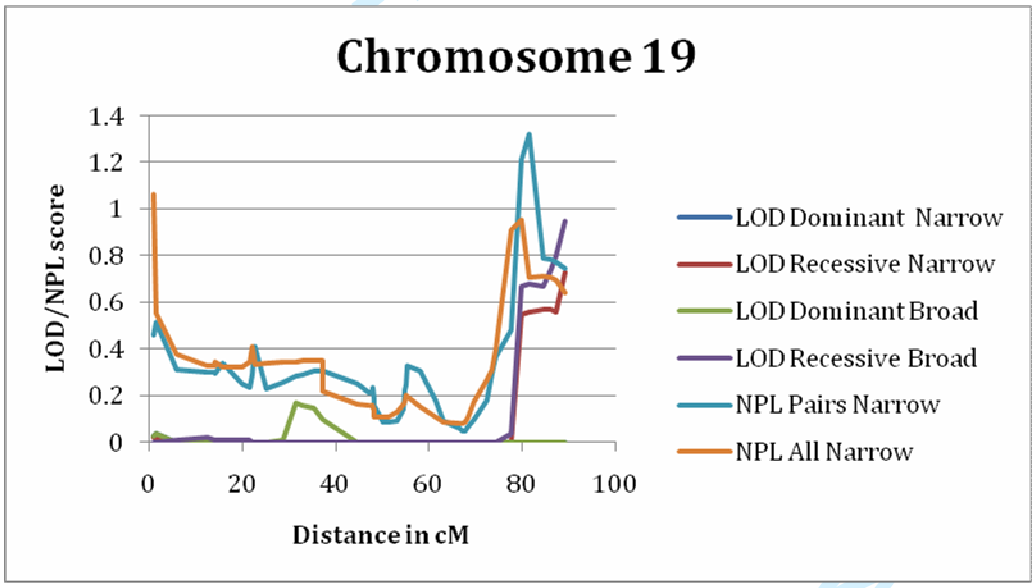

Chromosome 12
LOD/NPL score
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
200
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

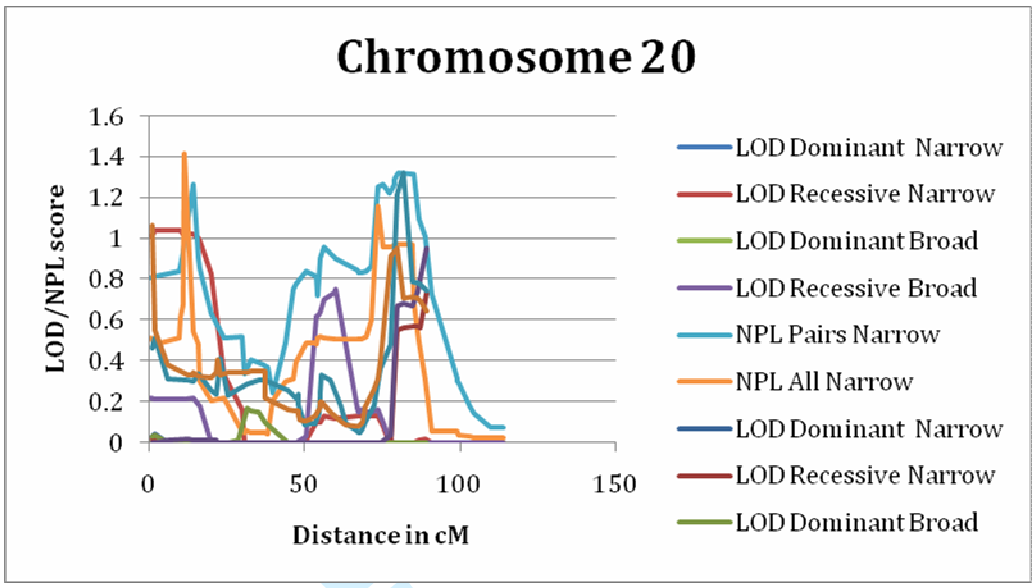

Chromosome 13
LOD/NPL score
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

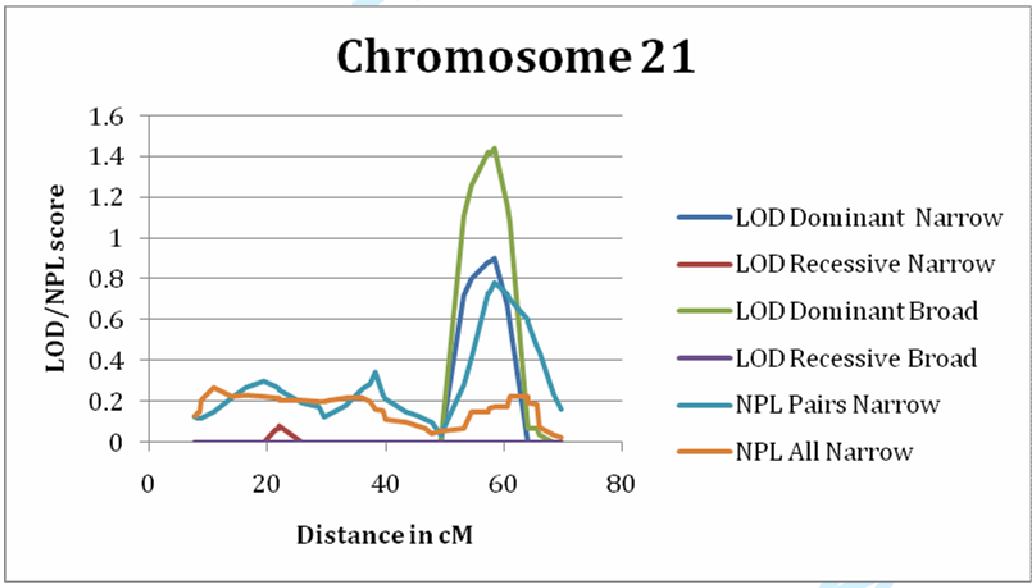

Chromosome 14
LOD/NPL score
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

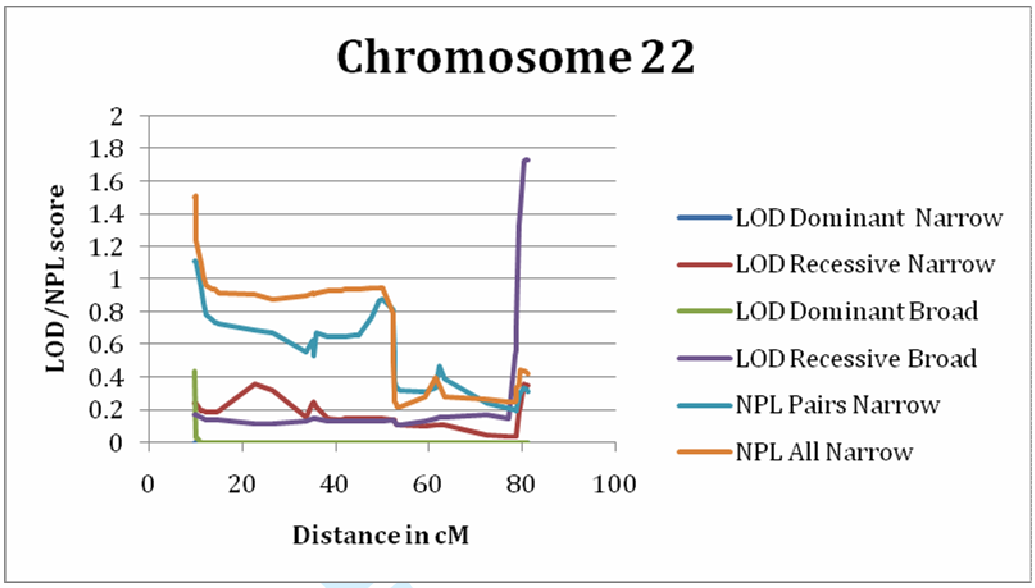

Chromosome 17
LOD/NPL score
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow


Chromosome 18
LOD/NPL score
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
50
100
150
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow

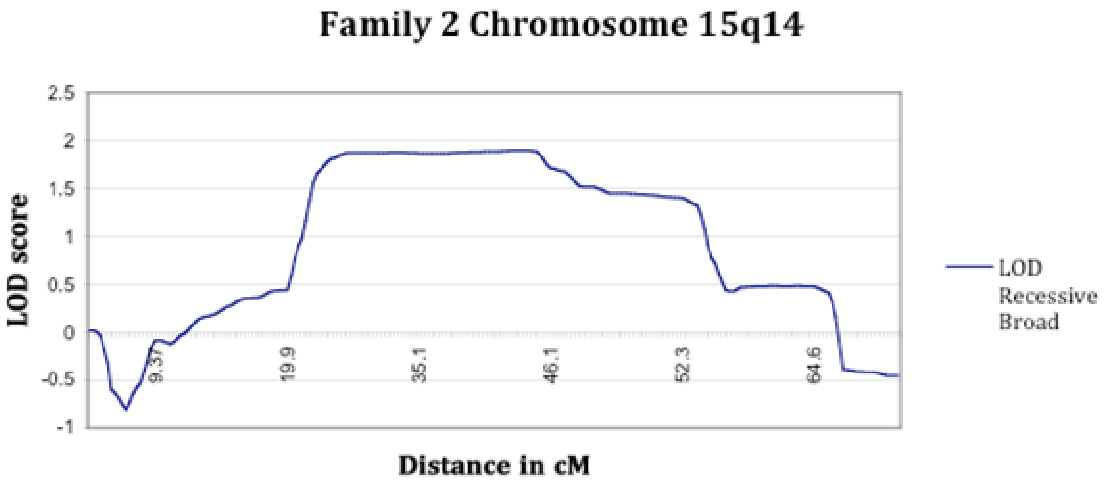

Chromosome 19
LOD/NPL score
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
20
40
60
80
100
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow


Chromosome 20
LOD/NPL score
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad


Chromosome 21
LOD/NPL score
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
20
40
60
80
Distance in cM
LOD Dominant Narrow
LOD Recessive Narrow
LOD Dominant Broad
LOD Recessive Broad
NPL Pairs Narrow
NPL All Narrow



Supplemental Figure 1: Plots of genome-wide LOD and NPL scores for all chromosomes except chromosome 1, 15, and 16, which are found in the main text. Dominant=autosomal dominant model. Recessive=autosomal recessive model. Narrow=narrow OCD phenotype. Broad=broad OCD phenotype.



Supplemental Figure 2: Plots of LOD scores for family 2 in the region of strongest linkage on chromosome 15q14. Recessive=autosomal recessive model. Narrow=narrow OCD Broad=broad OCD phenotype.