Psychopathology Scale between schizophrenia cases and control subjects using a two-tailed t-test and Welch's t-test. We compared sex between case and control groups using Fisher's exact test. Next, we compared d' in the CPT and CA, PEM, PEN, TE in the WCST between the case and control groups using a two-tailed t-test and Welch's test (Table S4).

Supporting Information

Method S1 Meta-analysis.
(DOC)

Table S1 Haplotype analysis of the eight SNPs of *NETO1*.
(DOC)

Table S2 Allele frequencies of the eight SNPs of *NETO1* in males.
(DOC)

Table S3 Allele frequencies of the eight SNPs of *NETO1* in females.
(DOC)

Table S4 Cognitive performance of three SNPs in *NETO1*.
(DOC)

Acknowledgments

We sincerely thank the patients and healthy volunteers for their participation in this study. We express our gratitude to Ryoko Ishihara, Ph.D., Hiromi Noma, Saori Yamashita, Ph.D., and Mami Yoshida for their technical assistance.

Author Contributions

Conceived and designed the experiments: MB BA NO. Performed the experiments: MB T. Koide T. Kikuchi KK YA NK. Analyzed the data: MB T. Koide BA KY IK MI TI TY NI NO. Contributed reagents/materials/analysis tools: T. Koide YA. Wrote the paper: MB BA NO.

References

1. Steinpreis R (1996) The behavioral and neurochemical effects of phencyclidine in humans and animals: some implications for modeling psychosis. Behav Brain Res 74: 45–55.
2. Kristiansen LV, Huerta I, Beneyto M, Meador-Woodruff JH (2007) NMDA receptors and schizophrenia. Current opinion in pharmacology 7: 48–55.
3. Stephan KE, Baldeweg T, Friston KJ (2006) Synaptic plasticity and dysconnection in schizophrenia. Biol Psychiatry 59: 929–939.
4. Belforte J, Zsiros V, Sklar E, Jiang Z, Yu G, et al. (2010) Postnatal NMDA receptor ablation in corticolimbic interneurons confers schizophrenia-like phenotypes. Nat Neurosci 13: 76–83.
5. Stohr H, Berger C, Frohlich S, Weber BH (2002) A novel gene encoding a putative transmembrane protein with two extracellular CUB domains and a low-density lipoprotein class A module: isolation of alternatively spliced isoforms in retina and brain. Gene 286: 223–231.
6. Li F, Tsien JZ (2009) Memory and the NMDA receptors. N Engl J Med 361: 302–303.
7. Ng D, Pitcher GM, Szilard RK, Sertie A, Kanisek M, et al. (2009) Neto1 is a novel CUB-domain NMDA receptor-interacting protein required for synaptic plasticity and learning. PLoS Biol 7: e41.
8. Tang TT, Yang F, Chen BS, Lu Y, Ji Y, et al. (2009) Dysbindin regulates hippocampal LTP by controlling NMDA receptor surface expression. Proc Natl Acad Sci U S A 106: 21395–21400.
9. Gao XM, Sakai K, Roberts RC, Conley RR, Dean B, et al. (2000) Ionotropic glutamate receptors and expression of N-methyl-D-aspartate receptor subunits in subregions of human hippocampus: effects of schizophrenia. Am J Psychiatry 157: 1141–1149.
10. Straub C, Hunt DL, Yamasaki M, Kim KS, Watanabe M, et al. (2011) Distinct functions of kainate receptors in the brain are determined by the auxiliary subunit Netol. Nature neuroscience.
11. Copits BA, Robbins JS, Frausto S, Swanson GT (2011) Synaptic Targeting and Functional Modulation of GluK1 Kainate Receptors by the Auxiliary Neuropilin and Tolloid-Like (NETO) Proteins. The Journal of neuroscience: the official journal of the Society for Neuroscience 31: 7334–7340.
12. O'Donnell L, Soileau B, Heard P, Carter E, Sebold C, et al. (2010) Genetic determinants of autism in individuals with deletions of 18q. Hum Genet 128: 155–164.
13. Li H, Wetten S, Li L, St Jean PL, Upmanyu R, et al. (2008) Candidate single-nucleotide polymorphisms from a genomewide association study of Alzheimer disease. Arch Neurol 65: 45–53.
14. Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, et al. (2009) Common variants on chromosome 6p22.1 are associated with schizophrenia. Nature 460: 753–757.
15. Ikeda M, Aleksic B, Kinoshita Y, Okochi T, Kawashima K, et al. (2011) Genome-wide association study of schizophrenia in a Japanese population. Biol Psychiatry 69: 472–478.
16. Shi G, Boerwinkle E, Morrison AC, Gu CC, Chakravarti A, et al. (2011) Mining gold dust under the genome wide significance level: a two-stage approach to analysis of GWAS. Genetic epidemiology 35: 111–118.
17. Yamada K, Iwayama Y, Hattori E, Iwamoto K, Toyota T, et al. (2011) Genome-wide association study of schizophrenia in Japanese population. PloS one 6: e20468.
18. Lin PI, Vance JM, Pericak-Vance MA, Martin ER (2007) No gene is an island: the flip-flop phenomenon. American journal of human genetics 80: 531–538.
19. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, et al. (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 460: 748–752.
20. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, et al. (2009) Finding the missing heritability of complex diseases. Nature 461: 747–753.
21. Anderson CA, Soranzo N, Zeggini E, Barrett JC (2011) Synthetic associations are unlikely to account for many common disease genome-wide association signals. PLoS Biol 9: e1000580.
22. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB (2010) Rare variants create synthetic genome-wide associations. PLoS Biol 8: e1000294.
23. Aleksic B, Kushima I, Ito Y, Nakamura Y, Ujike H, et al. (2010) Genetic association study of KREMEN1 and DKK1 and schizophrenia in a Japanese population. Schizophrenia research 118: 113–117.
24. Barrett J, Fry B, Maller J, Daly M (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263–265.
25. Skol AD, Scott LJ, Abecasis GR, Boehnke M (2006) Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet 38: 209–213.
26. Dudbridge F (2008) Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. Human heredity 66: 87–98.
27. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559–575.
28. Cornblatt BA, Risch NJ, Faris G, Friedman D, Erlenmeyer-Kimling L (1988) The Continuous Performance Test, identical pairs version (CPT-IP): I. New findings about sustained attention in normal families. Psychiatry Res 26: 223–238.
29. Heaton RK (1981) The Wisconsin Card Sorting Test (Manual). OdessaFL: Psychological Assessment Resources.
30. Kashima H, Handa T, Kato M, Sakuma K, Yokoyama N, et al. (1987) Neuropsychological investigation on chronic schizophrenia. Aspects of its frontal functions. Takahashi R, Flor-Henry P, Gruzelier J, Niwa S, eds. Amsterdam: Elsevier.
31. Hori H, Noguchi H, Hashimoto R, Nakabayashi T, Omori M, et al. (2006) Antipsychotic medication and cognitive function in schizophrenia. Schizophr Res 86: 138–146.
32. Suzuki H, Gen K, Inoue Y (2011) An unblinded comparison of the clinical and cognitive effects of switching from first-generation antipsychotics to aripiprazole, perospirone or olanzapine in patients with chronic schizophrenia. Prog Neuropsychopharmacol Biol Psychiatry 35: 161–168.
33. Heaton RK, Chelune GL, Talley JL, Kay GG, Curtiss G (1993) Wisconsin Card Sorting Test manual: revised and expanded. OclessaFL: Psychological Assessment Resourses.
34. Inagaki A, Inada T (2008) Dose equivalence of psychotropic drugs. Part XXI: dose equivalence of novel antipsychotics: Blonanserin. Rinsho Seishin Yakuri (Japanese Journal of Clinical Psychopharmacology). pp 887–890.
35. Inagaki A, Inada T (2010) Dose equivalence of psychotropic drugs. Part XXII: dose equivalence of depot antipsychotics III: risperidon long-acting injection. Rinsho Seishin Yakuri (Japanese Journal of Clinical Psychopharmacology). pp 1349–1353.
36. Kay SR, Fiszbein A, Opler LA (1987) The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull 13: 261–276.

Schizophrenia Bulletin Advance Access published November 1, 2010

Schizophrenia Bulletin
doi:10.1093/schbul/sbq118

Resequencing and Association Analysis of the *KALRN* and *EPHB1* Genes And Their Contribution to Schizophrenia Susceptibility

Itaru Kushima[1,2], Yukako Nakamura[1,2], Branko Aleksic[1,2], Masashi Ikeda[2,3], Yoshihito Ito[1,2], Tomoko Shiino[1,2], Tomo Okochi[2,3], Yasuhisa Fukuo[2,3], Hiroshi Ujike[4], Michio Suzuki[2,5], Toshiya Inada[6], Ryota Hashimoto[2,7,8], Masatoshi Takeda[7,8], Kozo Kaibuchi[2,9], Nakao Iwata[*,2,3], and Norio Ozaki[1,2]

[1]Department of Psychiatry, Graduate School of Medicine, Nagoya University, Nagoya, Japan; [2]Core Research for Evolutional Science and Technology, Japan Science and Technology Corporation, Tokyo, Japan; [3]Department of Psychiatry, School of Medicine, Fujita Health University, Toyoake, Aichi 470-1192, Japan; [4]Department of Neuropsychiatry, Graduate School of Medicine, Okayama University, Dentistry and Pharmaceutical Sciences, Okayama, Japan; [5]Department of Neuropsychiatry, Graduate School of Medicine and Pharmaceutical Sciences, Toyama University, Toyama, Japan; [6]Seiwa Hospital, Institute of Neuropsychiatry, Tokyo, Japan; [7]Molecular Research Center for Children's Mental Development, United Graduate School of Child Development, Osaka University, Kanazawa University and Hamamatsu University School of Medicine, Osaka, Japan; [8]Department of Psychiatry, Graduate School of Medicine, Osaka University, Osaka, Japan; [9]Department of Cell Pharmacology, Graduate School of Medicine, Nagoya University, Nagoya, Japan

[*]To whom correspondence should be addressed; tel: +81-562-93-9250; fax: +81-562-93-1831, e-mail: nakao@fujita-hu.ac.jp

Background: **Our genome-wide association study of schizophrenia found association signals at the Kalirin gene (*KALRN*) and EPH receptor B1 gene (*EPHB1*) in a Japanese population. The importance of these synaptogenic pathway genes in schizophrenia is gaining independent supports. Although there has been growing interest in rare (<1%) missense mutations as potential contributors to the unexplained heritability of schizophrenia, there are no population-based studies targeting rare (<1%) coding mutations with a larger effect size (eg, OR >1.5) in *KALRN* or *EPHB1*. *Methods and Results:*** **The present study design consisted of 3 phases. At the discovery phase, we conducted resequencing analyses for all exon regions of *KALRN* and *EPHB1* using a DNA microarray–based method. Seventeen rare (<1%) missense mutations were discovered in the first sample set (320 schizophrenic patients). After the prioritization phase based on frequencies in the second sample set (729 cases and 562 controls), we performed association analyses for each selected mutation using the third sample set (1511 cases and 1517 controls), along with a combined association analysis across all selected mutations. In *KALRN*, we detected a significant association between schizophrenia and P2255T (OR = 2.09, corrected $P = .048$, 1 tailed); this was supported in the combined association analysis (OR = 2.07, corrected $P = .006$, 1 tailed). We found no evidence of association of *EPHB1* with schizophrenia. *In silico* analysis indicated the functional relevance of these rare missense mutations.** ***Conclusion:*** **We provide evidence that multiple rare (<1%) missense mutations in *KALRN* may be genetic risk factors for schizophrenia.**

Key words: synaptogenic pathway/rare missense mutations/GWAS/Japanese population

Introduction

Schizophrenia is a genetically heterogeneous disorder with heritability estimated at up to 80%.[1] According to a recent simulation based on genome-wide association study (GWAS) datasets, a highly polygenic model involving a number of common variants of very small effect may explain more than one-third of the total variation in risk of schizophrenia.[2] On the other hand, interest has been growing in rare variants as potential contributors to the unexplained heritability of schizophrenia.[3] This is partly triggered by recent studies establishing an important role for rare genomic copy number variants (CNVs) in the etiology of schizophrenia.[4] Another potential genetic variation to explain the remaining heritability is rare missense mutations. Kryukov et al[5] reported that ~20% of new (de novo) missense mutations in humans result in a loss of function, whereas ~53% have mildly deleterious effects and ~27% are effectively neutral with respect to phenotype by a combined analysis of mutations causing human Mendelian diseases, mutations driving human-chimpanzee sequence divergence, and systematic data on human genetic variation. Their results were supported by an independent study.[6] Because the pressure of purifying selection acting on the mildly deleterious mutations is weak, their cumulative high frequency in the human population is being maintained

by "mutation-selection balance." This provides support to a speculation that the accumulation of mildly deleterious missense mutations in individual human genomes can be a genetic basis for complex diseases.[5] The importance of rare missense mutations in schizophrenia is demonstrated by a study of the *ABCA13* gene in which multiple rare (<1%) coding variants were associated with schizophrenia.[7]

We recently performed a GWAS for schizophrenia in a Japanese population.[8] Although single locus analysis did not reveal genome-wide support for any locus, a shared polygenic risk of schizophrenia between the Japanese and the Caucasian samples was confirmed. In our GWAS, association signals were detected at the regions of the Kalirin gene (*KALRN*) on 3q21.2 and the EPH receptor B1 gene (*EPHB1*) on 3q21-q23, both of which are in the same synaptogenic pathway[9] (supplementary figure S1). Associations of each gene with schizophrenia have recently received support from independent GWASs in different populations.[10,11] Furthermore, a rare de novo CNV overlapping with the *EPHB1* gene locus was detected in a patient with schizophrenia.[12]

KALRN is a large neuronal dual Rho guanine nucleotide exchange factor (GEF) that activates small guanosine triphosphate–binding proteins of the Rho family, including Rac1.[13] This activation enables KALRN to regulate neurite initiation, axonal growth, dendritic morphogenesis, and spine morphogenesis. Consistent with its biological function, KALRN is a key factor responsible for reduced densities of dendritic spines on pyramidal neurons in the dorsolateral prefrontal cortex (DLPFC)[14] observed in postmortem brains from schizophrenic patients. The messenger RNA expression level of *KALRN* is significantly reduced in DLPFC of patients with schizophrenia and strongly correlated with spine density.[15] In addition, *KALRN*-knockout mice not only exhibit spine loss and reduced glutamatergic transmission in the frontal cortex but also schizophrenia-like phenotypes including robust deficits in working memory, sociability, prepulse inhibition, and locomotor hyperactivity reversible by clozapine, an atypical antipsychotic.[16] These synaptic and behavioral dysfunctions are apparent during young adulthood in mice (12 weeks old), which coincides with the onset of schizophrenia in patients. Notably, Disrupted-in-Schizophrenia 1, a prominent schizophrenia risk factor, was shown to be involved in the maintenance of spine morphology and function by regulating access of KALRN to Rac1.[17] *EPHB1* belongs to a receptor tyrosine kinase family and controls multiple aspects of neuronal development, including synapse formation and maturation, as well as synaptic structural and functional plasticity. In neurons, activation of EphB receptors by its ligand B-type ephrins induces the rapid formation and enlargement of dendritic spines, as well as rapid synapse maturation. One of the downstream effectors of ephrinB/EphB signaling is KALRN. In young hippocampal neurons, KALRN is reported to play an important role in the maturation of synapses induced by trans-synaptic ephrinB/EphB signaling.[18]

According to the above-mentioned study,[5] most missense mutations with a frequency of <1% are mildly deleterious, indicating that a low frequency of missense mutation per se can serve as a strong predictor of a deleterious effect of variants. Therefore, the working hypothesis of the present study is that rare (<1%) missense or nonsense mutations with a larger effect size (eg, OR >1.5) in *KALRN* and *EPHB1* may be genetic risk factors for schizophrenia. Recently, a DNA microarray–based resequencing method has been developed to enable accurate and rapid resequencing analysis of candidate genes.[19] Using this system, we conducted resequencing analyses for all exon regions of *KALRN* and *EPHB1* in 320 schizophrenic patients and found evidence that rare (<1%) missense mutations in *KALRN* are significantly associated with schizophrenia using the 3-phase study design.

Methods and Materials

Subjects

Three sample sets were used in this study. The first sample set, comprising 320 schizophrenic patients (mean age, 54.2 ± 14.1 years, 49.1% male), with long-term hospitalization for severe symptoms, was used to search for rare missense or nonsense mutations. We used the first sample set for mutation screenings because patients with extreme phenotypes (severe symptoms) can be expected to carry more deleterious mutations.[20] The second sample set, including 729 cases (45.4 ± 15.1 years, 52.2% male) and 562 controls (44.0 ± 14.4 years, 49.8% male), was used to prioritize detected functional variants for subsequent association analyses. The third sample set, including 1511 cases (45.9 ± 14.0 years, 49.6% male) and 1517 controls (46.0 ± 14.6 years, 49.6% male), was used for association analyses. Age and gender were matched in the second and third sample sets, respectively. All patients were diagnosed according to *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*, criteria, and controls were evaluated using unstructured interviews to exclude individuals with history of mental disorders. Detailed information regarding diagnostic procedures is available elsewhere.[21] All subjects were ethnically Japanese and provided written informed consent. This study was approved by the ethics committees at each participating university.

Array Design for Resequencing Analyses

We used the Affymetrix GeneChip CustomSeq Resequencing Array (Affymetrix, Santa Clara, California) for exon sequencing in the first sample set. These arrays rely on allele-specific hybridization for determining DNA

sequence.[19] Each individual nucleotide of both the sense and the antisense DNA strands is interrogated with four 25-mer probes that differ only with respect to the central position (A, C, G, and T). According to Affymetrix's Custom-Seq Array Design Guide, we designed arrays covering all exon regions of *KALRN* and *EPHB1* (Ensembl release 52 [Human CCDS set]; Transcript: ENST00000360013, ENST00000240874, and ENST00000291478 for *KALRN*; ENST00000398015 for *EPHB1*). Because the principle of the resequencing arrays is based on hybridization, it is necessary to avoid cross-hybridization for accurate resequencing. For this purpose, we removed repetitive elements and highly homologous sequences from the array design.

Array-Based Resequencing

The experiments were conducted according to the manufacturer's instructions (supplementary figure S2). Genomic DNA was extracted from peripheral blood using standard methods. To generate enough target-enriched subject material for hybridization to the arrays, we generated 47 and 14 amplicons per sample for *KALRN* and *EPHB1*, respectively, using long-range polymerase chain reaction (PCR). The PCR conditions were as follows: 94°C for 2 minutes followed by 30 cycles consisting of 94°C for 15 seconds, 68°C for 3 minutes, followed by a final extension of 68°C for 8 minutes, using TaKaRa LA Taq™ (Takara Bio, Otsu, Shiga, Japan). Each PCR product was quantified using PicoGreen (Molecular Probes, Eugene, Oregon), pooled in an equimolar fashion. The PCR products were then purified, fragmented, labeled, and hybridized to the arrays, following the protocol. Finally, the arrays were washed and stained using the GeneChip Fluidics Station 450 (Affymetrix) and scanned using the GeneChip Scanner 3000 (Affymetrix). The data were analyzed using the GeneChip Operating Software (GCOS; Affymetrix), the GeneChip Sequence Analysis Software (GSEQ; Affymetrix), and SeqC (JSI Medical Systems, Kippenheim, Germany; http://www.jsi-medisys.de/html/products/SeqC/SeqC.htm) to automate the generation of sequence and genotype calls from the intensity data. In this study, around 17 kb was sequenced per sample, meaning that more than 5.4 Mb was sequenced in total. All missense mutations presented in this study were confirmed using both Sanger sequencing and Custom TaqMan SNP genotyping assays (Applied Biosystems, Foster City, California).

Association Analysis of Each Missense Mutation

Although the rare (<1%) missense mutations were originally discovered among 320 schizophrenic patients, it was possible that a portion of them might have neutral or protective effects.[5] In addition, it was necessary to reduce the number of statistical tests for multiple comparison problems. To accomplish this, we prioritized rare (<1%) deleterious variants for subsequent association analyses based on the frequencies in the second case-control sample set because rare deleterious variants relevant to schizophrenia can be assumed to have higher frequency in cases than in controls. The criteria for prioritization were as follows: (1) frequencies of mutations were <1% in controls and (2) frequencies of mutations were higher in cases (ie, OR > 1). Mutations not detected in the second sample set were not followed up in this analysis. The frequencies of such mutations can be so low (<0.0005) that the results of association analyses are unlikely to be statistically significant in our sample size. For mutations meeting the above criteria, we conducted association analyses with schizophrenia using the third sample set. Genotyping was conducted by Custom TaqMan SNP genotyping assays (Applied Biosystems). For quality control, samples with missing call rates of 10% or higher were excluded from the analyses.

Combined Association Analyses

In general, it is difficult to establish an association of a rare mutation with a phenotype because statistical power is limited by low population frequency and because the number of rare variants requires a strict multiple test correction. Therefore, we conducted combined association analyses across rare mutations observed in each gene in the third sample set, comparing the number of mutations in cases with the number in controls. The criteria for mutations included in these analyses were same as the above criteria with 1 exception: Mutations not detected in the second sample set were included in the combined association analyses.

In Silico Analysis

The potential influence of missense mutations was evaluated using PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/) and[22] PMut (http://mmb2.pcb.ub.es:8080/PMut/)[23] softwares. PolyPhen-2 uses 8 sequence-based and 3 structure-based predictive features and compares a property of the wild-type allele and the corresponding property of the mutant allele. PolyPhen-2 trained on HumDiv datasets is reported to achieve true positive prediction rates of 92% with a false-positive rate of 20%.[22] A mutation is appraised qualitatively as benign, possibly damaging, or probably damaging based on naive Bayes posterior probability that a given mutation is damaging. PMut also allows the fast and accurate prediction (~80% success rate in humans) of the pathological character of missense mutations based on the use of neural networks. The final output is a pathogenicity index ranging from 0 to 1 (indexes >0.5 signal pathological mutations).

We also examined evolutionary conservation of the mutated residues and surrounding amino acids. Multiple sequence alignment of human *KALRN* or *EPHB1* with 6 orthologs was performed for this purpose.

Power Calculation

Power calculation was performed with a power calculator called CaTS (http://www.sph.umich.edu/csg/abecasis/CaTS/).[24] Power was estimated under the following parameter assumptions with respect to association test statistics: genetic relative risk = 2, prevalence of disease = 0.01, risk allele frequency = the values frequency observed in controls, and $\alpha = .05$; a multiplicative model was used.

Statistical Analysis

For the association analysis of each variant, Fisher exact test was used to examine whether rare deleterious variants were significantly overrepresented in the patient group rather than the control group.

A combined association test was performed following a previous study.[7] In brief, to account for variable sample size, sample size was adjusted to $N=n/(\sum(1/Ni))$, where Ni is the sample size at the ith variant, and n is the number of variants. The number of observed variants was adjusted as $\sum(pi)\times N$, where pi is the frequency of the ith variant. Fisher exact test was used in this test as well to examine an overrepresentation of rare deleterious missense mutations in the patient group rather than control group.

All statistical tests were 1 tailed, and a P value less than 0.05 was considered significant. Bonferroni correction was used for solving multiple testing problems.

Results

Discovery of Mutations

We detected 12 and 6 missense mutations with a frequency of <5% in *KALRN* and *EPHB1*, respectively, among 320 cases in the first sample set (table 1). All but 2 mutations (N2973S in *KALRN* and T981M in *EPHB1*) were novel. All mutations were validated by both Sanger sequencing and Custom TaqMan SNP genotyping assays. In the first sample set, 2 patients were compound heterozygotes for rare missense mutations in the 2 genes. One patient had R410H in *KALRN* and R905C in *EPHB1*. The other had A2382V in *KALRN* and D375N in *EPHB1*. There were no clinical characteristics shared between these patients. No nonsense mutations were identified in this study.

Association Analysis of Each Missense Mutation

In the prioritization phase using the second sample set, T1207M and P2255T in *KALRN* and R637C and R905C in *EPHB1* showed a higher frequency in cases than in controls (table 1). Seven missense mutations (R410H, Q770K, and A2382V in *KALRN* and F151S, D375N, D577N, and T981M in *EPHB1*) were not detected. The frequency of P1695Q was more than 4% both in cases and in controls. Based on our criteria, we selected 4 missense mutations (T1207M and P2255T in *KALRN* and R637C and R905C in *EPHB1*) for subsequent association analyses using the third sample set.

In the third phase, P2255T showed a nominally significant association with schizophrenia (OR = 2.09, P = .012) in the third sample set (table 2). This remained significant after correction for multiple testing of 4 variants (corrected P = .048). T1207M in *KALRN* and R637C and R905C in *EPHB1* were also more frequent in cases, although differences were not significant.

We excluded mutations not detected in the second sample set from this analysis. This was supported by a power analysis showing that the third sample set had only 10% power in analysis of very rare mutations.

Combined Association Analysis

In addition to 4 mutations (T1207M and P2255T in *KALRN* and R637C and R905C in *EPHB1*), 7 very rare mutations (R410H, Q770K, and A2382V in *KALRN* and F151S, D375N, D577N, and T981M in *EPHB1*), which were not detected in the second samples set, were included in the combined association analysis. A global comparison of the frequencies of 5 selected mutations in *KALRN* between cases and controls in the third sample set showed a significant increase in frequency in schizophrenic patients (OR = 2.07, P = .003) (table 3). This remained significant after correction for multiple testing (corrected P = .006). On the other hand, a global comparison of the frequencies of 6 selected mutations in *EPHB1* did not show a significant difference (OR = 1.09, P = .438).

In Silico Analysis

Results of *in silico* analysis are shown in table 4. All missense mutations but A2382V in *KALRN* were predicted to have functional relevance by PolyPhen-2 or PMut software.

A multiple alignment of the region of *KALRN* or *EPHB1* containing rare missense mutations with 6 orthologs is shown in table 4. Most of the rare missense mutations showed a high degree of amino acid conservation in different species.

Discussion

In this study, we conducted resequencing analyses for the 2 synaptogenic pathway genes (*KALRN* and *EPHB1*) in schizophrenia using a DNA microarray–based method. After resequencing more than 5.4 Mb, we discovered 17 rare (<1%) missense mutations in *KALRN* or *EPHB1* and detected a significant association between schizophrenia and P2255T in *KALRN*, as well as in the combined association analysis for *KALRN*. These findings are consistent with an estimation that most rare (<1%) missense mutations are mildly deleterious and are associated with a heterozygous fitness loss.[5]

Table 1. *KALRN* And *EPHB1* Missense Mutations Identified in The First Sample Set And Their Frequencies in The Second Sample Set

Gene	Genomic Position	Base Change	dbSNP Reference	AA Change	First Sample Set		Second Sample Set				
							Genotype Counts		Mutation Frequency		
					Homo	Hetero	SCZ	CONT	SCZ	CONT	OR >1
KALRN	125527659	G → A	ss250607852	R410H	0	1	0/0/701	0/0/541	0	0	
KALRN	125531474	T → A	ss250607853	L452Q	0	1	0/1/709	0/2/541	0.0007	0.0018	
KALRN	125600376	C → A	ss250607854	Q770K	0	1	0/0/706	0/0/544	0	0	
KALRN	125656787	C → T	ss250607855	T1207M	0	1	0/2/705	0/1/542	0.0014	0.0009	+
KALRN	125764534	C → A	ss250607856	P1695Q	0	1	0/59/636	1/44/492	0.0425	0.0428	
KALRN	125764599	A → T	ss250607857	M1717L	0	1	0/0/705	0/1/540	0	0.0009	
KALRN	125860927	G → A	ss250607858	R2049K	0	1	0/1/696	0/1/540	0.0007	0.0009	
KALRN	125873259	C → A	ss250607859	P2255T	0	7	1/14/684	0/7/536	0.0114	0.0064	+
KALRN	125873289	C → T	ss250607860	P2265S	1	0	0/6/701	0/7/533	0.0042	0.0065	
KALRN	125873382	G → T	ss250607861	G2296C	0	1	0/1/703	0/1/542	0.0007	0.0009	
KALRN	125876103	C → T	ss250607862	A2382V	0	1	0/0/697	0/0/540	0	0	
KALRN	125920964	A → G	rs16835896	N2973S	0	3	0/3/698	0/6/538	0.0021	0.0055	
EPHB1	136153231	T → C	ss252863894	F151S	0	1	0/0/710	0/0/543	0	0	
EPHB1	136334407	G → A	ss252863895	D375N	0	1	0/0/708	0/0/544	0	0	
EPHB1	136368508	G → A	ss252863896	D577N	0	1	0/0/707	0/0/544	0	0	
EPHB1	136394134	C → T	ss252863897	R637C	0	2	1/1/707	0/2/541	0.0021	0.0018	+
EPHB1	136450890	C → T	ss252863898	R905C	0	3	0/9/695	0/1/543	0.0064	0.0009	+
EPHB1	136460639	C → T	rs56186270	T981M	0	2	0/0/706	0/0/541	0	0	

Note: Genomic position based on NCBI build 36, chromosome 3. Amino acid changes based on NCBI Reference Sequence NP_001019831.2 (2986 aa) for KALRN and NP_004432.1 (984 aa) for EPHB1. All but N2973S (rs16835896) and T981M (rs56186270) are novel. AA change, amino acid change; dbSNP, Single Nucleotide Polymorphism Database; Homo, homozygote; Hetero, heterozygote; SCZ, schizophrenia; CONT, control; NCBI, National Center for Biotechnology Information.

Schizophrenia is a genetically heterogeneous disorder, with both very rare variants with a high effect size (eg, CNVs in 1q21.1, 15q13.3) and common variants with a low effect size (eg, rs1344706 in *ZNF804A*) involved in its genetic architecture. In this frequency-effect size spectrum, P2255T (OR: ~2, risk allele frequency in controls: ~0.005) is located between the CNV in 1q21.1 (OR: ~10, frequency in controls: ~0.0001)[25] and rs1344706[T] in *ZNF804A* (OR: ~1.1, risk allele frequency in controls: ~0.6),[26] both of which have been recently associated with schizophrenia. The relatively modest effect size of P2255T compared with that of the above CNVs can be attributable to the difference in the effect of each variant on gene(s): Although CNVs strongly influence the expression of multiple genes, missense mutations in *KALRN* are presumed to have limited effects on KALRN function. P2255T is located in the evolutionally conserved proline-rich region between the C-terminal GEF and SH3 domains[27] and is surrounded by 2 nearby phosphorylation sites (S2237 and S2262), according to Human Protein Reference Database (http://www.hprd.org/index_html)[28] (figure 1). *In silico* analysis with PhosphoMotif Finder[29] shows that T2255 itself can be recognized and phosphorylated by many kinases, suggesting functional implications of P2255T (figure 1). In addition, *in silico* analysis predicts that phosphorylation of T2255 will induce that of nearby S2253. Thus, P2255T may greatly change the phosphorylation status in a narrow

Table 2. Association Analyses of Each Missense Mutation in the Third Sample Set

	AA Change	Third Sample Set					
		Genotype Counts		Mutation Frequency			
		SCZ	CONT	SCZ	CONT	OR	*P* Value
KALRN	T1207M	0/7/1477	0/3/1482	0.0024	0.0010	2.34	.171
KALRN	P2255T	0/31/1448	0/15/1473	0.0105	0.0050	2.09	.012
EPHB1	R637C	0/4/1477	0/4/1478	0.0014	0.0014	1.00	.636
EPHB1	R905C	0/15/1458	0/12/1466	0.0051	0.0041	1.26	.347

Note: Abbreviations are explained in the first footnote to table 1. *P* values were calculated by Fisher exact test (1 tailed).

Table 3. Combined Association Analysis in The Third Sample Set

Gene	AA Change	Third Sample Set				Combined Analysis	
		Genotype Counts		Mutation Frequency		Gene Based	
		SCZ	CONT	SCZ	CONT	OR	*P* value
KALRN	R410H	0/0/1481	0/0/1484	0	0	2.07	.003
KALRN	Q770K	0/0/1486	0/0/1490	0	0		
KALRN	T1207M	0/7/1477	0/3/1482	0.0024	0.0010		
KALRN	P2255T	0/31/1448	0/15/1473	0.0105	0.0050		
KALRN	A2382V	0/7/1473	0/4/1480	0.0024	0.0013		
EPHB1	F151S	0/0/1478	0/0/1484	0	0	1.09	.438
EPHB1	D375N	0/0/1483	0/0/1490	0	0		
EPHB1	D577N	0/0/1486	0/2/1483	0	0.000673		
EPHB1	R637C	0/4/1477	0/4/1478	0.0014	0.0014		
EPHB1	R905C	0/15/1458	0/12/1466	0.0051	0.0041		
EPHB1	T981M	0/5/1481	0/4/1484	0.0017	0.0013		

Note: Abbreviations are explained in the first footnote to table 1. *P* values were calculated by Fisher exact test (1 tailed).

region between the C-terminal GEF and SH3 domain. A protein with multiple phosphorylated sites like KALRN can be assumed to have an exponential number of phospho-forms, and individual phospho-forms may have distinct biological effects. The diffuse distribution of these phospho-forms at steady state enables the phosphoproteome to encode information and flexibly respond to varying demands.[30] Thus, it is conceivable that P2255T may influence such plasticity in KALRN by changing the number of phosphorylated sites. Interestingly, detailed examination of clinical information from the first sample set, which was uniquely available to us, revealed that congenital or early-onset vascular disease was observed in 5 of 7 cases with P2255T (supplementary table S1). Because *KALRN* may represent a candidate gene for vascular diseases,[31,32] it is tempting to speculate that P2255T may be a potential risk factor for vascular disease.

In addition to P2255T, we detected multiple rare (<1%) missense mutations in *KALRN* or *EPHB1*. Such variants are not sufficiently frequent to be covered by GWAS nor do they have sufficiently large effect sizes to be detected by linkage analysis in family studies. For modest effect sizes, it is suggested that association testing may require composite tests of overall mutational load,

Table 4. Results of *In Silico*/Conservation Analysis

KALRN							
Analysis		R410H	Q770K	T1207M	P2255T	A2382V	
PolyPhen-2		Probably damaging	Probably damaging	Probably damaging	Benign	Benign	
PMut		Pathological	Neutral	Pathological	Pathological	Neutral	
Conservation analysis	Human (NP_001019831.2)	LDERSTI	IFLQLRI	IHATEIR	RSQPARL	SILAPLT	
	Chimpanzee (XP_516703.2)	LDERSTI	IFLQLRI	IHATEIR	RSQPARL	SILAPLT	
	Dog (XP_535768.2)	LDERSTI	IFLQLRI	IHATEIR	RSQPSRV	SVLAPLT	
	Cattle (XP_001790302.1)	LDERSTI	IFLQLRI	IHATEIR	RSQPARV	SILTPLT	
	Mouse (XP_001481079.1)	LDERSTI	IFLQLRI	IHATEIR	RSQPPRV	SILAPLA	
	Rat (NP_114451.1)	LDERSTI	IFLQLRI	IHATEIR	RSQPPRV	SILAPLT	
EPHB1							
Analysis		F151S	D375N	D577N	R637C	R905C	T981M
PolyPhen-2		Benign	Probably damaging	Possibly damaging	Probably damaging	Probably damaging	Probably damaging
PMut		Pathological	Neutral	Neutral	Pathological	Pathological	Pathological
Conservation analysis	Human (NP_004432.1)	QVDFGGR	RCDDNVE	VYSDKLQ	YKGRLKL	LLDRSIP	QSPTAMA
	Chimpanzee (XP_001150963.1)	QVDFGGR	RCDDNVE	LLVEQWQ	YKGRLKL	LLDRSIP	QSPTAMA
	Dog (XP_542791.2)	QVDFGGR	RCDDNVE	VYSDKLQ	YKGRLKL	LLDRSIP	QSPTTMA
	Cattle (XP_614602.4)	QVDFGGR	RCDDNVE	VYSDKLQ	YKGRLKL	LLDRSIP	QSPTAMA
	Mouse (NP_775623.2)	QVDFGGR	RCDDNVE	AYSDKLQ	YKGRLKL	LLDRSIP	QSPSVMA
	Rat (XP_217250.1)	QVDFGGR	RCDDNVE	VYSDKLQ	YKGRLKL	LLDRSIP	QSPSVMA

Note: The bold are the mutated amino acids.

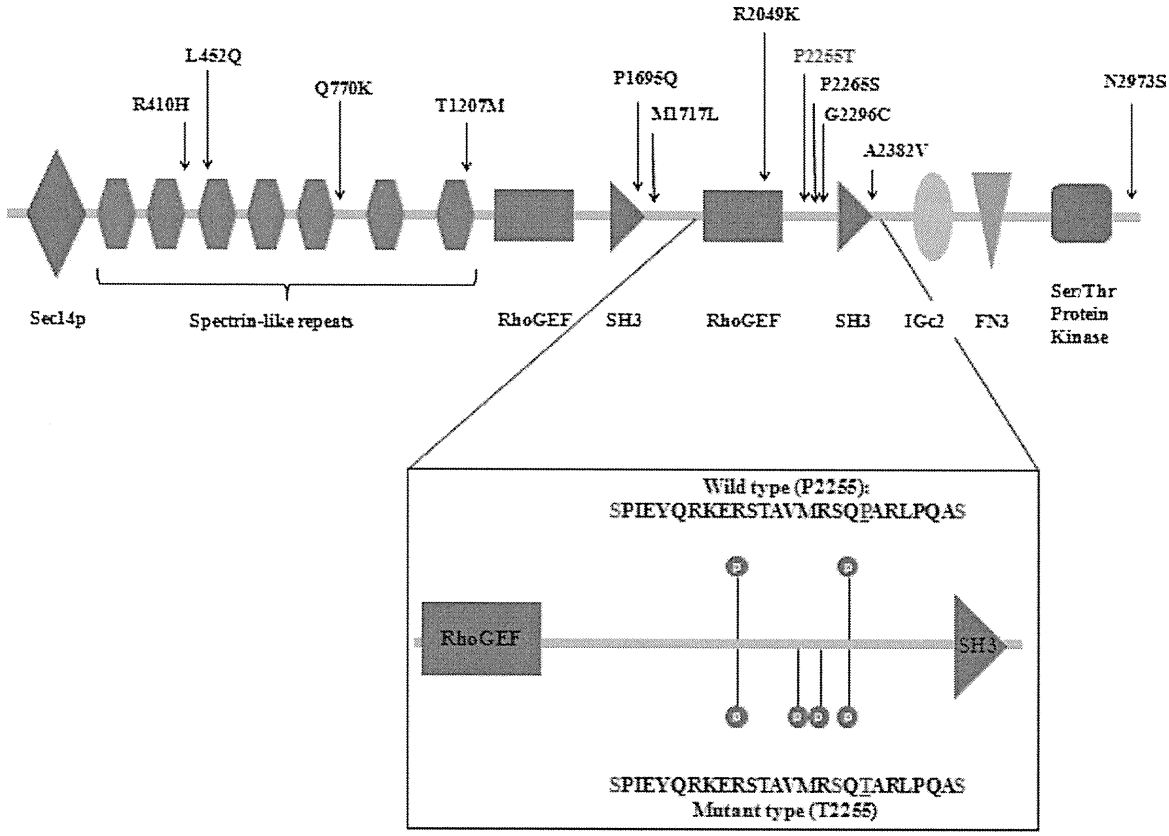

Fig. 1. Rare Missense Mutations in *KALRN* and Change in Phosphorylation Status by P2255T.

comparing frequencies of mutations of potentially similar functional effect in cases and controls. Thus, we also performed combined association analyses for *KALRN* or *EPHB1* and found evidence that multiple rare (<1%) missense mutations in *KALRN* as a whole are associated with schizophrenia. This finding is supported by *in silico* analyses showing that most of the mutations are predicted as being of functional relevance and that they are located in evolutionally conserved regions. In contrast, there were no significant differences in the cumulative frequencies of rare missense mutations in *EPHB1*. This might be due to a type II error. The cumulative frequency of rare mutations of *EPHB1* in controls is almost same as the one of *KALRN* in controls (0.0075 vs 0.0073), indicating that cumulative effect size of rare missense mutations in *EPHB1* may be smaller than the one in *KALRN*. In the mammalian genome, there are 5 different EphB receptors (EphB1, EphB2, EphB3, EphB4, and EphB6), with a high similarity at the amino acid level. Analysis of double and triple knockout mice lacking EphB1, EphB2, and EphB3 in different combinations revealed that EphBs have functional redundancy even though all these EphBs are responsible for spine morphogenesis and synapse formation to varying degrees.[33] This is in contrast with the drastic phenotypes observed in *KALRN*-knockout mice.[16] Therefore, biological effects of rare missense mutations in *EPHB1* may be compensated for by other intact *EPHBs*. This might lower the ORs of rare missense mutations in *EPHB1*. Given that all the mutations detected in *EPHB1* were predicted to have pathogenicity by PolyPhen-2 or PMut, a larger-scale case-control study with sufficient power may provide a significant result in a combined analysis for *EPHB1*.

One important aspect of the present study is that we found rare mutations associated with schizophrenia in the *KALRN* gene, in which GWASs detected association signals for schizophrenia. Several studies have recently reported the 1 gene may harbor both rare and common variants associated with the same diseases, including schizophrenia,[34] type 2 diabetes,[35] and hypertriglyceridemia.[36] Given that the cost of whole-genome sequencing is still high to search for rare mutations, resequencing analyses for genes with support from GWAS might be a better strategy for detection of rare mutations with larger effect size.

There are several limitations to this study. First, we could not conduct segregation analyses for mutations due to limited access to family members. Furthermore, given the modest risk (OR ~2), these mutations would show incomplete penetrance. In fact, it is reported that penetrance estimates of CNVs at 1q21.1 and 15q13.3,

both of which show higher ORs, are 0.061 and 0.074, respectively.[25] Therefore, a population-based study is a better choice to evaluate genetic associations for missense mutations with modest risk.[37] The second limitation is population stratification. Although a Japanese population is considered relatively homogenous, small population stratifications may have influenced our findings.[38] However, we believe that the recruitment of subjects in local regions minimized this concern. Third, we did not conduct functional analyses for detected missense mutations. The detailed effects of these mutations on the pathophysiology of schizophrenia need to be examined in a future study. Fourth, our resequencing analyses were not comprehensive in terms of the kind of variants and the number of genes. In other words, the present study did not cover indels or CNVs because of the methodological limitation of the DNA microarray–based method. Because these classes of variants could have a more profound effect on protein function, their genetic contribution to schizophrenia might be revealed in future studies. Also, as shown in *EPHB1*, it is assumed that a variety of molecules or pathways have a role in spine formation or synapse plasticity, which are impaired in patients with schizophrenia, to compensate for each other. A combined analysis of a large number of genes relevant for synaptic function might provide more robust evidence that rare missense mutations as a whole contribute to pathomechanisms of schizophrenia.

In conclusion, we provide the first evidence that multiple rare (<1%) missense mutations in *KALRN* may be genetic risk factors for schizophrenia. Further studies will be needed to examine the pathogenicity of these mutations from a biologic point of view.

Funding

This work was supported by research grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan; Ministry of Health, Labor and Welfare of Japan; Grant-in-Aid for Scientific Research on Pathomechanisms of Brain Disorders from the Ministry of Education, Culture, Sports, Science and Technology of Japan; Academic Frontier Project for Private Universities, Comparative Cognitive Science Institutes; Core Research for Evolutional Science and Technology.

Supplementary Material

Supplementary material is available at http://schizophreniabulletin.oxfordjournals.org.

Acknowledgments

We sincerely thank the patients and healthy volunteers for their participation in this study. We also thank Ryoko Ishihara and Junko Tsuda for technical assistance.

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

References

1. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry*. 2003;60:1187–1192.
2. Purcell SM, Wray NR, Stone JL, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. 2009;460:748–752.
3. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009;461:747–753.
4. Stefansson H, Rujescu D, Cichon S, et al. Large recurrent microdeletions associated with schizophrenia. *Nature*. 2008; 455:232–236.
5. Kryukov GV, Pennacchio LA, Sunyaev SR. Most rare missense alleles are deleterious in humans: implications for complex disease and association studies. *Am J Hum Genet*. 2007; 80:727–739.
6. Boyko AR, Williamson SH, Indap AR, et al. Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet*. 2008;4:e1000083.
7. Knight HM, Pickard BS, Maclean A, et al. A cytogenetic abnormality and rare coding variants identify ABCA13 as a candidate gene in schizophrenia, bipolar disorder, and depression. *Am J Hum Genet*. 2009;85:833–846.
8. Ikeda M, Aleksic B, Kinoshita Y, et al. Genome-wide association study of schizophrenia in a Japanese population [published online ahead of print September 10, 2010]. *Biol Psychiatry*. doi: 10.1016/j.biopsych.2010.07.010.
9. Penzes P, Beeser A, Chernoff J, et al. Rapid induction of dendritic spine morphogenesis by trans-synaptic ephrinB-EphB receptor activation of the Rho-GEF kalirin. *Neuron*. 2003; 37:263–274.
10. St Jean PL. Genes associated with schizophrenia identified using a whole genome scan. 2008. http://www.freepatentsonline.com/y2008/0176240.html. Accessed August 8, 2010.
11. Sullivan PF, Lin D, Tzeng JY, et al. Genomewide association for schizophrenia in the CATIE study: results of stage 1. *Mol Psychiatry*. 2008;13:570–584.
12. Xu B, Roos JL, Levy S, et al. Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat Genet*. 2008;40:880–885.
13. Rabiner CA, Mains RE, Eipper BA. Kalirin: a dual Rho guanine nucleotide exchange factor that is so much more than the sum of its many parts. *Neuroscientist*. 2005;11:148–160.
14. Glantz LA, Lewis DA. Decreased dendritic spine density on prefrontal cortical pyramidal neurons in schizophrenia. *Arch Gen Psychiatry*. 2000;57:65–73.
15. Hill JJ, Hashimoto T, Lewis DA. Molecular mechanisms contributing to dendritic spine alterations in the prefrontal cortex of subjects with schizophrenia. *Mol Psychiatry*. 2006; 11:557–566.
16. Cahill ME, Xie Z, Day M, et al. Kalirin regulates cortical spine morphogenesis and disease-related behavioral phenotypes. *Proc Natl Acad Sci U S A*. 2009;106:13058–13063.
17. Hayashi-Takagi A, Takaki M, Graziane N, et al. Disrupted-in-Schizophrenia 1 (DISC1) regulates spines of the glutamate synapse via Rac1. *Nat Neurosci*. 2010;13:327–332.

18. Penzes P, Jones KA. Dendritic spine dynamics—a key role for kalirin-7. *Trends Neurosci.* 2008;31:419–427.
19. Kothiyal P, Cox S, Ebert J, et al. An overview of custom array sequencing. *Curr Protoc Hum Genet.* 2009;61:7.17.1–7.17.11.
20. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet.* 2010;11:415–425.
21. Ikeda M, Aleksic B, Kirov G, et al. Copy number variation in schizophrenia in the Japanese population. *Biol Psychiatry.* 2010;67:283–286.
22. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010;7:248–249.
23. Ferrer-Costa C, Gelpi JL, Zamakola L, et al. PMUT: a web-based tool for the annotation of pathological mutations on proteins. *Bioinformatics.* 2005;21:3176–3178.
24. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet.* 2006;38:209–213.
25. Vassos E, Collier DA, Holden S, et al. Penetrance for copy number variants associated with schizophrenia. *Hum Mol Genet.* 2010;19:3477–3481.
26. O'Donovan MC, Craddock N, Norton N, et al. Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet.* 2008;40:1053–1055.
27. Johnson RC, Penzes P, Eipper BA, Mains RE. Isoforms of kalirin, a neuronal Dbl family member, generated through use of different 5'- and 3'-ends along with an internal translational initiation site. *J Biol Chem.* 2000;275:19324–19333.
28. Keshava Prasad TS, Goel R, Kandasamy K, et al. Human Protein Reference Database—2009 update. *Nucleic Acids Res.* 2009;37:D767–D772.
29. Amanchy R, Periaswamy B, Mathivanan S, et al. A curated compendium of phosphorylation motifs. *Nat Biotechnol.* 2007;25:285–286.
30. Thomson M, Gunawardena J. Unlimited multistability in multisite phosphorylation systems. *Nature.* 2009;460:274–277.
31. Wang L, Hauser ER, Shah SH, et al. Peakwide mapping on chromosome 3q13 identifies the kalirin gene as a novel candidate gene for coronary artery disease. *Am J Hum Genet.* 2007;80:650–663.
32. Krug T, Manso H, Gouveia L, et al. Kalirin: a novel genetic risk factor for ischemic stroke. *Hum Genet.* 2010;127:513–523.
33. Henkemeyer M, Itkis OS, Ngo M, Hickmott PW, Ethell IM. Multiple EphB receptor tyrosine kinases shape dendritic spines in the hippocampus. *J Cell Biol.* 2003;163:1313–1326.
34. Steinberg S, Mors O, Borglum AD, et al. Expanding the range of ZNF804A variants conferring risk of psychosis [published online ahead of print January 05, 2010]. Mol Psychiatry. doi: 10.1038/mp.2009.149.
35. Voight BF, Scott LJ, Steinthorsdottir V, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet.* 2010;42:579–589.
36. Johansen CT, Wang J, Lanktree MB, et al. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. *Nat Genet.* 2010;42:684–687.
37. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet.* 2008;40:695–701.
38. Yamaguchi-Kabata Y, Nakazono K, Takahashi A, et al. Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. *Am J Hum Genet.* 2008;83:445–456.

Copy Number Variation in Schizophrenia in the Japanese Population

Masashi Ikeda, Branko Aleksic, George Kirov, Yoko Kinoshita, Yoshio Yamanouchi, Tsuyoshi Kitajima, Kunihiro Kawashima, Tomo Okochi, Taro Kishi, Irina Zaharieva, Michael J. Owen, Michael C. O'Donovan, Norio Ozaki, and Nakao Iwata

Background: Copy number variants (CNVs) have been shown to increase the risk to develop schizophrenia. The best supported findings are at 1q21.1, 15q11.2, 15q13.3, and 22q11.2 and deletions at the gene neurexin 1 (*NRXN1*).

Methods: In this study, we used Affymetrix 5.0 arrays to investigate the role of rare CNVs in 575 patients with schizophrenia and 564 control subjects from Japan.

Results: There was a nonsignificant trend for excess of rare CNVs in schizophrenia ($p = .087$); however, we did not confirm the previously implicated association for very large CNVs (>500 kilobase [kb]) in this population. We provide support for three previous findings in schizophrenia, as we identified one deletion in a case at 1q21.1, one deletion within *NRXN1*, and four duplications in cases and one in a control subject at 16p13.1, a locus first implicated in autism and later in schizophrenia.

Conclusions: In this population, we support some of the previous findings in schizophrenia but could not find an increased burden of very large (>500 kb) CNVs, which was proposed recently. However, we provide support for the role of CNVs at 16p13.1, 1q21.1, and *NRXN1*.

Key Words: Deletion, duplication, *NRXN1*, 16p13.1, 1q21.1, schizophrenia

Copy number variations (CNVs) are deletions and duplications of DNA ranging from a kilobase (kb) to several megabases (Mb). Recently, rare CNVs were shown to play a role in the etiology of a number of neuropsychiatric disorders, particularly schizophrenia, autism, and mental retardation (1).

Several studies have reported a greater prevalence of rare CNVs in people with schizophrenia (2-4). However, some have found no such excess (5,6) and even among the positive studies, there is marked variation in the magnitude of the observed effect. For example, in the International Schizophrenia Consortium (ISC) study (4), cases had only a 1.15-fold excess of rare CNVs, rising to 1.67-fold for deletions greater than 500 kb. An increase only among very large CNVs (>1 Mb) in cases was found by Kirov *et al.* (7). Another study showed an odds ratio of 3.37 for CNVs, rising to 4.82 for early-onset schizophrenia (2). This may, in part, reflect differences in the sensitivity of CNV assays, definitions of low-frequency CNVs, or variation in the phenotypic composition of the samples, as cases with early onset or lower IQ were particularly enriched for CNVs in one study (2).

In addition to increased CNV burden, a number of specific CNVs have been associated with schizophrenia (4,7,8). There is strong replicated evidence for deletions at 1q21.1, 15q11.2, 15q13.3, and 22q11.2 and emerging evidence for duplications at 16p13.1 (4,7). Deletions of the neurexin 1 gene (*NRXN1*) have also been reported in multiple studies on schizophrenia (2,6,7,9,10). Given the discrepancy in estimates of the effect size of CNV burden as a risk factor for schizophrenia and in particular the absence of association in the only Asian sample reported to date (5), we aimed to test for an excess burden of CNVs in a population from Japan. We also sought supportive evidence for a contribution for the specific loci listed above.

Methods and Materials

We analyzed 1139 age- and gender-matched unrelated subjects of Japanese ethnicity (575 schizophrenic patients and 564 control subjects). Control subjects were members of the general public who had no personal history of mental disorders. This was ascertained during face-to-face interviews where subjects were asked if they had suffered an episode of depression, mania, or psychotic experiences or if they had received treatment for any psychiatric disorder. Patients were entered into the study if they 1) met DSM-IV criteria for schizophrenia; 2) were physically healthy and had normal routine laboratory tests; and 3) had no mood disorders, substance abuse, neurodevelopmental disorders, epilepsy, or known mental retardation. Consensus diagnoses were made by at least two experienced psychiatrists according to DSM-IV criteria on the basis of unstructured interviews with patients and families and review of medical records. After description of the study, written informed consent was obtained from each subject. This study was approved by the ethics committees of each participating university.

We used Affymetrix 5.0 Arrays (Affymetrix, Santa Clara, California), following the manufacturer's protocols (http://www.affymetrix.com). This array includes 470K single nucleotide polymorphism (SNP) probes and 420K nonpolymorphic probes. The CNVs discussed below in more detail (at *NRXN1*, 1q21.1, and 16p13.1) were validated using the Illumina HumanHap 660W- or 610-quad bead arrays (Illumina, San Diego, California), following the manufacturer's protocols (http://www.illumina.com).

Copy number variations were called using the Birdsuite program (http://www.broadinstitute.org/science/programs/medical-and-

From the Medical Research Council (MRC) Centre for Neuropsychiatric Genetics and Genomics (MI, GK, IZ, MJO, MCO), Department of Psychological Medicine and Neurology, School of Medicine, Cardiff University, Cardiff, United Kingdom; Department of Psychiatry (MI, YK, YY, TKit, KK, TO, TKis, NI), Fujita Health University School of Medicine, Toyoake, Aichi; Department of Psychiatry (BA, NO), Nagoya University Graduate School of Medicine, Nagoya; and Core Research for Environmental Science and Technology (CREST) (BA, YK, TKit, KK, TO, TKis, NO, NI), Japan Science and Technology Agency, Kawaguchi, Saitama, Japan.

Address correspondence to George Kirov, M.D., MRC Centre for Neuropsychiatric Genetics and Genomics, Department of Psychological Medicine and Neurology, School of Medicine, Cardiff University, Cardiff CF14 4XN, UK; E-mail: kirov@cardiff.ac.uk.

Received Jun 25, 2009; revised Aug 12, 2009; accepted Aug 31, 2009.

0006-3223/10/$36.00
doi:10.1016/j.biopsych.2009.08.034

Table 1. Global CNV Burden Analysis

		CNV Burden				CNVs Intersecting Genes			
CNV Type	Size	SCZ	CON	CNV Rate SCZ/CON	*p* Value	SCZ	CON	CNV Rate SCZ/CON	*p* Value
Deletions and Duplications	All	567	485	1.1/.95	.087	382	320	.74/.62	.084
	100–200 kb	285	229	.55/.45	.046	182	145	.35/.28	.074
	200–500 kb	221	192	.43/.37	.20	150	134	.29/.26	.30
	500 kb–1 Mb	48	52	.09/.10	.72	38	32	.07/.06	.31
	>1 Mb	13	12	.025/.023	.52	12	9	.023/.018	.35
Deletions Only	All	174	157	.34/.31	.30	91	87	.18/.17	.46
	100–200 kb	98	84	.19/.16	.26	52	47	.10/.09	.38
	200–500 kb	65	60	.13/.12	.42	29	35	.06/.07	.79
	500 kb–1 Mb	8	8	.015/.016	.62	8	3	.015/.006	.12
	>1 Mb	3	5	.006/.010	.86	2	2	.004/.004	.69
Duplications Only	All	393	328	.76/.64	.10	291	233	.56/.45	.075
	100–200 kb	187	145	.36/.28	.070	130	98	.25/.19	.071
	200–500 kb	156	132	.30/.26	.21	121	99	.23/.19	.18
	500 kb–1 Mb	40	44	.077/.086	.73	30	29	.058/.057	.53
	>1 Mb	10	7	.019/.014	.33	10	7	.019/.014	.33

p values are one-tailed and based on 10,000 permutations.
CNV, copy number variation; CON, control; kb, kilobase; Mb, megabase; SCZ, schizophrenia.

population-genetics/birdsuite/birdsuite-0) (11). The software first assigns copy number across regions of known copy number polymorphisms, then calls SNP genotypes (for samples and SNPs believed to have two copies of the locus), then searches for novel CNVs via a hidden Markov model, and generates an integrated sequence and copy number genotype at every locus. It takes into account genotypes within CNVs, e.g., A-null, AAB, and BBB, in addition to AA, AB, and BB calls (11).

We observed a batch effect, similar to what we reported in our previous study (7): arrays from different batches gave poor results if analyzed together. Therefore, we identified the batches and analyzed together samples within the same batch, as recommended in the Birdsuite manual (11). After initial filtering for quality control, using the standard criteria implemented in the Genotyping Console software (www.affymetrix.com), including quality control call rate (>86%), SNP call rate (>95%), and population stratification based upon principal components analysis, 1107 samples (560 cases and 547 control subjects) were retained for further analysis. They had 16,466 CNVs (eight subjects showed no CNVs). We then excluded low-confidence CNVs (logarithm of odds <10), CNVs <100 kb, and those with the lowest 1% density for probe coverage (52 segments). We removed 50 samples that had high sample-specific measures of noise (variance >2), as those had a mean of 175 CNV segments, indicating they were false-positives. We also removed 17 samples that had more than 20 apparent CNVs (the mean number of CNVs for these samples was 156), as such samples are also likely to be false-positives (4,7). The filtering left 1032 samples: 519 cases aged 43.4 ± 14.7 years (258 male and 261 female cases) and 513 control subjects aged 43.8 ± 14.5 years (252 male and 261 female control subject). They had a total of 5180 CNVs (~5 per person). Finally, following previous studies (4,7), we filtered common CNVs (found in >1% of the total sample), leaving 1052 rare and larger than 100 kb CNVs for the analysis (~1 per person). This filtering was also performed for CNVs found at >5% in the total sample, resulting in 2081 CNVs. All CNVs that passed filtering and were present in <1% of the samples are available as an University of California, Santa Cruz (UCSC)-friendly file in Supplement 1.

Copy number variations were considered to colocalize if they overlapped by at least 50% of their length, as implemented in PLINK ver1.0.4 (http://pngu.mgh.harvard.edu/~purcell/plink/) (12) as used for the analysis of CNV loci in previous datasets (4,7).

Results

The numbers of rare CNVs stratified by size in cases and control subjects are listed in Table 1. Overall, we found an excess of CNVs in subjects with schizophrenia (case-control ratio = 1.16). Although not significant ($p = .087$, one-tailed permutation test), this is similar to that reported by the largest CNV study (4) where the case-control ratio was 1.15. The effect in that study (4) was coming mostly from deletions >500 kb and duplications in the 100 kb to 200 kb range. No subcategory of CNV defined by size or nature (deletion or duplication) was significantly associated with disease in the current study. Copy number variations in the 100 kb to 200 kb range were more common in cases than in control subjects, ratio = 1.23, $p = .046$; however, this does not survive correction for the multiple testing of four size ranges and two types of CNVs. Duplications (but not deletions) within the same size range were the most significantly associated general category in the ISC study ($p = 1 \times 10^{-4}$) with virtually an identical effect (case-control ratio = 1.26). However, no specific duplications of this size overlapped between the two studies (4). We did not replicate the finding of an excess of large deletions (>500 kb) that was reported in the ISC study (4) or of deletions and duplications >1 Mb reported in the study by Kirov *et al.* (7).

Analysis of the burden of CNVs intersecting genes revealed no significant excess of genes disrupted in subjects with schizophrenia, either overall or for any size range, with similar trends to the results from the general burden analysis (Table 1).

We repeated the same analysis for CNVs <5% in the sample. This resulted in 388 and 368 deletions and 698 and 627 duplications in cases and control subjects, respectively. The trends between cases and control subjects were virtually identical to those in Table 1 (data not presented).

Although we found no enrichment of large CNVs in schizophrenia, we present the details of large CNVs (>1 Mb) in Table S1 in Supplement 2 because these have been most consistently implicated by others (4,7). Of those, one case but no control subjects had a deletion on 1q21.1, one of the most convincingly

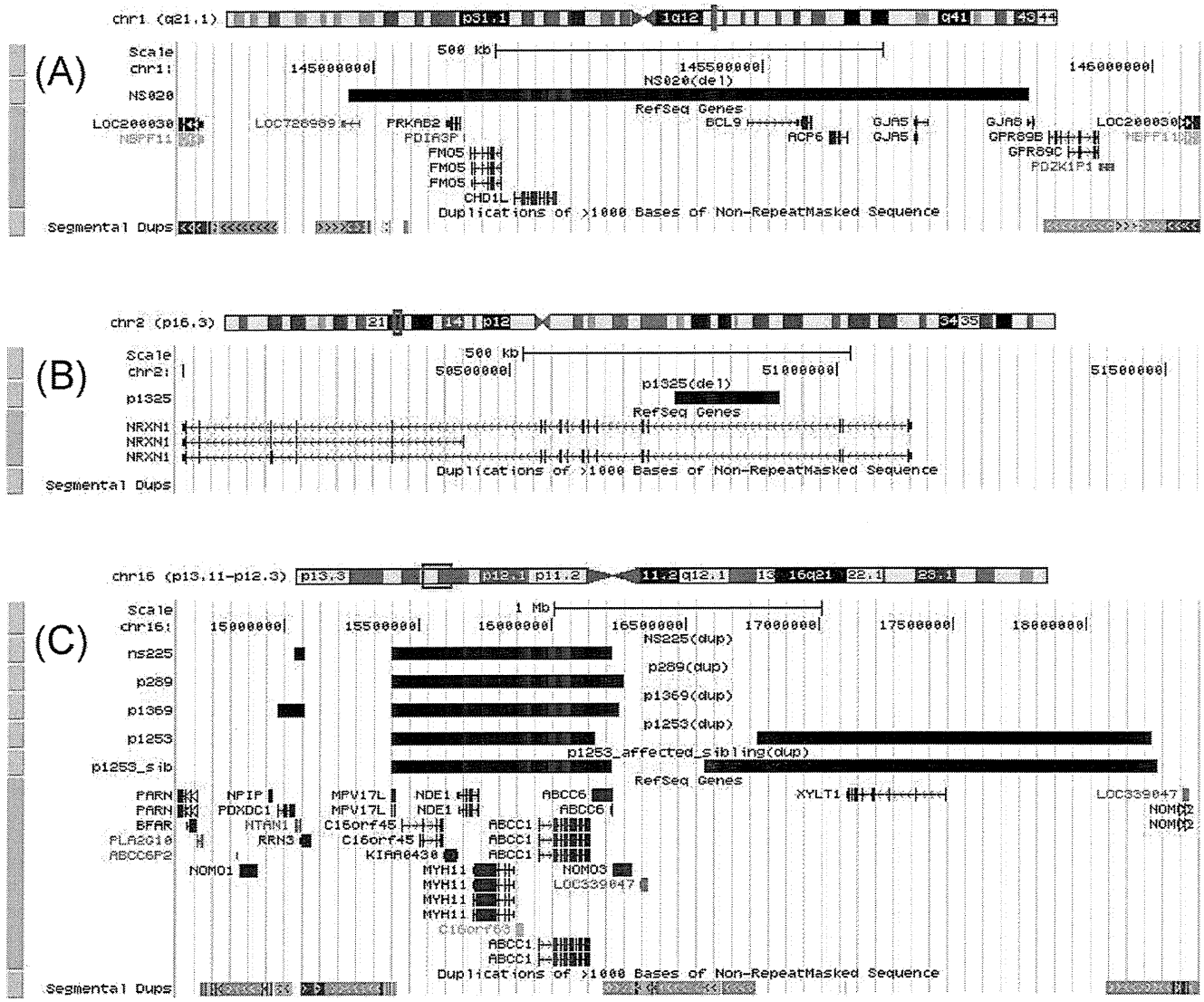

Figure 1. Positions of CNVs according to the validation experiments. CNV validation was undertaken using Illumina HumanHap 660W quad bead arrays (for CNVs at 1q21.1 and *NRXN1*) or 610-quad bead arrays (for CNVs at 16p13.1). Figures are produced on the UCSC Genome Browser according to NCBI Build 36.1, March 2006, hg18 (http://www.genome.ucsc.edu/) and indicate the positions of the CNVs: **(A)** 1q21.1; **(B)** *NRXN1*; and **(C)** 16p13.1: the last trace is that of the affected sibling of "p1253." CNV, copy number variation; NCBI, National Center for Biotechnology Information; UCSC, University of California, Santa Cruz.

implicated CNV risk factors for schizophrenia (4,8). Among large duplications, the most notable is that on 16p13.1, which was found in four cases and one control subject, while one more control subject had the reciprocal deletion (Fisher exact test p = .19, one-tailed). These CNVs in cases were confirmed using Illumina arrays (Figure S1 and Tables S1 and S2 in Supplement 2). One of the patients with 16p13.1 duplication had an affected sibling and unaffected mother who had also provided DNA. The duplication was found in the affected sibling but not the unaffected mother (DNA from the father was not available and there is no indication that he suffers with mental illness). The duplication in this family extends further on the centromeric side compared with the region usually included in CNVs of this region (Figure 1).

Of the remaining susceptibility loci reported in the recent studies (4,7,8), we found no deletions at 22q11.2 or 15q13.3. We also find no support for the 15q11.2 locus, where three deletions were found in control subjects and only one in a case (Fisher exact test p = .37, two-tailed, a trend in the opposite direction).

We also searched for CNVs that intersected genes and were present only in cases, reasoning as have others (2,3) that such CNVs are good candidates (Tables S3 and S4 in Supplement 2). One of the singleton deletions was in *NRXN1*, a gene implicated in previous studies (2,6,7,9,10) (Figure S1 and Table S2 in Supplement 2). Several more contain intriguing candidate genes (e.g., deletions in *PARK2*, *GRIK2*, *MAGEL2*, and *ATXN2L* and duplications in *CHRNA7* and *NRG4*), which have been implicated in neurodegenerative disorders or have possible functional relevance for neurodevelopment.

Discussion

In this study, we do not find a significant increase in the burden of CNVs in schizophrenia, either overall or for any

specific size range of CNVs, as proposed in previous studies (2–4,7). We did, however, find several trends in the same direction and of a similar magnitude as the largest global CNV survey of schizophrenia (4). Not all research has found such an increased burden, e.g., no evidence was obtained from a study in the Chinese population (5). It is possible that genuine population differences might drive this discrepancy between Caucasian and Asian samples, as might our exclusion of subjects with mental retardation or epilepsy. Sample size could also have played a role. Our sample had a modest power of ~.65 to detect a single CNV in a case for the following very strong candidate loci: 1q21.1, 15q13.3, and 22q11.2 and *NRXN1*, where approximately .2% of affected persons have deletions. In fact, we did find one deletion each in two of these loci (1q21.1 and *NRXN1*).

We found stronger support for association with duplications at 16p13.1, which contain the candidate gene *NDE1*. It is within the interval duplicated in all patients (Figure 1). Deletions and duplications of this region were implicated in autism (13) and schizophrenia (7), while deletions have been implicated in mental retardation (14). The most recent study surveying children with unexplained intellectual disability also reported significant association for both deletions and duplications at this locus ($p = 4.7 \times 10^{-5}$) (15), suggesting that this duplication is also pathogenic for a broad range of neuropsychiatric disorders. Our result for an excess of duplications in schizophrenic probands does not reach statistical significance; however, the frequency of the duplication is fourfold higher in cases than in control subjects (.8% vs. .2%), which is very similar to the rate found in our previous study from the United Kingdom (.6% vs. .2%) (7) and in the ISC study (.4% vs. .2%) (4). We found an identical duplication in an affected sibling. Larger CNVs in this locus, as in one of our probands, were also found in three cases and two control subjects in the ISC (4). The four probands in our study who carry 16p13.1 duplications do not appear to share any specific clinical features (Table S2 in Supplement 2).

We also found one deletion in a case at 1q21.1 and *NRXN1* and none in control subjects, which is close to the reported frequency of .2% in cases. Unlike those deletions of *NRXN1* that were associated with schizophrenia in a previous study (10), the CNV reported here does not intersect exons (10). However, it is large compared with most exon-sparing deletions reported in control subjects (10), and a new reanalysis of all *NRXN1* deletions shows that large (>100 kb) deletions in this gene might be almost as relevant as those affecting exons (16). The relevance to schizophrenia of the other CNVs found only in cases can only be assessed in future meta-analyses of such studies, but we note here that the three deletions we found in *PARK2* are of particular interest, as they have been implicated as a susceptibility factor for autism (17).

In summary, we provide support for the role of CNVs at 16p13.1, 1q21.1, and *NRXN1* in the etiology of schizophrenia. Although we find similar, but not significant, trends for an increased overall burden of CNVs, as well as for the involvement of duplications in the 100 kb to 200 kb range as proposed in the ICS study (4), in this population we could not find an increased burden of very large CNVs (>500 kb) in schizophrenia, which has been the main finding in recent studies (4,7). The discrepancy with previous studies could be due to our exclusion of patients with neurodevelopmental disorders, epilepsy, or known mental retardation, as such features are found in many of the carriers of large CNVs, e.g., 15q13.3 (15). Given the rarity of the CNVs that have been implicated so far in schizophrenia, there is a need for more large studies, studies in non-European populations, and meta-analyses.

This work was supported in part by research grant from the Japan Ministry of Education, Culture, Sports, Science and Technology; the Ministry of Health Labour and Welfare; the Core Research for Evolutional Science and Technology (CREST); the Health Sciences Foundation (Research on Health Sciences focusing on Drug Innovation); Medical Research Council (UK); and National Institute of Mental Health (USA) through a CONTE Center Grant (2 P50 MH066392-05A1). MI is a Japan Society for the Promotion of Science postdoctoral fellow for research abroad and is additionally supported by the Uehara Memorial Foundation and the Great Britain Sasakawa Foundation.

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

Supplementary material cited in this article is available online.

1. Cook EH Jr, Scherer SW (2008): Copy-number variations associated with neuropsychiatric conditions. *Nature* 455:919–923.
2. Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, *et al.* (2008): Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 320:539–543.
3. Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M (2008): Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat Genet* 40:880–885.
4. The International Schizophrenia Consortium (2008): Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* 455: 237–241.
5. Shi YY, He G, Zhang Z, Tang W, Zhang J Jr, Zhao Q, *et al.* (2008): A study of rare structural variants in schizophrenia patients and normal controls from Chinese Han population. *Mol Psychiatry* 13:911–913.
6. Need AC, Ge D, Weale ME, Maia J, Feng S, Heinzen EL, *et al.* (2009): A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genet* 5:e1000373.
7. Kirov G, Grozeva D, Norton N, Ivanov D, Mantripragada KK, Holmans P, *et al.* (2009): Support for the involvement of large CNVs in the pathogenesis of schizophrenia. *Hum Mol Genet* 18:1497–1503.
8. Stefansson H, Rujescu D, Cichon S, Pietilainen OP, Ingason A, Steinberg S, *et al.* (2008): Large recurrent microdeletions associated with schizophrenia. *Nature* 455:232–236.
9. Kirov G, Gumus D, Chen W, Norton N, Georgieva L, Sari M, *et al.* (2008): Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. *Hum Mol Genet* 17:458–465.
10. Rujescu D, Ingason A, Cichon S, Pietilainen OP, Barnes MR, Toulopoulou T, *et al.* (2009): Disruption of the neurexin 1 gene is associated with schizophrenia. *Hum Mol Genet* 18:988–996.
11. Korn JM, Kuruvilla FG, McCarroll SA, Wysoker A, Nemesh J, Cawley S, *et al.* (2008): Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat Genet* 40: 1253–1260.
12. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, *et al.* (2007): PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–575.
13. Ullmann R, Turner G, Kirchhoff M, Chen W, Tonge B, Rosenberg C, *et al.* (2007): Array CGH identifies reciprocal 16p13.1 duplications and deletions that predispose to autism and/or mental retardation. *Hum Mutat* 28:674–682.
14. Hannes FD, Sharp AJ, Mefford HC, de Ravel T, Ruivenkamp CA, Breuning MH, *et al.* (2008): Recurrent reciprocal deletions and duplications of 16p13.11: The deletion is a risk factor for MR/MCA while the duplication may be a rare benign variant. *J Med Genet* 46:223–232.
15. Mefford HC, Cooper GM, Zerr T, Smith JD, Baker C, Shafer N, *et al.* (2009): A method for rapid, targeted CNV genotyping identifies rare variants associated with neurocognitive disease. *Genome Res* 19:1579–1585.
16. Kirov G, Rujescu D, Ingason A, Collier DA, O'Donovan MC, Owen MJ (2009): Neurexin 1 (NRXN1) deletions in schizophrenia. *Schizophr Bull* 35:851–854.
17. Glessner JT, Wang K, Cai G, Korvatska O, Kim CE, Wood S, *et al.* (2009): Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* 459:569–573.

Identification of Novel Candidate Genes for Treatment Response to Risperidone and Susceptibility for Schizophrenia: Integrated Analysis Among Pharmacogenomics, Mouse Expression, and Genetic Case-Control Association Approaches

Masashi Ikeda, Yasuyuki Tomita, Akihiro Mouri, Minori Koga, Tomo Okochi, Reiji Yoshimura, Yoshio Yamanouchi, Yoko Kinoshita, Ryota Hashimoto, Hywel J. Williams, Masatoshi Takeda, Jun Nakamura, Toshitaka Nabeshima, Michael J. Owen, Michael C. O'Donovan, Hiroyuki Honda, Tadao Arinami, Norio Ozaki, and Nakao Iwata

Background: Pharmacogenomic approaches based on genomewide sets of single nucleotide polymorphisms (SNPs) are now feasible and offer the potential to uncover variants that influence drug response.

Methods: To detect potential predictor gene variants for risperidone response in schizophrenic subjects, we performed a convergent analysis based on 1) a genomewide (100K SNP) SNP pharmacogenetic study of risperidone response and 2) a global transcriptome study of genes with mRNA levels influenced by risperidone exposure in mouse prefrontal cortex.

Results: Fourteen genes were highlighted as of potential relevance to risperidone activity in both studies: *ATP2B2, HS3ST2, UNC5C, BAG3, PDE7B, PAICS, PTGFRN, NR3C2, ZBTB20, ST6GAL2, PIP5K1B, EPHA6, KCNH5,* and *AJAP1*. The SNPs related to these genes that were associated in the pharmacogenetic study were further assessed for evidence for association with schizophrenia in up to three case-control series comprising 1564 cases and 3862 controls in total (Japanese [JPN] 1st and 2nd samples and UK sample). Of 14 SNPs tested, one (rs9389370) in *PDE7B* showed significant evidence for association with schizophrenia in a discovery sample ($p_{allele} = .026$ in JPN_1st, two-tailed). This finding replicated in a joint analysis of two independent case-control samples ($p_{JPN\_2nd+UK} = .008$, one-tailed, uncorrected) and in all combined data sets ($p_{all} = .0014$, two-tailed, uncorrected and $p_{all} = .018$, two-tailed, Bonferroni correction).

Conclusions: We identified novel candidate genes for treatment response to risperidone and provide evidence that one of these additionally may confer susceptibility to schizophrenia. Specifically, *PDE7B* is an attractive candidate gene, although evidence from integrated methodology, including pharmacogenomics, pharmacotranscriptomic, and case-control association approaches.

Key Words: Expression: PDE7B, pharmacogenomics, risperidone, schizophrenia

Schizophrenia is a severe psychiatric disorder with a lifetime risk of approximately 1%. With its early onset, typically in late teens to early 20s, frequent relapse and chronic course, schizophrenia imposes a considerable burden on sufferers, their families, and society. Worldwide, it is a major source of morbidity, but it is often overlooked that it is also associated with a considerable truncation in life span, the mortality rate in individuals with schizophrenia being more than twice that of the age- and sex-matched population (1). A large number of antipsychotics have been developed as treatment agents. However, individual response to these drugs is highly variable, and identifying the optimal treatment for any patient is often a trial and error process that can span many years and even then, response is often poor. There is a pressing need both to identify new treatments and to attempt to improve the information based on which response to treatment can be predicted.

Genetic factors are generally assumed to contribute to variable treatment response (2), and on this basis, a number of pharmacogenetic studies have been performed. Here, the aim was to detect DNA sequence predictors for treatment response. Most studies have focused on genes encoding neurotransmitter receptors, such as dopamine or serotonin receptors, the logic being that antipsychotics usually have high affinities with members of these classes of receptor. Although a number of variants have been correlated with treatment response in several stud-

From the MRC, Centre for Neuropsychiatric Genetics and Genomics (MI, HJW, MJO, MCO), Department of Psychological Medicine and Neurology, School of Medicine, Cardiff University, Cardiff, United Kingdom; Department of Psychiatry (MI, TO, YY, YK, NI), Fujita Health University School of Medicine, Toyoake, Aichi, Japan; Department of Biotechnology (YT, HH), Nagoya University, Graduate School of Engineering, Nagoya, Japan; Department of Chemical Pharmacology (AM, TN), Graduate School of Pharmaceutical Sciences, Meijo University, Aichi, Japan; Department of Medical Genetics (MK, TA), Doctoral Program in Social and Environmental Medicine, Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan; Core Research for Environmental Science and Technology (CREST) (MK, TO, YK, RH, TA, NO, NI), Japan Science and Technology Agency, Kawaguchi, Saitama, Japan; Department of Psychiatry (RY, JN), University of Occupational and Environmental Health, Kitakyusyu, Fukuoka, Japan; The Osaka-Hamamatsu Joint Research Center for Child Mental Development (RH, MT), Osaka University, Graduate School of Medicine, Osaka, Japan; Department of Psychiatry (RH, MT), Osaka University, Graduate School of Medicine, Osaka, Japan; Department of Psychiatry (NO), Nagoya University, Graduate School of Medicine, Nagoya, Japan.

Address correspondence to Nakao Iwata, M.D., Ph.D., Department of Psychiatry, School of Medicine, Fujita Health University, Toyoake, Aichi, 470-1192, Japan; E-mail: nakao@fujita-hu.ac.jp.

Received Mar 3, 2009; revised Jul 29, 2009; accepted Aug 19, 2009.

0006-3223/10/$36.00
doi:10.1016/j.biopsych.2009.08.030

ies—for example, dopamine D2 and D3 receptor variants (2)—there are no definitive predictors of response.

Pharmacogenetics has been driven by a candidate gene approach. This approach has the disadvantage that targets for study are limited by our current understanding of the mechanisms of drugs, and therefore, this method cannot identify unsuspected predictor genes. Approaches that are independent of prior functional hypotheses of gene action based on genome-wide surveys of SNPs are, however, now feasible. The genome-wide approach has its disadvantages, but one of the most important is that, with effectively random sets of SNPs, the low prior probability that any is truly associated with disease requires a stringent type I error rate to control the enormous potential for reporting false positives. One way to address this issue is to use very large (and therefore highly powered) studies in which such stringent statistical support might realistically be achieved. Another approach that is more economical in genotyping costs is to undertake multistage analysis in which candidate variants from a screening sample are validated by replication in other data sets (3). However, because the samples for pharmacogenomics require a large amount of clinical data and are preferably prospective, large samples are difficult to collect.

Another approach is to try to enhance the prior probability for a given gene by integrating pharmacogenomic data with other sources of data—for example, from studies of gene expression (4). Under the hypothesis that genes related to drug response may be regulated by exposure to that drug, genes in which expression is altered in animals exposed to that drug have a higher probability of being genuinely associated in a pharmacogenetic study than any random gene. If this is correct, genes in which expression is altered by drug exposure require less stringent statistical support.

We aimed to detect predictor genes for risperidone response in schizophrenic patients using this convergent approach (4). Specifically, we compared data from a pharmacogenetic study based on first-episode, previously drug-naive subjects with schizophrenia who were treated with risperidone with data from a pharmacotranscriptomic study based on mice exposed to the same drug. Moreover, candidate variants from genes implicated by convergent data were also tested for evidence for association to schizophrenia per se because variants that are related to drug response may also be related to disease risk (5). Evidence that this occurs can be considered an additional independent line of circumstantial support that the convergence between the pharmacogenetic and transcriptomics does not merely reflect chance.

Methods and Materials

Subjects and Collection of Clinical Data

We performed an open-labeled pharmacogenetic study involving 108 first-episode, previously antipsychotic-naive schizophrenic patients. All received risperidone monotherapy after enrollment. Details are described elsewhere (6,7). Briefly, patients were entered into the study if they 1) met DSM-IV-TR criteria for schizophrenia (and then remained in follow-up to at least 6 months), 2) were physically healthy and had all laboratory parameters within normal limits, and 3) had neither a current nor a past DSM-IV-TR diagnosis of mood disorders or substance abuse. Consensus diagnoses were made by at least two experienced psychiatrists on the basis of unstructured interviews with patients and families and review of medical records. Duration of untreated psychosis (DUP) was defined as the period from the onset of psychotic symptoms to that of first antipsychotic exposure. Sixty subjects were recruited from outpatient clinics, and 48 subjects were treated as inpatients.

Subjects received risperidone monotherapy (starting dosage: .5–4 mg/day, mean starting dosage: 2.5 mg), and dosage was adjusted in accordance with symptomatic response by trained psychiatrists (1–8 mg/day, mean dosage: 3.4 mg at 8 weeks) for 8 weeks. Patients with insomnia were prescribed brotizolam, .25 mg or .5 mg, at bedtime. No other psychotropic drugs were permitted.

Clinical symptoms were evaluated at the first visit and after 8 weeks of treatment by the use of the Positive and Negative Syndrome Scale (PANSS). Evaluations were carried out by qualified psychiatrists and psychologists (the interrater reliability was measured by intraclass correlation coefficient was .90, unpublished data).

The clinical characteristics of subjects that we used as potential covariates were selected from another report (8): sex (57 male, 51 female), age (mean 30.2 ± 9.5 years), DUP (1.5–32 months, mean 7.6 ± 7.1 months), and baseline PANSS total score (mean 83.0 ± 22.9).

Samples used in the schizophrenia case-control association analysis consisted of three sets: (1) JPN_1st: this was used for identifying genes of potential interest and comprised 540 patients with schizophrenia (275 male and 265 female; aged 43.3 ± 15.0 years) and 425 healthy controls (236 male and 189 female; aged 36.3 ± 13.9 years) from the Japanese population; 2) JPN_2nd sample (used to follow up genes of interest) comprised 545 patients with schizophrenia (282 male and 263 female; aged 50.7 ± 14.9 years) and 500 controls (279 male and 221 female; aged 40.8 ± 15.4 years) from the Japanese population; 3) Additional follow-up data for SNPs of interest were extracted from a UK genomewide association study (GWAS) of schizophrenia comprising 479 patients with schizophrenia and 2937 controls from the UK population (9).

Controls in the Japanese population were screened for past history of mental disorders. All individuals were unrelated. After explanation of the study, written informed consent was obtained from each subject. This study was approved by the Ethics Committee at Fujita Health University, University of Occupational and Environmental Health, Nagoya University Graduate School of Medicine, Osaka University Graduate School of Medicine and by multiple ethics committees across the UK where sample recruitment was performed.

Microarray Experiments

See also Methods in Supplement 1.

SNP Chip. Genomewide genotyping was carried out using Illumina Sentrix human 1 Genotyping BeadChip (109,363 SNPs randomly distributed throughout the genome) according to the manufacturer's instructions (Illumina, San Diego, California). Details are given in the Supplement 1.

Mouse Expression Chip. We compared mRNA levels of the prefrontal cortex (PFC) between control ($n = 3$) and risperidone-exposed mice (2.4 mg/kg given orally, once a day for 21 days, $n = 3$). Affymetrix Mouse Gene 1.0 St. Array, which profiles the expression of 28,853 genes (Affymetrix, Santa Clara, California), was used to measure the amount of mRNA.

The procedures involving animals and their care were conducted in conformity with the international guidelines, Principles of Laboratory Animal Care (National Institutes of Health Publication 85-23, revised 1985).

Experimental Procedures and Statistical Analysis

Study 1: Pharmacogenomics. Quality control (QC) regarding population stratification (Figure S1 in the Supplement 1),

Hardy-Weinberg equilibrium (HWE), genotyping rate, and minor allele frequency was conducted by PLINK (10). Details are described in Supplement 1.

After QC, 99 samples (51 males and 48 females) and 62,935 autosomal SNPs (a mean call rate of 99.2%, indicating a high rate of successful genotyping) were analyzed to evaluate the effect of each SNP on antipsychotic response to risperidone.

To evaluate the effect of each SNP on antipsychotic response to risperidone, multiple regression analysis was carried out with a dependent variable [% PANSS change = 100 × ((PANSS at week 0) − (PANSS at Week 8))/PANSS at Week 0] and independent variables that included sex, age, duration of illness, initial PANSS score, and the genotype of each polymorphism. Each genotype was assessed using dominant, recessive, and multiplicative genetic models, respectively.

To calculate the best empiric p values based on the most significant result in each genetic model, we generated 1 million simulated data sets by randomizing the PANSS changes (the covariates stay with the genotypes) with respect to the GWAS data. This approach retains the linkage disequilibrium (LD) relationships between SNPs, and therefore allows for the appropriate degree of nonindependence in the data sets. The same multiple regression analysis model as applied to the real data were applied to each SNP in each permuted data set, and the empiric significance for a SNP was the proportion of the simulated data sets in which the test statistic was equal to, or greater than, that observed in the true data set (11–15).

SNPs were annotated to the closest genes with an up- and downstream span of 20 kb by WGAViewer (16).

Study 2: Mouse Expression Assay. In the mouse expression assay, data sets passing QC were normalized using GeneChip Operating Software (Affymetrix) and the raw intensity values exported for further analysis. Only genes called present based on Affymetrix detection p value for the presence of each gene on either chip were included. A t test was performed to assess the statistical significance of genes in which expression differed between control and risperidone-exposed mice. Power analysis was carried out using PowerAtlas (17). Our data set provides expected discovery rate (corresponding to power) of .37, an expected proportion of true positives (PTPs) of .72, and an expected proportion of true negatives of .80 at alpha set at .05. A major aim of this study was to prioritize genes showing convergent evidence in the pharmacogenomic study, thus we consider a high PTP optimal.

These data were submitted to CIBEX (http://cibex.nig.ac.jp/index.jsp, accession number: CBX77).

Study 3: Checking Overlap Results Between Pharmacogenomics (Study 1) and Mouse Expression Assay (Study 2). We checked candidate SNPs from the genes that showed convergent evidence for relevance to risperidone action from Study 1 and Study 2. Candidate genes were defined as follows: 1) genes for which there was at least one SNP with p values less than 5.0×10^{-4} in the pharmacogenomic study and in which expression significantly differed between groups at $\alpha < .05$ and 2) genes with much stronger evidence for $p < 1.0 \times 10^{-4}$ but that had weaker evidence for association in the pharmacogenomic study ($p < .05$).

Study 4: Case-Control Association Analysis of Strong Candidate Genes from Pharmacogenomics and Mouse Expression Assay. The candidate SNPs from Study 3 were further assessed for evidence for association with schizophrenia. These SNPs were genotyped by TaqMan assay (Applied Biosystems, California) in the Japanese case-control samples. Genotypes for the SNP in *PDE7B* in the UK samples were extracted from the Affymetrix GWAS data (9) after confirmation of good-quality cluster plots.

Genotype deviation from HWE was evaluated by a goodness of fit chi-square test. Marker-trait association was evaluated for allele/genotype-wise using standard contingency tables (SPSS 15.0, SPSS, Tokyo, Japan).

For SNPs analyzed in multiple samples, we conducted a meta-analysis using a random-effects model. Heterogeneity was measured using a Q statistic test in the combined studies. Odds ratios (ORs) were pooled using DerSimonian and Laird methods. The significance of the pooled OR was determined using a Z test. All data were analyzed using an R package, meta (http://www.r-project.org/index.html).

Results

Possible Predictor SNPs for Risperidone Treatment: From Pharmacogenomic Result (Study 1)

Among the 62,935 SNPs we examined in the pharmacogenomics study, 51,550 SNPs were annotated to 14,655 genes (annotation span: 5′ or 3′ ± 20 kb). For a number of genes, we had multiple SNPs with p values less than 5.0×10^{-4} because of the high LD among genotyped markers. Where this occurred, we list only the strongest associated SNP from that gene (the top 10 hits and SNPs with p value less than 5.0×10^{-4} in Table 1 and Table S1 in Supplement 1, respectively).

Table 1. Predictor Genes in the Pharmacogenomics (Top 10)

Ranking	SNP	Chr	Position[a]	Closest Gene[b]	p Value (Pharmacogenomics)
1	rs2289273	3	10,388,601	*ATP2B2*	1.60×10^{-5}
2	rs234091	1	183,186,172	*FAM129A*	2.00×10^{-5}
3	rs241202	8	28,689,604	*INTS9*	3.20×10^{-5}
4	rs4340422	19	48,604,802	*TEX101*	5.00×10^{-5}
5	rs6682786	1	23,615,883	*TCEA3*	7.30×10^{-5}
6	rs1001220	7	72,748,539	*WBSCR22*	7.70×10^{-5}
7	rs3829241	11	68,611,939	*TPCN2*	8.90×10^{-5}
8	rs460473	16	22,740,528	*HS3ST2*	1.03×10^{-4}
9	rs9792264	8	135,640,117	*ZFAT*	1.10×10^{-4}
10	rs6443999	3	186,056,249	*VPS8*	1.17×10^{-4}

Chr, chromosome; SNP, single nucleotide polymorphism.
[a]Based on Ensemble *Homo sapiens* Version 54.36p (NCBI36).
[b]SNPs are annotated to the closest genes with ± 20-kb span.

Table 2. Overlap Genes Based on the Pharmacogenomcis ($p < 5.0 \times 10^{-4}$) with Mouse Expression Assay ($p < .05$)

Ranking	SNP	Chr	Position[a]	Closest Gene[b]	*p* Value (Pharmacogenomics)	*p* Value[c] (Mouse Expression)	Fold Change
1	rs2289273	3	10,388,601	*ATP2B2*	1.60×10^{-5}	.000710	.504
8	rs460473	16	22,740,528	*HS3ST2*	1.03×10^{-4}	.00600	.259
28	rs3775003	4	96,390,234	*UNC5C*	2.20×10^{-4}	.0132	1.85
32	rs196290	10	121,398,061	*BAG3*	2.81×10^{-4}	.0283	1.33
35	rs9389370	6	136,472,958	*PDE7B*	2.88×10^{-4}	.00806	.710
53	rs1356787	4	57,012,104	*PAICS*	4.26×10^{-4}	.0368	.660
54	rs4641299	1	117,284,884	*PTGFRN*	4.27×10^{-4}	.00160	.283

Chr, chromosome; SNP, single nucleotide polymorphism.
[a]Based on Ensemble *Homo sapiens* Version 54.36p (NCBI36).
[b]SNPs are annotated to the closest genes with ± 20-kb span.
[c]Comparison between risperidone-treated mice ($n = 3$) and saline-treated mice ($n = 3$).

We also looked specifically in our data for support for genes recently suggested as associated with iloperidone based on the only other available antipsychotic GWAS data set (18) and candidate genes implicated in earlier studies (2) including *DRD2*, *DRD3*, *HTR2A*, and others (Tables S2 and S3 in Supplement 1). No strong evidence for association to any of these was found in our pharmacogenomics data set.

Genes Influenced by Risperidone Exposure in Mouse PFC (Study 2)

We examined 22,556 probes in 12,706 genes in RNA extracted from the PFC of mice treated with either risperidone or with vehicle. Of these, 754 (5.9%) and 2227 (17.5%) genes had at least one probe that showed nominally significant differences at $p <$ .01 and .05, respectively, a rate much higher than chance. The top genes with *p* value less than 5.0×10^{-4} are presented in Table S4 in Supplement 1.

Overlapping Genes Between Pharmacogenomic and Mouse Expression Assays (Study 3)

We looked to see whether the pharmacogenetic data (excluding 14 SNPs that could not be annotated to the closest gene) and expression overlapped. Seven genes containing nominally significant alteration in expression in mice also contained SNPs with *p* value less than 5×10^{-4} (Table 2). The relation between PANSS changes and physical locations of each SNP and the genotype effects to risperidone response can be seen in Figures S2 and S3 and Table S5 in Supplement 1. In addition, we found seven genes that met the more stringent threshold for expression change in the mouse and that had at least one significant SNP ($p < .05$) in the pharmacogenomic data (Table 3). It should be stressed these SNPs were not strongly associated with treatment response ($p = .0047$-$.0472$).

Consequently 14 SNPs were further assessed for case-control association analysis in Study 4.

Examining Candidate SNPs as Susceptibility Factor for Schizophrenia (Study 4)

The 14 candidate SNPs in genes showing convergent evidence from Study 3 were further tested for association with schizophrenia (Table 4). For rs242056, a proxy for rs2071999 in *AJAP1*, the genotypes significantly deviated from HWE in controls ($p = .0016$). This SNP was therefore excluded.

Of the remaining 13 SNPs, a single SNP (rs9389370) in *PDE7B* showed a nominally significant association in the JPN_1st case-control sample ($p_{\text{allele}} = .026$, two-tailed). In an attempt to extend this putative association, we used two other samples. In the second Japanese sample, we obtained significant evidence for association (second set, $P_{\text{allele}} = .02$, one-tailed) and a nonsignificant trend in the UK sample ($p_{\text{allele}} = .07$, one-tailed) (Table 4). Meta-analysis of the two replication data sets showed significant evidence for association ($p_{\text{JPN2nd+UK}} = .008$, one-tailed). As expected, in all data sets combined, the evidence was stronger than observed in the screening sample alone ($p_{\text{all}} = .0014$, two-tailed, uncorrected; $p = .018$, 13 times Bonferroni correction for number of SNPs tested in Study 4) with no evidence for heterogeneity ($p = .56$; Table 5).

Discussion

Combined Analysis as a Tool for Prioritizing Candidate Genes for Pharmacogenomics and Susceptibility

Genomewide approaches to pharmacogenomics have the capacity to provide novel insights into mechanisms and predictors of drug response. However, a major concern of this approach, which is not specific to pharmacogenomics, relates to balancing the need to set a stringent threshold for the type I error rate against the desire to achieve power to detect findings at that threshold. Unless the genetic effect sizes in pharmacogenetics are substantially greater than is typical for complex diseases (19), the sorts of sample sizes currently available for studies of

Table 3. Overlap Genes Based on the Mouse Expression Assay ($p < 1.0 \times 10^{-4}$) with the Pharmacogenomics ($p < .05$)

Ranking	Gene	Probe ID	Fold Change	*p* Value (Mouse Expression)	*p* Value (Pharmacogenomics)	SNP ID
2	*Nr3c2*	1435991-at	5.86	2.23×10^{-6}	.0297	rs2070951
3	*Zbtb20*	1439278-at	4.94	5.04×10^{-6}	.0230	rs9883949
4	*St6gal2*	1434819-at	.23	6.52×10^{-6}	.0102	rs1448110
7	*Pip5k1b*	1450389-s-at	2.46	1.02×10^{-5}	.0472	rs1414944
8	*Epha6*	1421527-at	3.18	1.46×10^{-5}	.0047	rs727229
9	*Kcnh5*	1441742-at	.44	2.72×10^{-5}	.0305	rs10141458
24	*Ajap1*	1438662-at	.66	9.21×10^{-5}	.0208	rs2071999

SNP, single nucleotide polymorphism.

Table 4. Case-Control Analysis of the Candidate SNPs from the Pharmacogenomics and Mouse Expression Data

SNP	Sample	Proxy SNPs	Phenotype	N	Genotype M/M	Genotype M/m	Genotype m/m	*p* Value Allele	*p* Value Genotype	*p* Value HWE	MAF
ATP2B2	JPN_1st		Case	536	275	208	53	.676	.184	.14	29.3
rs2289273			Control	417	209	179	29			.26	28.4
HS3ST2	JPN_1st		Case	538	163	263	112	.408	.682	.76	45.3
rs460473			Control	407	117	196	94			.50	47.2
UNC5C	JPN_1st		Case	540	336	178	26	.249	.136	.70	21.3
rs3775003			Control	406	231	159	16			.07	23.5
BAG3	JPN_1st		Case	539	234	243	62	.609	.877	.93	34.0
rs196290			Control	407	183	180	44			.98	32.9
PDE7B	JPN_1st		Case	535	259	222	54	**.0255**	.0738	.53	30.8
rs9389370			Control	422	229	165	28			.81	26.2
	JPN_2nd		Case	536	278	200	58	**.0214**[a]	.0966	.018	29.5
			Control	500	281	183	36			.41	25.5
	UK		Case	478	181	226	71	.0672[a]	.327	.97	38.5
			Control	2,932	1203	1348	381			.91	36.0
PAICS	JPN_1st		Case	540	181	274	85	.662	.808	.27	41.1
rs1356787			Control	424	134	223	67			.10	42.1
PTGFRN	JPN_1st		Case	535	319	180	36	.222	.219	.13	23.6
rs4641299			Control	413	255	141	17			.65	21.2
NR3C2	JPN_1st		Case	534	295	206	33	.284	.509	.71	25.5
rs2070951			Control	414	213	173	28			.37	27.7
ZBTB20	JPN_1st		Case	537	234	245	58	.494	.420	.61	33.6
rs9883949			Control	423	169	211	43			.05	35.1
ST6GAL2	JPN_1st	rs2241991	Case	533	224	246	63	.936	.814	.72	34.9
rs1448110		$r2 = 1$	Control	409	169	196	44			.25	34.7
PIP5K1B	JPN_1st		Case	536	173	275	88	.550	.764	.22	42.1
rs1414944			Control	420	145	208	67			.6	40.7
EPHA6	JPN_1st		Case	539	146	274	119	.877	.949	.65	47.5
rs727229			Control	419	110	217	92			.44	47.9
KCNH5	JPN_1st		Case	539	169	265	105	.224	.476	.95	44.1
rs10141458			Control	413	142	201	70			.94	41.3
AJAP1	JPN_1st	rs242056	Case	—	—	—	—	—	—	—	—
rs2071999		$r2 = .46$	Control	418	191	161	66			.0016	35.0

Bold numbers represent significant *p* value.

JPN_1st, first Japanese sample; JPN_2nd, second Japanese sample; MAF, minor allele frequency; HWE, Hardy-Weinberg equilibrium; M, major allele, m, minor allele; SNP, single nucleotide polymorphism; UK, United Kingdom sample.

[a]Based on one-tailed analysis.

antipsychotics have no realistic prospect of attaining the sorts of levels of significance suggested for genomewide significance (7.2×10^{-8} or 1×10^{-7}) (20,21). Although it is at least possible that the typical effects on gene expression of drugs may be much more substantial than that of SNPs on disease risk, broadly similar balances of power and type I error also apply to our genomewide expression study. Therefore, with the aim of prioritizing our findings, we attempted to cross-validate the top findings from our study using independent approaches as suggested (4). Our methods of prioritizing our findings were based on two hypotheses; one that the most highly significant sets of SNPs from our pharmacogenetic study of risperidone are likely to be enriched among genes whose expression is altered by that drug (and vice versa), the other that SNPs related to drug response may also be enriched among SNPs associated with disease. To what extent these hypotheses are correct is currently unknown.

From our data, we found 14 markers in genes that showed some degree of overlapping support in the pharmacogenomics

Table 5. Meta-Analysis of rs9389370 in PDE7B

Analysis	Sample	OR	95% CIs Lower Limit	95% CIs Upper Limit	*p* Value
	JPN_1st	1.26	1.03	1.54	**.0255**
	JPN_2nd	1.22	1.01	1.48	**.0214**[a]
	UK	1.11	.967	1.28	.0672[a]
Meta (Replication)	JPN_2nd+UK	1.15	1.03	1.29	**.0082**[a]
Meta (All)	JPN1st+JPN2nd+UK	1.17	1.06	1.30	**.0014**

Bold number represent significant *p* value.

CI, confidence interval; JPN_1st, first Japanese sample; JPN_2nd, second Japanese sample; OR, odds ratio; UK, United Kingdom Sample.

[a]Based on one-tailed analysis.

and mouse expression experiments. These top convergent candidate genes have no previous support for association with schizophrenia or risperidone response and are thus novel candidates for antipsychotic response. However, at present, they have no clinical utility in terms of predicting treatment response, and independent replication using other samples will be required. Moreover, even if replicated, the potential clinical utility for pharmacogenetics is questionable because the effect sizes in each case are small, although it is conceivable given the limited coverage of each gene that the true functional variants have much stronger effects.

Another method for prioritizing genes from genomewide data are to apply a gene ontology (GO) based approach to investigate whether sets of findings tend to converge on particular biological pathways or functions. Our previous experience of GO category analysis suggests that with respect to genetic data, these require large data sets (22). Nevertheless, in response to an anonymous reviewer's comments, for interested readers, we provide the results of our GO category analyses based on ALIGATOR (22) and David Bioinformatics Resources 2008 (http://david.abcc.ncifcrf.gov/) in Supplement 1. Although a number of categories were observed to be significant in each analysis (Tables S6 and S7 in Supplement 1), there is no overlap between the results of the two analytic approaches. Moreover, our favored approach based on ALIGATOR did not reveal any categories that were significant after correction for multiple testing, so it is likely that all of those findings represent chance positives.

Possible Predictor SNPs for Response to Risperidone

In this study, several genes were detected as possible novel predictors for treatment response to risperidone: *ATP2B2, HS3ST2, UNC5C, BAG3, PDE7B, PAICS, PTGFRN, NR3C2, ZBTB20, ST6GAL2, PIP5K1B, EPHA6, KCNH5,* and *AJAP1*. Because the multiple testing burden in SNPs is more severe, our primary analysis included selecting genes based on the more stringent thresholds in the pharmacogenomics data (Table 2) and were additionally shown to have altered expression in the mouse expression study. However, in response to review, we additionally provide data for much more weakly associated SNPs that have highly significant expression changes in the mouse brain (Table 3). Given the weak evidence for most of the latter group of SNPs, we think those are most likely to be chance positives but report the findings for others to test.

Among genes of particular interest in Table 2 is *ATP2B2*, which encodes one of four isoforms of the plasma membrane Ca^{2+} pumps of mammalian cells, showed both the strongest statistical association with treatment response ($p = 1.60 \times 10^{-5}$) and was among those genes that had the strongest association with differential expression because of exposure to risperidone ($p = .00071$). The product of this gene is thought to be involved in neurodevelopment (23) because of its influence on Ca^{2+} homeostasis and Ca^{2+} signaling. This in turn regulates multiple neuronal functions, including synaptic transmission, plasticity, and cell survival (24). Interestingly, several of the other genes with convergent evidence for a role in risperidone response might also be related to neurodevelopment via association with netrin (*UNC5C*) (25,26), interaction with heat shock proteins (*BAG3*) (27,28), cyclic adenosine monophosphate (cAMP) systems (*PDE7B*; details discussed later), glucocorticoids (*NR3C2*) (29), and ephirin (*EPHA6*) (30). Given the neurodevelopmental hypothesis of schizophrenia (31) and evidence that second-generation antipsychotics, including risperidone, have neurogenic actions in hippocampus and PFC (32), our findings suggest that genes involved in the regulation of neurodevelopment or neurogenesis are candidate genes for treatment response in schizophrenics, as well as for schizophrenia per se.

PDE7B Is Candidate Gene Either for Treatment Response and Susceptibility for Schizophrenia

We pursued the top findings from Study 3 to see whether the findings with best convergent evidence (human and mouse) for relevance to risperidone response might also influence susceptibility to schizophrenia. After correction for multiple testing, we found evidence for association between disease status and *PDE7B*, which was therefore the only gene supported across all study designs.

Phosphodiesterases (PDEs) are central in regulating degradation of cAMP and cyclic guanosine monophosphate (cGMP), which are important second messengers for many cellular functions (33). There are 21 known genes encoding PDEs in human, spread across 11 distinct PDE families (*PDE1* to *PDE11*). Among these, *PDE4B* has been reported as a candidate susceptibility gene for schizophrenia. This was on the basis of a translocation found in two affected members of a single pedigree and the observation that the protein interacts with Disrupted in Schizophrenia 1 (*DISC1*), itself another strong candidate gene for schizophrenia and affective disorders (34). Elevation of cellular cAMP leads to dissociation of *PDE4B* from DISC1 and an increase in *PDE4B* activity (34).

PDE7B degrades cAMP, but not cyclic guanosine monophosphate (cGMP), and is predominantly expressed in brain (33). To date, no direct evidence for association of *PDE7B* with schizophrenia has been reported; however, several findings provide some functional plausibility to our results. First, mRNA for *PDE7B* and dopamine D1, D2, and D3 receptors show a similar pattern of distribution, and it is thought that the dopamine D1 receptor activates *PDE7B* through the cAMP pathway (35). Second, *PDE7B* maps to 6q23-24, one of the most significant linkage regions for schizophrenia (OMIM %600511; SCZD3) (36). Lastly, association has recently been reported to the *Abelson Helper Integration Site 1* (AHI1) and *Family with sequence similarity 54 A* (*FAM54A*), which are respectively in the 5′ and 3′ regions of *PDE7B*. However, it should be noted that in those studies, SNPs in *PDE7B* were not associated with disease (37,38).

PDE inhibitors have recently emerged as being of interest as therapeutic agents for neuropsychiatric disorders, such as schizophrenia, depression, and dementia (33). Our results indicate that among these, drugs acting on *PDE7B* may be of particular value in schizophrenia, although particularly for clinical applications, our results should be treated with caution until independent replications have been reported.

Limitations and Conclusion

The major limitation in this study is that the sample sizes we used for the genomewide pharmacogenetics and gene expression studies are small. In particular, the phramacogenetics study is only highly powered to detect effects that are much larger than typical of common susceptibility alleles for diseases to date. This is less of a limitation with respect to one major goal of pharmacogenetics, namely, the identification of common markers with sufficiently large effects to be of value in guiding therapeutics. Our study suggests that in such large common effects may not exist, although being based on one of the earliest chips, the coverage of genes is incomplete, and it would be desirable to repeat this experiment with a denser set of SNPs. The extent to which clinical heterogeneity is likely to have an impact on treatment response, and therefore power to detect association to that response, is also currently