(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 March 2006 (02.03.2006)

PCT

(10) International Publication Number
WO 2006/023719 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2005/029541

(22) International Filing Date: 19 August 2005 (19.08.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/602,853 20 August 2004 (20.08.2004) US

(71) Applicant *(for all designated States except US)*: **ENH RESEARCH INSTITUTE** [US/US]; 1001 University Place, Evanston, IL 60202 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **DUAN, Jubao** [CN/US]; 8816 Crawford, Skokie, IL 60076 (US). **CROWE, Raymond** [US/US]; Department of Psychiatry, Room 1-292, MEB Bldg., University of Iowa, Iowa City, IA 52242-1000 (US). **MARTINEZ, Maria** [US/US]; ENH Research Institute, 1001 University Place, Evanston, IL 60201 (US). **MOWRY, Bryan** [AU/AU]; Queensland Centre for Schizophrenia Research, The Park, Centre for Mental Health, Wacol, Queensland QLD 4076 (AU). **LEVINSON, Douglas** [US/US]; University of Pennsylvania School of Medicine, 3535 Market Street, Room 4003, Philadelphia, PA 19104-3309 (US). **SANDERS, Alan** [US/US]; ENH Research Institute, 1001 University Place, Evanston, IL 60201 (US). **SILVERMAN, Jeremy** [US/US]; Department of Psychiatry, Box 1230, Mt. Sinai School of Medicine, One Gustave L. Levy Place, Bronx, NY 10468 (US). **GEJMAN, Pablo, V.** [US/US]; ENH Research Institute, 1001 University Place, Evanston, IL 60201 (US).

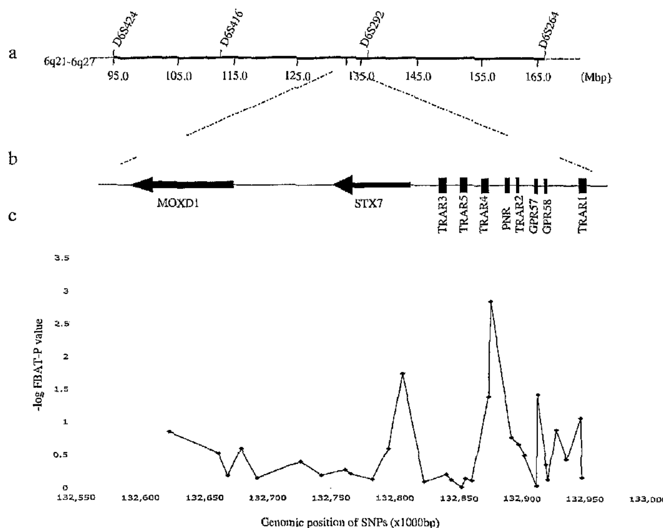
(74) Agent: **CONTRERA, Joseph, G.**; Jacobson Holman PLLC, 400 Seventh Street, N.W., Suite 600, Washington, DC 20004 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH,

[Continued on next page]

(54) Title: SCHIZOPHRENIA, SCHIZOAFFECTIVE DISORDER AND BIPOLAR DISORDER SUSCEPTIBILITY GENE MUTATION AND APPLICATIONS TO THEIR DIAGNOSIS AND TREATMENT



(57) Abstract: The present invention provides the identification of a number of SNPs that are associated schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders which were found to be strongly linked to individuals with the disease. The invention provides SNP locations on human chromosome 6, as well as methods of making PCR primers and assays for detecting the SNPs in tested individuals.

WO 2006/023719 A2

GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

SCHIZOPHRENIA, SCHIZOAFFECTIVE DISORDER AND BIPOLAR DISORDER SUSCEPTIBILITY GENE MUTATION AND APPLICATIONS TO THEIR DIAGNOSIS AND TREATMENT

BACKGROUND OF THE INVENTION

1. Field of Invention

[0001] The present invention relates to identifying a gene that codes for a receptor as being associated with schizophrenia and schizoaffective disorder known as TRAR4, and its use in the diagnosis and screening of therapeutic agents useful in the treatment of the disease.

2. Description of Prior Art

[0002] Schizophrenia (along with the closely related schizoaffective disorder) is a frequently chronic and devastating brain disorder that affects about 1% of the population worldwide (Jablensky et al. 1992). Typically it presents in adolescence or young adulthood and is characterized by major disruptions of thinking (delusions, disorganization), perception (hallucinations), mood, and behavior (Gottesman and Shields 1982). Schizophrenia and schizoaffective disorder are strongly familial, with a heritability of about 80%, but its etiology is hypothesized to involve both genetic and environmental factors (Sanders and Gejman 2001). Recently, encouraging evidence for several genes potentially involved in the etiology of schizophrenia has been reported,

namely dysbindin (DTNBP1) (Straub et al. 2002; Schwab et al. 2003), neuregulin 1 (NRG1) (Stefansson et al. 2002; Stefansson et al. 2003; Williams et al. 2003), proline dehydrogenase (oxidase) 1 (PRODH) (Jacquet et al. 2002; Liu et al. 2002), catechol-O-methyltransferase (COMT) (Li et al. 2000; Egan et al. 2001; Shifman et al. 2002), regulator of G-protein signaling 4 (RGS4) (Chowdari et al. 2002; Morris et al. 2004; Williams et al. 2004), D-amino acid oxidase activator (DAOA, previously called G72) (Chumakov et al. 2002; Schumacher et al. 2004), and D-amino-acid oxidase (DAO) (Chumakov et al. 2002; Schumacher et al. 2004). Most of the aforementioned genes are positional candidates with likely involvement with dopaminergic or NMDA brain mechanisms.

[0003] We have previously reported linkage of schizophrenia to chromosome 6q13-q26 (SCZD5; MIM 603175) (Cao et al. 1997), which since then accumulated wide, although not universal, support from converging evidence from a number of studies. In the first report of linkage to 6q (Cao et al. 1997) support for linkage was observed from D6S301 (located at 111 cM) to D6S305 (located at 170 cM), which was confirmed in subsequent studies (Kaufmann et al. 1998; Martinez et al. 1999; Levinson et al. 2000). All the families used in the studies underlying this patent application are from three datasets, which we call NIMH-IRP (National Institute of Mental Health - Intramural Research Program), NIMH-GI (NIMH-Genetics Initiative), and AU/US (Australia / United States), that previously have

been shown to yield evidence for linkage to chromosome 6q13-q26 (SCZD5; MIM 603175) (Cao et al. 1997; Martinez et al. 1999; Levinson et al. 2000). Analysis of 12 microsatellite markers on a 50 cM 6q region were reported in (Martinez et al. 1999); nonparametric ASP methods yielded P-values of 0.00018, 0.00095 and 0.013 for the NIMH-IRP, NIMH GI, and AU/US datasets, respectively. In a Palestinian/Israeli pedigree sample (Lerer et al. 2003), the linkage peak with an NPL of 4.61 was at D6S292 (~137 cM), and ~131 to ~144 cM contained the 1 NPL decrease portion of the linkage peak. Furthermore, another group found evidence for linkage in a Swedish pedigree from ~170 cM to ~180 cM (Lindholm et al. 2001). Whether these different results are better explained by the presence of more than one schizophrenia susceptibility gene in 6q or reflect typical peak variability in complex disorders (Hauser and Boehnke 1997; Hsueh et al. 2001), is currently unknown. Recently, bipolar disorder was reported to map to 6q with one study yielding a maximum LOD of 2.2 at 113 cM near D6S1021 (Dick et al. 2003b; Dick et al. 2003a) and another study reporting a maximum LOD of 3.56 at ~124 to ~126 cM near D6S1639 (Middleton et al. 2004), opening the prospect that a common gene for schizophrenia and bipolar disorder (and the intermediate, schizoaffective disorder) may be located in 6q.

[0004] We were interested in a MOXD1-STX7-TRARs gene cluster at 6q23.2 (132.8 cM) with prime candidates for schizophrenia (figure 1), harboring MOXD1

(monooxygenase, dopamine-β-hydroxylase-like 1) (Chambers et al. 1998), STX7 (syntaxin 7) (Wang et al. 1997), and all known human trace amine receptors (TRARs), namely, TRAR1, TRAR3, TRAR4, TRAR5, PNR (putative neurotransmitter receptor) (Zeng et al. 1998; Borowsky et al. 2001; Bunzow et al. 2001; Lee et al. 2001), and three TRAR pseudogenes (TRAR2 and G-protein coupled receptors, GPR57 and GPR58) (Liu et al. 1998; Borowsky et al. 2001; Bunzow et al. 2001; Lee et al. 2001).

[0005] Trace amines (TAs) are endogenous amine compounds chemically similar to classical biogenic amines such as dopamine, norepinephrine, serotonin, and histamine. Abnormalities involving the classical biogenic amines are the basis for a variety of biological hypotheses for a wide variety of disorders, including dystonias, Parkinson's disease, schizophrenia, drug addiction, and mood disorders. In mammals, TAs are present at low levels with no apparent dedicated synapses, but blockade of amine degradation leads to significant accumulations of trace amines suggesting high synthesis and turnover, as recently reviewed (Premont et al. 2001). TAs in mammals include tyramine (TYR), tryptamine, β-phenylethylamine (β-PEA), and octopamine (OCT) (Branchek and Blackburn 2003), and are all synthesized from amino acid precursors by the aromatic amino acid decarboxylase.

[0006] TAs were thought to be "false transmitters," which displace classical biogenic

amines from their storage and act on transporters in a similar fashion to the amphetamine (Parker and Cubeddu 1986), but the identification of brain receptors specific to TAs indicates that they also have effects of their own (Borowsky et al. 2001). This might explain the fact that although TYR, β-PEA, OCT, and amphetamine require the integrity of vesicular stores of dopamine if displacement of dopamine were their only mechanism of action, they (except OCT) are still active when dopamine is depleted (Baud et al. 1985). TRARs bind amphetamine, MDMA (3,4-methylenedioxymethamphetamine; "ecstasy"), and LSD (D-lysergic acid diethylamide) with high affinity. This suggests a direct link of TRARs with mechanisms of psychosis because the administration of amphetamine can induce a schizophrenia-like psychosis (Connell 1958; Snyder et al. 1967; Angrist et al. 1974; Laruelle and Abi-Dargham 1999) and psychedelic experiences induced by LSD can have remarkable similarity to schizophrenia (Vardy and Kay 1983; Gouzoulis et al. 1994). Furthermore, LSD can induce habituation deficits (the normal decrease in response magnitude to repeated stimuli over time), which are similar to those exhibited by schizophrenic patients (Geyer and Braff 1987; Braff et al. 1992).

[0007] MOXD1 is a homologue of dopamine-β-hydroxylase potentially involved with the biosynthesis of norepinephrine from dopamine (Chambers et al. 1998). Syntaxin 7 (STX7) is a critical component of the synaptic protein complex SNARE (receptor for soluble N-ethylmaleimide-

sensitive factor attachment proteins), which is involved in NMDA (N-methyl D-aspartate) and dopaminergic receptor function (Pei et al. 2004) and whose dysfunction has been suggested in schizophrenia (Honer et al. 2002). Specifically, syntaxins mediate vesicle fusion in vesicular transport processes (Teng et al. 2001). We investigated DNA polymorphisms in the MOXD1-STX7-TRARs cluster with schizophrenia using family-based association methods and present preliminary evidence of association between TRAR4 and schizophrenia.

[0008] There are two reports suggesting linkage of the same 6q chromosomal area to bipolar disorder raise the possibility that *TRAR4* might be involved in the pathophysiology of both schizophrenia and bipolar disorder (and the intermediate, schizoaffective disorder), and there is a precedent for a gene potentially involved in both disorders (Chumakov et al. 2002; Hattori et al. 2003; Chen et al. 2004; Schumacher et al. 2004).

SUMMARY OF THE INVENTION

[0009] Molecular genetic studies of schizophrenia have found several replicated linkages to various chromosomal regions (Owen et al. 2004), and association studies have recently pointed to several genes at some of those linkage regions with independent confirmations, including *NRG1* at 8p21-p12 (Stefansson et al. 2002; Stefansson et al. 2003; Williams et al. 2003), *DTNBP1* at 6p22.3 (Straub et

al. 2002; Schwab et al. 2003), *COMT* at 22q11.21 (Li et al. 2000; Egan et al. 2001; Shifman et al. 2002), *RGS4* at 1q23.3 (Chowdari et al. 2002; Morris et al. 2004; Williams et al. 2004) and *DAOA* at 13q33.2 (Chumakov et al. 2002; Schumacher et al. 2004). Although non-replications have been reported, it would be extremely unlikely that all the aforementioned results will end as false positives. We now propose that *TRAR4* is also a susceptibility gene for schizophrenia.

[0010] The identification of *TRAR4* as a susceptibility gene for schizophrenia, which is consistent with human and animal models of toxic psychosis and in agreement with the expression pattern of TRAR4 (expressed in frontal cortex, amygdala, and hippocampus), appears to substantiate the dopaminergic hypothesis of schizophrenia, but the exact mechanisms of disease mediated by *TRAR4* remain to be elucidated.

[0011] Several linkage studies across multiple population groups provide convergent support for chromosome 6q13-q26 as containing a susceptibility locus for schizophrenia, and more recently for bipolar disorder. We genotyped 192 European and African American (AA) schizophrenia pedigrees (these pedigrees -in particular the NIMH-IRP sample- have many cases with schizoaffective disorder, major recurrent depression and other bipolar spectrum disorder conditions, though our study only counted schizophrenia and schizoaffective disorder as

"affected") from samples that previously showed linkage evidence to 6q13-q26, focusing on the MOXD1-STX7-TRARs gene cluster at 6q23.2, which contains a number of prime candidate genes for schizophrenia. Thirty-one screening SNPs were selected, providing a minimum coverage of at least one SNP per 20 kb. The association observed with rs4305745 (P = 0.0014) within the TRAR4 (trace amine receptor 4) gene remained significant after correction for multiple testing. Evidence for association was proportionally stronger in the smaller AA sub-sample.

[0012] Through database searching and sequencing genomic DNA in a thirty probands sub-sample, we obtained a high-density map of twenty-three SNPs spanning 21.6 kb of this gene. Single SNP and also haplotype analyses revealed that rs4305745, and/or two other polymorphisms in perfect linkage disequilibrium (LD) with rs4305745, appear to be the most likely variants underlying the association of the TRAR4 region with schizophrenia. Comparative genomic analyses further revealed that rs4305745, and/or the associated polymorphisms in complete LD, could possibly affect gene expression. Moreover, RT-PCR studies of various human tissues including brain confirm that TRAR4 is preferentially expressed in those brain regions implicated in the pathophysiology of schizophrenia.

[0013] We now show that *TRAR4*, a gene that belongs to the trace amine receptor family contributes to susceptibility to schizophrenia in three data sets

with evidence of genetic linkage to 6q. Furthermore, the *TRARs* gene cluster at chromosome 6q23 is contained within a wide area of linkage detected in multiple other clinical samples (Bailer et al. 2000; Levinson et al. 2000; Lindholm et al. 2001; Lerer et al. 2003; Lewis et al. 2003). The linkage evidence for schizophrenia in 6q is not population specific as it has been gathered from multiple population groups: African Americans, European Ancestry, and Jews and Arabs from Israel. However, the evidence for association of *TRAR4* in our samples, although present in EA and in AA, appears higher in AA.

[0014] It is therefore an object of the invention to identify SNPs associated with Schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders.

[0015] It is another object of the invention to develop an assay in which the SNPs associated with schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders can be identified.

[0016] It is a further object of the invention to develop a PCR based assay that can identify SNPs associated with schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders on human chromosome 6.

[0017] It is also an object of the present invention to create a method for predicting a risk of

an individual to human schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders, said method comprising amplifying genomic DNA of said individual using oligonucleotide primers to human chromosome 6 to obtain an amplified PCR product, identifying the nucleotides present at the polymorphic sites at nucleotides 132,874,282, 132,874,294 and 132,874,335 of human chromosome 6 (UCSC Map Position, version of July 2003), and predicting the risk of the individual to schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders based upon the haplotype present at the polymorphic sites at nucleotides 132,874,282, 132,874,294 and 132,874,335 of human chromosome 6, wherein a G at position 132,874,282 human chromosome 6, or a deletion at position 132,874,294 of human chromosome 6, or a G at position 132,874,335 of human chromosome 6 haplotype is indicative of an increased risk of developing schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders, and wherein an A at position 132,874,282 human chromosome 6, or an A at position 132,874,294 of human chromosome 6, or an A at position 132,874,335 of human chromosome 6 haplotype is indicative of a decreased risk of developing affected phenotypes. Our results also open the possibility that allelic heterogeneity for bipolar disorder and other psychiatric disorders will be found.

[0018] It is also an object of the invention to provide a diagnostic kit for detection of

schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders associated SNP haplotypes (A/G at 132,874,282 position, an A/-deletion at 132,874,282 position or A/G at 132,874,335 position) comprising at least one primer selected from the group consisting of SEQ ID NOS: 27-270.

[0019] These and other objects of the invention, as well as many of the attendant advantages thereof, will become more readily apparent when reference is made to the following detailed description of the preferred embodiments.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] Figure 1 shows the genomic structure of the 6q23.2 gene cluster and the association mapping of the initial screening. The genomic positions are based on the UCSC July 2003 assembly of the human genome. (a) The relative position of the 6q23.2 gene cluster to the peak markers from various linkage studies. D6S424 (Cao et al. 1997; Martinez et al. 1999), D6S416 (Cao et al. 1997), D6S292 (Lerer et al. 2003) and D6S264 (Lindholm et al. 2001). (b) Genes in the 6q23.2 gene cluster. (c) The -log transformation of the FBAT *P*-value for the 31 SNP markers analyzed for the initial association screening. Each data point of the markers points to its relative position in the gene cluster shown in (b). The most significant marker is rs4305745 with a *P*-value of 0.0014; see Fig. 5, (supplementary table 5

(Fig. 9)) for FBAT *P*-values and other detailed information for all the initially selected markers, and table 1 for the single marker association results for all the additional *TRAR4* markers examined in the dense mapping effort.

[0021] Figure 2 shows pairwise LD in 192 founders for the *TRAR4* region. (a) Relative physical position of 23 markers in *TRAR4* region (shown in Table 1). ss28447862, ss28447876, ss28447865, rs8192624, and ss28447866 were excluded in the following LD measurements because of their low minor allele frequencies. (b) LD pattern in AA (left panel) and EA (right panel). The graph was generated by GOLD (Abecasis and Cookson 2000). In each LD pattern, D' value (upper left diagonal) and *P*-values (lower right diagonal, converted to log *P*-values) are calculated from the program ldmax of the GOLD package (Abecasis and Cookson 2000).

[0022] Figure 3 illustrates the expression pattern of *TRAR4* in human tissues. (a) *TRAR4* expression pattern in various human brain regions. Lane 1 is a 100 bp molecular weight standard ladder (Promega). Lanes 2-13 are human brain, human fetal brain, cerebellum, fetal liver, placental, spinal cord, control (no reverse-transcriptase added), basal ganglia, frontal cortex, substantia nigra, amygdala, and hippocampus. RT-PCR from total RNAs were presented, β-actin was used as internal control. (b) Quantatitive real-time PCR determined the relative abundance of the *TRAR4* transcript in various human

brain regions. (c) Comparison of gene expression of *TRAR4* with *TRAR1*. Samples S1-S6 in (b) and (c) are basal ganglia, frontal cortex, substantia nigra, amygdala, hippocampus and cerebellum.

[0023] Figure 4 (Supplementary Figure 1) shows conserved non-coding regions defined by VISTA (Couronne et al. 2003) and the relative position to associated markers. (a) *TRAR4* 3′-flanking conserved regions generated by comparing human genome with mouse, rat, and chimpanzee genomes (upper, middle, and lower plots, respectively). The region with sequence similarity reaching 70% is defined as a conserved region and such peak areas are shaded pink. (b) Transformed −log FBAT *P*-values of *TRAR4* SNPs versus their relative genomic positions. (a) and (b) are aligned according to the genomic position (UCSC July 2003 genome draft). (c) The local genomic sequence alignment (human-mouse) around rs4305745 and two other polymorphisms (ss28447873 and rs7452939) in perfect LD with rs4305745.

[0024] Figure 5 (Supplementary Table 1) shows data from genotyping a total of 827 individuals from 192 families (67 NIMH-IRP, 69 NIMH-GI, 56 AU/US).

[0025] Figure 6A-6D (Supplementary Table 2) shows the nucleotide sequences for the PCR primers, the FP-TDI and TaqMan probes, and related information for each marker identified in the study.

[0026] Figure 7 (Supplementary Table 3) identifies the primer sequences for TRAR4 amplicons.

[0027] Figure 8 (Supplementary Table 4) shows the results of linkage analyses of alleles sharing with individual SNPs from the MOXD1-STX7-TRARs genes cluster.

[0028] Figure 9 (Supplementary Table 5) presents data showing SNP markers of initial screening and FBAT analysis.

[0029] Figure 10 (Supplementary Table 6) identifies the mutations detected in TRAR4 from 30 schizophrenic patients.

[0030] Figure 11 (Supplementary Table 7) shows the single marker association via FBAT for all markers with ten or more informative families in the sample.

[0031] Figure 12 (Supplementary Table 8) provides a table comparing the coding variants detected in AA schizophrenia probands and AA controls.

[0032] Figure 13 (Supplementary Table 9) shows the two marker haplotype association analysis for TRAR4.

DETAILED DESCRIPTION AND PREFERRED EMBODIMENTS

[0033] In describing a preferred embodiment of the invention specific terminology will be resorted to for the sake of clarity. However, the invention is

not intended to be limited to the specific terms so selected, and it is to be understood that each specific term includes all technical equivalents which operate in a similar manner to accomplish a similar purpose.

Subjects and Phenotyping

[0034] Three samples were studied, which we call the NIMH-IRP, NIMH-GI, and AU/US collections. Ascertainment of the NIMH-IRP sample was described initially (Gershon et al. 1988), and the full sample from which the present sample of 67 pedigrees was drawn was described later (Cao et al. 1997; Gejman et al. 2001).

[0035] The collection of the NIMH-GI sample was described in a report of a genome scan of 71 pedigrees (Cloninger et al. 1998), and additional NIMH-GI families were subsequently included in the repository-based dataset (see electronic-database information section); 69 pedigrees were drawn for the present analysis and two previous ones (Cao et al. 1997; Martinez et al. 1999).

[0036] The AU/US sample was described initially in a report of a genome scan of 43 pedigrees (Levinson et al. 1998); full or partial trios for the present study were from 56 of the 71 pedigrees in the expanded sample utilized in linkage fine-mapping studies (Mowry et al. 2000) and in additional

analyses of this dataset (Martinez et al. 1999; Levinson et al. 2000).

[0037] For the present study we genotyped a total of 827 individuals from 192 families (67 NIMH-IRP, 69 NIMH-GI, 56 AU/US). Details are provided in supplementary table 1 (Fig. 5) and in the electronic database information section. Individuals were diagnosed as affected with schizophrenia (the majority of the affected subjects) or schizoaffective disorder via the criteria of the Diagnostic and Statistical Manual of Mental Disorders, third edition, revised (DSM-IIIR) (APA 1987). The Institutional Review Board of the Evanston Northwestern Healthcare Research Institute approved this study.

SNP Selection and Genotyping

[0038] SNPs were selected from public databases with the help of a bioinformatics tool, SNPper (Riva and Kohane 2002), and novel *TRAR4* SNPs were identified by direct sequencing. The DNA samples were genotyped using two methods: (1) template-directed dye-terminator incorporation with fluorescence-polarization detection (FP-TDI) (Chen et al. 1999) or (2) the TaqMan assay developed by Applied Biosystems (ABI). For the FP-TDI assays, briefly, after PCR amplification of genomic DNA, the AcycloPrime™-FP SNP detection kit (PerkinElmer) was used for post-PCR cleanup and the single base extension reaction, and we detected FP by either an

Analyst fluorescence reader (LJL Biosystems) or a Wallac-Victor3 (PerkinElmer), and FP data were converted to genotypes with the assistance of an automated genotype calling spreadsheet (Akula et al. 2002). PCR primers and probes for the FP-TDI assays were designed using Primer3 (Rozen and Skaletsky 2000). For the TaqMan assays, briefly, the genomic sequence flanking the SNP was submitted to ABI for developing an assay-by-design. Each unique TaqMan minor groove binding (MGB) allele specific probe was labeled by either a 5′-FAM or a 5′-VIC reporter dye. PCR amplification of genomic DNA was performed in a 384-well plate in an ABI Prism 7900 or a DNA Engine Tetrad 2 (MJ Research). After PCR, the allele discrimination was performed on an ABI Prism 7900 Sequencing Detection System using Sequence Detector Software (SDS) version 2.0. Standard genotype calling was converted by a customized spreadsheet. Nucleotide sequences for the PCR primers, the FP-TDI and TaqMan probes, and related information for each marker can be found in supplementary table 2 (Figs. 6A-6D).

[0039] The average completion rate of our experiments was 96%. To empirically check for errors in the genotyping method, we compared genotypes for marker rs4305745 from different methods; a difference rate between FP-TDI and TaqMan assay was about 0.25%. Using MERLIN (Multipoint Engine for Rapid Likelihood Inference) (Abecasis et al. 2002) (with all the SNPs at once), we checked Mendelian inconsistencies, blanked them as described below (sometimes for

individuals and sometimes for the family when the error could not be traced to a particular individual), and then addressed all unlikely recombinants. Genotyping errors were detected for 0.17% of genotypes (MERLIN) (95 errors out of 54,611 nonzero genotypes), including 26 Mendelian inconsistencies (0.047%) and 69 unlikely recombinants (0.12%). We did not change genotypes for unlikely recombinants unless MERLIN estimated a high probability of an individual genotype error, compared to other possible errors (see MERLIN documentation for details) and/or manual re-reading each genotype tracing or other raw genotyping output for the family/marker in question pinpointed a specific error for a particular individual. All genotype errors (all Mendelian inconsistencies or when specific errors were seen that resulted in unlikely recombinants) were blanked (zeroed) for the involved individuals, and we did not perform a second pass genotyping given our high genotyping completion rate and low genotyping error. Genotypes were read blindly of psychiatric status.

[0040] We checked Hardy-Weinberg equilibrium (HWE) on family founders (at least 200) for all 55 SNPs. Three MOXD1 SNPs were found not to be in HWE (rs2206064, rs1981187 and rs2275394, though rs2206064's HWE *P*-value did not remain significant after taking into account the number of markers examined for HWE). The minor allele frequency of rs2206064 was only 2%, which might explain the lack of HWE. Both rs1981187 and rs2275394 had higher

frequencies of homozygotes than expected (which was not a bias introduced by cleaning - only a handful genotypes were blanked and they were not primarily heterozygous genotypes).

Intermarker Linkage Disequilibrium Analysis

[0041] LD (linkage disequilibrium) between the SNPs was estimated with the program ldmax from the GOLD (Graphical Overview of Linkage Disequilibrium) package (Abecasis and Cookson 2000) using the genotypes from unrelated founders. Ldmax estimates haplotype frequencies from genotype data using an expectation-maximization algorithm (Excoffier and Slatkin 1995). The standard and normalized Lewontin's Disequilibrium coefficients (D, D') are derived. Association significance is assessed from a chi-squared distribution with $(n_1-1)(n_2-1)$ degrees of freedom (df), where n_1 and n_2 are the number of alleles at each marker locus.

Association Analysis

[0042] To detect LD with illness, we used the transmission disequilibrium test (TDT), as implemented in the Family Based Association Test (FBAT) program v1.5 (Laird et al. 2000; Rabinowitz and Laird 2000). The null hypothesis of interest, here, is the absence of association in presence of linkage. We thus employed the empirical-variance estimator (-e flag option in the FBAT program) to

account for SNP-genotype correlations among affected siblings due to linkage. The FBAT test statistic uses a score function, $Z = S_j - E(S_j) / Var(S_j)$, where S_j is the observed number of transmitted marker alleles j to affected offspring, and $E(S_j)$ and $Var(S_j)$ are the expected and variance values of S_j under the null hypothesis. Asymptotically, Z is assumed to follow a normal distribution with a mean and a variance equal to 0 and 1, respectively. The test statistic can also be expressed as Z^2, which follows a chi-squared distribution with 1 df. FBAT has the ability to deal with the transmission of multi-locus haplotypes, even when phase is unknown and parental genotypes may be missing. It can use both pedigrees and nuclear families, but pedigrees are broken down into all individual nuclear families, though it only includes informative families, i.e., those contributing to the test statistic. For the analyses of the screening SNPs, alleles and haplotypes were tested for association if there were at least 10 informative families; in our data this corresponds to not testing alleles and haplotypes rarer than 3%. This restriction, however, was not used when the investigation was limited to specific subsets of families in the secondary analyses. For multi-locus association analyses, FBAT provides global *P*-values, which assess the significance of transmission distortion for all the tested haplotypes. In the present analyses, we limited the number of multi-locus systems tested by using a stepwise procedure, and limiting the number of multi-locus tests to the combinations including the SNP with highest single Z

score value, as further detailed in the results. FBAT analyses were performed assuming an additive model for each SNP, and only one affection status model was used (affected subjects had schizophrenia or schizoaffective disorder). The additive model is expected to perform well even when the true model is non-additive.

Linkage Analysis

[0043] Model-free linkage analyses with the MOXD1-STX7-TRARs gene cluster were carried out using the lod score test from the affected-only sharing method (Kong and Cox 1997), as implemented in the MERLIN program (Abecasis et al. 2002). The likelihood of the observed marker information among affected relatives is maximized as a function of the marker alleles sharing parameter, and is compared, through a likelihood ratio test, with the likelihood of the marker data under the null hypothesis of no linkage. The resulting distribution of the allele sharing test (T) is a χ^2 with 1 degree of freedom, and the statistic can also be reported as a lod score = T / 2ln(10).

[0044] We performed additional analyses to account for putative genetic and/or allelic heterogeneity within our family sample by ethnic origin. Association and linkage tests were evaluated separately in the subset of AA (African American) and EA (European Ancestry) families.

Mutation Detection

[0045] Sequencing of TRAR4 was performed on ABI 3100 genetic analyzer. Purified PCR products from various amplicons of relevant genomic DNA fragments were used as templates in sequencing reactions with the chemistry of BigDye 3.1 (ABI). PCR primers were designed by Primer 3 (Rozen and Skaletsky 2000) and were also used as sequencing primers for forward and reverse sequencing. The primer sequences and product sizes are in supplementary table 3 (Fig. 7). We used software SeqScape Ver2.1 (ABI) to assist in mutation detection, and we visually verified each mutation. The reference sequence of TRAR4 used in the analysis was from the human genome draft of the UCSC July 2003 freeze.

[0046] For the non-human primates, DNAs were extracted from peripheral blood samples of two different chimpanzees (PTR-S109 and PTR-S286) from West Africa and from tissue samples of two different lowland gorillas (GGO-S110 and GGO-S249). The forward primer of amplicon one and the reverse primer for amplicon seven were used to PCR amplify the entire DNA segment by standard methods with annealing at 60°C; this product was then sequenced bi-directionally with the seven primer pairs detailed in supplementary table 3 (Fig. 7). PCR products were confirmed by 1.5% agarose gel electrophoresis, and purified using Micro Spin Columns (Amersham Biosciences). The purified PCR products were sequenced using the BigDye Terminator Cycle Sequencing Kit (ABI) on an ABI Prism 377/3100 DNA

sequencer. Sequence data were assembled by the phred/phrap program (Ewing et al. 1998) and also were checked manually using the consed program (Gordon et al. 1998). Sequence data with both strand reads and/or high quality (more than 30 quality value) were used and deposited into the DDBJ/EMBL/GenBank International Nucleotide Sequence Database.

Accession numbers and URLs for data presented herein are as follows:

[0047] Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/
National Institute of Mental Health (NIMH) Human Genetics Initiative for Schizophrenia, families used in this manuscript were 30101, 30103, 30104, 30106, 30108, 30110, 30111, 30112, 30113, 30114, 30116, 30119, 30122, 30123, 30124, 30126, 30127, 30128, 30130, 30131, 30132, 30133, 30134, 30136, 30140, 30142, 30146, 31102, 31107, 31108, 31109, 31114, 31115, 31118, 31119, 31129, 31130, 31135, 31137, 31139, 31155, 32108, 32109, 32200, 32201, 32202, 32203, 32204, 32205, 32206, 32209, 32211, 32212, 32217, 32218, 32303, 32304, 32306, 32307, 32309, 32310, 32311, 32312, 32313, 32315, 32319, 32320, 32402, and 32403, http://zork.wustl.edu/nimh/sz.html

[0048] University of California at Santa Cruz (UCSC) Genome Bioinformatics July 2003 assembly of the human genome, http://genome.ucsc.edu/

[0049] SNPper bioinformatics tool, http://snpper.chip.org/

[0050] SIFT bioinformatics tool, http://blocks.fhcrc.org/sift/SIFT.html
PolyPhen bioinformatics tool, http://tux.embl-heidelberg.de/ramensky/

[0051] VISTA bioinformatics tool, http://www-gsd.lbl.gov/vista/

[0052] NCBI's SNP database, http://www.ncbi.nlm.nih.gov/SNP/

[0053] NCBI's Entrez search engine, http://www.ncbi.nlm.nih.gov/Entrez/

[0054] Mfold RNA structure prediction, http://www.bioinfo.rpi.edu/applications/mfold/

[0055] G protein-coupled receptors database, http://www.gpcr.org/

[0056] dbSNP, NCBI's database of "Single Nucleotide Polymorphisms", was used to deposit the 18 novel SNPs we found, http://www.ncbi.nlm.nih.gov/SNP/

[0057] The dbSNP accession numbers are ss28447859 through ss28447876 and will become available to the public when NCBI releases the latest dbSNP build, and at that time will be incorporated into rs#'s (Reference Cluster ID #'s) as well.

[0058] dbSTS, NCBI's database of "Sequence Tagged Sites", was used to deposit the 18 novel SNPs we found, http://www.ncbi.nlm.nih.gov/dbSTS/

[0059] The GenBank accession numbers are BV154568 through BV154585.

[0060] The DDBJ/EMBL/GenBank International Nucleotide Sequence Database was used to deposit the gorilla and chimpanzee [0052] *TRAR4* region sequences, http://www.ncbi.nih.gov/Genbank/index.html

[0061] The DDBJ/EMBL/GenBank accession numbers for the chimpanzee and gorilla sequences are AB180397 through AB180400.

RT-PCR and Real-Time PCR

[0062] Total mRNAs from various brain tissues were purchased from either BD Biosciences or Ambion. Gene expression of TRAR4 was first confirmed with general RT-PCR with primer pairs used previously for amplification of segment 4 of TRAR4 shown in supplementary table 3 (Fig. 7). In brief, total mRNA was reverse-transcribed with TaqMan Reverse Transcription Reagents (ABI), and the synthesized first-strand cDNAs were then used as templates to amplify *TRAR4* with HotStart Taq polymerase (Qiagen). β-actin was used as internal control in the RT-PCR.

[0063] Reverse transcribed cDNAs were also used in real-time PCR on an ABI Prism 7900 Sequence Detection System according to the manufacturer's protocol. The TaqMan MGB probes and PCR primer pairs for gene expression assay for TRAR4, GAPD (glyceraldehyde-3-phosphate dehydrogenase), or TRAR1 were purchased as an Assay-On-Demand from ABI (Applied Biosystems; Foster City, CA). The relative gene expression in different brain tissues was normalized to GAPD expression by using the standard curve method as described by ABI.

Bioinformatic Tools for Prediction of Functional Effects of Genetic Polymorphisms

[0064] SIFT (http://blocks.fhcrc.org/sift/SIFT.html) and PolyPhen (http://tux.embl-heidelberg.de/ramensky/) were used to predict the potential functional effect of missense polymorphisms (Ramensky et al. 2002; Ng and Henikoff 2003). We used VISTA (http://www-gsd.lbl.gov/vista/) to predict the potential regulatory sequence through defining the conserved region among genomic sequences different species (Couronne et al. 2003).

[0065] We have studied 192 families with previous evidence of linkage. Thirty-three SNPs were initially selected for study, of which thirty-one were selected for analysis in the screening experiment since two of the MOXD1 SNPs, rs2206064 and rs7751860, had minor allele frequencies resulting in

less than 3% informative families. The screening SNPs spanned ~500 kb of the MOXD1-STX7-TRARs genes cluster, a prime set of positional and pathophysiological candidates for schizophrenia. We selected at least one common SNP for each gene with a minimum coverage of ≥1 SNP per 20 kb (the screening set of SNPs). Linkage analyses confirmed the presence of excess allele sharing in this region with individual SNPs from the MOXD1-STX7-TRARs genes cluster. Nine SNPs showed linkage *P*-values < 0.05 as seen in supplementary table 4 (Fig. 8), with the most significant one being rs6937506 (*TRAR4*) with a LOD = 1.76, *P* = 0.002.

[0066] The following examples will provide illustrations of the use of the invention.

[0067] <u>Example 1</u>. Association results are presented in table 1 (Fig. 5) and can also be seen in supplementary table 5 (Fig. 9). Of 31 screening SNPs, four SNPs spanning 106 kb, two located in TRAR4, and one each in STX7 and GPR57, yielded a *P*-value < 0.05. The most significant one was found in TRAR4: rs4305745 (*P* = 0.0014) (Figure 1c), which was the only SNP that remained significant after Bonferroni correction for 31 tests. Association rs4305745 is located 1,214 bp downstream from the stop codon of TRAR4. We therefore concentrated further laboratory efforts on TRAR4. We aimed for a high-density map of > 1 SNP per 2 kb by searching public SNP databases and sequencing genomic DNA. The TRAR4 gene was sequenced (~1 kb of the 5′ region, the

1,038 bp CDS, and ~1.5 kb of the 3'UTR, which spans rs4305745) in 30 probands selected from the NIMH-GI families: 16 European Ancestry (EA) and 14 African Americans (AA). Ten coding variants (26 total variants as seen in table 2 and supplementary table 6, Fig. 10) were found by sequencing TRAR4, including three previously found in 96 healthy EA individuals (Freudenberg-Hua et al. 2003). Five of the 7 novel variants are missense and present only in AA, shown in table 2. Twenty SNPs spanning 21.6 kb of the *TRAR4* gene (> 1 SNP per 2 kb) were identified for genotyping.

[0068] Two additional markers 3' to rs4305745 showed association with schizophrenia in the whole sample as seen in table 1 and supplementary table 7 (Fig. 11), rs6903874 ($P = 0.0026$) and rs6937506 ($P = 0.0052$); and two other additional markers (ss28447873, an insertion/deletion polymorphism, and rs7452939) were found to be in perfect LD with rs4305745 (they span 53 bp), and, hence, also associated to the same degree with illness as rs4305745.

[0069] Although similar association trends in EA and AA were observed, supplementary table 7 (Fig. 11) shows that the evidence for association was proportionally stronger in the AA sub-sample. The evidence for association with rs4305745 and rs6937506 was significant in both samples ($P = 0.035$ in AA and $P = 0.015$ in EA for rs4305745; $P = 0.025$ in AA and $P = 0.035$ in EA for rs6937506), however, a cluster of

three SNPs in the promoter of TRAR4 yielded significant association in the AA sub-sample: rs4473885 (*P* = 0.032), rs4085406 (*P* = 0.047), and rs6907909 (*P* = 0.019), but not in the EA sample as seen in supplementary table 7 (Fig. 11). These differences raise the possibility of allelic heterogeneity specific to the AA population. All new missense mutations detected by sequencing were also exclusive to the AA sub-sample as seen in table 2. Two, A518G (Tyr173Cys) and C683T (Ala228Val), were predicted by either PolyPhen or SIFT (Ramensky et al. 2002; Couronne et al. 2003; Ng and Henikoff 2003) to be non-conservative, as shown in table 2.

[0070] After genotyping the whole sample, it was noted that none of the newly found missense SNPs co-segregated with disease in a specific manner (data not shown). Furthermore, all the missense variants except for A518G (Tyr173Cys) were also found in a set of 48 AA subjects from the Coriell Human Variation AA DNA panel, shown in supplementary table 8 (Fig. 12); additionally, some of these missense variants were homozygous in some control individuals. Interestingly, the ratio of missense to synonymous mutations (9:3) is close to what is expected under neutral expectations, i.e., a pseudogene, which has an expected ratio close to 4:1. Indeed, in the gorilla TRAR4 has already become a pseudogene (with a nonsense mutation at codon 15: Tyr in chimpanzee and human, STOP in gorilla), though the human and chimpanzee TRAR4 versions have not yet become pseudogenes.

[0071] Example 2. The TRAR4 region was found to have two LD blocks, depicted in figure 2. Association rs4305745 (marker 16 in figure 2) is in the LD block constituted by 3'-flanking SNPs. The pattern suggests that association for TRAR4 originates from rs4305745. None of the 5'-flanking SNPs are in LD with rs4305745, which instead is in strong LD with markers 19 (rs6903874) and 21 (rs6937506) from the 3' LD block (and also shows a trend with marker 12, rs8192625). The LD pattern generated from the 31 initial screening markers indicated that the whole region of MOXD1-STX7-TRARs is separated into 4 major strong LD blocks, while the TRAR4 region represented by rs4305745 is not in strong LD with any of the major LD.

[0072] Example 3. Haplotype association analyses with all TRAR4 two-locus systems were conducted (n = 17, after excluding five markers with minor allele frequencies < 3%) that were derived from rs4305745. This association was chosen as the anchor because it had the most significant single locus association ($P = 0.0014$). For each such two-locus system, we derived the global $\chi 2$ value using only those haplotypes with frequencies > 3%, detailed in supplementary table 9 (Fig. 13). All 17 of these two-locus systems exhibited P-values < 0.05 (not corrected by multiple testing) and all harbored the "A" allele of rs4305745. None of the two-locus systems showed stronger association than rs4305745 alone as seen in supplementary table 9 (Fig. 13), and

the same results were obtained with up to five multilocus systems, with each haplotype system extended stepwise to contain the most significant previous smaller haplotype (data not shown). This shows that rs4305745, and/or the other two nearby polymorphisms, ss28447873 and rs7452939, in perfect LD (further confirmed by genotyping the whole sample) with rs4305745 (supplementary table 6 (Fig. 10)), are the mutations underlying the association of the TRAR4 region with schizophrenia.

[0073] To explore the possible functional effects of associated SNPs and their haplotypes, we first defined the conserved non-coding sequence (considered as a potential functional region) by comparative genomic analysis of TRAR4 genomic sequences of human, mouse, and rat using VISTA (Couronne et al. 2003). The cluster of three polymorphisms (rs4305745, ss28447873, and rs7452939 - all equally implicated as candidates by the association analysis) exhibiting the most significant association is very close to two conserved regions (sequence identity > 70% among human, mouse, and rat genomes) right after the stop codon. The sequence identity immediately around this SNP (rs4305745) is about 50% as seen in supplementary figure 1, suggesting that this SNP and/or other polymorphisms in perfect LD with rs4305745 may ultimately affect gene expression, a hypothesis we are currently testing. RT-PCR from various brain tissues also confirmed that rs4305745 was flanked by the 3'UTR of TRAR4 (data not shown), suggesting that rs4305745 or one of its haplotypes may affect the

gene expression at the post-transcriptional level. Another significant SNP, rs6903874, is also within a conserved region (supplementary figure 1), and it is also possible that this SNP and/or other SNPs nearby might be functional, though it is much farther from the stop codon of TRAR4 than rs4305745.

[0074] A comprehensive gene expression analysis of *TRAR4* will help to elucidate its potential functional roles in the pathophysiology and pharmacology of schizophrenia. TRAR4 expression was investigated in various human tissues by RT-PCR and found that TRAR4 was expressed at low abundance in various human brain tissues as well as in human fetal liver, but not in the cerebellum or placenta as seen in Fig 3a. A quantitative real-time PCR shown in Fig. 3b further revealed that TRAR4 has comparable levels of expression in basal ganglia, frontal cortex, substantia nigra, amygadala, and hippocampus, with highest expression in hippocampus and lowest expression in basal ganglia; these results are consistent with a previous expression study including TRAR4 (Borowsky et al. 2001). These regions have been implicated in the pathophysiology and pharmacology of schizophrenia (Grossberg 2000; Freedman 2003). The tissue distribution of TRAR4 gene expression is similar to the only well-characterized trace amine receptor, TRAR1 (Borowsky et al. 2001; Bunzow et al. 2001), however, further comparison of gene expression of TRAR4 and TRAR1 indicated that TRAR4 is overall more abundant than TRAR1, particularly in basal ganglia (~14 fold),

frontal cortex (~21 fold), and substantia nigra (~14 fold) (Figure 3c), suggesting TRAR4 plays a more important role than TRAR1.

[0075] Regulatory sequence disruption can affect protein expression and cause disease (Mitchison 2001). The associated SNPs in the 3'UTR of TRAR4 may contribute to the susceptibility for the disease by affecting the gene expression at the post-transcriptional level. Our RT-PCR experiment indicated that the TRAR4 3'UTR spanned the most associated SNP rs4305745; therefore, it is possible TRAR4 gene expression was affected at the post-transcriptional level by these 3'UTR SNPs (rs4305745 and/or ss28447873 and rs7452939, two SNPs in perfect LD with rs4305745). The chimpanzee and gorilla sequencing result indicated that ancestral allele for rs4305745 is A (supplementary table 6 (Fig. 10)). Because of the perfect LD among rs4305745, ss28447873, and rs7452939, the associated allele A of rs4305745 actually represents a human haplotype of A-A-A spanning rs4305745, ss28447873, and rs7452939 (supplementary figure 1c; also note ancestral haplotype appears to be A-A-G as seen in supplementary table 6 (Fig. 10)). Interestingly, the predicted TRAR4 mRNA structure exhibited a significant change for the over-transmitted haplotype A-A-A compared to haplotype G-del-G, and the same structure change can be generated by allele A of rs4305745 alone (data not shown), suggesting rs4305745 is most likely the causative SNP.

[0076] We have found that the mutation rate in coding region for TRAR4 (1 mutation per 100 bp) is well above the average (1 mutation per 346 bp) (Cargill et al. 1999), and there are more missense mutations in TRAR4 than synonymous mutations, 9 versus 3, as seen in table 2 and supplementary table 8 (Fig. 12). We did not find evidence of association with schizophrenia for missense SNPs, however, some missense mutations may be pharmacologically important, particularly A518G (Tyr173Cys) located in the putative extracellular domain of the receptor and hence may affect ligand binding as seen in table 2. Besides the changes in protein structure, these missense mutations may also possibly alter the gene expression by affecting mRNA folding structures as described for dopamine D2 receptor (DRD2) (Duan et al. 2003). Actually, A518G was predicted to have a remarkable effect on TRAR4 mRNA folding as predicted *in silico* by mFold (Zuker et al. 1999) (data not shown). It is also notable that this missense SNP and several others were only found in AA, and 173Cys was only detected in AA schizophrenia probands.

[0077] Having described the invention, many modifications thereto will become apparent to those skilled in the art to which it pertains without deviation from the spirit of the invention as defined by the scope of the appended claims.

What is claimed is:

1. A diagnostic kit for the detection of SNP haplotypes associated with human schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders comprising at least one nucleic acid consisting of a nucleic acid selected from the group consisting of SEQ ID NO: 1-26.

2. An oligonucleotide primer consisting of a sequence selected from the group consisting of:

SEQ ID NOS: 27-270 and complements thereof.

3. A method for predicting a risk of an individual to human schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders, said method comprising:

a) amplifying genomic DNA of said individual using oligonucleotide primers to human chromosome 6 to obtain an amplified PCR product;

b) identifying the nucleotides present at the polymorphic sites at nucleotides 132,874,282, 132,874,294 and 132,874,335 of human chromosome 6 (UCSC Map Position, version of July 2003; and

c) predicting the risk of the individual to schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders based upon the haplotype present at the polymorphic sites at

nucleotides 132,874,282, 132,874,294 and 132,874,335 of human chromosome 6, wherein a G at position 132,874,282 human chromosome 6, or a deletion at position 132,874,294 of human chromosome 6, or a G at position 132,874,335 of human chromosome 6 haplotype is indicative of an increased risk of developing schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders, and wherein an A at position 132,874,282 human chromosome 6, or an A at position 132,874,294 of human chromosome 6, or an A at position 132,874,335 of human chromosome 6 haplotype is indicative of a decreased risk of developing schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders.

4. The method as claimed in claim 3 wherein the primers are selected from the group consisting of:

SEQ ID NOS: 27-270 and complements thereof.

5. A diagnostic kit for detection of schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders associated SNP haplotypes (A/G at 132,874,282 position, an A/- deletion at 132,874,282 position or A/G at 132,874,335 position) comprising at least one primer selected from the group consisting of SEQ ID NOS: 27-270.

6. A method of detection of human chromosome 6 gene variants above said method comprises: (a) amplifying genomic DNA of schizophrenia, schizoaffective disorder, bipolar disorder and related mental

disorders patients and normal control individuals using the primers for human chromosome 6 (SEQ ID 27-270); (b) sequencing the amplified PCR product and identifying the sequence variation computationally by comparing it with the already existing sequence of human chromosome 6; (c) screening normal control individuals and schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders patients. (d) computing the frequency of A/G haplotypes at 132,874,282 position, A/- deletion haplotypes at 132,874,282 position and A/G haplotypes at 132,874,335 position; (e) establishing the association of G (at 132,874,282 position), A/-deletion (at 132,874,282 position) and G (at 132,874,335 position) haplotypes with schizophrenia and related disease based on their frequency distribution in normals and schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders; (g) predicting the risk or susceptibility schizophrenia, schizoaffective disorder, bipolar disorder and related mental disorders based on the haplotype present at the polymorphic sites in the individual tested, G (at 132,874,282 position), a deletion (at 132,874,282 position) and a G (at 132,874,335 position) haplotypes being at high risk and A (at 132,874,282 position), A (at 132,874,282 position) and A (at 132,874,335 position) haplotypes at low risk for the disease.

Figure 1

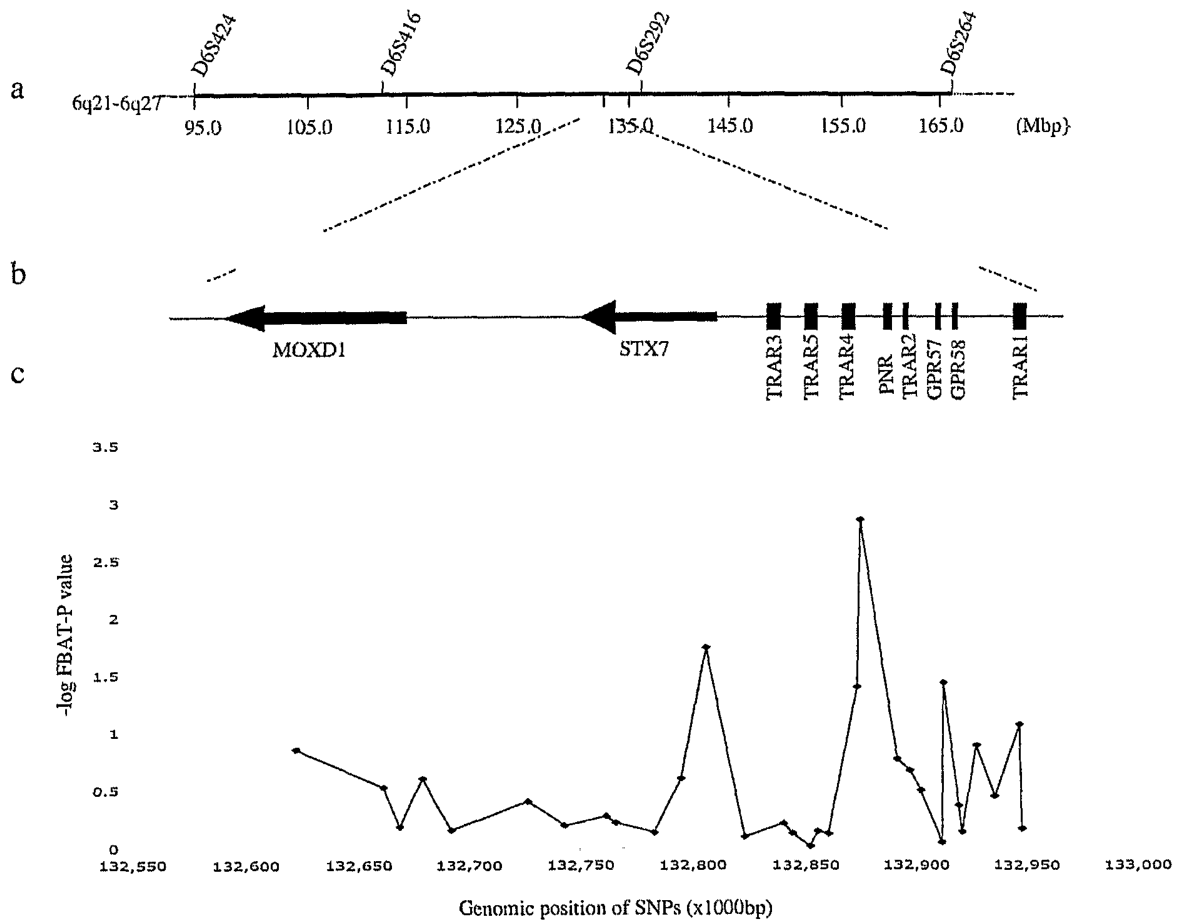


Figure 2

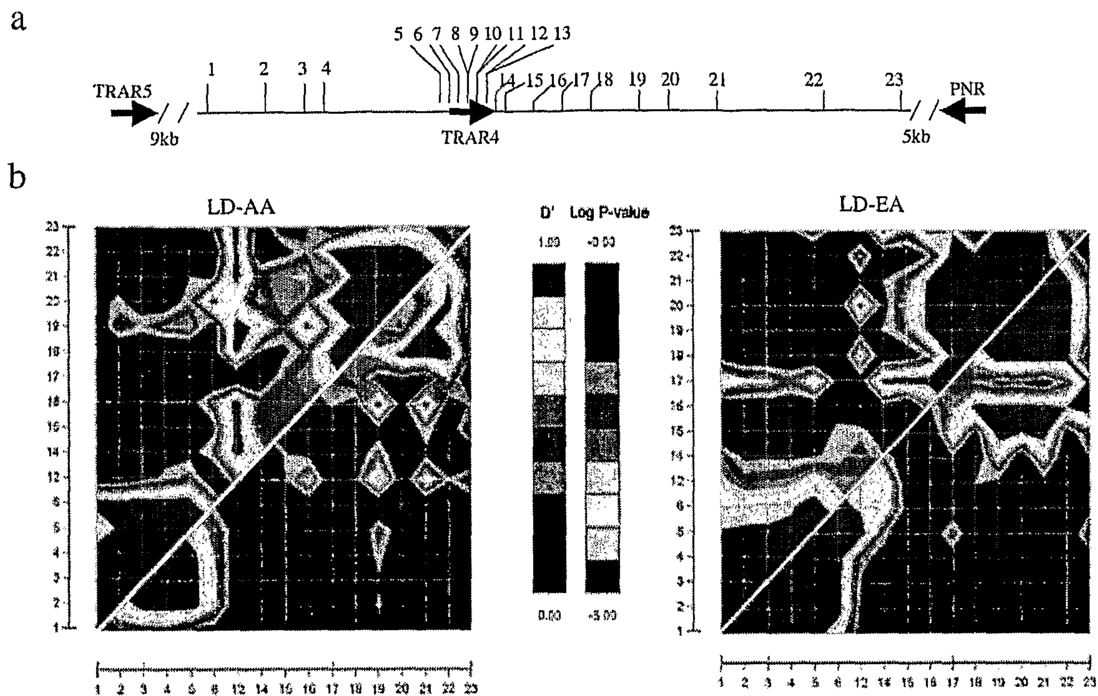
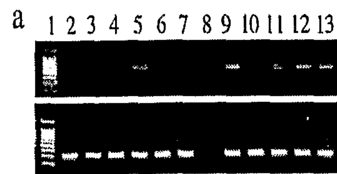
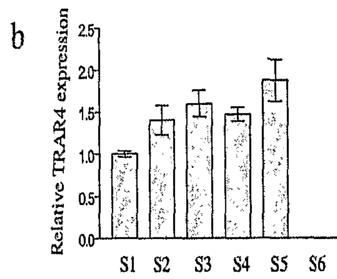
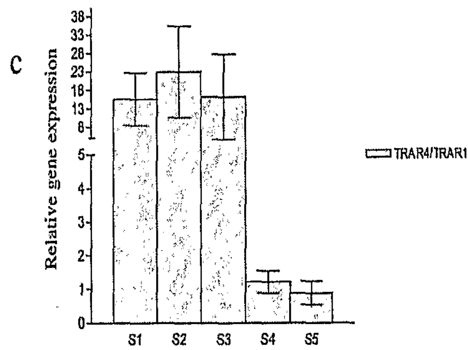


Figure 3







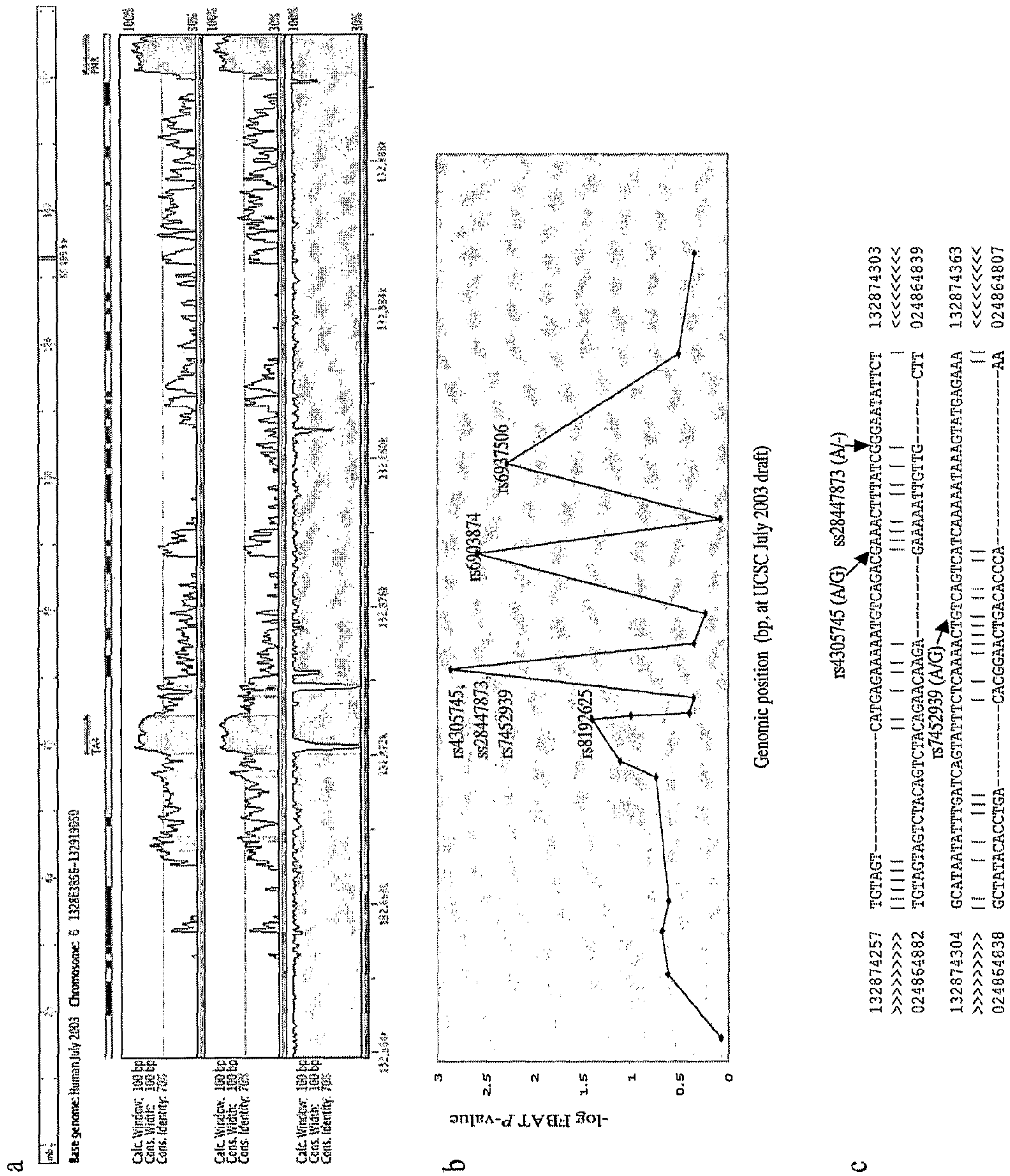


Figure 4
a
Base genome: Human July 2003 Chromosome: 6 132863856-132919050
Calc. Window: 100 bp
Cons. Width: 100 bp
Cons. Identity: 70%
100%
50%
b
-log FBAT P-value
3
2.5
2
1.5
1
0.5
0
rs4305745,
ss28447873,
rs7452939
rs8192625
rs6903874
rs6937506
Genomic position (bp, at UCSC July 2003 draft)
c
rs4305745 (A/G)
ss28447873 (A/-)
rs7452939 (A/G)
132874257 TGTAGT----------CATGAGAAAATGTCAGAGCGAACTTTATCGGGATATTCT 132874303
024864882 TGTAGTAGTCTACAGTCTACAGACAAGA-----------GAAAATTGTTG---------CTT 024864839
132874304 GCATAAATATTTGATCAGTATTTCTCAAAACTGTCAGTCATCAAAATAAAGTATGAGAAA 132874363
024864838 GCTATACACCTGA---------CACGGAACTGACACCCA-------------------AA 024864807

Figure 5

Supplementary Table 1
Subjects and Phenotyping

	Entire Sample									Individuals with DNA						
Collections Used	Subjects with DNA	Subjects without DNA	Total Subjects	Total Families	Full Trios	Founders with DNA	EA	AA	Other Race	Male	Female	SZ	SA	UD	Average Age at Psychosis Onset	Range of Ages at Psychosis Onset
NIMH-IRP only	325	117	442	67	39	108	88%	4%	7%	180	145	136	24	165	20.4	6 to 45
NIMH-GI only	321	39	360	69	38	106	55%	32%	13%	159	162	137	22	162	18.5	5 to 38
AU/US only	181	204	385	56	47	103	84%	13%	4%	100	81	79	-	102	23.1	14 to 42
All Collections	827	360	1,187	192	124	317	75%	17%	8%	439	388	352	46	429	20.3	5 to 45

Note: EA = European Ancestry; AA = African American; SZ = schizophrenia; SA = Schizoaffective disorder; UD = unknown diagnosis. Full trios consist of an affected offspring and both parents; the remainder of the families have partial trios with supplemental information from additional family members.

Figure 6

Supplementary Table 2

Assay Information

Order for All SNPs	Order for Screening SNPs	Gene	dbSNP Number	Allele 1	Allele 2	F PCR Primer (5'->3')	R PCR Primer (5'->3')
1	1	*MOXD1*	rs2206064	C	T	Acagtaatttctataattatcaaataatctcagctgaagaat (SEQ ID NO: 41)	Gagagagagggagagagaaatgaca (SEQ ID NO: 102)
2	2	*MOXD1*	rs599660	G	A	Catctctgatgtcacttttatgatac (SEQ ID NO: 42)	Tgagtgtcgctacaacacgaaa (SEQ ID NO: 103)
3	3	*MOXD1*	rs7751860	C	T	Ttgtaattcacaacctgcataagttttcaa (SEQ ID NO: 43)	Tttgttagatttcttcctttatctgccaaga (SEQ ID NO: 104)
4	4	*MOXD1*	rs1981187	C	T	Ctcgaattctcctcccaagatg (SEQ ID NO: 44)	Aaaaactctggcaatttgtatc (SEQ ID NO: 105)
5	5	*MOXD1*	rs1338387	C	T	Ggctctacagtaatctggcttcaag (SEQ ID NO: 45)	Gggagtgtttcctagaacatgaaaatttaag (SEQ ID NO: 106)
6	6	*MOXD1*	rs1538308	A	G	Gattttcaatatgtaaatgaaataacaatgtctgtagacaa (SEQ ID NO: 46)	Gaaatatgagttcagtgaaaaggcacag (SEQ ID NO: 107)
7	7	*MOXD1*	rs2275394	C	T	Ggaaatcttcctcaggcataca (SEQ ID NO: 47)	Aagctcagctagagtctttcaagc (SEQ ID NO: 108)
8	8	*MOXD1*	rs6937815	G	A	Ggctgctgtggtttgca (SEQ ID NO: 48)	Cttccaggtgcaggagaca (SEQ ID NO: 109)
9	9	*MOXD1*	rs3823288	A	T	Gcctgtaggccgctatgac (SEQ ID NO: 49)	Ccttagaagaggagtgagaggaaga (SEQ ID NO: 110)
10	10	*STX7*	rs1856352	A	G	Actttccaaaagtaattttgacctaaaaatattaaccatt (SEQ ID NO: 50)	Tcatctggttttgtgggccattatt (SEQ ID NO: 111)
11	11	*STX7*	rs3757299	A	C	Tgcattttccacattggcttctatg (SEQ ID NO: 51)	Tgttactggagcatcaatccatgtt (SEQ ID NO: 112)
12	12	*STX7*	rs1002799	A	G	Aaggacccagggctgaaaatc (SEQ ID NO: 52)	Tccactttcctcttgccttctttc (SEQ ID NO: 113)
13	13	*STX7*	rs2788942	G	A	Tcttccagaaagagattattgaacaaagca (SEQ ID NO: 53)	Tgggcatgggaggagagt (SEQ ID NO: 114)
14	14	*STX7*	rs2842884	G	C	Aaaggacagagtcatgttaaacaca (SEQ ID NO: 54)	Gggccctgccttataattgtt (SEQ ID NO: 115)
15	15	*STX7*	rs1591811	G	A	Gcggcattcctggacttg (SEQ ID NO: 56)	Agagcaagaaaaacaaataaaatgcatacaga (SEQ ID NO: 116)
16	16	*TRAR3*	rs2842899	A	T	Tggaaacttactggtcatgattgct (SEQ ID NO: 57)	Gcaatcagaaagtttgtaggtgtgt (SEQ ID NO: 117)
17	17	*TRAR3*	rs2788935	C	T	Ttcctatactttatgaatcaagtaatttgatgtttactga (SEQ ID NO: 58)	Gtaaggcactgtgctaggtattgaa (SEQ ID NO: 118)

Figure 6A

FP Probe Used (5'->3')	TaqMan VIC probe (5'->3')	TaqMan FAM probe (5'->3')	Allele 1 Detected by	Allele 2 Detected by
	Cattaatcttcacattttt (SEQ ID NO: 169)	Tacattaatcttcatattttt (SEQ ID NO: 220)	VIC	FAM
Gaaaaaggtaaaaaccatttccattt (SEQ ID NO: 161)			R110	TAMRA
	Attggccacgttcct (SEQ ID NO: 170)	Tggccacattcct (SEQ ID NO: 221)	VIC	FAM
Tgcacatgtgtgcaattaagagg (SEQ ID NO: 162)			R110	TAMRA
	Ttgaatagccgtggaatg (SEQ ID NO: 171)	Tgaatagccatggaatg (SEQ ID NO: 222)	VIC	FAM
	Ttaactgatctaaatttctca (SEQ ID NO: 172)	Ttaactgatctaaatctctca (SEQ ID NO: 223)	VIC	FAM
Aagagaaatacatcaacgcctagta (SEQ ID NO: 163)			R110	TAMRA
	Ccagagtctattcgcctgg (SEQ ID NO: 173)	Ccagagtctattcacctgg (SEQ ID NO: 224)	VIC	FAM
	Aaggattaaagataattttt (SEQ ID NO: 174)	Aaggattaaagatatttttt (SEQ ID NO: 225)	VIC	FAM
	tctataagacataaaggaaggta (SEQ ID NO: 175)	Aagacataagggaaggta (SEQ ID NO: 226)	VIC	FAM
	Ctatctgtaaaataaaacac (SEQ ID NO: 176)	Ctatctgtaaaatcaaacac (SEQ ID NO: 227)	VIC	FAM
	Tgggttatcccaaacag (SEQ ID NO: 177)	Tgggttatcccgaacag (SEQ ID NO: 228)	VIC	FAM
	Aaatcactgggtgtactt (SEQ ID NO: 178)	Taaaatcactggatgtactt (SEQ ID NO: 229)	VIC	FAM
Aaattgcatatttcctttgaaaac (SEQ ID NO: 164)			R110	TAMRA
	Atctgccacgtaactt (SEQ ID NO: 179)	Aatctgccatgtaactt (SEQ ID NO: 230)	VIC	FAM
	Cttcacttcaaacaac (SEQ ID NO: 180)	Ccttcacttctaacaac (SEQ ID NO: 231)	VIC	FAM
	Cttagcagcaagtgaaaa (SEQ ID NO: 181)	Ttagcagcaaatgaaaa (SEQ ID NO: 232)	VIC	FAM

Figure 6B

Supplementary Table 2

Assay Information

18	18	*TRAR5*	rs1933988	A	C	Atccttgactcttacctgcatcatg (SEQ ID NO: 59)	Accttgcttctctatattctgaaatgtgatac (SEQ ID NO: 119)
19	19	*TRAR5*	rs8192627	A	C	Gctctattttatccttggtttaggaaagc (SEQ ID NO: 60)	Tctaaaaataaactaatggttgatgaactagcct (SEQ ID NO: 120)
20	20	*TRAR5*	rs2840836	G	A	Tcaaaccatctccccactt (SEQ ID NO: 61)	Agaaggcccagctcacttca (SEQ ID NO: 121)
21		*TRAR4*	rs2840837	A	G	Ctgactgtcaggaaaggacttagag (SEQ ID NO: 62)	Gtttttaaagaaatgtcaggtctaagatggg (SEQ ID NO: 122)
22		*TRAR4*	rs1361280	A	G	Tccctgtggcaataagagatagga (SEQ ID NO: 63)	Ggtaagttttataggctaggaggacatg (SEQ ID NO: 123)
23		*TRAR4*	rs4473885	C	T	Tcaattttataccaacaaaaattccacaggtt (SEQ ID NO: 64)	Tctttagtccattccatttcttttgagatttca (SEQ ID NO: 124)
24		*TRAR4*	rs4085406	A	G	Atgtgagatactgctacaccctgta (SEQ ID NO: 65)	Ctaagccctcttcaacacttggta (SEQ ID NO: 125)
25		*TRAR4*	rs6907909	A	G	Ccaatgcagatggaaaattgattcttaaca (SEQ ID NO: 66)	Agccacccgggtttttgtt (SEQ ID NO: 126)
26		*TRAR4*	ss28447860	C	G	Gttgcacagagaaactcaaaaggtaaaaata (SEQ ID NO: 67)	Atttgacaaaaattatattgcacaagattattggaga (SEQ ID NO: 127)
27		*TRAR4*	ss28447862	T	C	Tgaatgggtcctgtgtgaaaatcc (SEQ ID NO: 68)	Tcaccaggaggtttccaaacac (SEQ ID NO: 128)
28		*TRAR4*	ss28447876	G	A	Tcagcgtgtcctggatcct (SEQ ID NO: 69)	Gataattcctccagcccatcgt (SEQ ID NO: 129)
29		*TRAR4*	ss28447863	A	G	Gatcctgcccctcatgtacag (SEQ ID NO: 70)	Cctcctatacagtttagggcatcag (SEQ ID NO: 130)
30		*TRAR4*	ss28447865	C	T	Actgggtgttgacagattttctatcc (SEQ ID NO: 71)	Gaggatgattctgtcttgctacca (SEQ ID NO: 131)
31		*TRAR4*	rs8192624	A	G	Caggagagagagaaaagcagctaaa (SEQ ID NO: 72)	Ggcatcaattaatgaatcaatgctatatggt (SEQ ID NO: 132)
32	21	*TRAR4*	rs8192625	A	G	Gattcattaattgatgcctttatgggcttt (SEQ ID NO: 73)	Gctgagttataataagcacaccaacag (SEQ ID NO: 133)
33		*TRAR4*	ss28447866	G	A	Cctttgatttatgctttattttacccatggt (SEQ ID NO: 74)	Caaattcatggttgctgaactgttct (SEQ ID NO: 134)
34		*TRAR4*	rs7772821	C	A	Agcaaccatgaatttgttttctgaaca (SEQ ID NO: 75)	Cgcttggtaattttaaaggtatcctgaac (SEQ ID NO: 135)
34		*TRAR4*	rs7772821	C	A	Agaacagttcagcaaccatgaatttg (SEQ ID NO: 76)	Gaagagcaatttatttgctattcattcatagtctt (SEQ ID NO: 136)
35		*TRAR4*	ss28447871	G	A	Cagcctgccaaaaatttca (SEQ ID NO: 77)	Caacagcctggtcaagatga (SEQ ID NO: 137)

Figure 6C

	Ccaaaagcattcagggtt (SEQ ID NO: 182)	Aaaagcatgcagggtt (SEQ ID NO: 233)	VIC	FAM
	Ttattttaagtggagatgtttta (SEQ ID NO: 183)	Ttaagtggagctgtttta (SEQ ID NO: 234)	VIC	FAM
Caattataaaagctcaactgaacac (SEQ ID NO: 165)			R110	TAMRA
	Caaaagcatggaagacaa (SEQ ID NO: 184)	Aagcatggaaggcaa (SEQ ID NO: 235)	VIC	FAM
	Tcagtgcattcatagaag (SEQ ID NO: 185)	Cagtgcattcgtagaag (SEQ ID NO: 236)	VIC	FAM
	Catgccgatttct (SEQ ID NO: 186)	Catgccaatttct (SEQ ID NO: 237)	VIC	FAM
	Acaacaaaaagaaatggcag (SEQ ID NO: 187)	Caacaaaaagagatggcag (SEQ ID NO: 238)	VIC	FAM
	Tattgatgcttactatttaca (SEQ ID NO: 188)	Ttgatgcttactgtttaca (SEQ ID NO: 239)	VIC	FAM
	Atgctcttaatttgataaaa (SEQ ID NO: 189)	Atgctcttaatttcataaaa (SEQ ID NO: 240)	VIC	FAM
	Aaagccaaacactatgtac (SEQ ID NO: 190)	Aagccaaacactgtgtac (SEQ ID NO: 241)	VIC	FAM
	Atgtacagcggtgctg (SEQ ID NO: 191)	Atgtacagcagtgctg (SEQ ID NO: 242)	VIC	FAM
	Cacaggtgtctatgacgat (SEQ ID NO: 192)	Caggtgtctgtgacgat (SEQ ID NO: 243)	VIC	FAM
	Tttctatctttttcgcctgtcg (SEQ ID NO: 193)	Ttctatctttttcacctgtcg (SEQ ID NO: 244)	VIC	FAM
	Tcacagtgatagcattt (SEQ ID NO: 194)	Acagtggtagcattt (SEQ ID NO: 245)	VIC	FAM
	Ccctgcctatatttat (SEQ ID NO: 195)	Cctgcctgtatttat (SEQ ID NO: 246)	VIC	FAM
	Ctgaccagttacaataa (SEQ ID NO: 196)	Ctgaccagttataataa (SEQ ID NO: 247)	VIC	FAM
	Ttcgtctatccaactgc (SEQ ID NO: 197)	Ttcgtctatacaactgc (SEQ ID NO: 248)	VIC	FAM
	Ttcgtctatccaactgc (SEQ ID NO: 198)	Ttcgtctatacaactgc (SEQ ID NO: 249)	VIC	FAM
Ttggtttcttttgctttggttatttt (SEQ ID NO: 166)			R110	TAMRA

Figure 6D

Supplementary Table 2

Assay Information

36	22	*TRAR4*	rs4305745	C	T	Tctatgtcctttcttccccaaa (SEQ ID NO: 78)	Tttcctgtagatcatgacagtttct (SEQ ID NO: 138)
36	22	*TRAR4*	rs4305745	C	T	Tcctttcttccccaaatccataacc (SEQ ID NO: 79)	Cctgtagatcatgacagtttctcatactttatt (SEQ ID NO: 139)
37		*TRAR4*	ss28447873	A	del	Caaatccataaccccatgtagtca (SEQ ID NO: 80)	Tcctgtagatcatgacagtttctcatact (SEQ ID NO: 140)
38		*TRAR4*	rs7452939	G	A	Tcctttcttccccaaatcca (SEQ ID NO: 81)	Cggatttcctcacttccccta (SEQ ID NO: 141)
39		*TRAR4*	rs7745308	T	G	Aaggagatgtattaaagaagtggcattca (SEQ ID NO: 82)	Tcccatcctttccctcctacaa (SEQ ID NO: 142)
40		*TRAR4*	rs6912930	A	C	Acctgtccaattttgtttaaatgtgcat (SEQ ID NO: 83)	Tggtgttccttaaatattttccttggacaa (SEQ ID NO: 143)
41		*TRAR4*	rs6903874	T	C	Gtttctctcataagtcagtggttcttga (SEQ ID NO: 84)	Gtgattgtgtgcatttttctagagaaaga (SEQ ID NO: 144)
42		*TRAR4*	rs7765655	G	A	Aaaacttcttaccaaattctctcacaggt (SEQ ID NO: 85)	Cctgttattttttccattctcttttctcagt (SEQ ID NO: 145)
43		*TRAR4*	rs6937506	G	A	Tcaggagctgtgcactgg (SEQ ID NO: 86)	Tgaaaaaccctgtgcgagtct (SEQ ID NO: 146)
44		*TRAR4*	rs4129284	C	T	Ccccaaacatgcccatatccta (SEQ ID NO: 87)	Ggtcctttgaccaagcaaggtaata (SEQ ID NO: 147)
45		*TRAR4*	rs9321354	A	C	Gctgctgttaatgtgacctcttgat (SEQ ID NO: 88)	Ctgtatcgcttctgattcactctgt (SEQ ID NO: 148)
46	23	*PNR*	rs3813354	G	A	Tgtggcatttgctgtgtccta (SEQ ID NO: 89)	Gagagcagcaggaagttggt (SEQ ID NO: 149)
46	23	*PNR*	rs3813354	G	A	Tgtggcatttgctgtgtccta (SEQ ID NO: 90)	Gggccagggagagcag (SEQ ID NO: 150)
47	24	*TRAR2*	rs4144146	G	A	Catggagaggatcaagaagttggt (SEQ ID NO: 91)	Ctgggcaacatgatcgtaatgattt (SEQ ID NO: 151)
48	25	*TRAR2*	rs4467795	G	C	Agtaactgtaggaggcatctacctc (SEQ ID NO: 92)	Ctgttctgtgcttctaggttctgat (SEQ ID NO: 152)
49	26	*GPR57*	rs4421218	C	T	Cttctctgtttgctactctcttcct (SEQ ID NO: 93)	Agggaaggagcatgaagagtagag (SEQ ID NO: 153)
50	27	*GPR57*	rs1081074	A	C	Aagagaagcttgacattgtgcctat (SEQ ID NO: 95)	Tctcctctgtctgattcactcaaga (SEQ ID NO: 154)
51	28	*GPR58*	rs8192646	G	A	Tccaccaaaataactattccagtcattaaaaga (SEQ ID NO: 96)	Accccgaaggcaaatgct (SEQ ID NO: 155)
52	29	*GPR58*	rs4451148	T	C	Gtgtaactcagagccctatgcaa (SEQ ID NO: 97)	Ctctgctcatagccttgtgtgt (SEQ ID NO: 156)

Figure 6E

Gcagaatattcccgataaagttt (SEQ ID NO: 167)			R110	TAMRA
	Aaaatgtcagacgaaact (SEQ ID NO: 199)	Atgtcagacaaaact (SEQ ID NO: 250)	VIC	FAM
	Actttatcgaggaatatt (SEQ ID NO: 200)	Ctttatcgggaatatt (SEQ ID NO: 251)	VIC	FAM
Cataatatttgatcagtatttctcaaaact (SEQ ID NO: 168)			R110	TAMRA
	Aaatccttcttaaaatttaag (SEQ ID NO: 201)	Aatccttcttaaactttaag (SEQ ID NO: 252)	VIC	FAM
	Tcatatcaaaagagtacccatg (SEQ ID NO: 202)	Tcatatcaaaagagtccccatg (SEQ ID NO: 253)	VIC	FAM
	Atttcaatgtgactatgatct (SEQ ID NO: 203)	Tcaatgtgactgtgatct (SEQ ID NO: 254)	VIC	FAM
	Tgtataaatccataaggaagtg (SEQ ID NO: 204)	Tataaatccataagaaagtg (SEQ ID NO: 255)	VIC	FAM
	Tatcccatctgttgtactag (SEQ ID NO: 205)	Ctatcccatctattgtactag (SEQ ID NO: 256)	VIC	FAM
	Ttcccagaacctgttact (SEQ ID NO: 206)	Attcccagaacttgttact (SEQ ID NO: 257)	VIC	FAM
	Ctctctacagaactaac (SEQ ID NO: 207)	Ctctaccgaactaac (SEQ ID NO: 258)	VIC	FAM
	Cttcaaagcgcttcac (SEQ ID NO: 208)	Cttcaaagcacttcac (SEQ ID NO: 259)	VIC	FAM
	Cttcaaagcgcttcac (SEQ ID NO: 209)	Cttcaaagcacttcac (SEQ ID NO: 260)	VIC	FAM
	Atcgctcacttcaagca (SEQ ID NO: 210)	Tcgctcactttaagca (SEQ ID NO: 261)	VIC	FAM
	Ccactacttctttccc (SEQ ID NO: 211)	Cactacttgtttccc (SEQ ID NO: 262)	VIC	FAM
	Ccctctgcctctcc (SEQ ID NO: 212)	Ccctctgtctctcc (SEQ ID NO: 263)	VIC	FAM
	Cacagaaaagaaggccca (SEQ ID NO: 213)	Acagaaaagcaggccca (SEQ ID NO: 264)	VIC	FAM
	Cttctatgttggtcggtcc (SEQ ID NO: 214)	Tacttctatgttagtcggtcc (SEQ ID NO: 265)	VIC	FAM
	Caacttgtcaaactgtca (SEQ ID NO: 215)	Aacttgtcaagctgtca (SEQ ID NO: 266)	VIC	FAM

Figure 6F

Supplementary Table 2

Assay Information

53	30	*GPR58*	rs4380767	T	C	Ctagcagcctgggaaaacagt (SEQ ID NO: 98)	Cccttagggtcttgtcttttggtta (SEQ ID NO: 157)
54	31	*TRAR1*	rs7739700	T	C	Gcaacaaagagaaccgggattttaa (SEQ ID NO: 99)	Cctggcatatggagacaagct (SEQ ID NO: 158)
55	32	*TRAR1*	rs8192620	C	T	Gattaaatgtagagttcaagtagccaaacc (SEQ ID NO: 100)	Acagtcatggaccctttttcttcac (SEQ ID NO: 159)
56	33	*TRAR1*	rs4897595	G	A	Tcctctcacttctaaaacatttacctattactct (SEQ ID NO: 101)	Aatgtccagctccaggtaaaagttt (SEQ ID NO: 160)

Note: rs3813354, rs4305745, and rs7772821 were genotyped by both FP and TaqMan. Previously known SNPs are indicated by their rs number; novel SNPs are indicated by their NCBI Assay ID number (ss number) that we obtained (data to be released upon the next build of dbSNP).

Figure 6G

	Ctttggattcagtcctcagaga (SEQ ID NO: 216)	Ttggattcagtccccagaga (SEQ ID NO: 267)	VIC	FAM
	Tttggaaccacgtcagc (SEQ ID NO: 217)	Ctttggaaccatgtcagc (SEQ ID NO: 268)	VIC	FAM
	Tttgaatgatgtgttgattt (SEQ ID NO: 218)	Ctttgaatgatgtattgattt (SEQ ID NO: 269)	VIC	FAM
	Ttggaagggactttat (SEQ ID NO: 219)	Tttggaaggaactttat (SEQ ID NO: 270)	VIC	FAM

Figure 7

Supplementary Table 3
Primer Sequences for *TRAR4* amplicons

Amplicon	Size (bp)	F PCR Primer (5'->3')	R PCR Primer (5'->3')
1	521	ggctcactgcccatttgt (SEQ ID NO: 27)	caatgggaaaatccctcgat (SEQ ID NO: 28)
2	573	gaataactttcctagtaatcactgttg (SEQ ID NO: 29)	gggatgaattgctgctcataa (SEQ ID NO: 30)
3	570	caacaaggacaaaacttctcca (SEQ ID NO: 31)	agcccatcgtcatagacacc (SEQ ID NO: 32)
4	578	aggaatttgcatcagcgtgt (SEQ ID NO: 33)	aaaacaaattcatggttgctga (SEQ ID NO: 34)
5	517	tttacccatggtttaggaaagc (SEQ ID NO: 35)	ccaataatttgataaaggctattcac (SEQ ID NO: 36)
6	526	gactcttccctctgctctgg (SEQ ID NO: 37)	gaaaagaaagctagagactgcac (SEQ ID NO: 38)
7	647	tcctttagggaggatatctttcaa (SEQ ID NO: 39)	tcctcacttcccctaatgtcc (SEQ ID NO: 40)

Figure 8

Supplementary Table 4
Linkage Analyses

Order for All SNPs	Order for Screening SNPs	Gene	dbSNP Number	All = 179 Families LOD	All = 179 Families *P*-value	AA = 27 Families LOD	AA = 27 Families *P*-value	EA = 123 Families LOD	EA = 123 Families *P*-value	NIMH-IRP = 67 Families LOD	NIMH-IRP = 67 Families *P*-value	NIMH-GI = 57 Families LOD	NIMH-GI = 57 Families *P*-value	AU/US = 55 Families LOD	AU/US = 55 Families *P*-value
1	1	*MOXD1*	rs2206064	0.26	0.14	0.27	0.13	-0.01	0.6	-0.01	0.6	0.21	0.2	0.20	0.2
2	2	*MOXD1*	rs599660	1.01	0.02	0.47	0.07	0.84	0.02	0.23	0.2	0.87	0.02	0.05	0.3
3	3	*MOXD1*	rs7751860	0.06	0.3	0.04	0.3	0.02	0.4	-0.01	0.6	0.31	0.12	0.00	0.5
4	4	*MOXD1*	rs1981187	0.04	0.3	0.45	0.08	0.17	0.2	0.02	0.4	0.00	0.5	0.14	0.2
5	5	*MOXD1*	rs1338387	0.11	0.2	0.06	0.3	0.30	0.12	0.01	0.4	0.06	0.3	0.24	0.15
6	6	*MOXD1*	rs1538308	0.03	0.3	-0.05	0.7	0.22	0.2	0.03	0.4	0.01	0.4	0.01	0.4
7	7	*MOXD1*	rs2275394	1.27	0.008	1.15	0.011	0.62	0.05	0.26	0.14	1.11	0.012	0.02	0.4
8	8	*MOXD1*	rs6937815	-0.13	0.8	-0.15	0.8	-0.11	0.8	0.00	0.5	-0.49	0.9	0.00	0.4
9	9	*MOXD1*	rs3823288	0.25	0.14	0.49	0.07	0.00	0.5	0.18	0.2	0.06	0.3	0.03	0.4
10	10	*STX7*	rs1856352	-0.14	0.8	-0.08	0.7	-0.04	0.7	-0.09	0.7	-0.15	0.8	-0.04	0.7
11	11	*STX7*	rs3757299	-0.03	0.6	-0.02	0.6	-0.08	0.7	-0.03	0.6	-0.02	0.6	0.13	0.2
12	12	*STX7*	rs1002799	-0.01	0.6	0.01	0.4	-0.11	0.8	0.00	0.4	-0.04	0.7	0.03	0.4
13	13	*STX7*	rs2788942	-0.03	0.6	-0.01	0.6	-0.01	0.6	-0.09	0.7	0.10	0.2	0.05	0.3
14	14	*STX7*	rs2842884	-0.04	0.7	-0.02	0.6	-0.07	0.7	-0.01	0.6	-0.14	0.8	0.20	0.2
15	15	*STX7*	rs1591811	0.12	0.2	-0.02	0.6	0.00	0.5	0.06	0.3	0.06	0.3	-0.01	0.6
16	16	*TRAR3*	rs2842899	0.01	0.4	-0.22	0.8	0.00	0.5	0.18	0.2	-0.21	0.8	0.01	0.4
17	17	*TRAR3*	rs2788935	0.13	0.2	0.07	0.3	0.01	0.4	0.21	0.2	0.00	0.5	0.12	0.2
18	18	*TRAR5*	rs1933988	0.20	0.2	0.52	0.06	0.00	0.5	0.20	0.2	0.01	0.4	0.14	0.2
19	19	*TRAR5*	rs8192627	0.04	0.3	0.09	0.3	0.03	0.4	0.03	0.3	0.00	0.4	0.01	0.4
20	20	*TRAR5*	rs2840836	-0.02	0.6	-0.13	0.8	-0.08	0.7	0.03	0.4	-0.18	0.8	-0.03	0.6
21		*TRAR4*	rs2840837	0.06	0.3	0.45	0.08	-0.06	0.7	0.19	0.2	-0.03	0.6	0.11	0.2
22		*TRAR4*	rs1361280	0.17	0.2	0.15	0.2	-0.02	0.6	0.26	0.14	0.00	0.5	0.32	0.11
23		*TRAR4*	rs4473885	0.27	0.13	0.10	0.2	-0.01	0.6	0.48	0.07	0.00	0.6	0.32	0.11
24		*TRAR4*	rs4085406	0.22	0.2	0.19	0.2	-0.01	0.6	0.27	0.13	0.00	0.5	0.32	0.11
25		*TRAR4*	rs6907909	0.13	0.2	0.10	0.2	0.00	0.5	0.23	0.15	0.00	0.5	0.22	0.2
26		*TRAR4*	ss28447860	-0.02	0.6	-0.02	0.6	0.00	0.5	0.05	0.3	-0.22	0.8	0.25	0.14
27		*TRAR4*	rs8192624	0.27	0.13	0.00	0.5	0.04	0.3	0.25	0.14	0.04	0.3	0.01	0.4
28	21	*TRAR4*	rs8192625	0.93	0.02	-0.07	0.7	0.86	0.02	0.71	0.04	0.23	0.2	0.01	0.4
29		*TRAR4*	rs7772821	0.19	0.2	0.88	0.02	-0.04	0.7	-0.04	0.7	0.72	0.03	0.16	0.2
30		*TRAR4*	ss28447871	0.08	0.3	0.00	0.5	0.02	0.4	0.26	0.14	-0.05	0.7	-0.02	0.6
31	22	*TRAR4*	rs4305745	0.23	0.2	0.86	0.02	0.00	0.6	0.03	0.3	0.25	0.14	0.01	0.4
32		*TRAR4*	rs7745308	0.55	0.06	0.27	0.13	0.65	0.04	0.37	0.1	0.14	0.2	0.10	0.3
33		*TRAR4*	rs6912930	0.07	0.3	0.09	0.3	0.00	0.5	0.88	0.02	-0.42	0.9	0.09	0.3
34		*TRAR4*	rs6903874	1.22	0.009	0.73	0.03	0.36	0.1	0.19	0.2	1.12	0.011	0.13	0.2
35		*TRAR4*	rs7765655	0.00	0.5	0.09	0.3	-0.05	0.7	0.32	0.11	-0.29	0.9	0.08	0.3
36		*TRAR4*	rs6937506	1.76	0.002	0.78	0.03	0.50	0.07	0.39	0.09	1.51	0.004	0.13	0.2
37		*TRAR4*	rs4129284	0.13	0.2	0.16	0.2	-0.02	0.6	1.18	0.01	-0.39	0.9	0.09	0.3
38		*TRAR4*	rs9321354	0.49	0.07	0.00	0.5	0.32	0.11	0.77	0.03	0.02	0.4	0.02	0.4
39	23	*PNR*	rs3813354	0.04	0.3	0.00	0.4	-0.03	0.6	0.40	0.09	-0.15	0.8	0.21	0.2
40	24	*TRAR2*	rs4144146	0.79	0.03	0.00	0.5	0.34	0.11	0.79	0.03	0.10	0.3	0.01	0.4
41	25	*TRAR2*	rs4467795	0.33	0.11	1.06	0.013	0.00	0.5	0.57	0.05	0.02	0.4	0.01	0.4
42	26	*GPR57*	rs4421218	-0.02	0.6	0.24	0.15	-0.03	0.7	0.00	0.5	-0.11	0.8	0.20	0.2
43	27	*GPR57*	rs1081074	0.05	0.3	0.09	0.3	-0.01	0.6	-0.02	0.6	0.16	0.2	0.12	0.2
44	28	*GPR58*	rs8192646	0.91	0.02	0.14	0.2	0.76	0.03	0.35	0.1	0.48	0.07	0.11	0.2
45	29	*GPR58*	rs4451148	0.06	0.3	0.43	0.08	0.46	0.07	0.06	0.3	-0.01	0.6	0.30	0.12
46	30	*GPR58*	rs4380767	0.89	0.02	-0.04	0.7	0.82	0.03	0.70	0.04	0.20	0.2	0.10	0.3
47	31	*TRAR1*	rs7739700	1.26	0.008	0.17	0.2	1.03	0.015	1.72	0.002	0.05	0.3	-0.02	0.6
48	32	*TRAR1*	rs8192620	0.40	0.09	0.44	0.08	0.57	0.05	0.44	0.08	0.00	0.5	0.23	0.2
49	33	*TRAR1*	rs4897595	-0.11	0.8	0.15	0.2	-0.06	0.7	-0.10	0.7	-0.05	0.7	0.05	0.3

Note: Previously known SNPs are indicated by their rs number; novel SNPs are indicated by their NCBI Assay ID number (ss number) that we obtained (data to be released upon the next build of dbSNP). The number of families above are those informative for linkage, i.e., those with at least two affected members genotyped.

Figure 9

Supplementary Table 5
SNP Markers of Initial Screening and FBAT Results

Gene	Reference SNP Number	Allele 1	Allele 2	MAF	Distance	Position	Position in Gene	Over Transmitted Allele	Z Score	*P* value
MOXD1	rs599660	G	A	0.45	38,891	132,622,245	intron 8 (boundary)	2	1.482	0.14
MOXD1	rs1981187	C	T	0.31	7,476	132,661,136	intron 3	2	1.037	0.30
MOXD1	rs1338387	C	T	0.27	10,414	132,668,612	intron 3	1	0.450	0.65
MOXD1	rs1538308	A	G	0.07	12,744	132,679,026	5' flanking	1	1.162	0.25
MOXD1	rs2275394	C	T	0.30	34,763	132,691,770	5' flanking	1	0.382	0.70
MOXD1	rs6937815	G	A	0.04	16,432	132,726,533	5' flanking	1	0.853	0.39
MOXD1	rs3823288	A	T	0.24	18,648	132,742,965	5' flanking	1	0.475	0.63
STX7	rs1856352	A	G	0.05	4,181	132,761,613	3' flanking	1	0.632	0.53
STX7	rs3757299	A	C	0.22	16,651	132,765,794	intron 8	2	0.524	0.60
STX7	rs1002799	A	G	0.25	12,162	132,782,445	intron 2	1	0.344	0.73
STX7	rs2788942	G	A	0.36	10,942	132,794,607	intron 2	1	1.148	0.25
STX7	rs2842884	G	C	0.36	17,417	132,805,549	intron 1	2	2.369	0.0178
STX7	rs1591811	G	A	0.35	17,213	132,822,966	5' flanking	1	0.263	0.79
TRAR3	rs2842899	A	T	0.28	3,755	132,840,179	Lys61Stop	2	0.507	0.61
TRAR3	rs2788935	C	T	0.29	8,289	132,843,934	3' flanking	1	0.332	0.74
TRAR5	rs1933988	A	C	0.36	3,161	132,852,223	promoter	2	0.048	0.96
TRAR5	rs8192627	A	C	0.06	4,541	132,855,384	Asp328Ala	1	0.361	0.72
TRAR5	rs2840836	G	A	0.23	12,977	132,859,925	3' flanking	2	0.316	0.75
TRAR4	rs8192625	A	G	0.09	1,380	132,872,902	Cys291Tyr	2	2.058	0.0396
TRAR4	rs4305745	C	T	0.47	16,922	132,874,282	3' flanking	2	3.190	0.0014
PNR	rs3813354	G	A	0.08	5,759	132,891,204	Ala64Ala	2	1.357	0.17
TRAR2	rs4144146	G	A	0.08	4,812	132,896,963	Phe62Phe	1	1.241	0.21
TRAR2	rs4467795	G	C	0.45	9,887	132,901,775	promoter	2	1.002	0.32
GPR57	rs4421218	C	T	0.05	365	132,911,662	promoter	2	0.141	0.89
GPR57	rs1081074	A	C	0.07	7,385	132,912,027	promoter	2	2.084	0.0372
GPR58	rs8192646	G	A	0.03	1,451	132,919,412	Trp123Stop	1	0.789	0.43
GPR58	rs4451148	T	C	0.26	6,429	132,920,863	promoter	1	0.355	0.72
GPR58	rs4380767	T	C	0.31	8,241	132,927,292	promoter	2	1.518	0.13
TRAR1	rs7739700	T	C	0.34	11,316	132,935,533	3' flanking	2	0.920	0.36
TRAR1	rs8192620	C	T	0.02	1,092	132,946,849	Val288Val	1	1.720	0.09
TRAR1	rs4897595	G	A	0.07	N/A	132,947,941	promoter	1	0.409	0.68

Note: Distance is bp to next SNP; Position in bp was derived from UCSC July 2003 freeze of chromosome 6. MAF = minor allele freqency.

Figure 10

Supplementary Table 6

Mutations Detected in *TRAR4* by Sequencing 30 Schizophrenia Probands

dbSNP Number	SNP	Ancestral Allele	Local Position	Position	*TRAR4* Position	MAF	Sequence Context (20bp-SNP-20bp)
rs6907909	A-713G	A	289	132,871,318	5' flanking	0.47	gccaagtattgatgcttact[a/g]tttacaccctattgtatctt (SEQ ID NO: 1)
ss28447859	G-285A	G	717	132,871,746	5' flanking	0.04	ggatatttaaaatcaaaag[g/a]aattttatcaaattaagagc (SEQ ID NO: 2)
ss28447860	C-276G	C	726	132,871,755	5' flanking	0.14	aaatcaaaagNaattttat[C/G]aaattaagagcatgagacat (SEQ ID NO: 3)
ss28447861	A-240C	A	762	132,871,791	5' flanking	0.04	acatttatcagttgaaaca[a/c]tctccaataatcttgtgcaa (SEQ ID NO: 4)
ss28447875	T-10A	T	992	132,872,021	5' flanking	0.11	aaacttctccatatgtaaN[a/t]aacagcgttatgagcagcaa (SEQ ID NO: 5)
rs8192622	C78T	C	1,079	132,872,108	Pro26Pro	0.05	ggtcctgtgtgaaaatccc[c/t]ttctcgccgggatcccgggt (SEQ ID NO: 6)
ss28447862	T110C	T	1,111	132,872,140	Ile37Thr	0.03	atcccgggtgattctgtaca[t/c]agtgtttggctttggggctg (SEQ ID NO: 7)
ss28447876	G493A	G	1,494	132,872,523	Gly165Ser	0.02	tcctgcccctcatgtacagc[g/a]gtgctgtgttctacacaggt (SEQ ID NO: 8)
ss28447863	A518G	A	1,519	132,872,548	Tyr173Cys	0.03	tgtgttctacacaggtgtct[a/g]tgacgatgggctggaggaat (SEQ ID NO: 9)
ss28447864	C630G	C	1,631	132,872,660	Thr210Thr	0.03	ctatccttctttatacctac[c/g]tttattatgataattctgta (SEQ ID NO: 10)
ss28447865	C683T	C	1,684	132,872,713	Ala228Val	0.10	tcttgtggctagacgacagg[c/t]gaaaaagatagaaaatactg (SEQ ID NO: 11)
ss28447867	A744G	A	1,745	132,872,774	Arg248Arg	0.02	tcagagagttacaaagccag[a/g]gtggccaggagagagagaaa (SEQ ID NO: 12)
rs8192624	G793A	G	1,794	132,872,823	Val265Ile	0.10	aaaccctgggggtcacagtg[g/a]tagcatttatgatttcatgg (SEQ ID NO: 13)
rs8192625	G872A	A	1,873	132,872,902	Cys291Tyr	0.13	gggctttataacccctgcct[g/a]tatttatgagatttgctgtt (SEQ ID NO: 14)
ss28447866	G976A	G	1,977	132,873,006	Val326Ile	0.03	ggaaagcaataaaagttatt[g/a]taactggtcaggttttaaag (SEQ ID NO: 15)
rs7772821	T1046G	G	2,047	132,873,076	3' flanking	0.23	tgaacatatataagcagttg[t/g]atagacgaagttcaggatac (SEQ ID NO: 16)
rs7752618	C1083G	C	2,084	132,873,113	3' flanking	0.04	atacctttaaaattaccaag[c/g]gaaatgagtttttaaaaatc (SEQ ID NO: 17)
ss28447868	C1350G	G	2,351	132,873,380	3' flanking	0.27	atttttcctaaaaatattt[g/c]tNtttttttttttatttatt (SEQ ID NO: 18)
ss28447869	G1352T	T	2,353	132,873,382	3' flanking	0.42	tttttcctaaaaatatttNt[t/g]tttttttttttatttattcc (SEQ ID NO: 19)
ss28447872	T1437A	T	2,438	132,873,467	3' flanking	0.03	ccaaaaatttcatttgtgaa[t/a]agcctttatcaaattattgg (SEQ ID NO: 20)
ss28447871	A1480G	A	2,481	132,873,510	3' flanking	0.23	tcttttgctttggttatttt[a/g]ccacaggagtccttttaggt (SEQ ID NO: 21)
ss28447870	T1544C	T	2,545	132,873,574	3' flanking	0.10	gggagagatctcagggtgta[t/c]ggggcaatttgcaaatgaag (SEQ ID NO: 22)
ss28447874	G1994C	G	2,995	132,874,024	3' flanking	0.03	ttgtatggaaatcagtgNta[g/c]atgccttagacacaggcata (SEQ ID NO: 23)
rs4305745	A2252G	A	3,253	132,874,282	3' flanking	0.39	tcatgagaaaaatgtcagac[a/g]aaactttatcgNggaatatt (SEQ ID NO: 24)
ss28447873	A2263del	A	3,264	132,874,294	3' flanking	0.39	atgtcagacNaaactttatcg[a/-]ggaatattctgcataatatt (SEQ ID NO: 25)
rs7452939	A2305G	G	3,306	132,874,335	3' flanking	0.39	gatcagtatttctcaaaact[a/g]tcagtcatcaaaaataaagt (SEQ ID NO: 26)

Note: MAF = minor allele freqency (from 60 chromosomes); EA = European Ancestry; AA = African American. SNPs were named according to their relative position to the first letter of start codon ATG and with the first base as the major allele; CDS is 1,038 bp. Position in bp was derived from UCSC July 2003 freeze of chromosome 6. Ancestral allele was determined according to the sequence comparison between the human and two chimpanzees (accession AB180397 and AB180398) and two gorillas (accession AB180399 and AB180400). Local position refers to the 3,539 bp sequenced fragment of *TRAR4*. Previously known SNPs are indicated by their rs number; novel SNPs are indicated by their NCBI Assay ID number (ss number) that we obtained (data to be released upon the next build of dbSNP). rs4305745 was in perfect LD with ss28447873 and rs7452939.

Figure 11

Supplementary Table 7

Single Marker Association via FBAT for all markers with 10 or more informative families in the whole sample

					All Families			AA Families			EA Families		
Order for All SNPs	Order for Screening SNPs	Gene	dbSNP Number	Associated Allele	Number Informative Families	O/E	*P*-value	Number Informative Families	O/E	*P*-value	Number Informative Families	O/E	*P*-value
1	2	*MOXD1*	rs599660	2	80	1.07	0.14	8	1.59	0.10	65	1.07	0.14
2	4	*MOXD1*	rs1981187	2	56	1.08	0.30	9	0.83	0.06	43	1.23	0.06
3	5	*MOXD1*	rs1338387	1	73	1.03	0.65	16	0.92	0.48	48	1.15	0.12
4	6	*MOXD1*	rs1538308	1	32	1.16	0.25	9	1.01	0.94	18	**1.45**	**0.029**
5	7	*MOXD1*	rs2275394	1	52	1.04	0.70	7	0.83	0.18	34	1.17	0.22
6	8	*MOXD1*	rs6937815	1	13	1.06	0.39	2	1.00	1.00	7	1.18	0.21
7	9	*MOXD1*	rs3823288	1	72	1.02	0.63	19	0.99	0.92	46	1.07	0.17
8	10	*STX7*	rs1856352	1	28	1.09	0.53	7	1.29	0.29	18	0.88	0.46
9	11	*STX7*	rs3757299	2	65	1.05	0.60	12	1.06	0.78	43	1.03	0.77
10	12	*STX7*	rs1002799	1	62	1.01	0.78	9	1.01	0.92	43	1.05	0.43
11	13	*STX7*	rs2788942	2	74	1.05	0.36	13	1.13	0.24	54	1.03	0.59
12	14	*STX7*	rs2842884	2	63	**1.17**	**0.018**	8	1.32	0.13	50	**1.17**	**0.022**
13	15	*STX7*	rs1591811	1	71	1.02	0.79	11	0.92	0.59	49	1.06	0.53
14	16	*TRAR3*	rs2842899	2	72	1.04	0.61	5	0.73	0.13	58	1.10	0.19
15	17	*TRAR3*	rs2788935	1	86	1.03	0.70	17	1.16	0.34	56	1.03	0.73
16	18	*TRAR5*	rs1933988	2	98	1.00	0.96	17	**1.32**	**0.043**	68	0.95	0.48
17	19	*TRAR5*	rs8192627	1	26	1.02	0.72	4	0.83	0.25	20	1.07	0.36
18	20	*TRAR5*	rs2840836	2	63	1.03	0.75	2	0.86	0.78	52	1.07	0.38
19		*TRAR4*	rs2840837	1	70	1.01	0.84	12	1.22	0.28	47	1.05	0.56
20		*TRAR4*	rs1361280	1	91	1.06	0.24	19	1.17	0.15	60	1.10	0.07
21		*TRAR4*	rs4473885	1	97	1.06	0.21	22	**1.21**	**0.032**	63	1.07	0.19
22		*TRAR4*	rs4085406	1	92	1.06	0.25	22	**1.23**	**0.047**	59	1.06	0.30
23		*TRAR4*	rs6907909	1	85	1.07	0.18	19	**1.28**	**0.019**	56	1.05	0.36
24		*TRAR4*	ss28447860	1	54	1.08	0.08	9	1.18	0.06	41	1.07	0.22
25		*TRAR4*	rs8192624	2	31	1.03	0.53	5	1.07	0.49	23	0.99	0.89
26	21	*TRAR4*	rs8192625	2	37	**1.10**	**0.040**	9	**1.24**	**0.011**	26	1.05	0.46
27		*TRAR4*	rs7772821	1	65	1.06	0.40	20	0.89	0.30	37	**1.21**	**0.041**
28		*TRAR4*	ss28447871	2	76	1.03	0.45	9	0.96	0.67	56	1.06	0.24
29	22	*TRAR4*	rs4305745	2	78	**1.17**	**0.0014**	17	**1.20**	**0.035**	53	**1.17**	**0.015**
30		*TRAR4*	rs7745308	2	21	1.12	0.44	10	1.23	0.25	10	0.91	0.71
31		*TRAR4*	rs6912930	1	86	1.04	0.59	15	1.17	0.24	60	1.00	0.96
32		*TRAR4*	rs6903874	1	75	**1.13**	**0.0026**	18	**1.27**	**0.0082**	49	1.09	0.07
33		*TRAR4*	rs7765655	1	64	1.02	0.83	10	1.29	0.16	49	0.98	0.81
34		*TRAR4*	rs6937506	1	77	**1.12**	**0.0052**	18	**1.24**	**0.025**	51	**1.11**	**0.035**
35		*TRAR4*	rs4129284	1	86	1.06	0.31	17	1.21	0.09	58	1.01	0.88
36		*TRAR4*	rs9321354	1	60	1.05	0.46	16	0.97	0.77	31	1.12	0.32
37	23	*PNR*	rs3813354	2	41	1.14	0.17	9	1.01	0.96	22	1.18	0.26
38	24	*TRAR2*	rs4144146	1	37	1.06	0.21	11	1.06	0.49	19	1.03	0.65
39	25	*TRAR2*	rs4467795	2	83	1.06	0.32	16	0.95	0.60	60	1.12	0.11
40	26	*GPR57*	rs4421218	2	23	1.03	0.89	5	1.57	0.14	14	0.65	0.14
41	27	*GPR57*	rs1081074	2	26	**1.12**	**0.037**	16	**1.20**	**0.025**	6	1.02	0.74
42	28	*GPR58*	rs8192646	1	14	1.07	0.43	3	1.00	1.00	11	1.08	0.40
43	29	*GPR58*	rs4451148	1	63	1.03	0.72	10	1.28	0.41	47	0.97	0.78
44	30	*GPR58*	rs4380767	2	80	1.07	0.13	5	0.86	0.17	67	1.07	0.13
45	31	*TRAR1*	rs7739700	2	90	1.06	0.36	15	0.98	0.91	65	1.08	0.26
46	32	*TRAR1*	rs8192620	1	66	1.17	0.09	12	1.05	0.85	50	1.19	0.10
47	33	*TRAR1*	rs4897595	1	33	1.02	0.68	6	0.94	0.62	23	1.03	0.63

Note: Previously known SNPs are indicated by their rs number; novel SNPs are indicated by their NCBI Assay ID number (ss number) that we obtained (data to be released upon the next build of dbSNP). *P*-values < 0.05 are bolded. For the analyses of these SNPs, alleles were tested for association if there were at least 10 informative families; in our data this corresponds to not testing alleles and haplotypes rarer than 3%. This restriction, however, was not used when the investigation was limited to specific subsets of families in the secondary analyses, i.e., the AA and EA ffamilies in this table.

Figure 12

Supplementary Table 8

Comparison of Coding Variants Deteced in AA Schizophrenia Probands and AA Controls

			Missense Variants										Synonymous Variants			
Type of Subject	Number Subjects	Method	T110C Ile37Thr	G162A Met54Ile	G493A Gly165Ser	A518G Tyr173Cys	C683T Ala228Val	T692C Ile231Thr	G793A Val265Ile	G872A Cys291Tyr	G976A Val326Ile	Missense Total	C78T Pro26Pro	C630G Thr210Thr	A744G Arg248Arg	Synonymous Total
Schizophrenia	14	sequencing	2/28	0/28	1/28	2/28	6/28	0/28	1/28	3/28	2/28	17/28	0/28	1/28	1/28	2/28
			7%	0%	4%	7%	21%	0%	4%	11%	7%	61%	0%	4%	4%	7%
Schizophrenia	18	genotyping	1/34	N/A	1/36	1/34	8/32	N/A	1/36	5/34	0/34	17/36	N/A	N/A	N/A	N/A
			3%	N/A	3%	3%	25%	N/A	3%	15%	0%	47%	N/A	N/A	N/A	N/A
Controls	48	sequencing	2/96	1/96	6/94	0/96	13/96	1/96	9/96	10/96	2/92	44/96	0/96	0/96	1/96	1/96
			2%	1%	6%	0%	14%	1%	9%	10%	2%	46%	0%	0%	1%	1%

Note: Minor allele counts are the numerator, and total chromosomes assayed is denominator, leading to a percentage below the counts.

Figure 13

Supplementary Table 9
Two-Marker Haplotype Association Analysis for *TRAR4* for Haplotypes Including rs4305745

Second Marker	Chi-squared	degrees freedom	Global *P*
rs2840837	10.906	3	0.0122
rs1361280	10.365	3	0.0157
rs4473885	12.720	3	0.0053
rs4085406	10.701	3	0.0135
rs6907909	11.983	3	0.0074
ss28447860	12.824	3	0.0050
rs8192624	11.318	3	0.0101
rs8192625	12.493	3	0.0059
rs7772821	13.302	3	0.0040
ss28447871	9.825	3	0.0201
rs7745308	11.689	3	0.0085
rs6912930	11.470	3	0.0094
rs6903874	9.110	2	0.0105
rs7765655	10.686	3	0.0136
rs6937506	10.261	3	0.0165
rs4129284	13.055	3	0.0045
rs9321354	9.847	3	0.0199

Note: Only markers with minor allele frequency > 3% are included, and only haplotypes at least 3% frequenct are analyzed. Previously known SNPs are indicated by their rs number; novel SNPs are indicated by their NCBI Assay ID number (ss number) that we obtained (data to be released upon the next build of dbSNP).