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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 September 2014 (25.09.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/152965 A2

(51) International Patent Classification:
G01N 33/53 (2006.01)

(21) International Application Number:
PCT/US2014/028429

(22) International Filing Date:
14 March 2014 (14.03.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/785,464 14 March 2013 (14.03.2013) US
61/788,509 15 March 2013 (15.03.2013) US

(71) Applicant: **THE CHILDREN'S HOSPITAL OF PHILADELPHIA** [US/US]; 34th Street & Civic Center Boulevard, Philadelphia, PA 19104 (US).

(72) Inventors: **HAKONARSON, Hakon**; 1877 Covered Bridge Road, Malvern, PA 19355 (US). **SLEIMAN, Patrick**; 2411F Delancey Street, Philadelphia, PA 19103 (US).

(74) Agents: **RIGAUT, Kathleen, D.** et al.; Dann, Dorfman, Herrell & Skillman, 1601 Market Street, Suite 2400, Philadelphia, PA 19103 (US).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

WO 2014/152965 A2

(54) Title: SCHIZOPHRENIA-ASSOCIATED GENETIC LOCI IDENTIFIED IN GENOME WIDE ASSOCIATION STUDIES AND USE THEREOF AS NOVEL THERAPEUTIC TARGETS

(57) Abstract: Compositions and methods for the identification of agents useful for the treatment of neurological disorders, including schizophrenia, are provided.

SCHIZOPHRENIA-ASSOCIATED GENETIC LOCI IDENTIFIED IN GENOME WIDE ASSOCIATION STUDIES AND USE THEREOF AS NOVEL THERAPEUTIC TARGETS

By

Hakon Hakonarson

Patrick Sleiman

This application claims priority to US Provisional application Nos. 61/785,464 and 61/788,509 filed March 14, 2013 and March 15, 2013, respectively. The entire contents of each of the aforementioned applications being incorporated herein by reference as though set forth in full.

FIELD OF THE INVENTION

This invention relates to the fields of genetics and the diagnosis and treatment of schizophrenia, bi-polar disorder and autism. More specifically, the invention provides newly identified genetic loci containing gene targets strongly associated with these devastating neurological disorders for use in screening assays to identify therapeutic agents useful for the treatment of the same.

BACKGROUND OF THE INVENTION

Several publications and patent documents are cited throughout the specification in order to describe the state of the art to which this invention pertains. Each of these citations is incorporated herein by reference as though set forth in full.

Schizophrenia is a chronic, severe, and disabling brain disorder that affects about 1.1 percent of the U.S. population. People with schizophrenia sometimes hear voices others don't hear, believe that others are broadcasting their thoughts to the world, or become convinced that others are plotting to harm them. These experiences can make them fearful and withdrawn and cause difficulties when they try to have relationships with others.

People with schizophrenia may not make sense when they talk, may sit for hours without moving or talking much, or may seem perfectly fine until they talk about what they are really thinking. Because many people with schizophrenia have difficulty holding a job or caring for themselves, the burden on their families and society is significant as well.

Available treatments can relieve many of the disorder's symptoms, but most people who have schizophrenia must cope with some residual symptoms as long as they live. Clearly, a need exists for improved compositions and methods for the diagnosis and treatment of this devastating neuronal disorder.

SUMMARY OF THE INVENTION

In accordance with the present invention, using genome wide association studies (GWAS), we have identified 7 genome-wide significant loci, two of which are novel (chr15q25.2 containing neuromedin B, and chr8q24.3, TSNARE1, containing SNARE domain containing protein) and a third (chr1q43 SDCCAG8), which has been reported to be associated with schizophrenia, bipolar disease and autism. The identification of these loci provide the means to screen agents which impact the activity of the proteins encoded by the genetic loci, thereby providing new therapeutics for the treatment of these neurological disorders. Such screening assays can be performed in vitro or in vivo.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows Manhattan plot of SCZ-BP-ASD datasets meta-analysed.

Figure 2 shows Manhattan plot of SCZ-BP-ASD datasets meta-analysed, using a heterogeneity filter >0.05.

Figure 3. Chr3 beta forest plot delineating individual studies and significance of the association.

Figure 4. Chr3 regional association plot. The association signal spans multiple genes.

Figure 5 shows a forest plot for the associated region on chr 6, followed by association plot for the region (Figure 6).

Figure 6. Chr6 regional association plot. The association signal spans across the MHC locus.

Figure 7. Chr7 beta forest plot delineating individual studies and significance of the association.

Figure 8. Chr7 regional association plot.

Figure 9. Chr15 beta forest plot delineating individual studies and significance of the association.

Figure 10. Chr15 regional association plot.

Figure 11. Manhattan plot of the SCZ-BP meta-analysis. The dotted line indicates genome wide significance threshold.

Figure 12. TSNARE1 regional association plot.

Figure 13. The overall LCC with the DLG1 mediated sub-network corresponding to the post-receptor segment of the NR pathway circled and boxed. Gene-wise P-value significance denoted on a green to red color scale. Circled nodes are genes with SNPs that had P-values < 1x10-3; all others are boxed.

DETAILED DESCRIPTION OF THE INVENTION

Schizophrenia is a devastating mental disorder characterized by reality distortion. Common features are positive symptoms of hallucinations, delusions, disorganized speech and abnormal thought process, negative symptoms of social deficit, lack of motivation, anhedonia and impaired emotion processing, and cognitive deficits with occupational dysfunction. Onset of symptoms typically occurs in late adolescence or early adulthood, with approximately 1.5% of the population affected.

The present invention provides several newly identified genetic loci which contain genes strongly associated with the presence of a neurological disorder including schizophrenia, bi-polar disease and autism. The provision of these new gene targets facilitates the development of screening assays for identifying agents which modulate the activities of the encoded proteins. Such agents should have therapeutic efficacy for the treatment of neurological diseases.

The term "genetic alteration" refers to a change from the wild-type or reference sequence of one or more nucleic acid molecules. Genetic alterations include without limitation, base pair substitutions, duplications, additions and deletions of at least one nucleotide from a nucleic acid molecule of known sequence.

The term "solid matrix" as used herein refers to any format, such as beads, microparticles, a microarray, the surface of a microtitration well or a test tube, a dipstick or a filter. The material of the matrix may be polystyrene, cellulose, latex, nitrocellulose, nylon, polyacrylamide, dextran or agarose.

The phrase "consisting essentially of" when referring to a particular nucleotide or amino acid means a sequence having the properties of a given SEQ ID NO:. For example, when used in

reference to an amino acid sequence, the phrase includes the sequence per se and molecular modifications that would not affect the functional and novel characteristics of the sequence.

"Target nucleic acid" as used herein refers to a previously defined region of a nucleic acid present in a complex nucleic acid mixture wherein the defined wild-type region contains at least one known nucleotide variation which may or may not be associated with schizophrenia. The nucleic acid molecule may be isolated from a natural source by cDNA cloning or subtractive hybridization or synthesized manually. The nucleic acid molecule may be synthesized manually by the triester synthetic method or by using an automated DNA synthesizer.

With regard to nucleic acids used in the invention, the term "isolated nucleic acid" is sometimes employed. This term, when applied to DNA, refers to a DNA molecule that is separated from sequences with which it is immediately contiguous (in the 5' and 3' directions) in the naturally occurring genome of the organism from which it was derived. For example, the "isolated nucleic acid" may comprise a DNA molecule inserted into a vector, such as a plasmid or virus vector, or integrated into the genomic DNA of a prokaryote or eukaryote. An "isolated nucleic acid molecule" may also comprise a cDNA molecule. An isolated nucleic acid molecule inserted into a vector is also sometimes referred to herein as a recombinant nucleic acid molecule.

With respect to RNA molecules, the term "isolated nucleic acid" primarily refers to an RNA molecule encoded by an isolated DNA molecule as defined above. Alternatively, the term may refer to an RNA molecule that has been sufficiently separated from RNA molecules with which it would be associated in its natural state (i.e., in cells or tissues), such that it exists in a "substantially pure" form.

By the use of the term "enriched" in reference to nucleic acid it is meant that the specific DNA or RNA sequence constitutes a significantly higher fraction (2-5 fold) of the total DNA or RNA present in the cells or solution of interest than in normal cells or in the cells from which the sequence was taken. This could be caused by a person by preferential reduction in the amount of other DNA or RNA present, or by a preferential increase in the amount of the specific DNA or RNA sequence, or by a combination of the two. However, it should be noted that "enriched" does not imply that there are no other DNA or RNA sequences present, just that the relative amount of the sequence of interest has been significantly increased.

It is also advantageous for some purposes that a nucleotide sequence be in purified form. The term "purified" in reference to nucleic acid does not require absolute purity (such as a homogeneous preparation); instead, it represents an indication that the sequence is relatively purer than in the natural environment (compared to the natural level, this level should be at least 2-5 fold greater, e.g., in terms of mg/ml). Individual clones isolated from a cDNA library may be purified to electrophoretic homogeneity. The claimed DNA molecules obtained from these clones can be obtained directly from total DNA or from total RNA. The cDNA clones are not naturally occurring, but rather are preferably obtained via manipulation of a partially purified naturally occurring substance (messenger RNA). The construction of a cDNA library from mRNA involves the creation of a synthetic substance (cDNA) and pure individual cDNA clones can be isolated from the synthetic library by clonal selection of the cells carrying the cDNA library. Thus, the process which includes the construction of a cDNA library from mRNA and isolation of distinct cDNA clones yields an approximately 10^6-fold purification of the native message. Thus, purification of at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated.

The term "substantially pure" refers to a preparation comprising at least 50-60% by weight the compound of interest (e.g., nucleic acid, oligonucleotide, etc.). More preferably, the preparation comprises at least 75% by weight, and most preferably 90-99% by weight, the compound of interest. Purity is measured by methods appropriate for the compound of interest.

The term "complementary" describes two nucleotides that can form multiple favorable interactions with one another. For example, adenine is complementary to thymine as they can form two hydrogen bonds. Similarly, guanine and cytosine are complementary since they can form three hydrogen bonds. Thus if a nucleic acid sequence contains the following sequence of bases, thymine, adenine, guanine and cytosine, a "complement" of this nucleic acid molecule would be a molecule containing adenine in the place of thymine, thymine in the place of adenine, cytosine in the place of guanine, and guanine in the place of cytosine. Because the complement can contain a nucleic acid sequence that forms optimal interactions with the parent nucleic acid molecule, such a complement can bind with high affinity to its parent molecule.

With respect to single stranded nucleic acids, particularly oligonucleotides, the term "specifically hybridizing" refers to the association between two single-stranded nucleotide molecules of sufficiently complementary sequence to permit such hybridization under pre-

determined conditions generally used in the art (sometimes termed "substantially complementary"). In particular, the term refers to hybridization of an oligonucleotide with a substantially complementary sequence contained within a single-stranded DNA or RNA molecule of the invention, to the substantial exclusion of hybridization of the oligonucleotide with single-stranded nucleic acids of non-complementary sequence. For example, specific hybridization can refer to a sequence which hybridizes to any schizophrenia specific marker gene or nucleic acid, but does not hybridize to other nucleotides. Also polynucleotide which "specifically hybridizes" may hybridize only to a neurospecific specific marker, such an schizophrenia-specific marker shown in the Tables contained herein. Appropriate conditions enabling specific hybridization of single stranded nucleic acid molecules of varying complementarity are well known in the art.

For instance, one common formula for calculating the stringency conditions required to achieve hybridization between nucleic acid molecules of a specified sequence homology is set forth below (Sambrook et al., Molecular Cloning, Cold Spring Harbor Laboratory (1989):

$$Tm = 81.5"C + 16.6Log [Na+] + 0.41(\% G+C) - 0.63 (\% formamide) - 600/\#bp in duplex$$

As an illustration of the above formula, using [Na+] = [0.368] and 50% formamide, with GC content of 42% and an average probe size of 200 bases, the Tm is 57°C. The Tm of a DNA duplex decreases by 1 - 1.5°C with every 1% decrease in homology. Thus, targets with greater than about 75% sequence identity would be observed using a hybridization temperature of 42°C. The stringency of the hybridization and wash depend primarily on the salt concentration and temperature of the solutions. In general, to maximize the rate of annealing of the probe with its target, the hybridization is usually carried out at salt and temperature conditions that are 20-25°C below the calculated Tm of the hybrid. Wash conditions should be as stringent as possible for the degree of identity of the probe for the target. In general, wash conditions are selected to be approximately 12-20°C below the Tm of the hybrid. In regards to the nucleic acids of the current invention, a moderate stringency hybridization is defined as hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 μg/ml denatured salmon sperm DNA at 42°C, and washed in 2X SSC and 0.5% SDS at 55°C for 15 minutes. A high stringency hybridization is defined as hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 μg/ml

denatured salmon sperm DNA at 42°C, and washed in 1X SSC and 0.5% SDS at 65°C for 15 minutes. A very high stringency hybridization is defined as hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 μg/ml denatured salmon sperm DNA at 42°C, and washed in 0.1X SSC and 0.5% SDS at 65°C for 15 minutes.

The term "oligonucleotide," as used herein is defined as a nucleic acid molecule comprised of two or more ribo or deoxyribonucleotides, preferably more than three. The exact size of the oligonucleotide will depend on various factors and on the particular application and use of the oligonucleotide. Oligonucleotides, which include probes and primers, can be any length from 3 nucleotides to the full length of the nucleic acid molecule, and explicitly include every possible number of contiguous nucleic acids from 3 through the full length of the polynucleotide. Preferably, oligonucleotides are at least about 10 nucleotides in length, more preferably at least 15 nucleotides in length, more preferably at least about 20 nucleotides in length.

The term "probe" as used herein refers to an oligonucleotide, polynucleotide or nucleic acid, either RNA or DNA, whether occurring naturally as in a purified restriction enzyme digest or produced synthetically, which is capable of annealing with or specifically hybridizing to a nucleic acid with sequences complementary to the probe. A probe may be either single stranded or double stranded. The exact length of the probe will depend upon many factors, including temperature, source of probe and use of the method. For example, for diagnostic applications, depending on the complexity of the target sequence, the oligonucleotide probe typically contains 15-25 or more nucleotides, although it may contain fewer nucleotides. The probes herein are selected to be complementary to different strands of a particular target nucleic acid sequence. This means that the probes must be sufficiently complementary so as to be able to "specifically hybridize" or anneal with their respective target strands under a set of pre-determined conditions. Therefore, the probe sequence need not reflect the exact complementary sequence of the target. For example, a non complementary nucleotide fragment may be attached to the 5' or 3' end of the probe, with the remainder of the probe sequence being complementary to the target strand. Alternatively, non complementary bases or longer sequences can be interspersed into the probe, provided that the probe sequence has sufficient complementarity with the sequence of the target nucleic acid to anneal therewith specifically.

The term "primer" as used herein refers to an oligonucleotide, either RNA or DNA, either single stranded or double stranded, either derived from a biological system, generated by restriction enzyme digestion, or produced synthetically which, when placed in the proper environment, is able to functionally act as an initiator of template-dependent nucleic acid synthesis. When presented with an appropriate nucleic acid template, suitable nucleoside triphosphate precursors of nucleic acids, a polymerase enzyme, suitable cofactors and conditions such as a suitable temperature and pH, the primer may be extended at its 3' terminus by the addition of nucleotides by the action of a polymerase or similar activity to yield a primer extension product. The primer may vary in length depending on the particular conditions and requirement of the application. For example, in diagnostic applications, the oligonucleotide primer is typically 15 25 or more nucleotides in length. The primer must be of sufficient complementarity to the desired template to prime the synthesis of the desired extension product, that is, to be able anneal with the desired template strand in a manner sufficient to provide the 3' hydroxyl moiety of the primer in appropriate juxtaposition for use in the initiation of synthesis by a polymerase or similar enzyme. It is not required that the primer sequence represent an exact complement of the desired template. For example, a non complementary nucleotide sequence may be attached to the 5' end of an otherwise complementary primer. Alternatively, non complementary bases may be interspersed within the oligonucleotide primer sequence, provided that the primer sequence has sufficient complementarity with the sequence of the desired template strand to functionally provide a template primer complex for the synthesis of the extension product.

Polymerase chain reaction (PCR) has been described in US Patents 4,683,195, 4,800,195, and 4,965,188, the entire disclosures of which are incorporated by reference herein.

The term "vector" relates to a single or double stranded circular nucleic acid molecule that can be infected, transfected or transformed into cells and replicate independently or within the host cell genome. A circular double stranded nucleic acid molecule can be cut and thereby linearized upon treatment with restriction enzymes. An assortment of vectors, restriction enzymes, and the knowledge of the nucleotide sequences that are targeted by restriction enzymes are readily available to those skilled in the art, and include any replicon, such as a plasmid, cosmid, bacmid, phage or virus, to which another genetic sequence or element (either DNA or RNA) may be attached so as to bring about the replication of the attached sequence or element. A nucleic acid

molecule of the invention can be inserted into a vector by cutting the vector with restriction enzymes and ligating the two pieces together.

Many techniques are available to those skilled in the art to facilitate transformation, transfection, or transduction of the expression construct into a prokaryotic or eukaryotic organism. The terms "transformation", "transfection", and "transduction" refer to methods of inserting a nucleic acid and/or expression construct into a cell or host organism. These methods involve a variety of techniques, such as treating the cells with high concentrations of salt, an electric field, or detergent, to render the host cell outer membrane or wall permeable to nucleic acid molecules of interest, microinjection, PEG-fusion, and the like.

The term "promoter element" describes a nucleotide sequence that is incorporated into a vector that, once inside an appropriate cell, can facilitate transcription factor and/or polymerase binding and subsequent transcription of portions of the vector DNA into mRNA. In one embodiment, the promoter element of the present invention precedes the 5' end of the schizophrenia specific marker nucleic acid molecule such that the latter is transcribed into mRNA. Host cell machinery then translates mRNA into a polypeptide.

Those skilled in the art will recognize that a nucleic acid vector can contain nucleic acid elements other than the promoter element and the schizophrenia specific marker gene nucleic acid molecule. These other nucleic acid elements include, but are not limited to, origins of replication, ribosomal binding sites, nucleic acid sequences encoding drug resistance enzymes or amino acid metabolic enzymes, and nucleic acid sequences encoding secretion signals, localization signals, or signals useful for polypeptide purification.

A "replicon" is any genetic element, for example, a plasmid, cosmid, bacmid, plastid, phage or virus, that is capable of replication largely under its own control. A replicon may be either RNA or DNA and may be single or double stranded.

An "expression operon" refers to a nucleic acid segment that may possess transcriptional and translational control sequences, such as promoters, enhancers, translational start signals (e.g., ATG or AUG codons), polyadenylation signals, terminators, and the like, and which facilitate the expression of a polypeptide coding sequence in a host cell or organism.

As used herein, the terms "reporter," "reporter system", "reporter gene," or "reporter gene product" shall mean an operative genetic system in which a nucleic acid comprises a gene that encodes a product that when expressed produces a reporter signal that is a readily measurable,

e.g., by biological assay, immunoassay, radio immunoassay, or by colorimetric, fluorogenic, chemiluminescent or other methods. The nucleic acid may be either RNA or DNA, linear or circular, single or double stranded, antisense or sense polarity, and is operatively linked to the necessary control elements for the expression of the reporter gene product. The required control elements will vary according to the nature of the reporter system and whether the reporter gene is in the form of DNA or RNA, but may include, but not be limited to, such elements as promoters, enhancers, translational control sequences, poly A addition signals, transcriptional termination signals and the like.

The introduced nucleic acid may or may not be integrated (covalently linked) into nucleic acid of the recipient cell or organism. In bacterial, yeast, plant and mammalian cells, for example, the introduced nucleic acid may be maintained as an episomal element or independent replicon such as a plasmid. Alternatively, the introduced nucleic acid may become integrated into the nucleic acid of the recipient cell or organism and be stably maintained in that cell or organism and further passed on or inherited to progeny cells or organisms of the recipient cell or organism. Finally, the introduced nucleic acid may exist in the recipient cell or host organism only transiently.

The term "selectable marker gene" refers to a gene that when expressed confers a selectable phenotype, such as antibiotic resistance, on a transformed cell.

The term "operably linked" means that the regulatory sequences necessary for expression of the coding sequence are placed in the DNA molecule in the appropriate positions relative to the coding sequence so as to effect expression of the coding sequence. This same definition is sometimes applied to the arrangement of transcription units and other transcription control elements (e.g. enhancers) in an expression vector.

The terms "recombinant organism," or "transgenic organism" refer to organisms which have a new combination of genes or nucleic acid molecules. A new combination of genes or nucleic acid molecules can be introduced into an organism using a wide array of nucleic acid manipulation techniques available to those skilled in the art. The term "organism" relates to any living being comprised of a least one cell. An organism can be as simple as one eukaryotic cell or as complex as a mammal. Therefore, the phrase "a recombinant organism" encompasses a recombinant cell, as well as eukaryotic and prokaryotic organism.

The term "isolated protein" or "isolated and purified protein" is sometimes used herein. This term refers primarily to a protein produced by expression of an isolated nucleic acid molecule of the invention. Alternatively, this term may refer to a protein that has been sufficiently separated from other proteins with which it would naturally be associated, so as to exist in "substantially pure" form. "Isolated" is not meant to exclude artificial or synthetic mixtures with other compounds or materials, or the presence of impurities that do not interfere with the fundamental activity, and that may be present, for example, due to incomplete purification, addition of stabilizers, or compounding into, for example, immunogenic preparations or pharmaceutically acceptable preparations.

A "specific binding pair" comprises a specific binding member (sbm) and a binding partner (bp) which have a particular specificity for each other and which in normal conditions bind to each other in preference to other molecules. Examples of specific binding pairs are antigens and antibodies, ligands and receptors and complementary nucleotide sequences. The skilled person is aware of many other examples. Further, the term "specific binding pair" is also applicable where either or both of the specific binding member and the binding partner comprise a part of a large molecule. In embodiments in which the specific binding pair comprises nucleic acid sequences, they will be of a length to hybridize to each other under conditions of the assay, preferably greater than 10 nucleotides long, more preferably greater than 15 or 20 nucleotides long.

"Sample" or "patient sample" or "biological sample" generally refers to a sample which may be tested for a particular molecule, preferably a schizophrenia specific marker molecule, such as a marker shown in the tables provided below. Samples may include but are not limited to cells, body fluids, including blood, serum, plasma, urine, saliva, tears, pleural fluid and the like.

The terms "agent" and "test compound" are used interchangeably herein and denote a chemical compound, a mixture of chemical compounds, a biological macromolecule, or an extract made from biological materials such as bacteria, plants, fungi, or animal (particularly mammalian) cells or tissues. Biological macromolecules include siRNA, shRNA, antisense oligonucleotides, peptides, peptide/DNA complexes, and any nucleic acid based molecule which exhibits the capacity to modulate the activity of the CNV containing nucleic acids described

herein or their encoded proteins. Agents are evaluated for potential biological activity by inclusion in screening assays described hereinbelow.

METHODS OF USING NEUROLOGICAL DISEASE-ASSOCIATED GENETIC LOCI FOR DEVELOPMENT OF THERAPEUTIC AGENTS

Since the genetic loci identified herein have been associated with the etiology of schizophrenia, bi-polar disease and autism, methods for identifying agents that modulate the activity of the genes and their encoded products should result in the generation of efficacious therapeutic agents for the treatment of this condition.

As can be seen from the data provided herein, several chromosomes contain regions which provide suitable targets for the rational design of therapeutic agents which modulate their activity. Specific organic molecules can thus be identified with capacity to bind to the active site of the proteins encoded by the identified nucleic acids based on conformation or key amino acid residues required for function. A combinatorial chemistry approach will be used to identify molecules with greatest activity and then iterations of these molecules will be developed for further cycles of screening. In certain embodiments, candidate agents can be screening from large libraries of synthetic or natural compounds. Such compound libraries are commercially available from a number of companies, including but not limited to Maybridge Chemical Co., (Trevillet,Cornwall, UK), Comgenex (Princeton, NJ), Microsour (New Milford, CT) Aldrich (Milwaukee, WI) Akos Consulting and Solutions GmbH (Basel, Switzerland), Ambinter (Paris, France), Asinex (Moscow, Russia) Aurora (Graz, Austria), BioFocus DPI (Switzerland), Bionet (Camelford, UK), Chembridge (San Diego, CA), Chem Div (San Diego, CA). The skilled person is aware of other sources and can readily purchase the same. Once therapeutically efficacious compounds are identified in the screening assays described herein, they can be formulated in to pharmaceutical compositions and utilized for the treatment of schizophrenia.

The polypeptides or fragments employed in drug screening assays may either be free in solution, affixed to a solid support or within a cell. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant polynucleotides expressing the polypeptide or fragment, preferably in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One

may determine, for example, formation of complexes between the polypeptide or fragment and the agent being tested, or examine the degree to which the formation of a complex between the polypeptide or fragment and a known substrate is interfered with by the agent being tested.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity for the encoded polypeptides and is described in detail in Geysen, PCT published application WO 84/03564, published on Sep. 13, 1984. Briefly stated, large numbers of different, small peptide test compounds, such as those described above, are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with the target polypeptide and washed. Bound polypeptide is then detected by methods well known in the art.

A further technique for drug screening involves the use of host eukaryotic cell lines or cells (such as described above) which have a nonfunctional or altered schizophrenia associated gene. These host cell lines or cells are defective at the polypeptide level. The host cell lines or cells are grown in the presence of drug compound. The rate of cellular metabolism of the host cells is measured to determine if the compound is capable of regulating the cellular metabolism in the defective cells. Host cells contemplated for use in the present invention include but are not limited to bacterial cells, fungal cells, insect cells, mammalian cells, and plant cells. The schizophrenia-associated DNA molecules may be introduced singly into such host cells or in combination to assess the phenotype of cells conferred by such expression. Methods for introducing DNA molecules are also well known to those of ordinary skill in the art. Such methods are set forth in Ausubel et al. eds., Current Protocols in Molecular Biology, John Wiley & Sons, NY, N.Y. 1995, the disclosure of which is incorporated by reference herein.

A wide variety of expression vectors are available that can be modified to express the novel DNA sequences of this invention. The specific vectors exemplified herein are merely illustrative, and are not intended to limit the scope of the invention. Expression methods are described by Sambrook et al. Molecular Cloning: A Laboratory Manual or Current Protocols in Molecular Biology 16.3-17.44 (1989). Expression methods in Saccharomyces are also described in Current Protocols in Molecular Biology (1989).

Suitable vectors for use in practicing the invention include prokaryotic vectors such as the pNH vectors (Stratagene Inc., 11099 N. Torrey Pines Rd., La Jolla, Calif. 92037), pET vectors (Novogen Inc., 565 Science Dr., Madison, Wis. 53711) and the pGEX vectors

(Pharmacia LKB Biotechnology Inc., Piscataway, N.J. 08854). Examples of eukaryotic vectors useful in practicing the present invention include the vectors pRc/CMV, pRc/RSV, and pREP (Invitrogen, 11588 Sorrento Valley Rd., San Diego, Calif. 92121); pcDNA3.1/V5&His (Invitrogen); baculovirus vectors such as pVL1392, pVL1393, or pAC360 (Invitrogen); and yeast vectors such as YRP17, YIP5, and YEP24 (New England Biolabs, Beverly, Mass.), as well as pRS403 and pRS413 Stratagene Inc.); Picchia vectors such as pHIL-D1 (Phillips Petroleum Co., Bartlesville, Okla. 74004); retroviral vectors such as PLNCX and pLPCX (Clontech); and adenoviral and adeno-associated viral vectors.

Promoters for use in expression vectors of this invention include promoters that are operable in prokaryotic or eukaryotic cells. Promoters that are operable in prokaryotic cells include lactose (lac) control elements, bacteriophage lambda (pL) control elements, arabinose control elements, tryptophan (trp) control elements, bacteriophage T7 control elements, and hybrids thereof. Promoters that are operable in eukaryotic cells include Epstein Barr virus promoters, adenovirus promoters, SV40 promoters, Rous Sarcoma Virus promoters, cytomegalovirus (CMV) promoters, baculovirus promoters such as AcMNPV polyhedrin promoter, Picchia promoters such as the alcohol oxidase promoter, and Saccharomyces promoters such as the gal4 inducible promoter and the PGK constitutive promoter, as well as neuronal-specific platelet-derived growth factor promoter (PDGF), the Thy-1 promoter, the hamster and mouse Prion promoter (MoPrP), and the Glial fibrillar acidic protein (GFAP) for the expression of transgenes in glial cells.

In addition, a vector of this invention may contain any one of a number of various markers facilitating the selection of a transformed host cell. Such markers include genes associated with temperature sensitivity, drug resistance, or enzymes associated with phenotypic characteristics of the host organisms.

Host cells expressing the schizophrenia-associated nucleic acids of the present invention or functional fragments thereof provide a system in which to screen potential compounds or agents for the ability to modulate the development of neurological disease. Thus, in one embodiment, the nucleic acid molecules of the invention may be used to create recombinant cell lines for use in assays to identify agents which modulate aspects of cellular metabolism associated with neuronal signaling and neuronal cell communication and structure. Also

provided herein are methods to screen for compounds capable of modulating the function of proteins encoded by the nucleic acids identified herein.

Another approach entails the use of phage display libraries engineered to express fragment of the polypeptides encoded by the nucleic acids of the invention on the phage surface. Such libraries are then contacted with a combinatorial chemical library under conditions wherein binding affinity between the expressed peptide and the components of the chemical library may be detected. US Patents 6,057,098 and 5,965,456 provide methods and apparatus for performing such assays.

The goal of rational drug design is to produce structural analogs of biologically active polypeptides of interest or of small molecules with which they interact (e.g., agonists, antagonists, inhibitors) in order to fashion drugs which are, for example, more active or stable forms of the polypeptide, or which, e.g., enhance or interfere with the function of a polypeptide in vivo. See, e.g., Hodgson, (1991) Bio/Technology 9:19-21. In one approach, discussed above, the three-dimensional structure of a protein of interest or, for example, of the protein-substrate complex, is solved by x-ray crystallography, by nuclear magnetic resonance, by computer modeling or most typically, by a combination of approaches. Less often, useful information regarding the structure of a polypeptide may be gained by modeling based on the structure of homologous proteins. An example of rational drug design is the development of HIV protease inhibitors (Erickson et al., (1990) Science 249:527-533). In addition, peptides may be analyzed by an alanine scan (Wells, (1991) Meth. Enzym. 202:390-411). In this technique, an amino acid residue is replaced by Ala, and its effect on the peptide's activity is determined. Each of the amino acid residues of the peptide is analyzed in this manner to determine the important regions of the peptide.

It is also possible to isolate a target-specific antibody, selected by a functional assay, and then to solve its crystal structure. In principle, this approach yields a pharmacore upon which subsequent drug design can be based.

One can bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original molecule. The anti-id could then be used to identify and isolate peptides from banks of

chemically or biologically produced banks of peptides. Selected peptides would then act as the pharmacore.

Thus, one may design drugs which have, e.g., improved polypeptide activity or stability or which act as inhibitors, agonists, antagonists, etc. of polypeptide activity. By virtue of the availability of schizophrenia associated nucleic acid sequences described herein, sufficient amounts of the encoded polypeptide may be made available to perform such analytical studies as x-ray crystallography. In addition, the knowledge of the protein sequence provided herein will guide those employing computer modeling techniques in place of, or in addition to x-ray crystallography.

In another embodiment, the availability of schizophrenia-associated nucleic acids enables the production of strains of laboratory mice carrying the schizophrenia-associated genes of the invention. Transgenic mice expressing the schizophrenia-associated nucleic acids of the invention provide a model system in which to examine the role of the protein encoded by the nucleic acid in the development and progression towards schizophrenia. Methods of introducing transgenes in laboratory mice are known to those of skill in the art. Three common methods include: 1. integration of retroviral vectors encoding the foreign gene of interest into an early embryo; 2. injection of DNA into the pronucleus of a newly fertilized egg; and 3. the incorporation of genetically manipulated embryonic stem cells into an early embryo. Production of the transgenic mice described above will facilitate the molecular elucidation of the role that a target protein plays in various cellular metabolic, neuronal and cognitive processes. Such mice provide an in vivo screening tool to study putative therapeutic drugs in a whole animal model and are encompassed by the present invention.

The term "animal" is used herein to include all vertebrate animals, except humans. It also includes an individual animal in all stages of development, including embryonic and fetal stages. A "transgenic animal" is any animal containing one or more cells bearing genetic information altered or received, directly or indirectly, by deliberate genetic manipulation at the subcellular level, such as by targeted recombination or microinjection or infection with recombinant virus. The term "transgenic animal" is not meant to encompass classical cross-breeding or in vitro fertilization, but rather is meant to encompass animals in which one or more cells are altered by or receive a recombinant DNA molecule. This molecule may be specifically targeted to a defined genetic locus, be randomly integrated within a chromosome, or it may be extrachromosomally

replicating DNA. The term "germ cell line transgenic animal" refers to a transgenic animal in which the genetic alteration or genetic information was introduced into a germ line cell, thereby conferring the ability to transfer the genetic information to offspring. If such offspring, in fact, possess some or all of that alteration or genetic information, then they, too, are transgenic animals.

The alteration of genetic information may be foreign to the species of animal to which the recipient belongs, or foreign only to the particular individual recipient, or may be genetic information already possessed by the recipient. In the last case, the altered or introduced gene may be expressed differently than the native gene. Such altered or foreign genetic information would encompass the introduction of schizophrenia-associated nucleotide sequences.

The DNA used for altering a target gene may be obtained by a wide variety of techniques that include, but are not limited to, isolation from genomic sources, preparation of cDNAs from isolated mRNA templates, direct synthesis, or a combination thereof.

A preferred type of target cell for transgene introduction is the embryonal stem cell (ES). ES cells may be obtained from pre-implantation embryos cultured in vitro (Evans et al., (1981) Nature 292:154-156; Bradley et al., (1984) Nature 309:255-258; Gossler et al., (1986) Proc. Natl. Acad. Sci. 83:9065-9069). Transgenes can be efficiently introduced into the ES cells by standard techniques such as DNA transfection or by retrovirus-mediated transduction. The resultant transformed ES cells can thereafter be combined with blastocysts from a non-human animal. The introduced ES cells thereafter colonize the embryo and contribute to the germ line of the resulting chimeric animal.

One approach to the problem of determining the contributions of individual genes and their expression products is to use isolated schizophrenia-associated genes as insertional cassettes to selectively inactivate a wild-type gene in totipotent ES cells (such as those described above) and then generate transgenic mice. The use of gene-targeted ES cells in the generation of gene-targeted transgenic mice was described, and is reviewed elsewhere (Frohman et al., (1989) Cell 56:145-147; Bradley et al., (1992) Bio/Technology 10:534-539).

Techniques are available to inactivate or alter any genetic region to a mutation desired by using targeted homologous recombination to insert specific changes into chromosomal alleles. However, in comparison with homologous extrachromosomal recombination, which occurs at a frequency approaching 100%, homologous plasmid-chromosome recombination was originally

reported to only be detected at frequencies between 10^{-6} and 10^{-3}. Nonhomologous plasmid-chromosome interactions are more frequent occurring at levels 10^{5}-fold to 10^{2} fold greater than comparable homologous insertion.

To overcome this low proportion of targeted recombination in murine ES cells, various strategies have been developed to detect or select rare homologous recombinants. One approach for detecting homologous alteration events uses the polymerase chain reaction (PCR) to screen pools of transformant cells for homologous insertion, followed by screening of individual clones. Alternatively, a positive genetic selection approach has been developed in which a marker gene is constructed which will only be active if homologous insertion occurs, allowing these recombinants to be selected directly. One of the most powerful approaches developed for selecting homologous recombinants is the positive-negative selection (PNS) method developed for genes for which no direct selection of the alteration exists. The PNS method is more efficient for targeting genes which are not expressed at high levels because the marker gene has its own promoter. Non-homologous recombinants are selected against by using the Herpes Simplex virus thymidine kinase (HSV-TK) gene and selecting against its nonhomologous insertion with effective herpes drugs such as gancyclovir (GANC) or (1-(2-deoxy-2-fluoro-B-D arabinofluranosyl)-5-iodou- racil, (FIAU). By this counter selection, the number of homologous recombinants in the surviving transformants can be increased. Utilizing schizophrenia-associated nucleic acid as a targeted insertional cassette provides means to detect a successful insertion as visualized, for example, by acquisition of immunoreactivity to an antibody immunologically specific for the polypeptide encoded by schizophrenia-associated nucleic acid and, therefore, facilitates screening/selection of ES cells with the desired genotype.

As used herein, a knock-in animal is one in which the endogenous murine gene, for example, has been replaced with human schizophrenia-associated gene of the invention. Such knock-in animals provide an ideal model system for studying the development of schizophrenia.

As used herein, the expression of a schizophrenia-associated nucleic acid, fragment thereof, or an schizophrenia-associated fusion protein can be targeted in a "tissue specific manner" or "cell type specific manner" using a vector in which nucleic acid sequences encoding all or a portion of schizophrenia-associated nucleic acid are operably linked to regulatory sequences (e.g., promoters and/or enhancers) that direct expression of the encoded protein in a particular tissue or cell type. Such regulatory elements may be used to advantage for both in vitro

and in vivo applications. Promoters for directing tissue specific proteins are well known in the art and described herein.

The nucleic acid sequence encoding the schizophrenia-associated genes of the invention may be operably linked to a variety of different promoter sequences for expression in transgenic animals. Such promoters include, but are not limited to a prion gene promoter such as hamster and mouse Prion promoter (MoPrP), described in U.S. Pat. No. 5,877,399 and in Borchelt et al., Genet. Anal. 13(6) (1996) pages 159-163; a rat neuronal specific enolase promoter, described in U.S. Pat. Nos. 5,612,486, and 5,387,742; a platelet-derived growth factor B gene promoter, described in U.S. Pat. No. 5,811,633; a brain specific dystrophin promoter, described in U.S. Pat. No. 5,849,999; a Thy-1 promoter; a PGK promoter; a CMV promoter; a neuronal-specific platelet-derived growth factor B gene promoter; and Glial fibrillar acidic protein (GFAP) promoter for the expression of transgenes in glial cells.

Methods of use for the transgenic mice of the invention are also provided herein. Transgenic mice into which a nucleic acid containing the schizophrenia-associated gene or its encoded protein have been introduced are useful, for example, to develop screening methods to screen therapeutic agents to identify those capable of modulating the development of schizophrenia.

PHARMACEUTICALS AND PEPTIDE THERAPIES

The elucidation of the role played by the schizophrenia associated genes described herein in neuronal signaling and brain structure facilitates the development of pharmaceutical compositions useful for treatment and diagnosis of schizophrenia. These compositions may comprise, in addition to one of the above substances, a pharmaceutically acceptable excipient, carrier, buffer, stabilizer or other materials well known to those skilled in the art. Such materials should be non-toxic and should not interfere with the efficacy of the active ingredient. The precise nature of the carrier or other material may depend on the route of administration, e.g. oral, intravenous, cutaneous or subcutaneous, nasal, intramuscular, intraperitoneal routes.

Whether it is a polypeptide, antibody, peptide, nucleic acid molecule, small molecule or other pharmaceutically useful compound according to the present invention that is to be given to an individual, administration is preferably in a "prophylactically effective amount" or a

"therapeutically effective amount" (as the case may be, although prophylaxis may be considered therapy), this being sufficient to show benefit to the individual.

The following examples are provided to illustrate certain embodiments of the invention. They are not intended to limit the invention in any way.

5

EXAMPLE I

Schizophrenia is a neuropsychiatric disorder with world-wide prevalence of 1%, causing a huge social/economic burden. While the etiology of schizophrenia remains unknown, genetic factors play a key role as the disease is highly heritable. Here, we used genome-wide association
10 (GWAS) approach to uncover variants that associate with schizophrenia, uncovering novel biological pathways that may lead to new treatments.

We have total of 18,069 schizophrenia, bipolar and autism samples together with 47,440 control samples, all with GWAS data and were meta-analyzed together. See the Manhattan plots shown in Figures 1 and 2. We identified several genome wide significant loci on chromosomes
15 15, 8, and 7.

Genome-wide significant loci

- chr15q25.2 neuromedin B – novel
  - Neuropeptide, associated with obesity and eating behavior.
  - Association between SCZ and obesity reported by Kraepelin as early as 1919. 40-60% of SCZ population obese.
- chr8q24.3 TSNARE1 – novel
  - SNARE domain containing protein
- chr7p22.3 MAD1L1 – not previously GWS
  - MAD1L1 is a component of the mitotic spindle-assembly checkpoint that prevents the onset of anaphase until all chromosome are properly aligned at the metaphase plate
  - NUDT1 Misincorporation of oxidized nucleoside triphosphates into DNA/RNA during replication and transcription can cause mutations that may result in carcinogenesis or neurodegeneration. The protein encoded by this gene is an enzyme that hydrolyzes oxidized purine nucleoside triphosphates, such as 8-oxo-dGTP, 8-oxo-dATP, 2-hydroxy-dATP, and 2-hydroxy rATP, to monophosphates, thereby preventing misincorporation. The encoded protein is localized mainly in the cytoplasm, with some in the mitochondria, suggesting that it is involved in the sanitization of nucleotide pools both for nuclear and mitochondrial genomes.
- chr6p22.1 MHC Locus previously associated with SCZ BP and mood disorders
- chr3p21.1 PBRM1/ITIH3 Locus previously associated with SCZ BP and mood disorders
- chr1q43 SDCCAG8 Locus previously associated with SCZ BP and mood disorders not associated with autism

We also identified several targets on chromosome 3. See Figures 3 and 4. These targets included The CACNA1D (calcium channel, voltage-dependent, L type, alpha 1D subunit) gene

– a member of the CACN family of genes showing significance in previous CNV analyses. Other genes in the region that potentially also associate include: NT5DC2 5'-nucleotidase domain containing 2; ITH1 Homo sapiens inter-alpha-trypsin inhibitor heavy chain 1; NEK4 NIMA (never in mitosis gene a)-related kinase 4 (NEK4), transcript variant 2; GNL3 guanine nucleotide binding protein-like 3; PB1 polybromo 1 isoform 4; and GLT8D1 Glycosyltransferase 8 domain-containing protein 1.

We also identified targets on chromosome 6. See figures 5 and 6.

Several targets were identified on chromosome 7. See Figures 7 and 8. These genes include MAD1L1 Mitotic spindle assembly checkpoint protein MAD1; FTSJ2 S-adenosylmethionine-binding protein involved in processing and modification of rRNA; NUDT1 nudix-type motif 1 involved in the sanitization of nucleotide pools both for nuclear and mitochondrial genomes; and SNX8 sorting nexin 8 involved in several stages of intracellular trafficking.

Finally targets were also identified on chromosome 15. See Figures 9 and 10. The association signal on chromosome 15 spans several genes, including NMB neuromedin B (bombesin) which can signal satiety and modulate 5-HT. Other genes in the associated region include SEC11A signal peptidase; SCAND2 non coding RNA; ZSCAN2 zinc finger protein 29 transcriptional regulation; ALPK3 alpha-kinase 3; and PDE8A phosphodiesterase 8A which hydrolyzes the second messenger cAMP.

All of these genes provide targets for screening assays to identify therapeutic agents for the treatment of neurological disorders. Taken together, we have identified 7 genome-wide significant loci, two of which are novel (chr15q25.2 containing neuromedin B, and chr8q24.3 TSNARE1, containing SNARE domain containing protein) and a third (chr1q43 SDCCAG8), which has been reported to be associated with schizophrenia and bipolar disease, but we show for the first time is also associated with autism.

Current pharmacological treatment of schizophrenia is limited to the typical and atypical antipsychotics, which almost exclusively targeted dopamine and serotonin receptors. Our inventions yield new insights in to the genetics of schizophrenia and neuropsychiatric diseases and new targets for therapeutic and diagnostic developments and products.

EXAMPLE II

A meta analysis of 13,394 schizophrenia and biploar cases and 34,676 controls, from 16 cohorts, was carried out to identify novel psychosis susceptibility loci. Following meta analysis, 40 variants at 6 loci surpassed genome wide significance. Five of the 6 loci have been associated with SCZ or BP. Two of the genome wide significant variants that mapped to one locus, TSNARE1, had not previously been shown to associate with either schizophrenia or bipolar disorder. The function of TSNARE1 is unclear, however, bio-informatic predictions based on phylogenetic ancestry indicate it may have a vertebrate specific function in intracellular protein transport and synaptic vesicle exocytosis.

Methods

Study cohorts. The study included 13,394 schizophrenia and bipolar cases and 34,676 controls (Table 1) including 3,182 schizophrenia cases and 1,032 bipolar cases collected from 28 clinical trials conducted by Janssen Research & Development, LLC, these were matched to 15,277 and 8,000 controls respectively from the Children's Hospital of Philadelphia (CHOP). All JNJ cases were genotyped on the Illumina 1M and CHOP controls on either the Illumina HH550 or 610 Quad arrays.

In addition, 351 schizophrenia cases and 2,107 control subjects from the University of Pennsylvania (UPenn) were included, along with 806 schizophrenia cases from Mount Sinai School of Medicine and Sheba Medical Center. Both cohorts were genotyped on the Affymetrix 6.0 array at The Children's Hospital of Philadelphia (CHOP) as previously described(1).
The remaining 8,023 schizophrenia cases and 9,292 controls formed part of the Schizophrenia Psychiatric Genome-Wide Association Study Consortium (PGC), as previously described(2), and were downloaded from the NIMH website as schizophrenia distribution 9 (www.nimhgenetics.org/).

Collection	Case	Control	Platform
JNJ SCZ / CHOP control	3182	15277	Illumina 1M / Illumina 550K
PGC MGS2	2681	2653	Affymetrix 6.0
UPENN / MS	1157	2107	Affymetrix 6.0
JNJ BP / CHOP control	1032	8000	Illumina 1M / Illumina 550K

PGC Aberdeen	720	699	Affymetrix 5.0
PGC UCLA	705	637	Illumina 550K
PGC Bulgaria	527	609	Affymetrix 6.0
PGC UCL	521	494	Affymetrix 5.0
PGC Cardiff / 58BC	475	1494	Affymetrix 500K
PGC catie	410	391	Affymetrix 500K
PGC Sweden1&2	558	396	Affymetrix 5.0 / Affymetrix 6.0
PGC Edinburgh	368	284	Affymetrix 6.0
PGC Potugal	346	216	Affymetrix 5.0
PGC Dublin	272	860	Affymetrix 6.0
PGC TOP3 (Norway)	248	369	Affymetrix 6.0
PGC Zucker Hillside	192	190	Affymetrix 500K
Total	13394	34676	

Table 1

Description of JNJ Samples. The unrelated schizophrenia (SZ), schizoaffective (SA), or bipolar I (BP) patients were from 28 clinical trials (Table 1) conducted by Janssen Research & Development, LLC to assess the efficacy and safety of risperidone, paliperidone and an investigative compound (R209130). The diagnoses of SZ, SA, and BP were based on clinician rated DSM IV criteria. Detailed descriptions of these clinical trials can be found at ClinicalTrials.gov as well as in published works (4-33) thus are not repeated here.

A total of 5544 DNA samples from 5431 patients and 49 quality control (QC) samples were genotyped on the Illumina Human1M-Duo. DNA samples from all patients who participated in these clinical trials and consented to the genetic study were genotyped for 21 out of the 28 clinical trials. A small number of DNA samples from the remaining 7 clinical trials were also genotyped (Table 2). The DNA samples were genotyped in 2 batches, with 3102 samples in the first batch and 2491 samples in the second batch. Genotype data were successfully generated on 5508 samples. A few sample QC steps were performed to remove the duplicated and/or problematic samples. First, gender discrepancies were examined using both the heterozygosity rate of the X-chromosome SNPs and the call rate of the Y-chromosome SNPs. Samples with discrepant and ambiguous gender information were excluded. Second, the relatedness of the genotyped samples was examined using pairwise Identity-by-State. Planned but not confirmed duplicates as well as unplanned duplicates with discrepant phenotype data were excluded from subsequent analyses. For each pair of samples that were planned and

confirmed duplicates, unplanned duplicates with consistent phenotype data, or samples of related individuals, the sample with a smaller standard deviation of the LogR-ratio (LRR) was retained. After the sample QC, there were 4962 samples (3251 SZ, 377 SA, and 1334 BP) remaining. Table 3 summarizes the basic demographic information of these patients.

5

Trial	Trial ID	ClinicalTrials.gov Ident-ifier	Drug*	Dis-ease	# of Patients Geno-typed	All Samples Geno-typed	Geno-typing Batch	Publi cation	PMID
R076477-SCH-101	CR 004273	NCT 00791232	Pali ER, Ris	SZ	1	No	1	(4)	
R076477-SCH-301	CR 004384	NCT 00086320	Pali ER, Pbo	SZ	187	Yes	1	(5)	1722 4706
R076477-SCH-302	CR 004381	NCT 00085748	Pali ER, Pbo	SZ	93	Yes	1	(6)	1816 5460
R076477-SCH-303	CR 003379	NCT 00078039	Pali ER, Olz, Pbo	SZ	473	Yes	1	(7), (10)	1709 2691, 1846 6043
R076477-SCH-304	CR 004378	NCT 00077714	Pali ER, Olz, Pbo	SZ	296	Yes	1	(8), (10)	1760 1495, 1846 6043
R076477-SCH-305	CR 004375	NCT 00083668	Pali ER, Olz, Pbo	SZ	333	Yes	1	(9), (10)	1746 6492, 1846 6043
R076477-SCH-3015	CR 010501	NCT 00334126	Pali ER, Quet	SZ	220	Yes	1	(11)	1941 1369
R076477-SCA-3001	CR 010498	NCT 00397033	Pali ER, Pbo	SA	173	Yes	2	(12, 14)	2049 2853, 2095 7127
R076477-	CR	NCT	Pali ER,	SA	187	Yes	2	(13,	2081

SCA-3002	013099	00412373	Pbo					14)	4330, 2095 7127
R076477-BIM-3001	CR 010834	NCT 00299715	Pali ER, Pbo	BP	310	Yes	2	(15)	2062 4657
R076477-BIM-3002	CR 010858	NCT 00309699	Pali ER, Quet, Pbo	BP	350	Yes	2	(16)	2056 5430
R076477-BIM-3003	CR 010855	NCT 00309686	Pali ER, Pbo	BP	214	Yes	2	(17)	2094 7174
R092670-SCH-201	CR 004357	NCT 00074477	Pali Palm, Pbo	SZ	168	Yes	1	(18)	1994 1696
R092670-PSY-3001	CR 004198	NCT 00111189	Pali Palm, Pbo	SZ	14	No	1	(19), (20)	1995 9339, 2169 6265
R092670-PSY-3002	CR 004195	NCT 00210717	Pali Palm, Consta	SZ	493	Yes	1	(21)	2177 7507
R092670-PSY-3003	CR 002353	NCT 00210548	Pali Palm, Pbo	SZ	249	Yes	1	(22)	2038 9255
R092670-PSY-3004	CR 003562	NCT 00101634	Pali Palm, Pbo	SZ	404	Yes	1	(23)	2055 5312
R092670-PSY-3005	CR 002350	NCT 00119756	Pali Palm	SZ	17	No	1	(24)	1948 1579
R092670-PSY-3007	CR 012550	NCT 00590577	Pali Palm, Pbo	SZ	468	Yes	2	(25), (26)	2047 3057, 2156 9242
RIS-SCH-401	CR 002899	NCT 00297388	Ris	SZ or	148	Yes	2	(27)	1696 5196

				SA					
RIS-SCP-402	CR 002890	NCT 00061802	Ris, Pbo	SZ or SA	62	Yes	1	(28)	1705 4789
RIS-BIM-301	CR 003631	NCT 00076115	Ris, Pbo	BP	120	Yes	2	(29)	1983 9994
RIS-INT-85			Ris	SZ	8	No	1	(30)	1520 1572
RIS-INT-69	CR 006049	NCT 00253162	Ris, Halo, Pbo	BP	233	Yes	1	(31)	1557 2276
RIS-USA-239	CR 006052	NCT 00257075	Ris, Pbo	BP	186	Yes	2	(32)	1516 9694
RIS-USA-259	CR 002761	NCT 00034775	Ris	SZ	16	No	1	(33)	1532 3593
R209130-SCH-201			R209130, Pbo	SZ	7	No	1		
R209130-SCH-202	CR 004342	NCT 00063297	R209130, Pbo	SZ	1	No	1		

Table 2: Summary of the JNJ clinical trials. * Pali ER: Paliperidone ER OROS; Pali Palm: Paliperidone Palmitate; Ris: Risperidone; Consta: Risperdal Consta; Olz: Olanzapine; Quet: Quetiapine; Halo:Haloperidol; Pbo: Placebo

	Schizophrenia (N=3251)	Schizoaffective (N=377)	Bipolar (N=1344)
Sex, n (%)			
F	1240 (38.1)	152 (40.3)	629 (47.2)
M	2011 (61.9)	225 (59.7)	705 (52.8)
Age, years			
Mean (SD)	40.2 (12)	38.7 (9.5)	37.8 (13.5)
Median (Range)	40 (17, 81)	39 (19, 61)	39 (10, 77)

Race, n (%)			
Asian	117 (3.6)	52 (13.8)	37 (2.8)
Black or African America	703 (21.6)	86 (22.8)	247 (18.5)
White	2360 (72.6)	228 (60.5)	1021 (76.5)
Other	71 (2.2)	11 (2.9)	29 (2.2)

Table 3: Basic demographic information of the JNJ SZ, SA, and BP patients

Preliminary quality control. For each cohort, we excluded from further analysis any sample that had missing genotypes for more than 2% of the SNPs on the array, further we only included SNPs with genotype missing rate < 5%, minor allele frequency > 0.01, as well as HWE-p value > 0.0001.

Duplicate samples and cryptic relatedness. We generated pairwise IBD values for all samples using the plink genome command, excluding one sample from any pair with a PI_HAT value exceeding 0.3.

Population stratification. Principal components were generated on each cohort using smartPCA, eigenvectors were included as covariates in a logistic regression to control for population stratification as required. To determine the genomic inflation for each case control set we carried an association analysis on the genotyped data using plink prior to imputation. If genomic inflation exceeded 1.03, principal components were included as covariates in the post-imputation GWAS.

Prephasing. For each cohort, samples were prephased for imputation using the SHAPEIT package. Each chromosome was prephased separately. For case control sets that were typed on different arrays, such as the Johnson and Johnson and CHOP set, the prephasing was carried for each chip type separately and prior to imputation the haplotypes were restricted to SNPs common to both arrays.

Imputation. We used the Impute2 package to impute unobserved genotypes in each cohort using the reference haplotypes in release 2 of the HapMap 3 that included approximately 1.5 million variants from 1,011 individuals from Africa, Asia, Europe and the Americas.

Genotype concordance. Internal cross validation was carried out automatically by Impute2, the calculation is performed by masking one variant at a time in the study data, imputing the masked variant and comparing the result to the original genotype. Average concordance for all datasets was > 90%.

Post-imputation association analysis. Case control association was carried out using the snptest package. We applied an additive model on the genotype dosages generated by Impute2 including the proportion of missing data and gender were as covariates for all cohorts. In the presence of population stratification, we also included the first 10 principal components from the smartPCA analysis as covariates.

Meta analysis. Inverse variance fixed effects meta analysis was carried out using the metal package controlling for genomic inflation during the meta analysis. A final round of genomic control was also applied to the results of the meta analysis. Random effects meta analysis was carried out using the RE2 model in the METASOFT package.

Results

Following the meta analysis 40 variants remained significant after Bonferroni correction (P-values < $5x10^{-8}$) (Table 4). The 40 SNPs mapped to 6 loci, 5 of which had been previously associated with susceptibility to SCZ and / or BP (Figure 11). Two SNPs mapped to a novel locus containing a gene of unknown function, TSNARE1 (t-SNARE domain containing 1) which maps to chr8q24.3.In addition to the two genome-wide significant SNPs multiple other SNPs in LD showed a trend towards association at the locus (Figure 12). Odds ratios for the most significantly associated SNP (rs10098073, P-value $9.05x10^{-9}$) across the 16 cohorts ranged from 0.76 to 0.99 (SD 0.06) with one outlier, the Dublin cohort, crossing 1 at 1.02.

While the function of the TSNARE1gene remains unknown, a recent publication suggests it may have evolved, within the vertebrate lineage, from the harbinger transposon superfamily

(3). Bio-informatic predictions based on phylogenetic ancestry indicate it may bind SNARE (soluble N-ethylmaleimide-sensitive factor attached protein receptor) proteins and have SNAP receptor activity. TSNARE1 may therefore have a vertebrate specific function in intracellular protein transport and synaptic vesicle exocytosis.

5

Gene	rsID	CHR	BP (hg17)	Freq	Pval	Direction effect	A1	A2	HetPVal
SDCCAG8	rs10927013	1	241655779	0.4368	$8.09x10^{-9}$	++++++++++++++++	G	T	0.91
	rs2039839	1	241665293	0.5645	$1.28x10^{-8}$	----------	A	G	0.93
	rs6703335	1	241675590	0.5627	$3.71x10^{-9}$	----------+	A	G	0.91
ITIH1	rs2710323	3	52790945	0.4669	$2.97x10^{-8}$	-++-++-+++++++++	C	T	0.39
MHC	rs9379858	6	26475668	0.1069	$1.42x10^{-8}$	+++-++-++-++++++	C	T	0.35
	rs3799380	6	26575161	0.2201	$2.80x10^{-8}$	+++-++-+++++++++	C	T	0.82
	rs6456728	6	26585758	0.7797	$2.36x10^{-8}$	---+--+---------	A	G	0.82
	rs7749305	6	27554545	0.1004	$2.57x10^{-8}$	+++-++++++++++++	C	T	0.29
	rs6904596	6	27599278	0.8963	$3.06x10^{-8}$	---+----------	A	G	0.27
	rs17750424	6	27809101	0.0847	$1.99x10^{-8}$	++++++++++++++++	C	T	0.38
	rs17693963	6	27818144	0.9034	$1.61x10^{-9}$	+++-++++++++++++	A	C	0.36
	rs34706883	6	27913234	0.9096	$3.14x10^{-8}$	++++++++++++++++	A	C	0.39
	rs13212651	6	27914964	0.9096	$3.34x10^{-8}$	++++++++++++++++	A	G	0.39
	rs13194781	6	27923618	0.9095	$3.51x10^{-8}$	++++++++++++++++	A	G	0.42
	rs13199772	6	27942064	0.9096	$3.40x10^{-8}$	++++++++++++++++	A	G	0.42
MAD1L1	rs10275045	7	1887352	0.4035	$2.16x10^{-8}$	-+-+-------+---+	C	T	0.36
	rs12537914	7	1914885	0.3647	$2.03x10^{-8}$	-+-+-------+----	C	T	0.47
	rs4721184	7	1917310	0.3698	$2.74x10^{-8}$	+-+-+++++++-+++-	C	T	0.47
	rs2056480	7	1920827	0.6321	$1.54x10^{-8}$	---+-------+---+	A	G	0.56
	rs4721190	7	1921258	0.5959	$3.83x10^{-8}$	---+-------+---+	A	G	0.69
	rs12699477	7	1935479	0.3523	$8.78x10^{-9}$	+-++++++++-++++	C	T	0.67
	rs2280550	7	1943082	0.6389	$3.80x10^{-8}$	-+-+-------+---+	A	G	0.49
	rs12666575	7	1970947	0.3578	$8.27x10^{-10}$	------+----+---+	C	T	0.19
	rs4721295	7	2003195	0.3778	$6.23x10^{-10}$	+-++++-++++-+++-	G	T	0.07
	rs1107592	7	2007958	0.62	$5.64x10^{-9}$	+-++++-++++-+++-	A	G	0.04
	rs3778969	7	2106516	0.6329	$6.43x10^{-9}$	------+----+---+	A	G	0.16
	rs10224497	7	2116493	0.5742	$1.09x10^{-8}$	++++++-++++-+++-	A	G	0.13
	rs10239050	7	2124916	0.6149	$1.11x10^{-8}$	+-++++-++++-+++-	A	G	0.12
	rs3800913	7	2129763	0.6036	$4.03x10^{-8}$	++++++-++++-+++-	A	G	0.15
	rs3800917	7	2134465	0.64	$2.29x10^{-9}$	-+----+----+---+	A	G	0.22
	rs3778991	7	2138981	0.656	$2.21x10^{-9}$	------+----+----	A	G	0.20
	rs3778994	7	2142381	0.6165	$2.53x10^{-9}$	------+----+----	A	C	0.21
	rs4721441	7	2150586	0.3564	$9.41x10^{-10}$	++++++-++++-++++	C	T	0.29
	rs3779003	7	2151428	0.3471	$2.33x10^{-9}$	------+----+----	C	T	0.23

	rs3757440	7	2239462	0.647	$1.98x10^{-8}$	++++++++++--+++-	A	G	0.28
	rs7787274	7	2242519	0.6506	$1.08x10^{-8}$	----------+---+	A	G	0.21
	rs7799006	7	2244752	0.3503	$1.24x10^{-8}$	----------+---+	C	T	0.23
CSMD1	rs6558872	8	4225547	0.5707	$4.75x10^{-8}$	----+----+-----+	A	G	0.46
TSNARE1	rs10098073	8	143307411	0.5351	$9.05x10^{-9}$	++-++++++++++++	A	C	0.93
	rs4129585	8	143310840	0.5659	$2.38x10^{-8}$	++++++++++++++++	A	C	0.99

Table 4: Genome wide significant variants following SCZ meta analysis.

EXAMPLE III

Protein – protein interactions (PPI) show association of specific gene pathways with schizophrenia and bipolar disorder.

To test if any gene pathways are enriched in schizophrenia/BP we constructed a PPI network based on the results of the scz/BP GWAS meta-analysis including 13,394 cases and 34,676 controls (as previously described). Gene-level P-values were calculated from the meta-analysis data, 2,998 genes were significant with P- values below 0.05 from the 17,693 genes included in the analysis. The protein-protein interaction (PPI) network was constructed based on indices of protein interactions derived from primary interaction databases including BIND, BioGRID, CORUM, DIP, HPRD, InnateDB, IntAct, MatrixDB, MINT, MPact, MPIDB, MPPI and OPHID as compiled by iRefindex. The network algorithm was retrained using human-human protein interactions supported by at least two publications listed in Pubmed. The largest connected component (LCC, 540 nodes and 768 edges) was used for further analysis. Pathway analysis of the genes in the LCC highlighted several enriched functional pathways in schizophrenia and bipolar disease including the src kinase pathway, circled in Table 5, below.

Table 5

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	KEGG_PATHWAY	Proteasome	RT		13	2.4	2.1E-6	2.5E-4
	KEGG_PATHWAY	Renal cell carcinoma	RT		15	2.7	6.9E-6	4.1E-4
	KEGG_PATHWAY	MAPK signaling pathway	RT		32	5.8	7.4E-6	2.9E-4
	KEGG_PATHWAY	Pathways in cancer	RT		35	6.3	3.1E-5	9.1E-4
	KEGG_PATHWAY	GnRH signaling pathway	RT		16	2.9	9.1E-5	2.2E-3
	KEGG_PATHWAY	Long-term potentiation	RT		13	2.4	1.2E-4	2.3E-3
	KEGG_PATHWAY	Ubiquitin mediated proteolysis	RT		19	3.4	1.4E-4	2.4E-3
	KEGG_PATHWAY	Non-small cell lung cancer	RT		11	2.0	2.8E-4	4.2E-3
	KEGG_PATHWAY	Progesterone-mediated oocyte maturation	RT		14	2.5	3.0E-4	4.0E-3
	KEGG_PATHWAY	Neurotrophin signaling pathway	RT		17	3.1	4.1E-4	4.9E-3
	KEGG_PATHWAY	Regulation of actin cytoskeleton	RT		24	4.3	4.1E-4	4.5E-3
	KEGG_PATHWAY	Long-term depression	RT		12	2.2	5.5E-4	5.5E-3
	KEGG_PATHWAY	T cell receptor signaling pathway	RT		15	2.7	9.0E-4	8.2E-3
	KEGG_PATHWAY	Prostate cancer	RT		13	2.4	1.5E-3	1.2E-2
	KEGG_PATHWAY	Focal adhesion	RT		21	3.8	2.4E-3	1.9E-2
	KEGG_PATHWAY	Pancreatic cancer	RT		11	2.0	2.9E-3	2.1E-2
	KEGG_PATHWAY	Glioma	RT		10	1.8	3.8E-3	2.6E-2
	KEGG_PATHWAY	ErbB signaling pathway	RT		12	2.2	3.8E-3	2.5E-2
	KEGG_PATHWAY	TGF-beta signaling pathway	RT		12	2.2	3.8E-3	2.5E-2
	KEGG_PATHWAY	Fc epsilon RI signaling pathway	RT		11	2.0	5.2E-3	3.2E-2

References

1. Glessner JT, Reilly MP, Kim CE, Takahashi N, Albano A, Hou C, et al. Strong synaptic transmission impact by copy number variations in schizophrenia. Proc Natl Acad Sci U S A. 2010;107(23):10584-9. PMCID: 2890845.

2. Genome-wide association study identifies five new schizophrenia loci. Nature genetics. 2011;43(10):969-76. PMCID: 3303194.

3. Smith JJ, Sumiyama K, Amemiya CT. A living fossil in the genome of a living fossil: Harbinger transposons in the coelacanth genome. Mol Biol Evol. 2012;29(3):985-93. PMCID: 3278475.

4. Cleton A, Rossenu S, Talluri K, Francetic I, Remmerie B, Janssens L, et al. Evaluation of the pharmacokinetics of an extended-release formulation of paliperidone with an immediate-release formulation of risperidone. Annual Meeting of the American Society for Clinical Pharmacology and Therapeutics; March 21-24, 2007; Anaheim, California, USA: Clinical Pharmacology and Therapeutics; 2007. p. S62.

5. Kramer M, Simpson G, Maciulis V, Kushner S, Vijapurkar U, Lim P, et al. Paliperidone extended-release tablets for prevention of symptom recurrence in patients with schizophrenia: a randomized, double-blind, placebo-controlled study. J Clin Psychopharmacol. 2007;27(1):6-14.

6. Tzimos A, Samokhvalov V, Kramer M, Ford L, Gassmann-Mayer C, Lim P, et al. Safety and tolerability of oral paliperidone extended-release tablets in elderly patients with schizophrenia: a double-blind, placebo-controlled study with six-month open-label extension. Am J Geriatr Psychiatry. 2008;16(1):31-43.

7. Kane J, Canas F, Kramer M, Ford L, Gassmann-Mayer C, Lim P, et al. Treatment of schizophrenia with paliperidone extended-release tablets: a 6-week placebo-controlled trial. Schizophr Res. 2007;90(1-3):147-61.

8. Marder SR, Kramer M, Ford L, Eerdekens E, Lim P, Eerdekens M, et al. Efficacy and safety of paliperidone extended-release tablets: results of a 6-week, randomized, placebo-controlled study. Biol Psychiatry. 2007;62(12):1363-70.

9. Davidson M, Emsley R, Kramer M, Ford L, Pan G, Lim P, et al. Efficacy, safety and early response of paliperidone extended-release tablets (paliperidone ER): results of a 6-week, randomized, placebo-controlled study. Schizophr Res. 2007;93(1-3):117-30.

10. Meltzer HY, Bobo WV, Nuamah IF, Lane R, Hough D, Kramer M, et al. Efficacy and tolerability of oral paliperidone extended-release tablets in the treatment of acute schizophrenia: pooled data from three 6-week, placebo-controlled studies. J Clin Psychiatry. 2008;69(5):817-29.

11. Canuso CM, Dirks B, Carothers J, Kosik-Gonzalez C, Bossie CA, Zhu Y, et al. Randomized, double-blind, placebo-controlled study of paliperidone extended-release and quetiapine in inpatients with recently exacerbated schizophrenia. Am J Psychiatry. 2009;166(6):691-701.

12. Canuso CM, Lindenmayer JP, Kosik-Gonzalez C, Turkoz I, Carothers J, Bossie CA, et al. A randomized, double-blind, placebo-controlled study of 2 dose ranges of paliperidone extended-release in the treatment of subjects with schizoaffective disorder. J Clin Psychiatry. 2010;71(5):587-98.

13. Canuso CM, Schooler N, Carothers J, Turkoz I, Kosik-Gonzalez C, Bossie CA, et al. Paliperidone extended-release in schizoaffective disorder: a randomized, controlled study comparing a flexible dose with placebo in patients treated with and without antidepressants and/or mood stabilizers. J Clin Psychopharmacol. 2010;30(5):487-95.

14. Canuso CM, Turkoz I, Fu DJ, Bossie CA. Role of paliperidone extended-release in treatment of schizoaffective disorder. Neuropsychiatr Dis Treat. 2010;6:667-79.

15. Berwaerts J, Xu H, Nuamah I, Lim P, Hough D. Evaluation of the efficacy and safety of paliperidone extended-release in the treatment of acute mania: a randomized, double-blind, dose-response study. J Affect Disord. 2012;136(1-2):e51-60.

16. Vieta E, Nuamah IF, Lim P, Yuen EC, Palumbo JM, Hough DW, et al. A randomized, placebo- and active-controlled study of paliperidone extended release for the treatment of acute manic and mixed episodes of bipolar I disorder. Bipolar Disord. 2010;12(3):230-43.

17. Berwaerts J, Lane R, Nuamah IF, Lim P, Remmerie B, Hough DW. Paliperidone extended-release as adjunctive therapy to lithium or valproate in the treatment of acute mania: a randomized, placebo-controlled study. J Affect Disord. 2011;129(1-3):252-60.

18. Kramer M, Litman R, Hough D, Lane R, Lim P, Liu Y, et al. Paliperidone palmitate, a potential long-acting treatment for patients with schizophrenia. Results of a randomized, double-blind, placebo-controlled efficacy and safety study. Int J Neuropsychopharmacol. 2010;13(5):635-47.

19. Hough D, Gopal S, Vijapurkar U, Lim P, Morozova M, Eerdekens M. Paliperidone palmitate maintenance treatment in delaying the time-to-relapse in patients with schizophrenia: a randomized, double-blind, placebo-controlled study. Schizophr Res. 2010;116(2-3):107-17.

20. Kozma CM, Slaton T, Dirani R, Fastenau J, Gopal S, Hough D. Changes in schizophrenia-related hospitalization and ER use among patients receiving paliperidone palmitate: results from a clinical trial with a 52-week open-label extension (OLE). Curr Med Res Opin. 2011;27(8):1603-11.

21. Fleischhacker WW, Gopal S, Lane R, Gassmann-Mayer C, Lim P, Hough D, et al. A randomized trial of paliperidone palmitate and risperidone long-acting injectable in schizophrenia. Int J Neuropsychopharmacol.1-12.

22. Gopal S, Hough DW, Xu H, Lull JM, Gassmann-Mayer C, Remmerie BM, et al. Efficacy and safety of paliperidone palmitate in adult patients with acutely symptomatic schizophrenia: a randomized, double-blind, placebo-controlled, dose-response study. Int Clin Psychopharmacol. 2010;25(5):247-56.

23. Nasrallah HA, Gopal S, Gassmann-Mayer C, Quiroz JA, Lim P, Eerdekens M, et al. A controlled, evidence-based trial of paliperidone palmitate, a long-acting injectable antipsychotic, in schizophrenia. Neuropsychopharmacology. 2010;35(10):2072-82.

24. Hough D, Lindenmayer JP, Gopal S, Melkote R, Lim P, Herben V, et al. Safety and tolerability of deltoid and gluteal injections of paliperidone palmitate in schizophrenia. Prog Neuropsychopharmacol Biol Psychiatry. 2009;33(6):1022-31.

25. Pandina GJ, Lindenmayer JP, Lull J, Lim P, Gopal S, Herben V, et al. A randomized, placebo-controlled study to assess the efficacy and safety of 3 doses of paliperidone palmitate in adults with acutely exacerbated schizophrenia. J Clin Psychopharmacol. 2010;30(3):235-44.

26. Bossie CA, Sliwa JK, Ma YW, Fu DJ, Alphs L. Onset of efficacy and tolerability following the initiation dosing of long-acting paliperidone palmitate: post-hoc analyses of a randomized, double-blind clinical trial. BMC Psychiatry. 2011;11:79.

27. Simpson GM, Mahmoud RA, Lasser RA, Kujawa M, Bossie CA, Turkoz I, et al. A 1-year double-blind study of 2 doses of long-acting risperidone in stable patients with schizophrenia or schizoaffective disorder. J Clin Psychiatry. 2006;67(8):1194-203.

28. Gharabawi GM, Greenspan A, Rupnow MF, Kosik-Gonzalez C, Bossie CA, Zhu Y, et al. Reduction in psychotic symptoms as a predictor of patient satisfaction with antipsychotic medication in schizophrenia: data from a randomized double-blind trial. BMC Psychiatry. 2006;6:45.

29. Haas M, Delbello MP, Pandina G, Kushner S, Van Hove I, Augustyns I, et al. Risperidone for the treatment of acute mania in children and adolescents with bipolar disorder: a randomized, double-blind, placebo-controlled study. Bipolar Disord. 2009;11(7):687-700.

30. Turner M, Eerdekens E, Jacko M, Eerdekens M. Long-acting injectable risperidone: safety and efficacy in stable patients switched from conventional depot antipsychotics. Int Clin Psychopharmacol. 2004;19(4):241-9.

31. Smulevich AB, Khanna S, Eerdekens M, Karcher K, Kramer M, Grossman F. Acute and continuation risperidone monotherapy in bipolar mania: a 3-week placebo-controlled trial followed by a 9-week double-blind trial of risperidone and haloperidol. Eur Neuropsychopharmacol. 2005;15(1):75-84.

32. Hirschfeld RM, Keck PE, Jr., Kramer M, Karcher K, Canuso C, Eerdekens M, et al. Rapid antimanic effect of risperidone monotherapy: a 3-week multicenter, double-blind, placebo-controlled trial. Am J Psychiatry. 2004;161(6):1057-65.

33. Lindenmayer JP, Eerdekens E, Berry SA, Eerdekens M. Safety and efficacy of long-acting risperidone in schizophrenia: a 12-week, multicenter, open-label study in stable patients switched from typical and atypical oral antipsychotics. J Clin Psychiatry. 2004;65(8):1084-9.

While the invention has been described in detail and with reference to specific examples thereof, it will be apparent to one skilled in the art that various changes and modifications can be made therein without departing from the spirit and scope thereof.

What is claimed is:

1. A method for identifying agents which bind a protein encoded by a target sequence associated with neurological disease, comprising;

a) incubating said protein in the presence and absence of said agent, one of said protein or agent being detectably labeled; and

b) determining whether said agent forms a complex with said protein, thereby identifying agents which bind proteins encoding by target sequences associated with neurological disease.

2. The method of claim 1, wherein said neurological disease is selected from the group consisting of schizophrenia, bi-polar disorder and autism.

3. The method of claim 1, wherein said protein is selected from the group consisting of neuromedin B, TSNARE1, MAD1L1, CACNA1D, NT5DC2, ITH1, NEK4 NIMA, GNL3, PB1, GLT8D1, FTSJ2, NUDT1, SNX8, SEC11A, SCAND2, ZSCAN2, ALPK3, PDE8A and those presented hereinabove including those in Figure 11 and Table 4.

4. A method for identifying agents which bind a protein encoded by a target sequence associated with neurological disease, comprising;

a) incubating a cell expressing said protein in the presence and absence of said agent, and

b) determining whether said agent interacts with said protein, thereby identifying agents which bind proteins encoding by target sequences associated with neurological disease.

5. The method of claim 4, wherein said neurological disease is selected from the group consisting of schizophrenia, bi-polar disorder and autism.

6. The method of claim 4, wherein said protein is selected from the group consisting of neuromedin B, TSNARE1, MAD1L1, CACNA1D, NT5DC2, ITH1, NEK4 NIMA, GNL3, PB1,

GLT8D1, FTSJ2, NUDT1, SNX8, SEC11A, SCAND2, ZSCAN2, ALPK3, PDE8A and those presented hereinabove including those in Figure 11 and Table 4.

7. The method of claim 6, wherein said cell is a neuronal cell and said agent alters a parameter selected from the group consisting of neuronal cellular morphology, neuronal signaling, and proliferation.

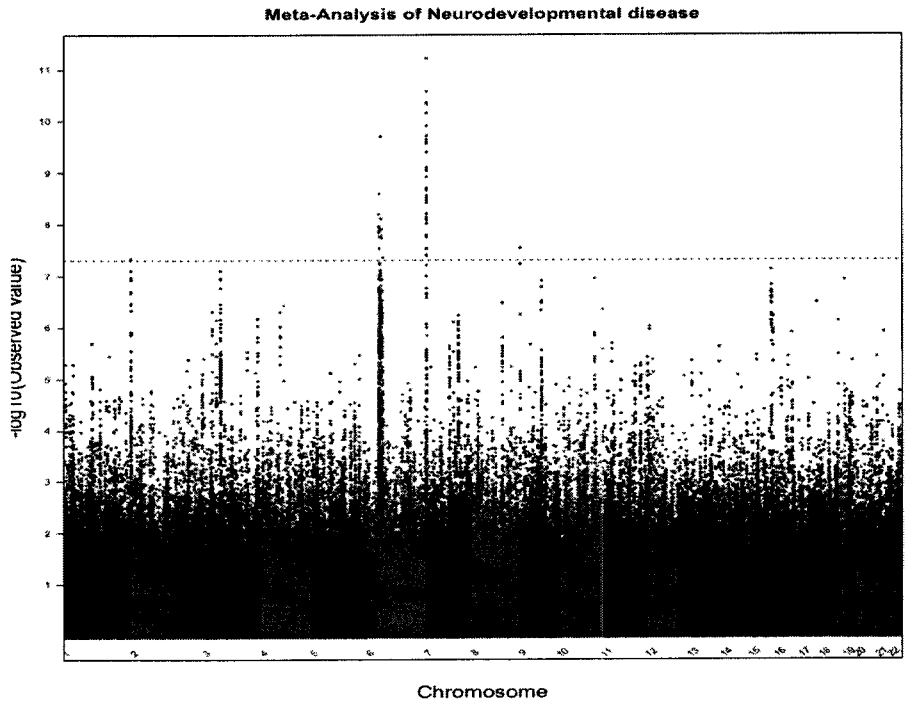

Figure 1

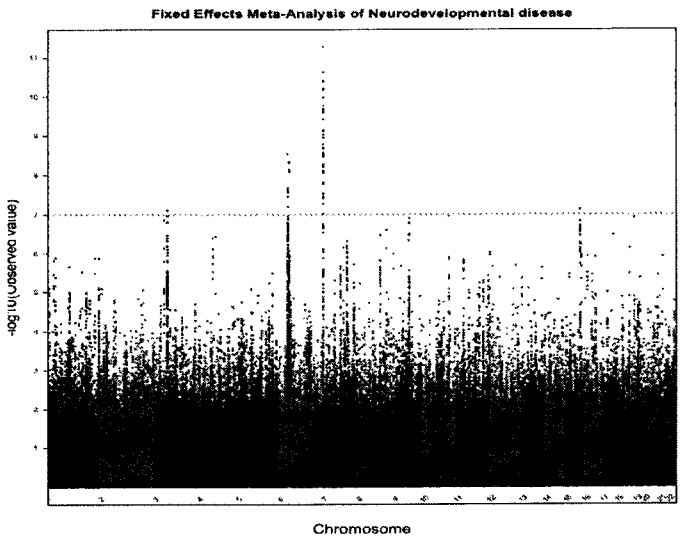

Figure 2

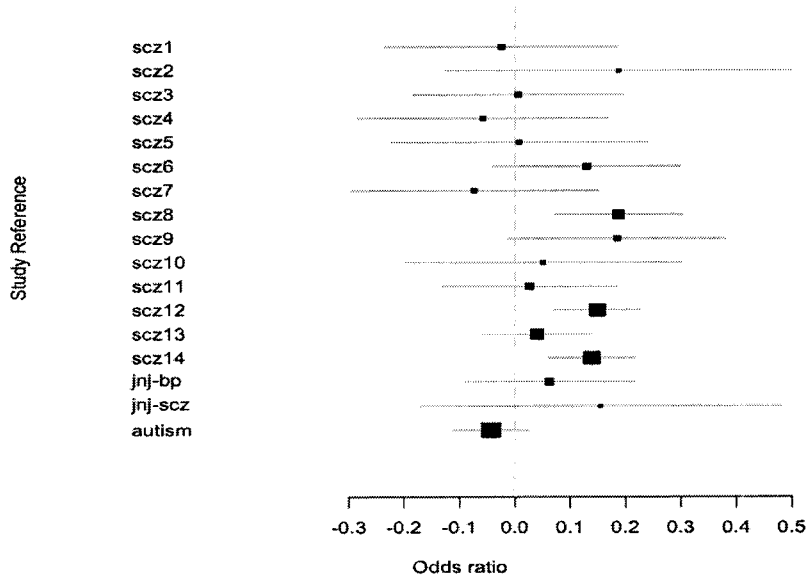

Figure 3

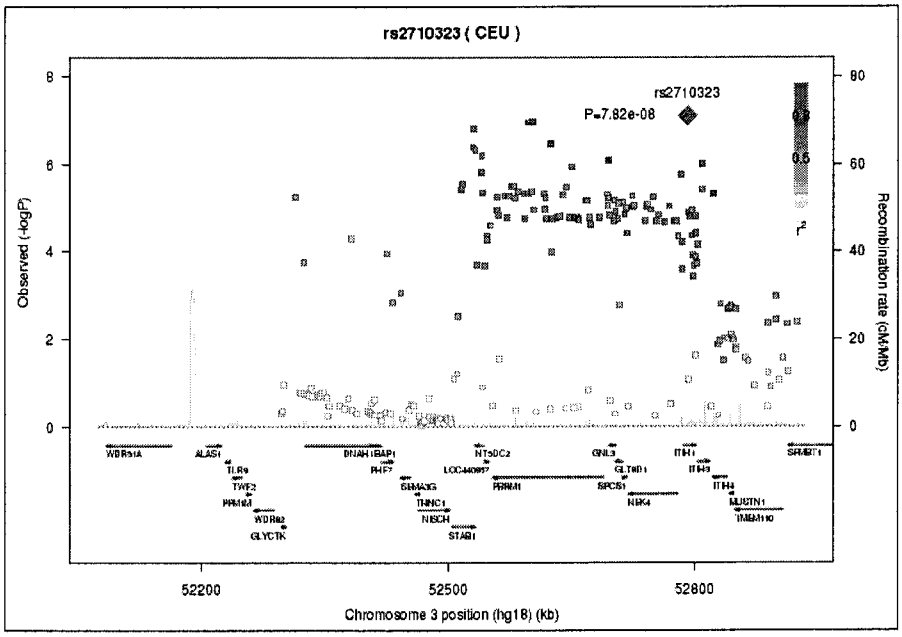

Figure 4

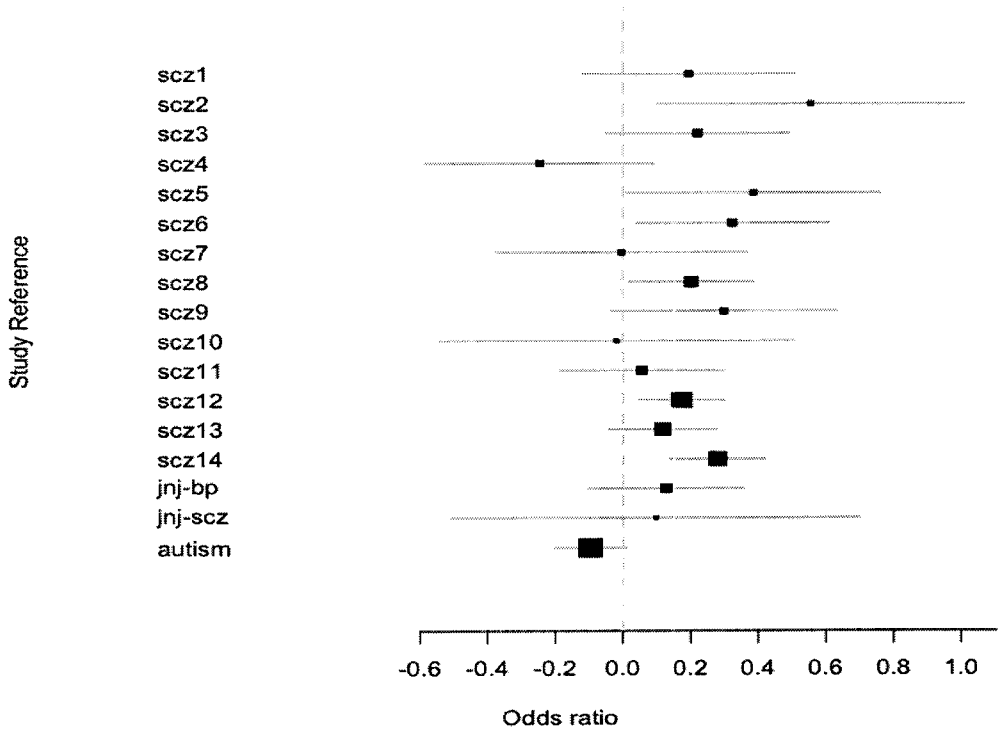

Figure 5

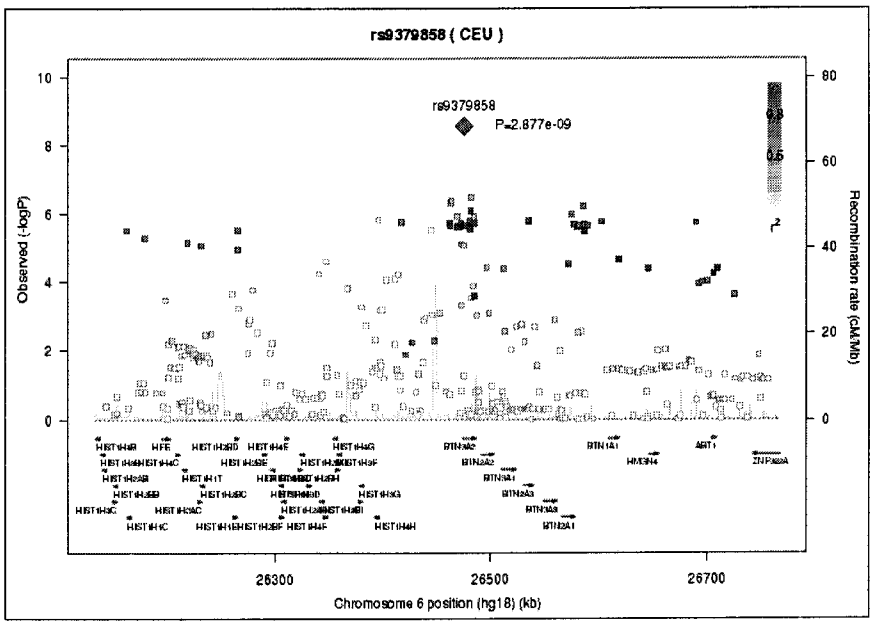

Figure 6

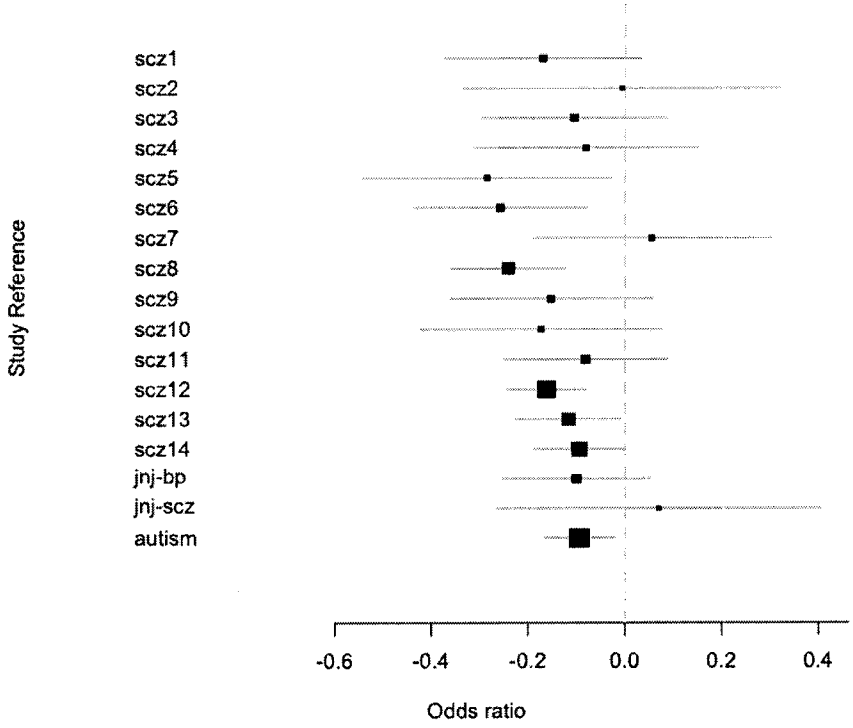
Study Reference
scz1
scz2
scz3
scz4
scz5
scz6
scz7
scz8
scz9
scz10
scz11
scz12
scz13
scz14
jnj-bp
jnj-scz
autism
-0.6
-0.4
-0.2
0.0
0.2
0.4
Odds ratio

Figure 7

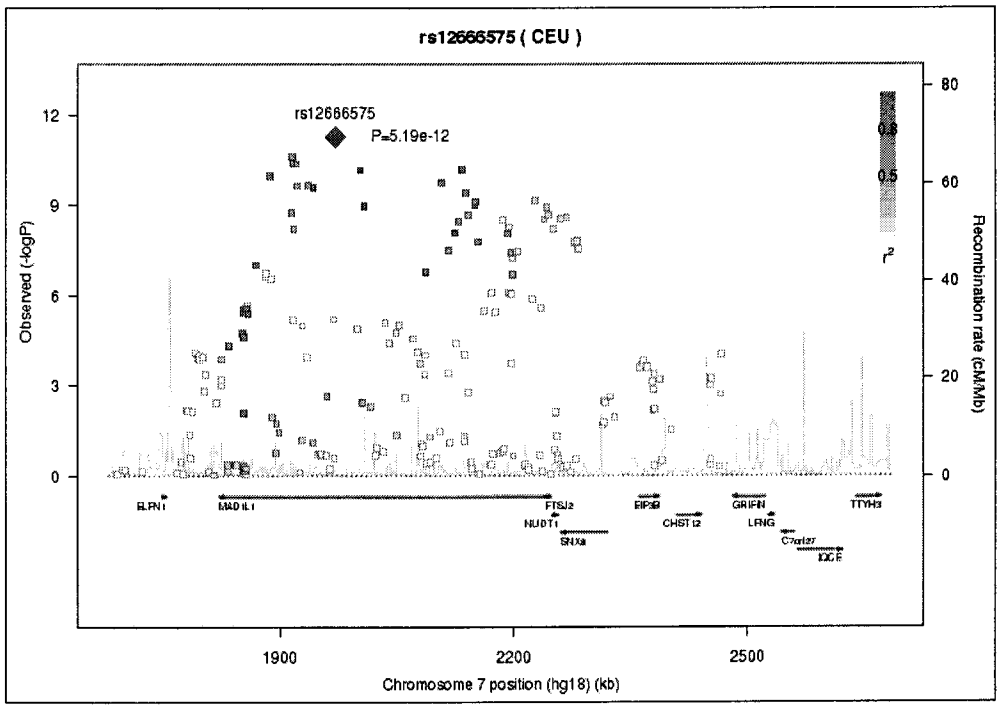

Figure 8

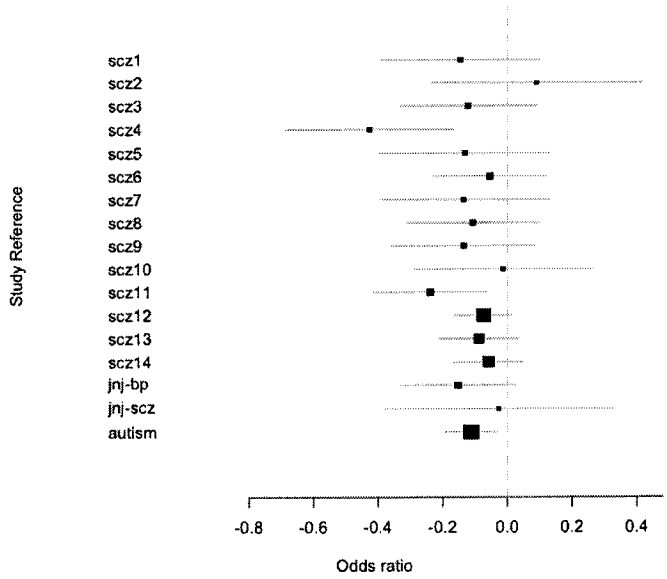

Figure 9

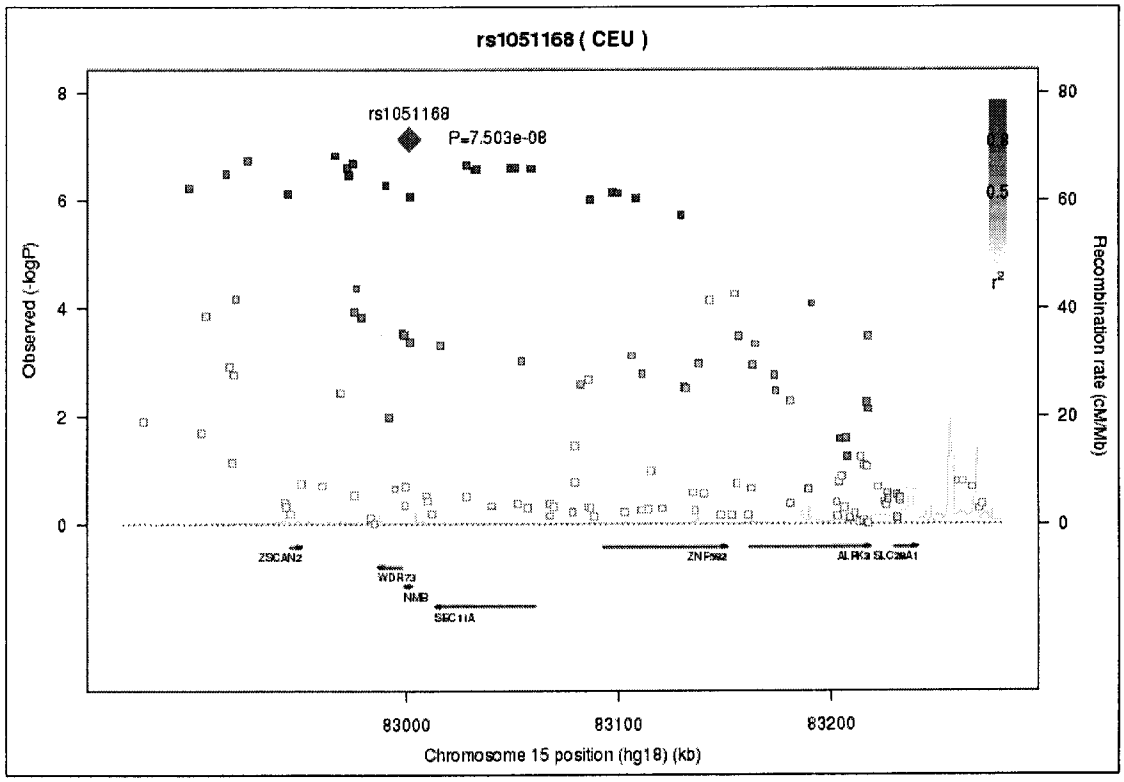
rs1051168 (CEU)
rs1051168
P=7.503e-08
Observed (-logP)
Recombination rate (cM/Mb)
r²
0.5
ZSCAN2
WDR73
NMB
SIC11A
ZNF592
ALPK3
SLC28A1
83000
83100
83200
Chromosome 15 position (hg18) (kb)

Figure 10

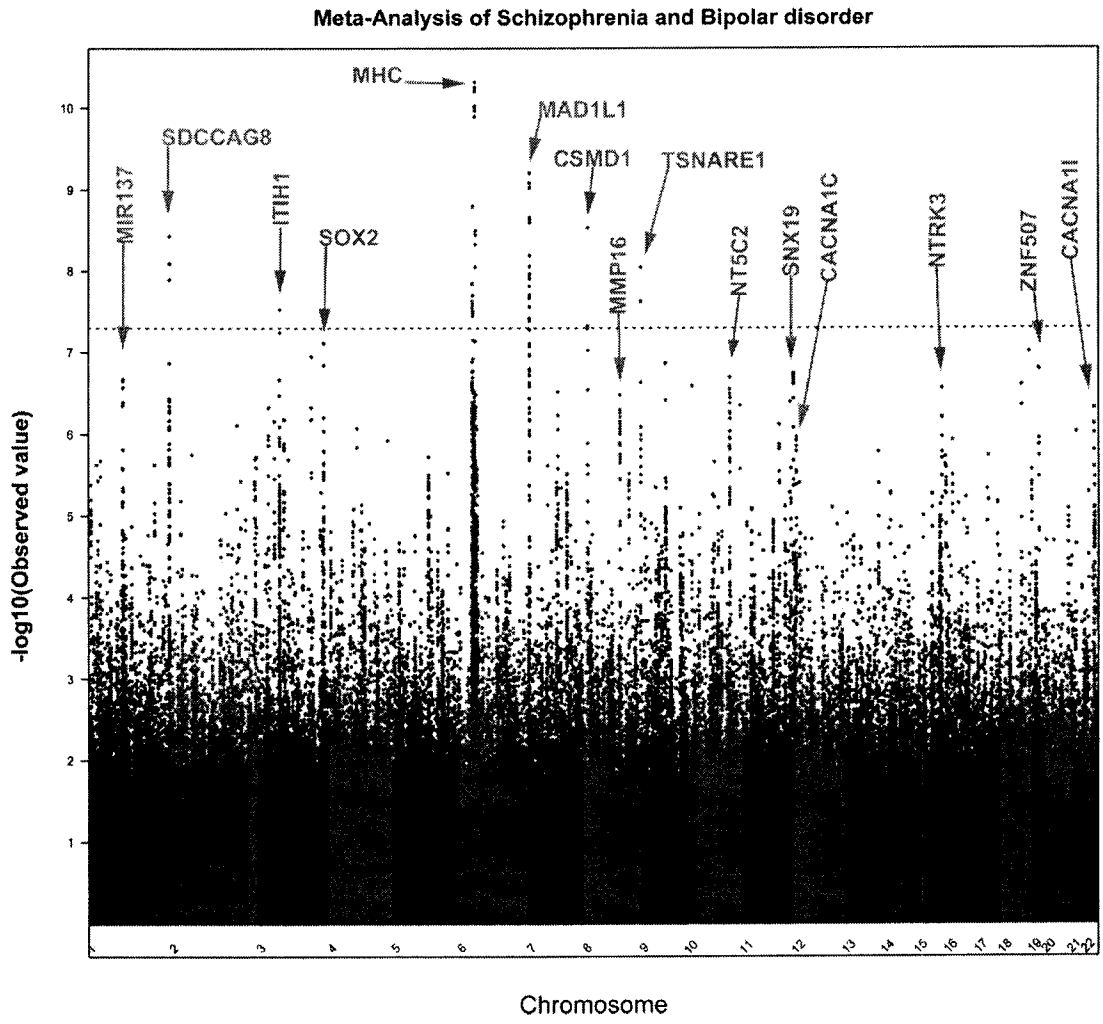
Meta-Analysis of Schizophrenia and Bipolar disorder
MHC
MAD1L1
SDCCAG8
CSMD1
TSNARE1
MIR137
ITIH1
SOX2
MMP16
NT5C2
SNX19
CACNA1C
NTRK3
ZNF507
CACNA1I
-log10(Observed value)
Chromosome

Figure 11

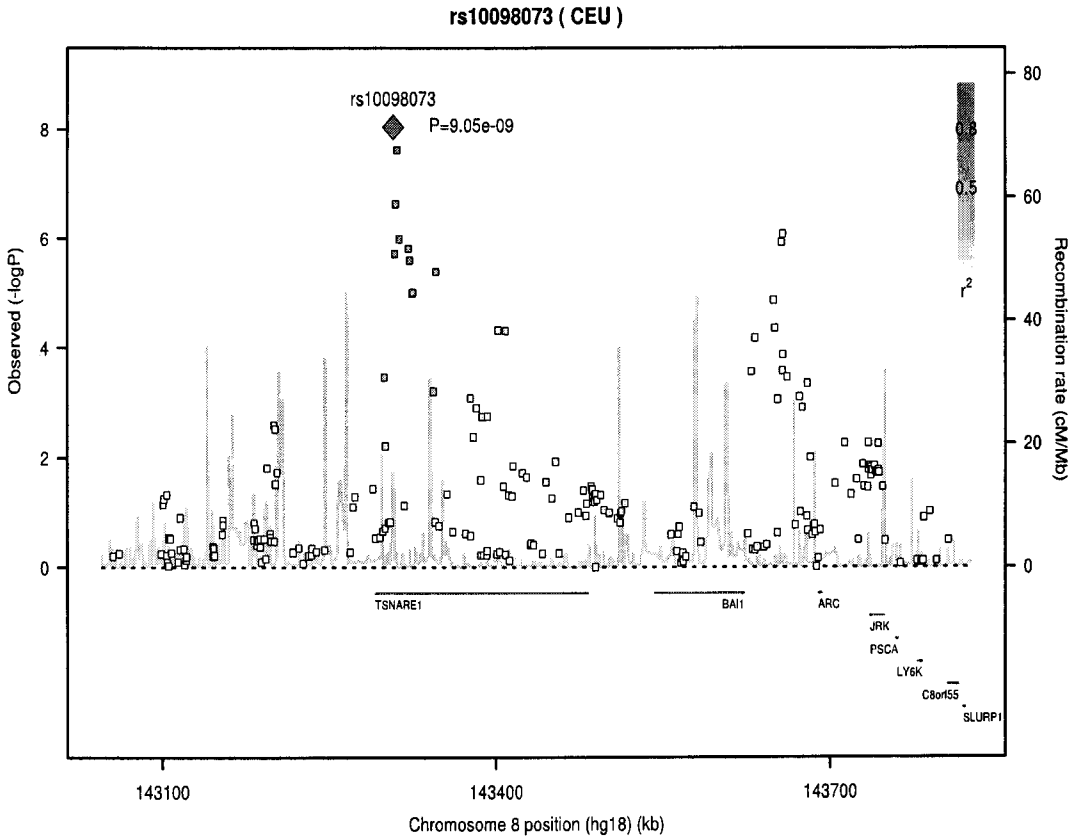
rs10098073 (CEU)
rs10098073
P=9.05e-09
Observed (-logP)
Recombination rate (cM/Mb)
r²
0.8
0.5
TSNARE1
BAI1
ARC
JRK
PSCA
LY6K
C8orf155
SLURP1
143100
143400
143700
Chromosome 8 position (hg18) (kb)

Figure 12
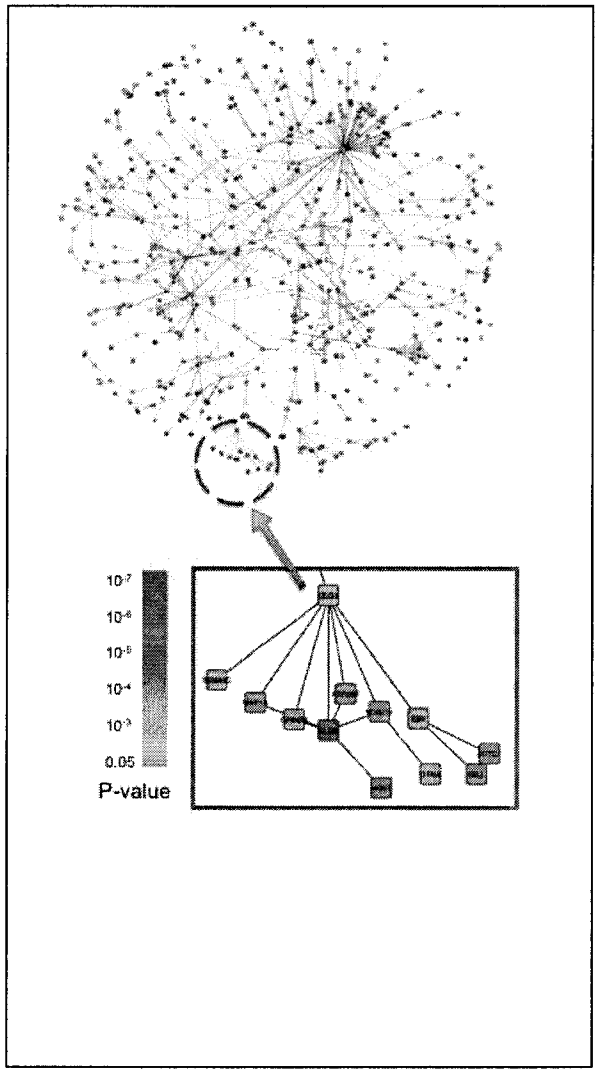
Figure 13