Published in final edited form as:

Schizophr Res. 2011 February ; 125(2-3): 201–208. doi:10.1016/j.schres.2010.12.002.

Application of systems biology approach identifies and validates *GRB2* as a risk gene for schizophrenia in the Irish Case Control Study of Schizophrenia (ICCS) sample

Jingchun Sun^{a,b}, Chunling Wan^a, Peilin Jia^{a,b}, Ayman H. Fanous^{c,d,e,f}, Kenneth S. Kendler^{d,g}, Brien P. Riley^{d,g}, and Zhongming Zhao^{a,b,*}

^a Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, TN 37232, USA

^b Department of Psychiatry, Vanderbilt University School of Medicine, Nashville, TN 37232, USA

^c Washington VA Medical Center, Washington, DC 20422, USA

^d Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA 23298, USA

^e Department of Psychiatry, Georgetown University School of Medicine, Washington, DC 20007, USA

^f Department of Psychiatry, Keck School of Medicine of the University of Southern California, Los Angeles, CA 90089, USA

^g Department of Human and Molecular Genetics, Virginia Commonwealth University, Richmond, VA 23298, USA

Abstract

Recently, we prioritized 160 schizophrenia candidate genes (SZGenes) by integrating multiple lines of evidence and subsequently identified twenty-four pathways in which these 160 genes are overrepresented. Among them, four neurotransmitter-related pathways were top ranked. In this study, we extended our previous pathway analysis by applying a systems biology approach to identifying candidate genes for schizophrenia. We constructed protein-protein interaction subnetworks for four neurotransmitter-related pathways and merged them to obtain a general neurotransmitter network, from which five candidate genes stood out. We tested the association of four genes (*GRB2*, *HSPA5*, *YWHAG*, and *YWHAZ*) in the Irish Case-Control Study of Schizophrenia (ICCS) sample (1021 cases and 626 controls). Interestingly, six of the seven tested SNPs in *GRB2* showed significant signal, two of which (rs7207618 and rs9912608) remained significant after permutation test or Bonferroni correction, suggesting that *GRB2* might be a risk gene for schizophrenia in Irish population. To our knowledge, this is the first report of *GRB2*

*Address for correspondence: Zhongming Zhao, Ph.D., Department of Biomedical Informatics, Vanderbilt University School of Medicine, 2525 West End Avenue, Suite 600, Nashville, TN 37203, USA, Phone: (615) 343-9158, Fax: (615) 936-8545, zhongming.zhao@vanderbilt.edu.

Contributors: Authors JS, ZZ, and BPR designed the study. JS, PJ, ZZ, and AHF carried out computational analyses and candidate gene selection. BPR carried out genotyping and JS performed statistical analysis of the genotyping data. KSK provided ICCS sample and guidance on this project. JS, CW and ZZ wrote the first draft of the paper. JS, CW, PJ, AHF, BPR, KSK and ZZ commented and contributed to the subsequent revisions.

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

being significantly associated with schizophrenia in a specific population. Our results suggest that the systems biology approach is promising for identification of candidate genes and understanding the etiology of complex diseases.

Keywords

Schizophrenia; systems biology; association; GRB2; GWAS; neurotransmitter

1. Introduction

Schizophrenia is a severe mental disorder with a high degree of heritability. Recent studies suggest that this heritability may be due to many genes of very small effect, interacting with each other and with environmental risk factors (Jia et al. 2010; Purcell et al. 2009; Ruano et al. 2010). It is also likely that rare alleles of genetic variants, even specific to single patients or families, might be highly penetrant and contribute the disease risk with more substantial effect (McClellan et al. 2007). Such rare variants can be either single nucleotide polymorphisms (SNPs), copy number variations (CNVs) (Walsh et al. 2008), or both. Therefore, detection of specific causal genes/loci for this disease remains a great challenge but is essential for understanding the pathogenesis of schizophrenia (Ross et al. 2006). Traditionally, candidate gene selection in an association study is based on prior knowledge of physiological, biochemical or functional aspects of candidate gene products. Recent advances in genomics technologies, especially genome-wide association (GWA) and expression studies performed on microarrays have generated numerous unbiased, genome-wide datasets. The combination of such genomic, transcriptomic, proteomic, and metabolomic data may provide us a new paradigm to search for disease candidate genes (Giegling et al. 2008).

We hypothesized that many genes, each of which might contribute a small or moderate risk to schizophrenia, may contribute major risk through their interaction and combined effects. Accordingly, we attempted to select novel candidate genes from the networks/pathways constructed by the genes that have been implicated for schizophrenia. Recently, we prioritized 160 schizophrenia candidate genes (SZGenes) by a multi-dimensional evidence-based candidate gene prioritization approach (Sun et al. 2009). Our follow up pathway enrichment analysis suggested that 24 pathways are significantly overrepresented in this set of 160 genes (Sun et al. 2010). Among the 24 pathways, four neurotransmitter-related pathways were top ranked; they are glutamate receptor signaling (ranked 1st), serotonin receptor signaling (2nd), GABA receptor signaling (5th) and dopamine receptor signaling (7th). This result reflects investigators' decades-long investigation of the role of various neurotransmitters in the etiology of schizophrenia (Miyamoto et al. 2003). We next mapped SZGenes included in these four neurotransmitter-related pathways into the whole human protein-protein interaction (PPI) network (Sun et al. 2010) and then reconstructed a subnetwork for each pathway using the Steiner minimal tree algorithm (Klein and Ravi 1995). The four subnetworks that reflect individual neurotransmitters were subsequently merged into one general neurotransmitter network. For better visualization, this neurotransmitter network was graphically presented by using software Cytoscape (Shannon et al. 2003). Figure 1 shows that this network comprises 50 nodes (i.e., proteins encoded by genes) and 61 PPI pairs.

Among the 50 nodes, 32 (labeled in red in Fig. 1) were from the list of the 160 SZGenes, suggesting that this approach is effective in clustering informative genes. The over-representation of SZGenes in the neurotransmitter network is statistically significant from the remaining human protein-coding genes collected in the human PPI network (χ^2 test, $P <$

10^{-4}). The remaining 18 nodes (labeled in grey) were considered potential candidate genes, since they interact strongly with the genes with prior evidence. Among the 18 genes, five were selected based on their prevalence in the four neurotransmitter related subnetworks: growth factor receptor-bound protein 2 (GRB2, 17p24-q23, MIM: 108355) appeared in all four subnetworks; heat shock 70kDa protein 5 (HSPA5, 9q34, MIM: 138120), tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide (YWHAZ, 8q22.3, MIM: 601288), and protein kinase C, beta (PRKCB, 16p11.2, MIM: 176970) appeared in three subnetworks; and tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide (YWHAG, 7q11.23, MIM: 605356) appeared in two subnetworks. The overlap between subnetworks might imply that these five proteins are more functionally important than the others in the neurotransmitter related biological processes. We describe these five proteins and their genes below.

GRB2 is central to the general neurotransmitter network we constructed (Fig. 1). It is involved in neuron signal transduction and regulates neuron morphology during neural development (Lowenstein et al. 1992). GRB2 is widely known as an adaptor molecule that mediates protein-protein interactions (Lowenstein et al. 1992; Takenawa et al. 1998). GRB2 was recently found to interact with the protein encoded by Disrupted-in-Schizophrenia 1 (*DISC1*), one of prominent schizophrenia susceptibility genes (Shinoda et al. 2007). In humans, this interaction has not been reported. Our literature survey found only one previous association study of *GRB2* gene with schizophrenia, performed in a Japanese population, but the result failed to reach statistical significance (Ikeda et al. 2008).

HSPA5, a heat-shock protein-70 (HSP70) family member, is involved in protein folding and assembly in the endoplasmic reticulum (ER). Non-competitive N-methyl-D-aspartate (NMDA) receptor antagonists, which can induce schizophrenia-like psychosis in humans, induce HSP70 in the posterior cingulate and retrosplenial cortex of rat brain (Hashimoto et al. 1996). The expression of HSP70 induced by MK-801 (dizocilpine, a NMDA receptor antagonist) was reversed under the antipsychotic drug treatment in rat C6 glioma cells (Roh et al. 2008). These studies suggest that HSP70 might play an important role in schizophrenia pathophysiology. We found a previous report of HSPA5 as a genetic risk factor of bipolar disorder in a Japanese population (Kakiuchi et al. 2005).

YWHAG and YWHAZ belong to the family of YWHA (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation) proteins, which includes seven molecules (YWHAB, YWHAE, YWHAG, YWHAH, YWHAQ, YWHAS, and YWHAZ) in mammals (Fu et al. 2000). YWHA genes are abundantly expressed in human brain and mediate signal transduction through binding to phosphoserine-containing proteins (Umahara and Uchiyama 2010). YWHA genes have been widely studied for the association with schizophrenia, because their products are involved in many biological processes, especially in neurotransmission (Berg et al. 2003). According to the SchizophreniaGene database (<http://www.schizophreniaforum.org/res/sczgene/>), five genes (*YWHAB*, *YWHAE*, *YWHAG*, *YWHAH*, and *YWHAZ*) have been studied for schizophrenia, three of which (*YWHAE*, *YWHAH*, and *YWHAZ*) had positive association results (Bell et al. 2000; Ikeda et al. 2008; Wong et al. 2005).

PRKCB, a member of the protein kinase C (PKC) family, is thought to be involved in neuronal signaling and mediating appropriate formation of the neural tube in mouse (Cogram et al. 2004). There are no reports of association studies of *PRKCB* and schizophrenia yet, but *PRKCB* was recently reported to be associated with autistic disorder and to have reduced expression at superior temporal gyrus in autistic patients (Lintas et al. 2009).

To demonstrate the effectiveness of our network/pathway-based candidate gene selection approach, we selected four of the above genes (*GRB2*, *HSPA5*, *YWHAG*, and *YWHAZ*) to test for association with schizophrenia in our Irish Case-Control Study of Schizophrenia (ICCSS) sample. We had limited resources for genotyping in this study and so did not include *PRKCB* due to its large size (384.6 kb).

2. Materials and methods

2.1. Signals of the general neurotransmitter network in GWAS datasets

The general neurotransmitter network includes 50 nodes. It was important to examine genetic signals (e.g. nominal *P* values at the SNP and gene levels) from those genes from multiple GWAS datasets. We used three available schizophrenia GWAS datasets: GAIN (Genetic Association Information Network) (Manolio et al. 2007), nonGAIN (Shi et al. 2009), and CATIE (Clinical Antipsychotic Trials of Intervention Effectiveness) (Sullivan et al. 2008). We used PLINK software to conduct the GWA data analysis to obtain nominal *P*-values (Purcell et al. 2007). The detailed procedure was provided in our recent work (Jia et al. 2010). To obtain a gene-wise *P*-value for each gene, we chose its smallest nominal *P*-value among all the markers mapped to the gene region to represent the association significance of the gene with schizophrenia. We did not use ISC, SGENE, or PGC GWAS datasets because they were not publicly available.

2.2. The Irish Case-Control Study of Schizophrenia (ICCSS) sample

We used the ICCSS sample, which includes 1021 cases (68% males and 32% females) and 626 controls (55% males and 45% females), in our association study. The ICCSS sample has been described in detail in previous reports (e.g., Chen et al. 2007; Riley et al. 2009). Briefly, cases were ascertained from in- and out-patient psychiatric facilities in the Republic of Ireland and Northern Ireland. Subjects were eligible for inclusion if 1) they had a diagnosis of schizophrenia or poor-outcome schizoaffective disorder by DSM-III-R criteria that was confirmed by a blind expert diagnostic review and 2) reported all four grandparents as being born in Ireland or the United Kingdom. Detailed personal interviews and hospital record rating forms were completed for each proband. Controls were recruited from donors from the Northern Ireland Blood Transfusion Service (N=554), the Irish national police (N=38) and army reserve (N=34). Controls were briefly screened and were eligible for inclusion if they denied any history of psychotic illness and reported all four grandparents as being born in Ireland or the United Kingdom. Recorded sex was verified by X/Y genotypes. This sample has $\geq 78\%$ power to detect effects with minor allele frequency (MAF) $\geq 20\%$ and a genotype relative risk ≥ 1.3 (Riley et al. 2009). All participating individuals gave appropriate informed consent for inclusion in the study. Ireland has had a stable, homogeneous population for approximately 15,000 years, with a genetic structure minimally influenced by human migrations over the last three millennia (Cavalli-Sforza et al. 1994; Riley et al. 2009). According to the experience of the blood bank staff, non-Irish donors are very rare, in agreement with the history of minimal in-migration to Ireland. Those non-Irish donors would have been excluded on the basis of questions about their grandparents.

2.3. Marker selection and genotyping

To select tags for genotyping in our sample, we first retrieved genotype data of the markers located in gene regions including exons, introns, and 10kb of flanking sequence both upstream and downstream of the genes from the HapMap phase II European genotype dataset (release # 22, <http://hapmap.ncbi.nlm.nih.gov/>). Then, we selected tag SNPs using the computer program Haploview v4.2 (Barrett et al. 2005). Specifically, we first selected SNP markers with MAF > 0.1 and then used pair-wise tagging in the Tagger module implemented in the program to select SNPs that could capture $> 80\%$ of the markers with r^2

> 0.8. We obtained a total of 15 tag SNPs including 7 SNPs in *GRB2*, 2 SNPs in *HSPA5*, 3 SNPs in *YWHAG*, and 3 SNPs in *YWHAZ*, respectively. We genotyped these SNPs using the TaqMan method, as detailed in Pham et al. (2009). We genotyped 41 samples (11 cases and 30 controls) in duplicate and used discordant genotypes in these duplicates to estimate genotyping error rates.

2.4. Statistical analysis

SNPs and haplotypes were analyzed in Haploview v4.2 (Barrett et al. 2005). We excluded individuals missing >50% of genotypes, SNPs showing significant ($P < 0.001$) departures from Hardy-Weinberg Equilibrium (HWE), or SNPs with minor allele frequency (MAF) less than 0.05. We further compared the linkage disequilibrium (LD) of ICCSS data with that of HapMap-CEU data. Haplotype analysis was performed within blocks defined by the default confidence-interval method implemented in Haploview v4.2 (Gabriel et al. 2002). We conducted permutation tests using 100,000 replicates to assess the empirical significance of cases and controls.

3. Results

3.1. Survey of association signal in the general neurotransmitter network

As shown in Figure 1, the general neurotransmitter network comprised 32 SZGenes and 18 non-SZGenes. To evaluate our strategy on identifying novel candidate genes from this combined network, we surveyed their genetic association signal from previous studies or other sources. First, among the 32 SZGenes, 8 (*COMT*, *DRD2*, *DRD4*, *GABRB2*, *GRIK3*, *GRIN2B*, *HTR2A*, and *TPH1*) belonged to the top genes in the SchizophreniaGene database. The database listed 40 top genes according to the most updated data (<http://www.schizophreniaforum.org/res/sczgene/>, as of March 13, 2010), which were selected based on meta-analysis and then ranked by epidemiological credibility of genetic association studies developed by the Human Genome Epidemiology Network (Ioannidis et al. 2008). In addition, other 16 genes had at least 2 positive reports based on the association studies collected in the SchizophreniaGene database. On the contrary, among the 18 non-SZGenes, only 3 had positive association results with schizophrenia (*GRIK2*: 1 study; *YWHAZ*: 3 studies; *EGFR*: 1 study).

Second, we examined gene-wise association signal of the 50 genes in the network using three independent GWAS datasets for schizophrenia (GAIN, nonGAIN, and CATIE). Supplementary Table 1 summarizes their smallest nominal P -values. Among the 50 genes, 47 were included in the three GWAS datasets. Among them, 38 had at least one SNP with P -value less than 0.05. Specifically, 12 genes (25.5%) had gene-wise P -values less than 0.05 in three datasets, 9 (19.1%) in two datasets, and 17 (36.2%) in one dataset. We also compared the proportion of genes with P -values less than 0.05, 0.01, and 0.001 among all genes from three GWAS datasets with that of the 50 subnetwork genes (Supplementary Fig. 1). The network has more genes with P -values less than 0.05, 0.01 and 0.001 compared to the overall genome in three GWAS datasets, respectively. These observations suggest that the general neurotransmitter network might enrich genes with small P -values in the human genome. However, this information should be used with caution because of the bias of gene size. The 50 genes (average gene length: 192.8 kb) were overall much longer than the remaining human genes (60.7 kb). Thus, the subnetwork genes might have a better chance of obtaining small P -values.

Specifically, for the four genes that we selected for genotyping in the present study, there were 22 SNPs in *GRB2* that had been tested in at least one of the three GWA studies, 4 SNPs in *HSPA5*, 10 SNPs in *YWHAG*, and 10 SNPs in *YWHAZ*. The P -values of those SNPs

in the three GWAS datasets were not promising: only one SNP in *GRB2* and one SNP in *HSPA5* had nominal *P*-value less than 0.05 in any GWAS dataset (*GRB2*: rs12603538, *P*=0.0272, CATIE; *HSPA5*: rs3739554, *P*=0.0397, CATIE; see Supplementary Table 1). Thus, genetic signals from GWA studies provided very weak support of these four candidate genes. In the next subsections, we tested association of the four genes with schizophrenia using our ICCSS sample.

3.2. Genotyping in the ICCSS sample: missingness and error rates

In total, 34/1647 samples (2.06%, 20 cases, 14 controls) were excluded due to missing data. Results are based on 1001 cases and 612 controls (*N* = 1613), 1235 samples (74.98%) missing zero, 220 (13.36%) missing one, 85 (5.16%) missing two, 33 (2.00%) missing three, 20 (1.21%) missing four, 20 (1.21%) missing 5 to 7 markers. By marker, average genotyping completion was 95.74% (93.45-97.03%). We genotype 15 SNPs $\times$ 41 samples in duplicate (*N* = 615 genotype pairs); 12 were discordant, estimating our genotyping error rate at 1.95%. This rate is slightly higher than normal in our lab, but we noted that two markers in *YWHAG* (rs11765693 and rs17149177) accounted for 5 of the discordant genotypes and had higher rates of missing data than other SNPs studied. Neither of these two markers contributed association evidence, and when they were excluded, 7/533 (13 $\times$ 41) duplicate pairs were discordant (1.31%). All markers satisfied HWE criteria.

3.3. Single marker analysis in the ICCSS sample

In this study, we genotyped 15 tag SNPs in four candidate genes *GRB2*, *HSPA5*, *YWHAG*, and *YWHAZ* in the ICCSS sample. Table 1 shows the results of single marker analysis. We observed seven SNPs showing significant association with schizophrenia, six of which are in *GRB2* and one in *YWHAZ*. We detect no significant association signal from markers in *HSPA5* or *YWHAG*. Six out of the seven genotyped SNPs in *GRB2* (rs7207618, rs12600908, rs9912608, rs4789176, rs4350602, and rs4788891) showed evidence of association and one (rs4789172) missed this significance threshold marginally. Among them, rs7207618 (C/T) had the strongest association with schizophrenia, with C allele frequency of 0.80 in patients and 0.75 in controls ($\chi^2 = 13.904$, *P* = 0.0003, odds ratio (OR) = 1.37). Furthermore, two SNPs in *GRB2* (rs7207618 and rs9912608) remained significant (*P* = 0.0035 and 0.0276, respectively) after 100,000 permutations. These two SNPs were also significant after Bonferroni correction. The two SNPs are in strong linkage disequilibrium (*D'* = 0.996) and, thus, their association tests with schizophrenia are non-independent. In *YWHAZ*, one SNP (rs4734497) was found to be associated with schizophrenia at nominal significance level (*P* = 0.0224, Table 1), but did not remain significant after permutation testing.

3.4. Haplotype analysis of *GRB2* and *YWHAZ* in the ICCSS sample

We examined LD and haplotype structures of *GRB2* and *YWHAZ* by software Haploview using the genotype data from the ICCSS sample. In *GRB2*, SNPs 1-2 (rs7207618 and rs12600908) and SNPs 3-6 (rs9912608, rs4789172, rs4789176, and rs4350602) constitute two LD blocks (Fig. 2). The LD pattern observed in the ICCSS sample was similar to that observed in the HapMap CEU sample. In *YWHAZ*, SNPs 2-3 (rs4734497 and rs1470764) constitute one LD block.

Table 2 shows the results of haplotype analysis. We observed evidence of association between schizophrenia and five *GRB2* haplotypes and two *YWHAZ* haplotypes at nominal significance level of *P* < 0.05. However, after 100,000 permutations, only one haplotype, rs7207618- rs12600908 in *GRB2*, remained significant ($\chi^2 = 13.94$, *P* = 0.0003, permutation *P* = 0.0035, OR = 1.34). Note that the significance of this haplotype was largely driven by the SNP rs7207618, which is the most significant *GRB2* marker and it alone could reach a similar significance level (OR = 0.73, *P* = 0.0003, Table 1).

4. Discussion

In this study, we proposed a unique approach to selecting novel candidate genes for schizophrenia based on gene network and pathway analysis. Our approach is different from the traditional candidate gene screening, which employs prior biological knowledge to select novel candidate genes. The hypothesis underlying this approach is that multiple genes, each of which might contribute a small effect by itself, interact with each other to increase risk of schizophrenia. These genes might be functionally linked through the same or related pathways/networks. Further investigation of these pathways/networks, especially the novel genes out of them, may accelerate our discovery of susceptibility genes for complex diseases and our understanding of disease etiology at the cellular level. Based on this approach, we selected four genes that are functionally important in our constructed general neurotransmitter network (Fig. 1) for validation in the ICCSS sample. These four genes have not been well studied for the association with schizophrenia and their association signal was weak in our survey of three GWAS datasets (see section 3.1). We confirmed *GRB2* to be associated with schizophrenia using our ICCSS sample.

However, only one of the four novel genes we tested in this study was detected to be significantly associated with schizophrenia using our ICCSS sample. While this is still promising, several factors might have limited its effectiveness. First, in this study we tested only a few markers in three genes (*HSPA5*, *YWHAG*, and *YWHAZ*). Second, for the purpose of proof of concept, we selected the candidate genes that have not been studied (*HSPA5* and *YWHAG*) or were reported negative or weak association (*GRB2* and *YWHAZ*) with schizophrenia. Third, sample size in our study might not be large enough to gain power to detect association signal of the alleles that have moderate or small risk to a complex disease such as schizophrenia. Finally, biological data such as protein-protein interactions is still incomplete and with false-positives.

To further examine the association signal observed in our ICCSS sample, we performed a meta-analysis of the seven nominally significant SNPs (6 in *GRB2* and 1 in *YWHAZ*, Table 1) with three GWA studies (GAIN, nonGAIN, and CATIE). None of the seven SNPs were included in any of the three GWAS datasets. To estimate their genotypes in each GWA study, we performed an imputation analysis using the software IMPUTE v2.1.2 (Marchini et al. 2007) based on the HapMap phase II European ancestry genotype dataset (release #22). A genotype was accepted when its probability was >0.90. Then, we used the additive model implemented in software SNPTTEST v2 to obtain statistical association values for each SNP in each GWA study. We next performed meta-analysis using the inverse variance method (based on random-effects model) implemented in software META v1.2 (Marchini et al. 2007) for each SNP whose imputed genotype data existed in at least 90% individuals. No SNP had a combined *P*-value less than 0.05. Our further examination found all the six SNPs in *GRB2* presented flip-flop of their risk alleles in the four datasets (three GWAS datasets and ICCSS dataset), which substantially weakened the overall association. Furthermore, the *P*-values of those six SNPs in *GRB2* varied greatly among the four datasets, especially SNPs rs7207618 and rs9912608, whose heterogeneity tests were significant (Cochran's Q-test *P*-values were 0.0061 and 0.0257, respectively). Population stratification is a critical issue in detecting risk genes/loci (Oquendo et al. 2010). Our preliminary meta-analysis indicated that phenotypes and population substructures among the four studies (ICCSS, GAIN, nonGAIN, and CATIE) might be complicated and needs further cleaning. Other possible reasons of meta-analysis failure include: 1) risk allele(s) in *GRB2* might depend on population and our results were Irish-specific; 2) our results in ICCSS sample were false positives; and 3) Imputing process generated inaccurate genotypes in the samples of the three GWA studies.

We also tested the epistatic interactions of the four genes using the genotype data in ICCSS sample and software PLINK v1.07 (Purcell et al. 2007). Among the 105 possible SNP-SNP interaction pairs we tested, only three (rs13254653-rs4789172, rs1470764-rs4789172, and rs13254653-rs391957) had nominal *P*-values less than 0.05, none of which could survive after Bonferroni multiple test correction.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Ms. Cuie Sun for her help in the genotyping, Mr. Tim B. Bigdeli and Drs. Xiangning Chen and Qi Chen for their help with genotyping data analysis. The dataset(s) used for the analyses described in this manuscript were obtained from the database of Genotype and Phenotype (dbGaP) found at <http://www.ncbi.nlm.nih.gov/gap> through dbGaP accession number [GAIN: phs000021.v2.p1, nonGAIN: phs00167.v1.p1]. For GAIN dataset, the genotyping of samples was provided through the Genetic Association Information Network (GAIN). Samples and associated phenotype data for the Linking Genome-Wide Association Study of Schizophrenia were provided by P. German. The CATIE trial was funded by a grant from the National Institute of Mental Health (N01 MH900001) along with MH074027 (PI PF Sullivan). Genotyping was funded by Eli Lilly and Company. The CATIE dataset was approved to use in this analysis through our application.

Role of the funding source: This work was supported by National Institutes of Health Grant Nos. AA017437 and MH083094, NARSAD Maltz Investigator Award to Z.Z., Thomas F. and Kate Miller Jeffress Memorial Trust Fund grant No. J-900, and Department of Psychiatry at the Vanderbilt University. The funding agencies had no further role in the design, implementation, or generation of this research report.

References

- Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–265. [PubMed: 15297300]
- Bell R, Munro J, Russ C, Powell JF, Bruinvels A, Kerwin RW, Collier DA. Systematic screening of the 14-3-3 eta (eta) chain gene for polymorphic variants and case-control analysis in schizophrenia. *Am J Med Genet* 2000;96:736–743. [PubMed: 11121172]
- Berg D, Holzmann C, Riess O. 14-3-3 proteins in the nervous system. *Nat Rev Neurosci* 2003;4:752–762. [PubMed: 12951567]
- Cavalli-Sforza, LL.; Menozzi, P.; Piazza, A. *The History and Geography of Human genes*. Princeton, NJ: Princeton University Press; 1994.
- Chen X, Wang X, Hossain S, O'Neill FA, Walsh D, van den Oord E, Fanous A, Kendler KS. Interleukin 3 and schizophrenia: the impact of sex and family history. *Mol Psychiatry* 2007;12:273–282. [PubMed: 17179997]
- Cogram P, Hynes A, Dunlevy LP, Greene ND, Copp AJ. Specific isoforms of protein kinase C are essential for prevention of folate-resistant neural tube defects by inositol. *Hum Mol Genet* 2004;13:7–14. [PubMed: 14613966]
- Fu H, Subramanian RR, Masters SC. 14-3-3 proteins: structure, function, and regulation. *Annu Rev Pharmacol Toxicol* 2000;40:617–647. [PubMed: 10836149]
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The structure of haplotype blocks in the human genome. *Science* 2002;296:2225–2229. [PubMed: 12029063]
- Giegling I, Hartmann AM, Genius J, Benninghoff J, Moller HJ, Rujescu D. Systems biology and complex neurobehavioral traits. *Pharmacopsychiatry* 2008;41 1:S32–36. [PubMed: 18756418]
- Hashimoto K, Tomitaka S, Narita N, Minabe Y, Iyo M, Fukui S. Induction of heat shock protein (HSP)-70 in posterior cingulate and retrosplenial cortex of rat brain by dizocilpine and phencyclidine: lack of protective effects of sigma receptor ligands. *Addict Biol* 1996;1:61–70. [PubMed: 12893487]

- Ikeda M, Hikita T, Taya S, Uruguchi-Asaki J, Toyo-oka K, Wynshaw-Boris A, Ujike H, Inada T, Takao K, Miyakawa T, Ozaki N, Kaibuchi K, Iwata N. Identification of YWHAE, a gene encoding 14-3-3epsilon, as a possible susceptibility gene for schizophrenia. *Hum Mol Genet* 2008;17:3212–3222. [PubMed: 18658164]
- Ioannidis JP, Boffetta P, Little J, O'Brien TR, Uitterlinden AG, Vineis P, Balding DJ, Chokkalingam A, Dolan SM, Flanders WD, Higgins JP, McCarthy MI, McDermott DH, Page GP, Rebbeck TR, Seminara D, Khoury MJ. Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol* 2008;37:120–132. [PubMed: 17898028]
- Jia P, Wang L, Meltzer HY, Zhao Z. Common variants conferring risk of schizophrenia: A pathway analysis of GWAS data. *Schizophr Res* 2010;122:38–42. [PubMed: 20659789]
- Kakiuchi C, Ishiwata M, Nanko S, Kunugi H, Minabe Y, Nakamura K, Mori N, Fujii K, Umekage T, Tochigi M, Kohda K, Sasaki T, Yamada K, Yoshikawa T, Kato T. Functional polymorphisms of HSPA5: possible association with bipolar disorder. *Biochem Biophys Res Commun* 2005;336:1136–1143. [PubMed: 16168956]
- Klein P, Ravi R. A nearly best-possible approximation algorithm for node-weighted Steiner trees. *J Algorithms* 1995;19:104–115.
- Lintas C, Sacco R, Garbett K, Mirnics K, Militeri R, Bravaccio C, Curatolo P, Manzi B, Schneider C, Melmed R, Elia M, Pascucci T, Puglisi-Allegra S, Reichelt KL, Persico AM. Involvement of the PRKCB1 gene in autistic disorder: significant genetic association and reduced neocortical gene expression. *Mol Psychiatry* 2009;14:705–718. [PubMed: 18317465]
- Lowenstein EJ, Daly RJ, Batzer AG, Li W, Margolis B, Lammers R, Ullrich A, Skolnik EY, Bar-Sagi D, Schlessinger J. The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell* 1992;70:431–442. [PubMed: 1322798]
- Manolio TA, Rodriguez LL, Brooks L, Abecasis G, Ballinger D, Daly M, Donnelly P, Faraone SV, Frazer K, Gabriel S, Gejman P, Guttmacher A, Harris EL, Insel T, Kelsoe JR, Lander E, McCowin N, Mailman MD, Nabel E, Ostell J, Pugh E, Sherry S, Sullivan PF, Thompson JF, Warram J, Wholley D, Milos PM, Collins FS. New models of collaboration in genome-wide association studies: the Genetic Association Information Network. *Nat Genet* 2007;39:1045–1051. [PubMed: 17728769]
- Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007;39:906–913. [PubMed: 17572673]
- McClellan JM, Susser E, King MC. Schizophrenia: a common disease caused by multiple rare alleles. *Br J Psychiatry* 2007;190:194–199. [PubMed: 17329737]
- Miyamoto S, LaMantia AS, Duncan GE, Sullivan P, Gilmore JH, Lieberman JA. Recent advances in the neurobiology of schizophrenia. *Mol Interv* 2003;3:27–39. [PubMed: 14993436]
- Oquendo MA, Canino G, Lehner T, Licinio J. Genetic repositories for the study of major psychiatric conditions: what do we know about ethnic minorities' genetic vulnerability? *Mol Psychiatry* 2010;15:970–975. [PubMed: 20177407]
- Pham X, Sun C, Chen X, van den Oord EJ, Neale MC, Kendler KS, Hettema JM. Association study between GABA receptor genes and anxiety spectrum disorders. *Depress Anxiety* 2009;26:998–1003. [PubMed: 19842164]
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559–575. [PubMed: 17701901]
- Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, Sklar P. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009;460:748–752. [PubMed: 19571811]
- Riley B, Kuo PH, Maher BS, Fanous AH, Sun J, Wormley B, O'Neill FA, Walsh D, Zhao Z, Kendler KS. The dystrobrevin binding protein 1 (DTNBP1) gene is associated with schizophrenia in the Irish Case Control Study of Schizophrenia (ICCSS) sample. *Schizophr Res* 2009;115:245–253. [PubMed: 19800201]
- Roh K, Roh S, Yang BH, Lee JS, Chai YG, Choi MR, Park YC, Kim DJ, Kim D, Choi J, Kim SH. Effects of haloperidol and risperidone on the expression of heat shock protein 70 in MK-801-

- treated rat C6 glioma cells. *Prog Neuropsychopharmacol Biol Psychiatry* 2008;32:1793–1797. [PubMed: 18721842]
- Ross CA, Margolis RL, Reading SA, Pletnikov M, Coyle JT. Neurobiology of schizophrenia. *Neuron* 2006;52:139–153. [PubMed: 17015232]
- Ruano D, Abecasis GR, Glaser B, Lips ES, Cornelisse LN, de Jong AP, Evans DM, Davey Smith G, Timpson NJ, Smit AB, Heutink P, Verhage M, Posthuma D. Functional gene group analysis reveals a role of synaptic heterotrimeric G proteins in cognitive ability. *Am J Hum Genet* 2010;86:113–125. [PubMed: 20060087]
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13:2498–2504. [PubMed: 14597658]
- Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I, Dudbridge F, Holmans PA, Whittemore AS, Mowry BJ, Olincy A, Amin F, Cloninger CR, Silverman JM, Buccola NG, Byerley WF, Black DW, Crowe RR, Oksenberg JR, Mirel DB, Kendler KS, Freedman R, Gejman PV. Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 2009;460:753–757. [PubMed: 19571809]
- Shinoda T, Taya S, Tsuboi D, Hikita T, Matsuzawa R, Kuroda S, Iwamatsu A, Kaibuchi K. DISC1 regulates neurotrophin-induced axon elongation via interaction with Grb2. *J Neurosci* 2007;27:4–14. [PubMed: 17202467]
- Sullivan PF, Lin D, Tzeng JY, van den Oord E, Perkins D, Stroup TS, Wagner M, Lee S, Wright FA, Zou F, Liu W, Downing AM, Lieberman J, Close SL. Genomewide association for schizophrenia in the CATIE study: results of stage 1. *Mol Psychiatry* 2008;13:570–584. [PubMed: 18347602]
- Sun J, Jia P, Fanous AH, Oord E, Chen X, Riley BP, Amdur RL, Kendler KS, Zhao Z. Schizophrenia gene networks and pathways and their applications for novel candidate gene selection. *PLoS ONE* 2010;5:e11351. [PubMed: 20613869]
- Sun J, Jia P, Fanous AH, Webb BT, van den Oord EJ, Chen X, Bukszar J, Kendler KS, Zhao Z. A multi-dimensional evidence-based candidate gene prioritization approach for complex diseases—schizophrenia as a case. *Bioinformatics* 2009;25:2595–2602. [PubMed: 19602527]
- Takenawa T, Miki H, Matuoka K. Signaling through Grb2/Ash-control of the Ras pathway and cytoskeleton. *Curr Top Microbiol Immunol* 1998;228:325–342. [PubMed: 9401211]
- Umahara T, Uchihara T. 14-3-3 proteins and spinocerebellar ataxia type 1: from molecular interaction to human neuropathology. *Cerebellum* 2010;9:183–189. [PubMed: 20155408]
- Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, Nord AS, Kusenda M, Malhotra D, Bhandari A, Stray SM, Rippey CF, Rocanova P, Makarov V, Lakshmi B, Findling RL, Sikich L, Stromberg T, Merriman B, Gogtay N, Butler P, Eckstrand K, Noory L, Gochman P, Long R, Chen Z, Davis S, Baker C, Eichler EE, Meltzer PS, Nelson SF, Singleton AB, Lee MK, Rapoport JL, King MC, Sebat J. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 2008;320:539–543. [PubMed: 18369103]
- Wong AH, Likhodi O, Trakalo J, Yusuf M, Sinha A, Pato CN, Pato MT, Van Tol HH, Kennedy JL. Genetic and post-mortem mRNA analysis of the 14-3-3 genes that encode phosphoserine/threonine-binding regulatory proteins in schizophrenia and bipolar disorder. *Schizophr Res* 2005;78:137–146. [PubMed: 16054338]

Fig. 1.

A merged protein-protein interaction network based on four neurotransmitter receptor signaling pathways. P1 represents glutamate receptor signaling, P2 represents GABA receptor signaling, P3 represents serotonin receptor signaling, and P4 represents dopamine receptor signaling. Nodes in red denote SZGenes (schizophrenia candidate genes from our previous work, see text) and nodes in grey denote novel candidate genes recruited by network construction. Node size is relative to its frequency appearing in the four neurotransmitter-related pathways (e.g., GRB2 appeared in all four pathways). A green edge indicates its interaction appeared 3 times in the four pathways, a blue edge indicates twice, and a grey edge indicates once.

Table 1
Allele distribution of the polymorphisms from four potential candidate genes in ICCSS sample.

Gene	Chr.	SNP	Position (bp)	Allele	HWE <i>P</i>	Case frequency	Control frequency	OR (95% CI)	<i>P</i> -value	Perm. <i>P</i> -value
<i>GRB2</i>	17	rs7207618	70833282	T/C	0.1294	0.20/0.80	0.25/0.75	0.73 (0.61-0.87)	0.0003	0.0035
		rs12600908	70839969	G/A	0.6186	0.91/0.09	0.89/0.11	1.28 (1.00-1.63)	0.0299	0.3682
		rs9912608	70848626	G/C	0.0332	0.18/0.82	0.22/0.78	0.78 (0.65-0.93)	0.0025	0.0276
		rs4789172	70853307	T/C	0.5396	0.56/0.44	0.53/0.47	1.15 (1.00-1.33)	0.0513	0.4745
		rs4789176	70864463	T/A	0.8397	0.87/0.13	0.84/0.16	1.22 (0.99-1.48)	0.0373	0.1827
<i>HSPA5</i>	9	rs4350602	70867364	T/C	0.4654	0.77/0.23	0.74/0.26	1.22 (1.03-1.45)	0.0216	0.3014
		rs4788891	70902740	G/A	0.7514	0.86/0.14	0.82/0.18	1.25 (1.03-1.53)	0.0184	0.0588
		rs12009	127037124	G/A	0.0566	0.49/0.51	0.25/0.51	1.00 (0.87-1.16)	0.9466	0.9466
<i>YWHAQ</i>	7	rs391957	127043845	T/C	0.0027	0.43/0.57	0.43/0.57	0.99 (0.86-1.15)	0.9123	0.9123
		rs2961037	75790148	C/A	0.9607	0.47/0.53	0.44/0.56	0.89 (0.78-1.03)	0.1291	0.2980
		rs11765693	75823309	G/A	0.7425	0.69/0.31	0.68/0.32	0.96 (0.82-1.11)	0.5317	0.8441
<i>YWHAZ</i>	8	rs17149177	75823488	A/G	0.3475	0.80/0.20	0.78/0.22	1.14 (0.96-1.37)	0.2070	0.4464
		rs13254653	101996871	C/T	0.0379	0.39/0.61	0.37/0.63	1.07 (0.92-1.25)	0.3639	0.6616
		rs4734497	102004147	T/C	0.0181	0.37/0.63	0.33/0.67	0.84 (0.73-0.98)	0.0224	0.0630
		rs1470764	102019214	T/C	0.2397	0.62/0.38	0.61/0.39	0.96 (0.83-1.11)	0.7755	0.9820

ICCSS, Irish Case-Control Study of Schizophrenia; SNP, single-nucleotide polymorphism; HWE *P*, *P*-value for Hardy-Weinberg equilibrium (HWE); OR, odds ratio; CI, confidence interval; Perm. *P*, permutation *P*-value.

Table 2

Haplotype analysis of genes *GRB2* and *YWHAZ* in ICCSS sample.

Haplotype	Case frequency	Control frequency	χ^2	OR (95% CI)	P-value	Perm. P-value
<i>GRB2</i> block 1: rs7207618 - rs12600908						
CG	0.803	0.746	13.937	1.34 (1.12-1.59)	0.0003	0.0035
TG	0.110	0.141	6.223	0.78 (0.63-0.98)	0.0187	0.1536
TA	0.086	0.111	5.819	0.77 (0.60-0.98)	0.0166	0.1921
<i>GRB2</i> block 2: rs9912608 - rs4789172 - rs4789176 - rs4350602						
CTTT	0.562	0.524	4.145	1.15 (0.99-1.34)	0.0272	0.3519
CCTT	0.209	0.208	0.014	1.02 (0.85-1.23)	0.9632	1.0000
GCAC	0.132	0.154	3.314	0.78 (0.63-0.96)	0.0751	0.6029
GCTC	0.050	0.070	5.254	0.72 (0.53-0.98)	0.0184	0.1825
CCTC	0.043	0.034	0.924	1.42 (0.94-2.15)	0.3466	0.9937
<i>YWHAZ</i> : rs4734497 - rs1470764						
CC	0.377	0.384	0.180	0.99 (0.85-1.16)	0.6717	0.9112
TT	0.365	0.323	5.422	1.18 (1.01-1.38)	0.0199	0.0544
CT	0.251	0.287	4.522	0.83 (0.71-0.98)	0.0335	0.0869

OR, odds ratio; CI, confidence interval; Perm. P, permutation P-value.