Genome-Wide Association Study of Schizophrenia in a Japanese Population

Masashi Ikeda, Branko Aleksic, Yoko Kinoshita, Tomo Okochi, Kunihiro Kawashima, Itaru Kushima, Yoshihito Ito, Yukako Nakamura, Taro Kishi, Takenori Okumura, Yasuhisa Fukuo, Hywel J. Williams, Marian L. Hamshere, Dobril Ivanov, Toshiya Inada, Michio Suzuki, Ryota Hashimoto, Hiroshi Ujike, Masatoshi Takeda, Nick Craddock, Kozo Kaibuchi, Michael J. Owen, Norio Ozaki, Michael C. O'Donovan, and Nakao Iwata

Background: Genome-wide association studies have detected a small number of weak but strongly supported schizophrenia risk alleles. Moreover, a substantial polygenic component to the disorder consisting of a large number of such alleles has been reported by the International Schizophrenia Consortium.

Method: We report a Japanese genome-wide association study of schizophrenia comprising 575 cases and 564 controls. We attempted to replicate 97 markers, representing a nonredundant panel of markers derived mainly from the top 150 findings, in up to three data sets totaling 1990 cases and 5389 controls. We then attempted to replicate the observation of a polygenic component to the disorder in the Japanese and to determine whether this overlaps that seen in UK populations.

Results: Single-locus analysis did not reveal genome-wide support for any locus in the genome-wide association study sample (best $p = 6.2 \times 10^{-6}$) or in the complete data set in which the best supported locus was *SULT6B1* (rs11895771: $p = 3.7 \times 10^{-5}$ in the meta-analysis). Of loci previously supported by genome-wide association studies, we obtained in the Japanese support for *NOTCH4* (rs2071287: $p_{meta} = 5.1 \times 10^{-5}$). Using the approach reported by the International Schizophrenia Consortium, we replicated the observation of a polygenic component to schizophrenia within the Japanese population ($p = .005$). Our trans Japan–UK analysis of schizophrenia also revealed a significant correlation (best $p = 7.0 \times 10^{-5}$) in the polygenic component across populations.

Conclusions: These results indicate a shared polygenic risk of schizophrenia between Japanese and Caucasian samples, although we did not detect unequivocal evidence for a novel susceptibility gene for schizophrenia.

Key Words: Genome-wide association study, *NOTCH4*, polygenic component, schizophrenia, *SULT6B1*

Epidemiologic studies show that genetic factors account for more than 80% of the population variance in susceptibility for schizophrenia; however, as with virtually all other relatively common disorders, it has historically proven difficult to identify the specific genetic variants involved (1).

The application of genome-wide association technology to large case–control samples of mainly European ancestry has recently implicated a number of risk loci for which the evidence is strong. These include loci defined by single nucleotide polymorphisms (SNPs) in which the effects are weak (odds ratios [ORs] 1.1–1.25) among which the strongest supported loci are *zinc finger protein 804 A* (*ZNF804A*) (2–5), a broad region including the major histocompatibility complex (MHC) on chromosome 6p21.3–22.1 (6–8), *neurogranin* (*NRGN*), and *transcription factor 4* (*TCF4*) (8).

Although the robust support for a number of recently implicated loci represents something of a break from the past inconsistencies, little of the genetic variance of schizophrenia can be explained by the loci identified thus far. One explanation for this is that much of the risk is conferred by common but weak genetic effects that require larger samples. Another explanation is that most of the risk cannot be readily detected by genome-wide association studies (GWAS), the missing genetic component being conferred by mutations that exert substantial individual effects that are rare or even unique to individual pedigrees.

Although the relative contributions of these classes of variant awaits empiric resolution, the GWAS of the International Schizophrenia Consortium (ISC) provided strong support for a substantial polygenic contribution (at least 30%) to the population risk of schizophrenia, much of which is conferred by common alleles with small effect sizes (6,9,10). The basic principle of their analysis was that in the presence of a substantial common polygenic component, although most of the individual genetic effects will not be

From the Department of Psychiatry (MI, BA, YK, ToO, KuK, IK, YI, TK, TaO, YF, NI), School of Medicine, Fujita Health University, Toyoake, Aichi, Japan; Medical Research Council (MI, HJW, MLH, DI, NC, MJO, MCO), Centre for Neuropsychiatric Genetics and Genomics, Department of Psychological Medicine and Neurology, School of Medicine, Cardiff University, Cardiff, United Kingdom; Department of Psychiatry (BA, IK, YI, YN, NO), Graduate School of Medicine, Nagoya University, Nagoya, and Japan Science and Technology Agency (BA, YK, ToO, KuK, IK, YI, YN, TK, TaO, YF, MS, RH, MT, NO, NI), the Core Research for Evolutional Science and Technology, Kawaguchi, Japan; Biostatistics and Bioinformatics Unit (MLH, DI), School of Medicine, Cardiff University, Cardiff, United Kingdom; Seiwa Hospital (TI), Institute of Neuropsychiatry, Tokyo; Department of Neuropsychiatry (MS), Graduate School of Medicine, University of Toyama, Toyama; Department of Psychiatry (RH, MT), Graduate School of Medicine, Osaka University, Osaka; Molecular Research Center for Children's Mental Development (RH, MT), United Graduate School of Child Development, Osaka University, Kanazawa University, and Hamamatsu University School of Medicine, Osaka; Department of Neurospychiatry (HU), Okayama University, Graduate School of Medicine, Dentistry and Pharmaceutical Science, Okayama; and Department of Cell Pharmacology (KoK), Graduate School of Medicine, Nagoya University, Nagoya, Japan.

Authors MI and BA contributed equally to this work.

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 May 20, 2010; revised Jul 9, 2010; accepted Jul 9, 2010.

0006-3223/$36.00
doi:10.1016/j.biopsych.2010.07.010

detectable in current sample sizes, the sum of many such effects across multiple SNPs might differ between cases and controls. After discounting the influence of various potential sources of bias, the authors concluded that the findings were best explained by the existence of an important polygenic component to the disorder comprising a large number of common alleles, although some contribution from low-frequency alleles was not excluded or deemed unlikely (6).

There were two additional striking findings in the ISC article (6). The first was that those alleles selected as "risk" alleles for schizophrenia were also enriched in people with bipolar disorder, supporting the hypothesis of shared genetic susceptibility between these disorders (11,12). The second was that sets of "risk" alleles defined from white individuals of European origin were better at predicting affected status in other white European subjects than they were in African Americans, although an attenuated effect was seen in an African American sample. This may be attributable to differences in allele frequencies and linkage disequilibrium between Europeans and African Americans, although genetic heterogeneity remains a possibility. In this article describing a study that sought novel susceptibility variants, we report the first GWAS for schizophrenia in a Japanese sample. Although the Japanese population is considered relatively homogeneous (13), GWAS studies in other populations strongly suggest that our study of 575 cases and 564 controls is underpowered to detect any findings at genome-wide levels of significance. Thus, we attempted to enhance power by following up the top 150 of the most strongly supported SNPs from the GWAS in an independent sample of 1511 cases and 1517 controls drawn from the Japanese population as well as 479 cases and 2938 controls from the United Kingdom (2). We also sought to examine whether the Japanese population shares with Europeans a polygenic component for schizophrenia and bipolar disorder using schizophrenia and bipolar case–control samples from the United Kingdom that have been previously subjected to GWAS (2,14). Because it is unlikely that stratification effects would bias the allele distributions en masse in samples ascertained in Japan in the same direction as in a European sample, confirmation of a shared polygenic effect argues strongly against the idea that residual uncontrolled stratification is responsible for the effect. Moreover, because rare alleles of large effect are expected to reflect an ongoing process of new mutation (to compensate for their removal by selection), the existence of transcontinental effects also argue against the idea that rare alleles alone can drive this effect, it being unlikely that relatively new variants would be carried on the same ancestral haplotypes in both populations.

Methods and Materials

Participants

We selected 575 patients with schizophrenia (43.5 ± 14.8 years) and 564 healthy controls (44.0 ± 14.4 years) for genome-wide association analysis (our screening GWAS: [JPN_GWAS]). All subjects were unrelated, living in the Tokai area of the mainland of Japan, and self-identified as Japanese. The details of the sample and copy number variation analysis of this GWAS data set have been reported previously (15), and see also Supplement 1.

For follow-up studies, we used an independent Japanese sample comprising 1511 cases (aged 45.9 ± 14.0 years) and 1517 controls (aged 46.0 ± 14.6 years) diagnosed and ascertained in the same way as the GWAS data set. These samples were recruited from three areas on the Japanese mainland, comprising the Kansai and Chugoku areas in addition to the Tokai area. To enhance the sample in the replication analysis, data were added from 934 Japanese controls genotyped by Illumina550 (Illumina, San Diego, California) as part of the Japanese Single Nucleotide Polymorphisms (JSNP) project (http://snp.ims.u-tokyo.ac.jp/index.html). If SNP data were available in the JSNP sample, we merged the two sample sets to form a final Japanese replication sample (we refer this as "Rep_JPN") comprising 1511 cases and 2451 controls (SNPs genotyped in both samples can be seen in Table S1 in Supplement 2).

We additionally included data from a UK schizophrenia GWAS data set of 479 cases and 2938 controls genotyped using the Affymetrix 500K array (Santa Clara, California), details of which have been reported before (2,14).

For the polygenic component analysis, we also included the Wellcome Trust Case-Control Consortium (WTCCC) bipolar disorder data set of 1868 cases and 2938 shared controls, details of which are reported elsewhere (2,14).

After complete description of the study to the subjects, written informed consent was obtained. This study was approved by the ethics committees of each university participating in this project.

GWAS and Quality Control

Genotyping was performed using the Affymetrix Genome-Wide Human SNP Array 5.0 according to the manufacturer's protocol. After applying several quality control (QC) criteria (e.g., call rate ≥ 95%, autosomal chromosomes, Hardy–Weinberg equilibrium (HWE) ≥ .0001 and minor allele frequency [MAF] ≥ 5%; Supplement 1), the final GWAS consisted of 1108 samples (560 cases and 548 controls) and 297,645 SNPs (MAF ≥ 5%).

Q-Q plots were generated on the basis of allele-wise analysis of SNPs that passed QC (Supplement 1), and our observed value of λ is consistent with those generally reported in well-matched samples ($\lambda = 1.065$ and $\lambda_{1000} = 1.117$).

Follow-Up Genotyping

Follow-up genotyping in our independent Japanese case–control sample was performed by Sequenom (San Diego, California) using the Sequenom iPLEX Gold System. Markers that could not be assayed on this platform were genotyped using a TaqMan assay (Applied Biosystems, Foster City, California).

Candidate SNPs were selected for replication as follows. First, the top 200 SNPs were identified (corresponding to $p \sim < 5 \times 10^{-4}$). Highly correlated markers based on $r^2 > .9$ to a more significant marker within 100 kb (r^2 was based on HapMap information [release Number 24, October 2008] and our own GWAS from controls) were then removed. From this list, we included the following: 1) SNPs with $p < 5 \times 10^{-5}$ ($n = 15$ after 11 redundant SNPs removed. Total number = 26. Of these, two SNPs failed for primer design. 2) Under the premise that in GWAS analysis, power favors more common alleles and that the enrichment for true associations is greater in this category of alleles (6), SNPs with MAF ≥ 10% surpassing a more relaxed threshold ($P < \sim 3.5 \times 10^{-4}$) were selected, corresponding to the top 150 SNPs ($n = 76$ after 12 low MAF SNPs and 36 redundant SNPs removed. This resulted in a total of 124. Of these, 5 SNPs failed primer design. We additionally included 13 SNPs that ranked from 151st to 200th on the grounds that they could be included in the Sequenom panels of markers without compromising the design of the higher-priority SNPs. Consequently, 97 SNPs were genotyped in the replication sample, of which 5 did not pass QC on the basis of genotype call rate (> .95) and HWE ($p > .001$). All genotype calls were confirmed by visual inspection of cluster plots.

SNP-Based Association Analysis

Consistent with most other GWAS, our study is based upon allele-wise association analysis which assumes an additive model.

Genomic control adjusted p values were also calculated based upon median chi-square statistics. This was performed using PLINK v1.07 (16).

Combined analysis across data sets (Meta_JPN: JPN_GWAS + Rep_JPN, Meta_ALL: JPN_GWAS + Rep_JPN + UK schizophrenia) were conducted using the Cochran–Mantel Haenszel (CMH) approach conditioned by sample as implemented in PLINK v. 1.07.

Polygenic Component Analysis

Discovery (for selecting "score alleles" based on association statistics) and targeting (for calculation of polygenic score) samples are summarized in Table S2 in Supplement 1. Briefly, we examined five discovery and target pairs:

1. Japanese: A set of 280 cases and 274 controls were selected for discovery, and the results were tested in an additional set of 280 cases and 274 controls. The discovery/target samples were selected at random (on the basis of random number generation) from the Japanese GWAS data set. This procedure was repeated 1000 times to ensure the results of this analysis were representative of random divisions of the data set.

2, 3. Each of the UK schizophrenia (479 schizophrenia and 2938 controls) (2) and bipolar (1868 cases and 2938 controls) (14) samples were used separately as a discovery data set to generate lists of "risk" alleles that were tested in the full Japanese GWAS sample.

4, 5. The full Japanese GWAS sample was used as a discovery data set to generate lists of "risk" alleles that were tested in the UK schizophrenia and bipolar data sets.

For the UK data sets, we used the QC criteria applied in the primary manuscripts (2,14) in which SNPs that deviated from HWE ($p < 1 \times 10^{-5}$ in cases or .001 in control) and had a low call rate (< 97%) were excluded. Note that the criteria for HWE exclusion in the UK data set is slightly different from that in the Japanese GWAS. The precise choice of HWE filter is arbitrary, but we note that both data sets criteria are on the more stringent side of customary practice.

Following the ISC (6), we reduced the set of SNPs by removing SNPs that are in linkage disequilibrium (LD) using the same criteria applied by the ISC (r^2 threshold at .25, window size 200 SNPs). In the tests of the split Japanese data set, we used LD-pruned SNPs selected on the basis of the metrics in the full set of Japanese controls. For all comparisons between Japanese and European data sets, we pruned SNPs sequentially first on the basis of the LD metrics in the discovery data set and second on those in the target data set. Polygenic score was calculated by weighting scores for "risk" alleles by the logOR observed in the discovery data set according to the method used by the ISC (6).

Nominally associated alleles were selected on the basis of the genomic-control adjusted p value in the allele-wise association analysis from the discovery samples at the following liberal significance thresholds (P_T) ($P_T < .5$, $P_T < .4$, $P_T < .3$, $P_T < .2$ and $P_T < .1$). The polygenic score was calculated using PLINK v. 1.07. Nagaelkerke's pseudo R^2 (a measure of variance explained by a particular factor) was calculated by logistic regression analysis using R (http://www.r-project.org) with covariation for "nonmissing SNPs" according to the ISC study (6).

Results

Single Marker Association Analysis

A summary plot of the GWAS (MAF ≥ 5%) is presented in Figure S1 in Supplement 1. We did not observe any associations at a widely used approximate benchmark for genome-wide significance ($p = 7.2 \times 10^{-8}$) (17). The strongest associations were observed at rs12218361, which maps to chromosome 10 at 126.06 Mb and is 3′ of *ornithine aminotransferase* (*OAT*, $p_{allele} = 6.2 \times 10^{-6}$, two-tailed), and rs11895771, which maps to chromosome 2 at 37.27 Mb within *sulfotransferase family, cytosolic, 6 B, member1* (*SULT6B1*, $p_{allele} = 8.0 \times 10^{-6}$, two-tailed). The most significant 200 markers are given in Table S1 in Supplement 2.

We genotyped 97 LD-pruned SNPs mainly from the top 150 GWAS findings in an independent Japanese replication sample (1511 cases and 1517 controls). For 22 of these, it was possible to expand the control sample size using data from the Japanese population based on the public database (JSNP). Data for 81 SNPs were also available in the UK data set (Affymetrix 500 K chip) and were included in the association analysis. On the basis of the replication sample from Japanese (Rep_JPN) alone, rs9880957 showed the most significant association ($p = 2.8 \times 10^{-3}$, two-tailed, OR = 1.2), but the associated allele was not the same as in the GWAS. Additionally, we undertook set-based analysis (using PLINK) to investigate whether there was an excess of association signals for these top GWAS findings in the replication data set that surpassed nominal p thresholds (e.g., $p < .1, .05, .01, .001$) in the Rep_JPN and UK data sets (10,000 permutation without lambda correction for all SNPs that passed the p threshold). However, no significant enrichment was observed (data not shown). That finding is compatible with the polygenic analysis we describe subsequently and with the now widely accepted hypothesis that common alleles that might be detectable in principle by GWAS exert effects that are too weak to be substantially enriched for associations that surpassed the threshold we specified for follow-up.

In the CMH analysis of the complete Japanese sample (Meta_JPN: JPN_GWAS + Rep_JPN), the best p was found at rs1011131 in LOC392288 ($p = 1.2 \times 10^{-4}$, two-tailed), which is weaker than in the initial GWAS ($p = 2.5 \times 10^{-5}$, two-tailed). Further expanding the sample size by including UK samples (Meta-ALL: JPN_GWAS + Rep–JPN + UK schizophrenia) did not provide convincing support for any locus (Table S1 in Supplement 2). The strongest association signal in Meta_ALL was rs11895771 ($p = 3.7 \times 10^{-5}$, two-tailed) in *SULT6B1*, which had been ranked second in the screening GWAS (Table 1).

Excluding *ZNF804A* (the Japanese data for which were included in the paper by O'Donovan *et al.*) (2), we additionally tested regions containing schizophrenia candidate loci supported by genome-wide significant associations in previous GWAS data sets (6–8). Specifically, we focused on three regions: the MHC region (Chr6 25 ~ 33 Mb), *NRGN*, and *TCF4*. In this analysis, we first imputed ungenotyped SNPs in these regions (boundaries ± 1 Mb) for fine mapping (the imputation method is presented in Supplement 1). None of the specific SNPs at these loci that have been reported by others (6–8) as genome-wide significant were imputable in our Japanese GWAS sample (Figures S2–S4 in Supplement 1). However, interestingly, we did observe a strong, fairly well circumscribed association signal on chromosome 6 in the region of *NOTCH4* (Figure S2 in Supplement 1). Furthermore, genetic association within *NOTCH4* has been reported (18) in another Japanese study (nonoverlapping with the present sample) at rs2071287 (Figure S2 in Supplement 1), which is in complete LD ($D' = 1$, $r^2 = .56$) with rs2071286, the best SNP tested in our GWAS data. Because that previously supported SNP (rs2071287) is also associated in our GWAS ($p = 2.1 \times 10^{-3}$), we then followed up this SNP in the Rep_JPN sample; rs2071287 was again significantly associated ($P_{allele} = .018$, two-tailed, Figure S5 in Supplement 1; note: we could not impute this SNP with high confidence in the UK schizophrenia

Table 1. Top Single Nucleotide Polymorphisms Based on GWAS and Meta-Analysis

CHR	SNP	BP	Closest Gene	Meta_ALL (JPN_GWAS+Rep_JPN+UK_SCZ)							Meta_JPN (JPN_GWAS+Rep_JPN)				JPN_GWAS		Rep_JPN		UK_SCZ	
				A1	MAF	A2	P_{CMH}	OR[a]	L95	U95	P_{CMH}	OR[a]	L95	U95	P_{allele}	OR[a]	P_{allele}	OR[a]	P_{allele}	OR[a]
2	rs11895771	37266439	*SULT6B1*	T	.49	G	3.7×10^{-5}	.84	.77	.91	4.1×10^{-4}	.84	.76	.92	8.0×10^{-6}	.64	.14	.92	.033	.84
7	rs1011131	19474460	*LOC392288*	G	.07	C	1.2×10^{-4}	1.30	1.14	1.48	1.2×10^{-4}	1.31	1.14	1.50	2.5×10^{-5}	1.78	.054	1.17	.63	1.14
14	rs1176970	40505514	*LOC644919*	G	.15	C	1.4×10^{-4}	1.22	1.10	1.35	3.0×10^{-4}	1.27	1.12	1.44	3.2×10^{-4}	1.58	.041	1.17	.14	1.14
1	rs4908274	103162502	*COL11A1*	A	.28	T	3.1×10^{-4}	1.20	1.09	1.32	3.1×10^{-4}	1.20	1.09	1.32	1.1×10^{-4}	1.45	.067	1.12	NA	NA
6	rs2294424	11860537	*C6orf105*	T	.41	C	5.0×10^{-4}	1.15	1.06	1.24	5.0×10^{-4}	1.17	1.28	1.07	1.2×10^{-4}	1.40	.081	1.1	.41	1.08
2	rs13010889	40617519		A	.15	C	.0011	.85	.77	.94	.0016	.85	.77	.94	8.7×10^{-5}	.67	.17	.92	.40	.84
2	rs17026152	40611159		A	.26	G	.0012	.85	.77	.94	.0012	.85	.77	.94	1.3×10^{-4}	.69	.15	.92	NA	NA
6	rs2787566	101985455	*GRIK2*	A	.04	G	.0014	1.34	1.12	1.61	.0014	1.39	1.1	1.7	2.8×10^{-4}	2.03	.15	1.19	.49	1.16
6	rs2071286	32287874	*NOTCH4*	T	.19	C	.0014[b]	.87	.79	.95	.0049[b]	.86	.78	.96	3.3×10^{-4}	.68	.23[b]	.93	.13	.87
8	rs17462248	29426926		G	.2	T	.0017	1.16	1.06	1.27	.020	1.14	1.0	1.3	2.1×10^{-4}	1.52	.60	1.04	.030	1.2

p values were calculated on the basis of the allele-wise test (two-tailed).
A1, minor allele based on whole sample; A2, major allele based on whole sample; BP, base position; CHR, chromosome(hg18); GWAS, genome-wide association study JPN_GWAS: our screening GWAS; L95, lower bound of 95% confidence interval for OR; MAF, minor allele frequency based on whole sample; NA, not analyzed; OR, odds ratio; SNP, single nucleotide polymorphism; U95, upper bound of 95% confidence interval for odds ratio, UK-SCZ: UK schizophrenia.
[a]OR was calculated on the basis of A1 in Meta-ALL as reference.
[b]Controls from Japanese SNPs (JSNP) were merged into the replication sample.

data set because of the high missing rate of 12%). Next we conducted a meta-analysis based on Meta_JPN (imputed data from JPN_GWAS was down-weighted using PROPER-INFO from SNPTEST by METAL: http://www.sph.umich.edu/csg/abecasis/metal/) and the sample of Tochigi (18). This provided fairly strong evidence for association ($P_{meta} = 5.1 \times 10^{-5}$, two-tailed, Figure S5 in Supplement 1).

Polygenic Component Analysis

p values and pseudo-R^2 statistics (Nagaelkerke's R^2) for the analysis based on the split Japanese sample are presented in Figure 1 and in Table S3 in Supplement 1. The polygenic scores in the target data were higher in the cases than the controls and, in most cases, significantly so. As in the ISC study, the evidence became stronger and the pseudo-R^2 improved at more liberal P_T values. The most significant correlation was found at $P_T < .5$ ($p = .005$). In this condition, the pseudo-R^2 was slightly lower ($R^2 = .021$) compared with the ISC study (6) in which $R^2 \leq .032$ were observed in the Caucasian samples (Figure 1), although we note that the ISC study used information from a greater number of SNPs, with the larger sample available to that group allowing the inclusion of SNPs with MAF as low as 2%.

The results of the analysis based on discovery in the UK schizophrenia data set and targeting the JPN_GWAS are shown in Figure 2 (Table S3 in Supplement 1). Again, as with the ISC data, the signal and predictive power improved at the more liberal thresholds, with only the most relaxed threshold (corresponding to the optimal threshold from the ISC study) attaining significance ($p = .029$). However, the analysis using the WTCCC bipolar sample for discovery and the Japanese as the target did not reveal significant support for shared risk across disorders (Figure 2 and Table S3 in Supplement 1).

Following are the results of the analyses based on discovery in the JPN_GWAS and testing in the UK schizophrenia and bipolar data sets. Alleles trained in this direction were highly significant, but weakly predictive, of schizophrenia status in the UK sample ($p_{min} = 7.0 \times 10^{-5}$) than those analyses based on training in the UK data sets. Again, no significant effect was observed for bipolar disorder. In the schizophrenia analysis, we observed no clear relationship

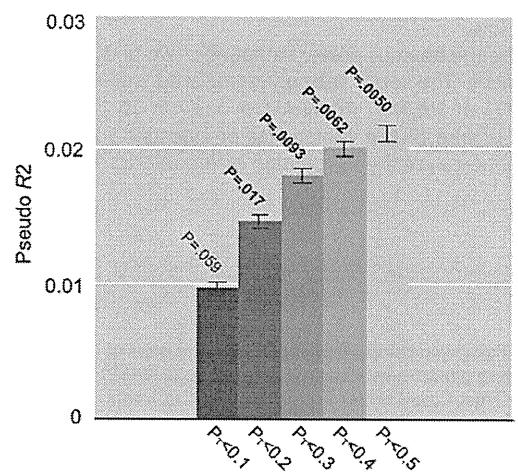

Figure 1. Polygenic component analysis for the pair within screening genome-wide association studies samples. $p_T = p$ threshold. Pseudo R^2 and *p* values represent the mean and median values, respectively, from 1000 random divisions of the data set. Error bars represent the 95% confidence intervals for R^2 from those repeat analyses. Bold numbers represent significant *p* values (< .05).

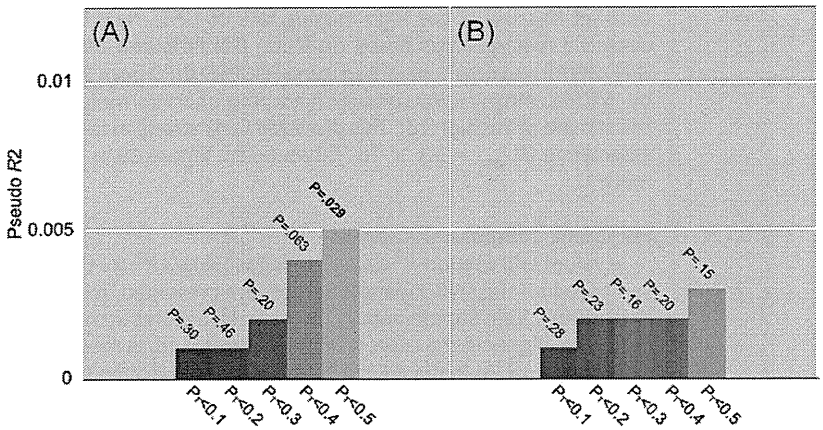

Figure 2. Polygenic component analysis for the pairs of Wellcome Trust Case-Control Consortium (WTCCC) data sets/screening genome-wide association studies (GWAS). **(A)** UK schizophrenia/screening GWAS discovery/target pair. **(B)** WTCCC bipolar/screening GWAS discovery/target pair. $p_T = p$ threshold. Bold numbers represent significant p values (< .05).

between the test allele significance threshold (P_T) and either the statistical support or the pseudo-R^2 (Figure 3 and Table S3 in Supplement 1).

Discussion

In this study, we did not detect unequivocal evidence for a novel susceptibility gene for schizophrenia, although our results do provide weak support for association between *SULT6B1* and schizophrenia, and our analyses of previously implicated regions and candidate genes provide support for the hypothesis that previous findings at the MHC region of chromosome 6 may point to *NOTCH4*. The absence of association at genome-wide levels of significance is not surprising given the relatively small size of our GWAS. Recent large-scale GWAS of schizophrenia suggest that the effect sizes of common risk alleles are small (ORs < 1.25). Power analysis suggests that our GWAS has only .18% power under an additive model to detect at $\alpha = 7.2 \times 10^{-8}$, a susceptibility variant with an allele frequency of .3 conferring an OR of 1.25. Clearly, with power like this, it would be extremely unlikely that any one locus would be detected at strong levels of support; however, in the presence of a thousand or more loci as has been suggested (6), the power to detect at least one of these would be considerably greater, albeit the subsequent power to replicate that specific locus would once again be low.

Despite the obvious power limitations, two findings are worthy of comment. The most strongly associated individual SNP was rs11895771 at *SULT6B1* (Meta-ALL $p = 3.7 \times 10^{-5}$). *SULT6B1* is a member of one of the subfamilies of cytosolic sulfotransferases (SULT) that catalyze the sulfonation of xenobiotics, hormones, and neurotransmitters, including 17β-estradiol and corticosterone (19), functions that are at least plausibly related to schizophrenia (20–22), and brain function (23–25) more widely.

The second locus of interest was *NOTCH4*. *NOTCH4* has been reported to be associated with schizophrenia in a small UK sample (26) (not overlapping with the present sample), but replication data from candidate gene studies have not been strongly supportive. However, a recent synthesis of GWASs as well as a large number of additional subjects reported a genome-wide significant association at rs3131296 (8), which is located within *NOTCH4* (Figure S2 in Supplement 1), although the extensive LD across the MHC region makes pinpointing the source of that signal to a specific gene impossible. It is therefore of interest in our evaluation of the MHC region that the signal clearly maximized to the *NOTCH4* region (Figure S2 in Supplement 1), lending support to the hypothesis that this may be the relevant susceptibility gene in the region. We are unable to evaluate the specific SNP (rs3131296) reported in the SGENE study for the Japanese population because of the failure of imputation. In the Japanese population, the MAF of rs3131296 differs considerably from that in Europeans (MAF = 10% and 2.3% for CEU and JPT populations, respectively, in HapMap Phase 3 data, 13% reported in SGENE), which means the ability of this marker to tag a common functional variant is likely to differ significantly between populations. Given the evidence for association observed in our study and the prior genetic evidence for *NOTCH4*, this locus warrants further detailed analysis in larger and more ethnically diverse samples.

This study provides the first independent (of the samples used by the ISC) replication of the polygenic score analysis reported by

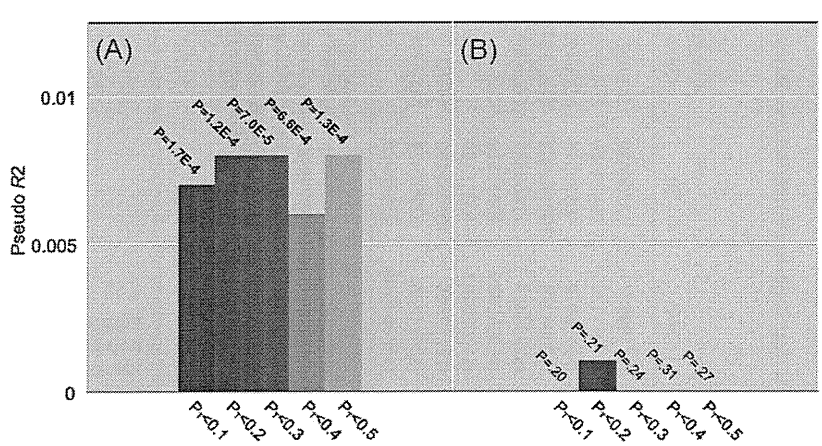

Figure 3. Polygenic component analysis for the pairs of the screening genome-wide association studies (GWAS)/ Wellcome Trust Case-Control Consortium (WTCCC) data sets. **(A)** Screening GWAS/UK schizophrenia discovery/ target pair. **(B)** Screening GWAS/ WTCCC bipolar discovery/target pair. $p_T = p$ threshold. Bold numbers represent significant p values (< .05).

the ISC (6). Although our sample is low powered (power is .6 for our full sample and .56 for half of the sample to detect at an alpha level of .5, a weak genetic effect [OR 1.1] conferred by an allele with a frequency of .3), the set of "risk" alleles (in quotation marks to emphasize that most are not likely to be true risk alleles) derived from half of the Japanese sample was significantly correlated with affection status in the other half of the samples. One possible important confounding factor to consider is an effect of population stratification. To check for this as a possible effect, we used 1) principal components analysis–adjusted (the first 10 principal components) discovery statistics for the selection of SNPs and 2) the first 10 principal component vectors as covariates in calculating the polygenic score in the target sample. However, the application of either or both of these did not lead to a material difference in the results (Table S4 in Supplement 1), indicating that stratification is not likely to explain our replication of the ISC findings.

Our Japan–UK analyses also suggests this effect is unlikely to be due to stratification (this was also convincingly argued in the ISC study) because the Japanese and UK schizophrenia samples are ascertained directionally for the same stratification biases and because the UK schizophrenia sample, but not the UK bipolar sample, would be unlikely to be stratified in that manner. Instead, those data point to a shared genetic component to schizophrenia susceptibility across major ethnic groups, as predicted by an effect driven by common "risk" alleles rather than rare alleles, although not excluding an effect of rare alleles, which are much more likely to reside on different haplotype backgrounds in different populations. However, there is also evidence for population differences in risk. Thus, the analyses restricted to the Japanese population showed much higher maximal estimates for R^2 (.021) compared with the analyses of schizophrenia between populations ($R^2 = .005 \sim .008$) and was more similar to the estimates of R^2 when the analyses were performed within European populations (6). The ISC also undertook one cross-population analysis, between Caucasian and African Americans. As in our study, R^2 was much lower between the ethnic groups (.004) than within the European populations. These results suggest that although at least some "risk" alleles are shared across populations, there are also differences in those "risk" alleles or at least in the extent to which they are tagged by markers at the density currently provided by the arrays we have studied. At a practical level, this means that failures to replicate findings across ethnic groups, even with respect to common alleles, should be treated with considerable caution.

One intriguing finding was our failure to find evidence that "risk" alleles for bipolar disorder in the European sample predict risk of schizophrenia in the Japanese sample (or vice versa). One likely explanation is that there is only a partial overlap between "risk" alleles for schizophrenia and bipolar disorder and that this, together with the additionally reduced R^2 because of ethnic differences, has affected our ability to demonstrate an effect. This interpretation is at least partially consistent with the ISC study in which the measures of R^2 that were observed in bipolar data sets were less than those observed in the schizophrenia data sets. A more interesting but speculative interpretation is that the Japanese sample represents a phenotypically purer form of schizophrenia than the European samples. These hypotheses require further evaluation in larger Japanese samples, exploration of aspects of the schizophrenia phenotype in the European samples, and transdiagnostic polygenic score analyses within Japanese samples.

This work was supported in part by research grants from the Japan Ministry of Education, Culture, Sports, Science and Technology; the Ministry of Health, Labor and Welfare; the Core Research for Evolutional Science and Technology; and the Health Sciences Foundation (Research on Health Sciences focusing on Drug Innovation). The UK research was supported by grants from the Medical Research Council and the Wellcome Trust.

The authors MCO and NI are joint last authors.

Dr. Ikeda reports receiving support from the Japan Society for the Promotion of Science postdoctoral fellowship for research abroad and is also supported by the Uehara Memorial Foundation and the Great Britain Sasakawa Foundation. The other authors report no biomedical financial interests or potential conflicts of interest.

Supplementary material cited in this article is available online.

1. O'Donovan MC, Craddock NJ, Owen MJ (2009): Genetics of psychosis; insights from views across the genome. *Hum Genet* 126:3–12.
2. O'Donovan MC, Craddock N, Norton N, Williams H, Peirce T, Moskvina V, *et al.* (2008): Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet* 40:1053–1055.
3. Steinberg S, Mors O, Borglum AD, Gustafsson O, Werge T, Mortensen PB, *et al.* (2010): Expanding the range of ZNF804A variants conferring risk of psychosis [published online ahead of print January 5]. *Mol Psychiatry.*
4. Riley B, Thiselton D, Maher BS, Bigdeli T, Wormley B, McMichael GO, *et al.* (2010): Replication of association between schizophrenia and ZNF804A in the Irish Case–Control Study of Schizophrenia sample. *Mol Psychiatry* 15:29–37.
5. Williams HJ, Norton N, Dwyer S, Moskvina V, Nikolov I, Carroll L, *et al.* (2010): Fine mapping of ZNF804A and genome-wide significant evidence for its involvement in schizophrenia and bipolar disorder [published online ahead of print April 6].
6. International Schizophrenia Consortium, 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.
7. Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I, *et al.* (2009): Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 460:753–757.
8. Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D, *et al.* (2009): Common variants conferring risk of schizophrenia. *Nature* 460:744–747.
9. Wray NR, Visscher PM (2009): Narrowing the boundaries of the genetic architecture of schizophrenia. *Schizophr Bull* 36:14–23.
10. Evans DM, Visscher PM, Wray NR (2009): Harnessing the information contained within genome-wide association studies to improve individual prediction of complex disease risk. *Hum Mol Genet* 18:3525–3531.
11. Craddock N, O'Donovan MC, Owen MJ (2009): Psychosis genetics: Modeling the relationship between schizophrenia, bipolar disorder, and mixed (or "schizoaffective") psychoses. *Schizophr Bull* 35:482–490.
12. Craddock N, O'Donovan MC, Owen MJ (2007): Symptom dimensions and the Kraepelinian dichotomy. *Br J Psychiatry* 190:361; author reply: 361–362.
13. Yamaguchi-Kabata Y, Nakazono K, Takahashi A, Saito S, Hosono N, Kubo M, *et al.* (2008): 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* 83:445–456.
14. Wellcome Trust Case-Control Consortium (2007): Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447: 661–678.
15. Ikeda M, Aleksic B, Kirov G, Kinoshita Y, Yamanouchi Y, Kitajima T, *et al.* (2009): Copy number variation in schizophrenia in the Japanese population. *Biol Psychiatry* 67:283–286.
16. 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.
17. Dudbridge F, Gusnanto A (2008): Estimation of significance thresholds for genomewide association scans. *Genet Epidemiol* 32:227–234.
18. Tochigi M, Zhang X, Ohashi J, Hibino H, Otowa T, Rogers M, *et al.* (2007): Association study between the TNXB locus and schizophrenia in a Japanese population. *Am J Med Genet B Neuropsychiatr Genet* 144B: 305–309.

19. Cao H, Agarwal SK, Burnside J (1999): Cloning and expression of a novel chicken sulfotransferase cDNA regulated by GH. *J Endocrinol* 160:491–500.
20. Eaton WW, Byrne M, Ewald H, Mors O, Chen CY, Agerbo E, Mortensen PB (2006): Association of schizophrenia and autoimmune diseases: Linkage of Danish national registers. *Am J Psychiatry* 163:521–528.
21. Baumgartner A, Pietzcker A, Gaebel W (2000): The hypothalamic-pituitary-thyroid axis in patients with schizophrenia. *Schizophr Res* 44:233–243.
22. Kulkarni J, de Castella A, Fitzgerald PB, Gurvich CT, Bailey M, Bartholomeusz C, Burger H (2008): Estrogen in severe mental illness: A potential new treatment approach. *Arch Gen Psychiatry* 65:955–960.
23. Cerqueira JJ, Pego JM, Taipa R, Bessa JM, Almeida OF, Sousa N (2005): Morphological correlates of corticosteroid-induced changes in prefrontal cortex-dependent behaviors. *J Neurosci* 25:7792–7800.
24. Enkel T, Koch M (2009): Chronic corticosterone treatment impairs trace conditioning in rats with a neonatal medial prefrontal cortex lesion. *Behav Brain Res* 203:173–179.
25. Gogos A, Nathan PJ, Guille V, Croft RJ, van den Buuse M (2006): Estrogen prevents 5-HT1A receptor-induced disruptions of prepulse inhibition in healthy women. *Neuropsychopharmacology* 31:885–889.
26. Wei J, Hemmings GP (2000): The NOTCH4 locus is associated with susceptibility to schizophrenia. *Nat Genet* 25:376–377.

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.

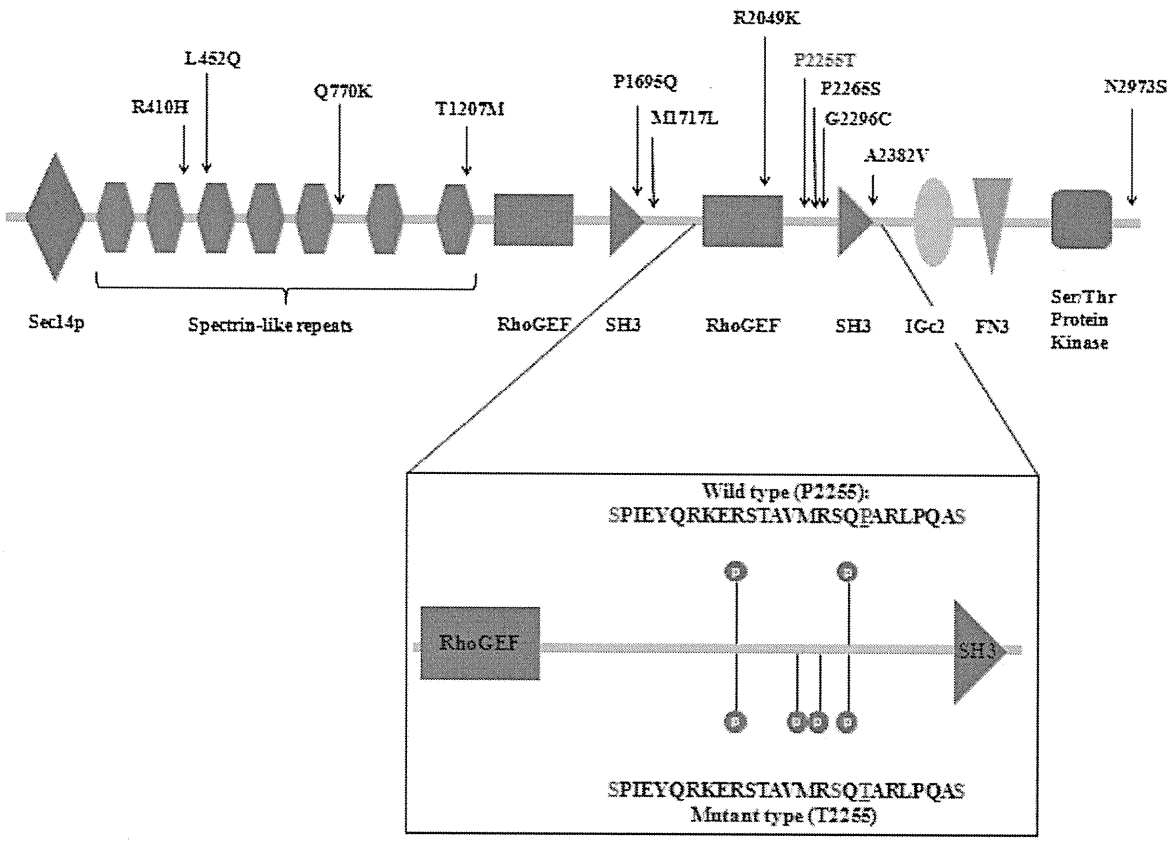

Fig. 1. Rare Missense Mutations in *KALRN* and Change in Phosphorylation Status by P2255T.

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.

OPEN ACCESS Freely available online

PLoS one

Pallidal Hyperdopaminergic Innervation Underlying D2 Receptor-Dependent Behavioral Deficits in the Schizophrenia Animal Model Established by EGF

Hidekazu Sotoyama, Yingjun Zheng, Yuriko Iwakura, Makoto Mizuno, Miho Aizawa, Ksenia Shcherbakova, Ran Wang, Hisaaki Namba, Hiroyuki Nawa*

Department of Molecular Neurobiology, Brain Research Institute, Niigata University, Niigata, Japan

Abstract

Epidermal growth factor (EGF) is one of the ErbB receptor ligands implicated in schizophrenia neuropathology as well as in dopaminergic development. Based on the immune inflammatory hypothesis for schizophrenia, neonatal rats are exposed to this cytokine and later develop neurobehavioral abnormality such as prepulse inhibition (PPI) deficit. Here we found that the EGF-treated rats exhibited persistent increases in tyrosine hydroxylase levels and dopamine content in the globus pallidus. Furthermore, pallidal dopamine release was elevated in EGF-treated rats, but normalized by subchronic treatment with risperidone concomitant with amelioration of their PPI deficits. To evaluate pathophysiologic roles of the dopamine abnormality, we administered reserpine bilaterally to the globus pallidus to reduce the local dopamine pool. Reserpine infusion ameliorated PPI deficits of EGF-treated rats without apparent aversive effects on locomotor activity in these rats. We also administered dopamine D1-like and D2-like receptor antagonists (SCH23390 and raclopride) and a D2-like receptor agonist (quinpirole) to the globus pallidus and measured PPI and bar-hang latencies. Raclopride (0.5 and 2.0 μg/site) significantly elevated PPI levels of EGF-treated rats, but SCH23390 (0.5 and 2.0 μg/site) had no effect. The higher dose of raclopride induced catalepsy-like changes in control animals but not in EGF-treated rats. Conversely, local quinpirole administration to EGF-untreated control rats induced PPI deficits and anti-cataleptic behaviors, confirming the pathophysiologic role of the pallidal hyperdopaminergic state. These findings suggest that the pallidal dopaminergic innervation is vulnerable to circulating EGF at perinatal and/or neonatal stages and has strong impact on the D2-like receptor-dependent behavioral deficits relevant to schizophrenia.

Citation: Sotoyama H, Zheng Y, Iwakura Y, Mizuno M, Aizawa M, et al. (2011) Pallidal Hyperdopaminergic Innervation Underlying D2 Receptor-Dependent Behavioral Deficits in the Schizophrenia Animal Model Established by EGF. PLoS ONE 6(10): e25831. doi:10.1371/journal.pone.0025831

Editor: Kazutaka Ikeda, Tokyo Metropolitan Institute of Medical Science, Japan

Received August 3, 2011; **Accepted** September 12, 2011; **Published** October 12, 2011

Funding: This study was supported by a Grant-in-Aid for Basic Scientific Research (B), Health and Labor Sciences Research Grants, Core Research for Evolutional Science and Technology from the JST Agency, Takeda Science Foundation, and grants for Promotion of Niigata University Research Projects. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: hnawa@bri.niigata-u.ac.jp

Introduction

Epidermal growth factor (EGF) and structurally related EGF-like peptides (such as neuregulin-1; NRG1) regulate GABAergic and dopaminergic development [1–5]. Genetic studies suggest that the genomic mutation or polymorphism for EGF, NRG1 and their receptors (ErbBs) is associated with schizophrenia risk [6–9]. Changes in the expression levels of EGF, NRG1 and ErbBs are also found in postmortem brains and peripheral blood of schizophrenia patients [10–13]. Both EGF and NRG1 are known to influence their own receptor binding and provoke signal cross-talks of other ErbB subtype(s) via ErbB1-4 receptor heterodimirization [14], [15]. Accordingly, abnormal ErbB signaling might be one of key features in schizophrenia neuropathology and/or etiology, although the pathophysiological nature of ErbB signaling in schizophrenia is largely unresolved [16], [17].

Among many environmental factors implicated in schizophrenia etiology, maternal viral infection and obstetric complications are suggested to play an important role in regulating vulnerability to schizophrenia [18–21]. These inflammatory insults often evoke abnormal cytokine signaling and perturb normal brain development [18], [21]. For instance, EGF levels in the amniotic fluid can sometimes reach the order of a microgram per liter, which is sufficient to occupy almost all EGF receptors (ErbB1) in the human fetus and lead to unfavorable uterine contractions and pre-term labor [22], [23]. To evaluate the impact of prenatal and perinatal ErbB hyper–signaling on neurobehavioral development, various ErbB ligands such as EGF and NRG1 were subcutaneously administered to rodent pups as their neurodevelopmental period matches the second trimester of human fetus having immature blood–brain barrier and initiating glial proliferation [24]. We found that ErbB ligands can penetrate the immature blood–brain barrier and reach brain neurons, resulting in various behavioral impairments at the post-pubertal stage [25]. For example, rats challenged with EGF as neonates display behavioral abnormalities in acoustic prepulse inhibition (PPI), latent inhibition of learning, social interaction, and methamphetamine sensitivity [26], [27]. The magnitude of the behavioral deficits, however, depends upon the type of ErbB ligands administered and the genetic background of the host animals [4], [27], [28]. In rodents, therefore, abnormal ErbB signaling in the prenatal and/or perinatal stage results in the neurobehavioral deficits [26–28] and/or dopaminergic abnormal-

ities relevant to schizophrenia [4], although the neuropathologic mechanism underlying the individual deficits remains to be clarified.

Here, we prepared the animal model for schizophrenia by subcutaneously injecting EGF to newborn rats and studied the mechanism for their PPI abnormality. Based upon our latest finding on the neurotrophic interactions between EGF signaling and dopamine [5], [29], [30], we mainly characterized dopaminergic neuropathology of these rats using neurochemical and anatomical approaches. Furthermore we explored the pathologic mechanism underlying their behavioral deficits by pharmacologically manipulating local dopamine transmission.

Results

Upregulation of dopaminergic markers in the globus pallidus continuing until adulthood

At the early postnatal stage of rats, EGF is verified to reach the brain through the immature blood-brain barrier and promote phenotypic development of midbrain dopaminergic neurons, leading to neurobehavioral abnormalities relevant to schizophrenia [26], [31]. However, there was a question of whether the neurotrophic influences on dopaminergic neurons can continue until the post-pubertal stage when rats develop the neurobehavioral abnormalities [26]. To address this question, here, we quantitated and compared protein levels of tyrosine hydroxylase (TH; a rate-limiting enzyme for dopamine synthesis) in various brain regions of EGF-treated and control rats at their adult stage using enzyme immunoassay (ELISA) [32], [33] (Fig. 1A). In the globus pallidus, TH protein levels were significantly higher in EGF-treated rats than in control rats ($P = 0.035$). There were no significant differences in the other brain regions in this ELISA (Fig. 1A). In the globus pallidus, there was concomitant elevation of tissue dopamine content ($P = 0.030$) and its metabolites, DOPAC ($P = 0.014$) (Fig. 1B, C) and HVA ($P = 0.019$, data not shown). We confirmed the pallidal hyperdopaminergic state by immunoblotting; protein levels of TH and VMAT2, but not those of DβH (a maker for noradrenergic neurons), were significantly elevated in EGF-treated rats at the adult stage (Fig. 1D).

To determine whether the TH increase in the globus pallidus accompanies morphological alterations, we investigated the neuroanatomical features of dopaminergic fibers and terminals in EGF-treated rats (Fig. 2). In contrast to the neurochemical alterations, the anatomical difference in TH immunoreactivity was less marked. The density of TH-immunoreactive fibers was higher in EGF-treated rats only in the lateral area of medial and caudal globus pallidus neighboring the striatum. In rostral globus pallidus, however, we failed to detect the difference (data not shown).

Antipsychotic effects on extracellular dopamine levels in the globus pallidus of EGF-treated rats

Sensorimotor gating involves pallidal function and dopaminergic transmission [34]. We next monitored extracellular dopamine levels in the globus pallidus using a microdialysis technique and estimated a link between local dopamine concentration and sensorimotor gating (Fig. 3A). In agreement with the previous results [26], neonatal EGF treatment reduced PPI scores at the adult stage and an atypical antipsychotic agent ameliorated the PPI reduction (Fig. 3B). EGF-treated animals also exhibited an increase in pulse alone-startle amplitude compared with control animals but not with risperidone-treated animals [$F(2,33) = 3.61$, $P = 0.038$] (Fig. 3C). We found that basal extracellular dopamine levels (1.52 ± 0.19 nM) in pallidal dialysates from EGF-treated rats were significantly elevated than in those from controls

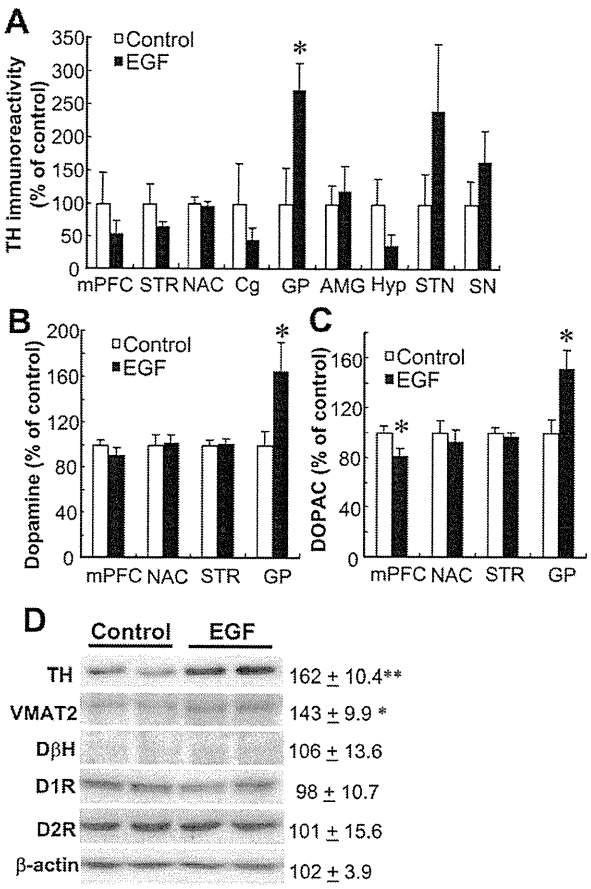

Figure 1. Effects of neonatal EGF challenge on tyrosine hydroxylase and dopamine metabolism at the adult stage. EGF or cytochrome c (control) was daily injected (s.c.) into neonatal rats from PND2 to PND10, and rats were bred until adulthood (PND 56). (A) Levels of tyrosine hydroxylase (TH) immunoreactivity in brain tissue homogenates were determined by ELISA and presented as a ratio of control levels (mean ± SEM). Tissue contents of dopamine (B) and its metabolite DOPAC (C) were measured by HPLC. (D) Tissue lysates from the globus pallidus of EGF rats and controls (8 weeks old, N = 4) were subjected to immunoblotting for antibodies directed against TH, vesicular monoamine transporter 2 (VMAT2), dopamine beta hydroxylase (DβH), dopamine D1 receptor (D1R), dopamine D2 receptor (D2R), and β-actin (an internal control). Immunoreactivity was measured by densitometric analysis and its percentage ratio to that of control rats was calculated (mean ± SEM). Abbreviations; medial prefrontal cortex (prelimbic cortex; mPFC), striatum (STR), nucleus accumbens (NAC), cingulate cortex (Cg), globus pallidus (GP), amygdala (AMG), hypothalamus (Hyp), subthalamic nucleus (STN), and substantia nigra (SN). *P<0.05, **P<0.01 by two-tailed t-test.
doi:10.1371/journal.pone.0025831.g001

(0.85 ± 0.06 nM; $P < 0.001$, $N = 11$–13) but decreased by subchronic treatment with risperidone (0.67 ± 0.08 nM; $P < 0.001$, $N = 11$) [$F(2,33) = 14.8$, $P < 0.001$] (Fig. 3D). Following potassium depolarization stimuli, extracellular dopamine levels were also elevated in EGF-treated rats but not in risperidone-administered EGF-treated rats [$F(2,33) = 5.07$, $P = 0.008$, $N = 11$–13]. The release difference at both the basal and evoked states suggests that EGF altered the capacity of net dopamine release but not the excitability of dopamine terminals. Of note, the overall changes in extracellular dopamine levels exhibited an opposite trend to that of the changes in PPI levels in the above experiments. These

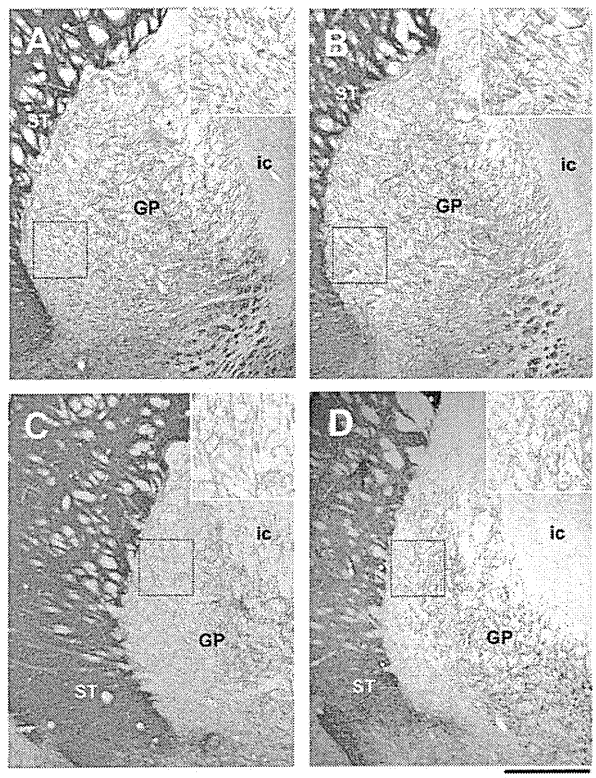

Figure 2. Immunohistochemical analysis of tyrosine hydroxylase-positive fibers and terminals in the globus pallidus. Coronal sections of the striatum containing medial and caudal globus pallidus (1.2 and 1.6 mm posterior from the bregma) were prepared from control and EGF-treated rats and immunostained with anti-TH antibody (N = 3–4 rats per group). TH-immunoreactive fibers in medial (A, B) and caudal (C, D) globus pallidus of control (A, C) and EGF rats (B, D) are shown. A lateral area of the globus pallidus is marked with a window, enlarged 2-fold, and presented in the top-right corner of each picture. ST; striatum, GP; globus pallidus, ic; internal capsule. Scale bar, 500 μm.
doi:10.1371/journal.pone.0025831.g002

results raise the hypothesis that enhanced pallidal dopamine release might be associated with the PPI decrease in this model.

Local reserpine administration to the globus pallidus ameliorates prepulse inhibition deficits

To confirm our hypothesis, we administered reserpine (an inhibitor of VMAT) or vehicle to reduce local dopamine content in synaptic vesicles and measured prepulse inhibition (Fig. 4A). A three-way ANOVA with EGF treatment and reserpine challenge as the between-subjects factors and prepulse intensity as the within-subjects factor revealed a significant interaction between EGF treatment and reserpine challenge [$F_{(1,29)} = 7.81$, $p = 0.009$]. Repeated Fisher's LSD revealed that reserpine challenge significantly improved PPI scores in EGF rats ($P = 0.012$). In contrast, reserpine did not alter PPI scores of control rats or pulse-alone startles of both groups of rats. Following the behavior tests, we monitored local dopamine contents to ascertain the regional specificity of the reserpine action (Fig. 4B). The dose of reserpine reduced the dopamine pool in the globus pallidus of EGF rats ($P<0.001$) but not in the neighboring striatum of EGF rats or in these regions of control rats.

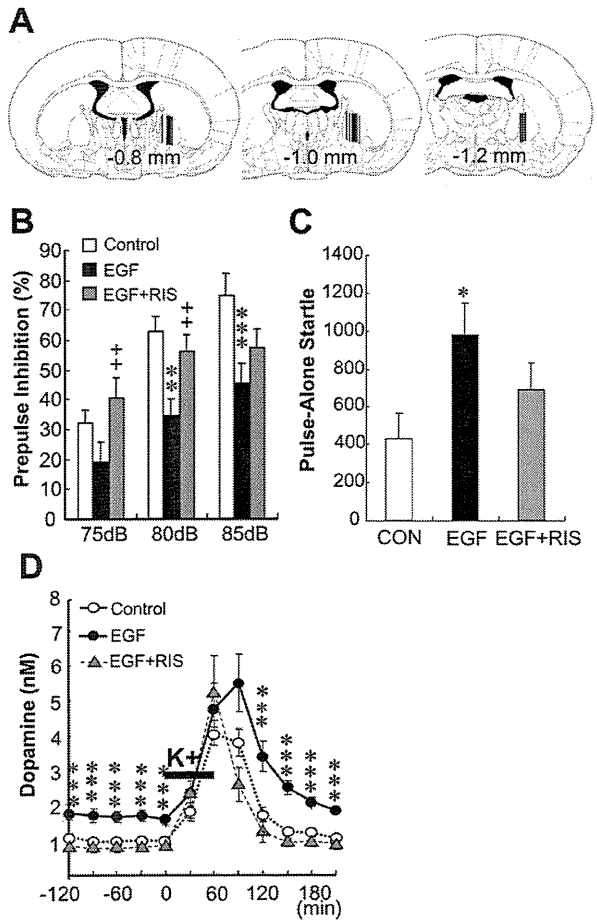

Figure 3. Pallidal dopamine release enhanced in EGF-treated rats and normalized with risperidone. Neonatal rats were treated with EGF or cytochrome c (control) as described in Figure 1. Risperidone (1 mg/kg, i.p.) was administered to some of the EGF-treated rats for 14 days at the adult stage. (A) The location of dialysis probe was examined and is shown in rat brain atlas. Six rats were excluded for incorrect probe placement. The digit represents the distance from the bregma. (B) PPI levels were monitored in control, EGF-treated and EGF+risperidone-treated animals following microdialysis. (C) Pulse-alone startle responses and PPI levels were monitored in control, EGF-treated and EGF+risperidone-treated animals following microdialysis. (D) Basal concentrations of dopamine in dialysates were monitored for 150 min, dopamine release was evoked by perfusion of 80 mM KCl over 60 min (solid bar), and then monitored over 150 min. Data represent dopamine concentrations in 30-min fractions (nM, mean ± SEM, N = 11–13 rats per group). There was a significant interaction between time and dopamine release [$F(12,198) = 2.72$, $P = 0.002$]. *$P<0.05$, ***$P<0.001$, compared with controls and ++$P<0.01$, compared with EGF-treated rats by Fisher's LSD.
doi:10.1371/journal.pone.0025831.g003

There was a significant difference in pulse-alone startle amplitude between EGF-treated and control rats [$F_{(1,28)} = 16.4$, $P<0.001$ for EGF]. Our current design to evaluate reserpine action on %PPI might be inappropriate due to the basal difference in pulse-alone startle amplitudes of EGF and control rats [35]. We prepared another set of animals and gave the lower intensity of tone stimuli (110 dB for main pulses) to EGF-treated rats (Fig. 4C, D). The PPI difference between EGF and control groups as well as the risperidone effect still remained, even though there was no significant difference in the amplitudes of pulse-alone startles among groups.

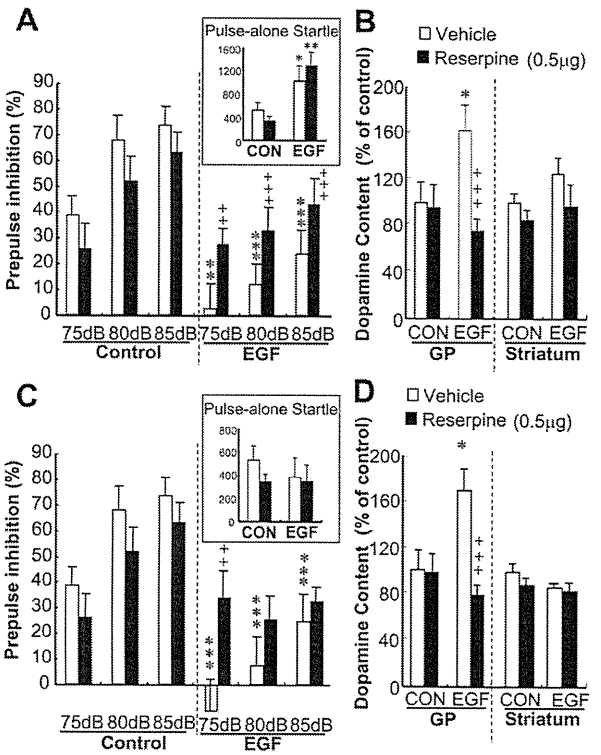

Figure 4. Effects of pallidal reserpine infusion on prepulse inhibition. (A) Startle responses of EGF and control (CON) rats were triggered with 120-dB tone pulse and PPI scores were measured with 75-, 80-, and 85-dB prepulse stimuli 120 min after local vehicle or reserpine infusion to both hemispheres of the globus pallidus. Pulse-alone startle responses to a 120-dB tone were measured in arbitrary units and are shown in the inset. (C) To match pulse-alone startle responses between EGF and control rats, 110-dB and 120-dB pulses were given to EGF and control rats, respectively, as shown in the insert. Following PPI test in (A) and (C), dopamine content was measured in the globus pallidus (GP) and neighboring striatum and shown in (B) and (D), respectively. Bars indicate mean ± SEM (N = 8–9 each). *P<0.05, **P<0.01, ***P<0.001, compared with vehicle-infused control rats; ++P<0.01, +++P<0.001, compared with vehicle-infused EGF rats by Fisher's LSD.
doi:10.1371/journal.pone.0025831.g004

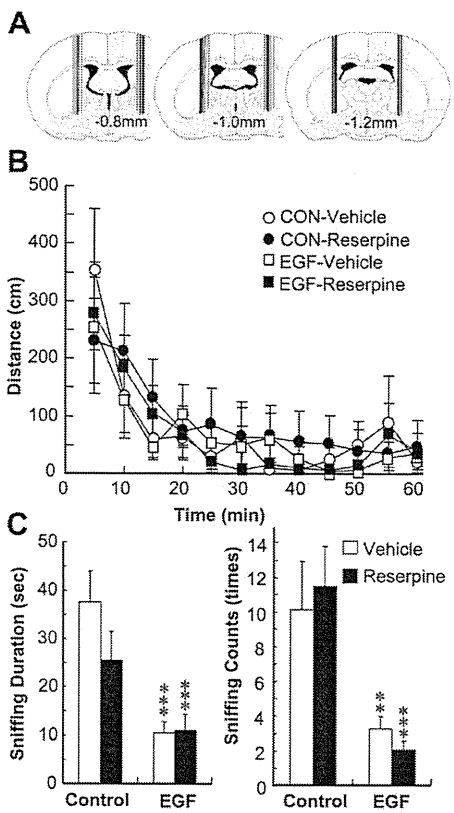

Figure 5. Influences of pallidal reserpine infusion on locomotor activity and social interaction. EGF and control (CON) rats received local reserpine- or vehicle-infusion to both hemispheres of the GP. (A) Cannula placement was confirmed in fixed brains, and two out of 36 rats were excluded from the final data analysis due to incorrect cannula placement. (B) Two hours after pallidal infusion, rats were placed in the automated activity monitoring chamber for 60 min. Data represent horizontal movement (cm) for every 5 min (mean ± SEM, N = 7–9 for each group). (C) Following locomotor test, an unfamiliar male rat was placed in the same chamber. The number and duration of sniffing behaviors of operated rats were counted for 10 min. Bars indicate mean ± SEM (N = 8–9 for each group). **p<0.01, ***p<0.001 by Fisher's LSD, compared with vehicle-infused controls.
doi:10.1371/journal.pone.0025831.g005

To control the potential aversive effects of reserpine injection, we assessed the exploratory motor behaviors of reserpine-injected animals (Fig. 5A, B). We found no significant influence of reserpine on horizontal movement. As EGF-treated rats are known to exhibit a decrease in social interaction scores [26], we also evaluated the effects of reserpine on social behaviors following the above exploratory locomotor test (Fig. 5C). Two-way ANOVA for sniffing duration revealed a significant main effect of EGF [$F(1,30) = 17.8$, $P<0.001$] but no main effect of reserpine or no interaction between EGF and reserpine. The same statistical conclusion was drawn for sniffing counts as well. These results indicate that reserpine-induced pallidal dopamine reduction affected PPI levels of EGF rats, but did not influence their motor function or deficits in social behaviors.

Effects of pallidal dopamine D2-like receptor blockade and activation on prepulse inhibition

To test the possibility that the reserpine effects on PPI might result from its influences on the noradrenergic or serotonergic systems, we manipulated local dopaminergic transmission using dopamine receptor antagonists. We bilaterally administered a dopamine D1-like receptor antagonist, SCH23390 (0.5 μg and 2 μg per site), or a dopamine D2-like receptor antagonist, raclopride (0.5 μg and 2 μg per site), to the globus pallidus (Fig. 6A). SCH23390 failed to affect PPI scores at any dose in both groups of rats (Fig. 6B). In contrast, raclopride had differential effects in EGF and control rats [$F(2,43) = 3.91$ for EGF×raclopride dose, $P = 0.028$] (Fig. 6C). Post-hoc testing revealed that raclopride injection significantly ameliorated the PPI deficits of EGF rats ($P = 0.019$ and 0.043 for 0.5 μg and 2 μg raclopride, respectively) but not affected PPI levels of control rats. Similar to the results of the reserpine experiment, vehicle-treated control rats failed to react to the D2-like receptor antagonist at any dose.

To address whether a hyperdopaminergic state in the globus pallidus is sufficient to induce PPI deficits, we examined the effects of local stimulation of D2-like receptors. The bilateral challenge of naïve rats with quinpirole, a D2-like receptor agonist (5 μg per site), markedly disrupted PPI [$F(1,16) = 6.50$, $P = 0.022$] without altering pulse-alone startle (Fig. 7). These results support our