Molecular Psychiatry (2007) 12, 1026–1032

www.nature.com/mp

ORIGINAL ARTICLE

Pituitary adenylate cyclase-activating polypeptide is associated with schizophrenia

R Hashimoto[1,2,3,11], H Hashimoto[1,4,11], N Shintani[4,11], S Chiba[2,3], S Hattori[3], T Okada[3], M Nakajima[4], K Tanaka[4], N Kawagishi[4], K Nemoto[5], T Mori[3,5], T Ohnishi[3,5], H Noguchi[3], H Hori[3], T Suzuki[6], N Iwata[6], N Ozaki[7], T Nakabayashi[8], O Saitoh[8], A Kosuga[9], M Tatsumi[9], K Kamijima[9], DR Weinberger[10], H Kunugi[3] and A Baba[4]

[1]*The Osaka-Hamamatsu Joint Research Center for Child Mental Development, Osaka University Graduate School of Medicine, Suita, Osaka, Japan;* [2]*Department of Psychiatry, Osaka University Graduate School of Medicine, Suita, Osaka, Japan;* [3]*Department of Mental Disorder Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Kodaira, Tokyo, Japan;* [4]*Laboratory of Molecular Neuropharmacology, Graduate School of Pharmaceutical Sciences, Osaka University, Suita, Osaka, Japan;* [5]*Department of Radiology, National Center Hospital of Mental, Nervous, and Muscular Disorders, National Center of Neurology and Psychiatry, Kodaira, Tokyo, Japan;* [6]*Department of Psychiatry, Fujita Health University School of Medicine, Toyoake, Aichi, Japan;* [7]*Department of Psychiatry, Nagoya University Graduate School of Medicine, Nagoya, Aichi, Japan;* [8]*Department of Psychiatry, National Center Hospital of Mental, Nervous, and Muscular Disorders, National Center of Neurology and Psychiatry, Kodaira, Tokyo, Japan;* [9]*Department of Psychiatry, Showa University School of Medicine, Shinagawaku, Tokyo, Japan and* [10]*Genes, Cognition, and Psychosis Program, Clinical Brain Disorders Branch, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA*

Pituitary adenylate cyclase-activating polypeptide (PACAP, ADCYAP1: adenylate cyclase-activating polypeptide 1), a neuropeptide with neurotransmission modulating activity, is a promising schizophrenia candidate gene. Here, we provide evidence that genetic variants of the genes encoding PACAP and its receptor, PAC1, are associated with schizophrenia. We studied the effects of the associated polymorphism in the PACAP gene on neurobiological traits related to risk for schizophrenia. This allele of the PACAP gene, which is overrepresented in schizophrenia patients, was associated with reduced hippocampal volume and poorer memory performance. Abnormal behaviors in PACAP knockout mice, including elevated locomotor activity and deficits in prepulse inhibition of the startle response, were reversed by treatment with an atypical antipsychotic, risperidone. These convergent data suggest that alterations in PACAP signaling might contribute to the pathogenesis of schizophrenia.
Molecular Psychiatry (2007) **12,** 1026–1032; doi:10.1038/sj.mp.4001982; published online 27 March 2007

Keywords: schizophrenia; PACAP; SNP; hippocampus; memory; PPI

Introduction

Schizophrenia is a common neuropsychiatric disorder affecting 0.5–1% of the general population worldwide. This disease is characterized by psychosis and profound disturbances of cognition, emotion and social functioning. The pathophysiology of schizophrenia is still unclear; however, this disease is highly heritable[1] and several intermediate phenotypes such as neurocognitive dysfunction, abnormal brain morphology and deficits in prepulse inhibition (PPI) of the startle response are known to be useful to identify susceptibility genes for schizophrenia.[2,3]

The adenylate cyclase-activating polypeptide 1 (ADCYAP1) gene encodes pituitary adenylate cyclase-activating polypeptide (PACAP), a neuropeptide, which is a member of the vasoactive intestinal peptide (VIP)/secretin/glucagon family. It exerts multiple activities as a neurotransmitter or neuromodulator via three heptahelical G-protein-linked receptors, one PACAP-specific (PAC1) receptor and two receptors that are shared with VIP (VPAC1 and VPAC2).[4–6] PACAP induces cyclic AMP accumulation through activation of these receptors.[4–6] We generated mice lacking the PACAP gene ($PACAP^{-/-}$); these mice had profound behavioral abnormalities including hyperactivity and explosive jumping in an open field, increased novelty-seeking behavior and deficits in PPI.[7,8] In addition, the PACAP gene is located on 18p11, which linkage studies have suggested as a locus for schizophrenia and bipolar disorder.[9] Although previous studies indicated that the PACAP gene could be a good candidate gene for schizophrenia, only one preliminary study has examined a

Correspondence: Dr R Hashimoto, The Osaka-Hamamatsu Joint Research Center for Child Mental Development, Osaka University Graduate School of Medicine, D3, 2-2, Yamadaoka, Suita, Osaka, 565-0871, Japan.
E-mail: hashimor@psy.med.osaka-u.ac.jp
[11]These authors contributed equally to this work.
Received 2 October 2006; revised 17 January 2007; accepted 20 February 2007; published online 27 March 2007

possible association with schizophrenia and reported negative results.[10] Here, we present data demonstrating a possible association between PACAP-PAC1 signaling and schizophrenia, using a multidisciplinary approach in both humans and rodents.

Materials and methods

Subjects

Subjects for the clinical association study were 804 patients with schizophrenia (51.1% males with a mean age of 44.2 years (s.d. 14.5) and a mean age of onset of 24.8 years (s.d. 8.8)) and 967 healthy controls (47.7% males with a mean age of 40.4 years (s.d. 16.1)). All the subjects were biologically unrelated Japanese. Three hundred and fifty-one patients with schizophrenia and 518 controls were from Tokyo Metropolitan (the east part of Japan), and 453 patients with schizophrenia and 449 controls were from Aichi prefecture (the central part of Japan). Patients were recruited at the National Center Hospital of Mental, Nervous, and Muscular Disorders; Nagoya University Hospital; Showa University Hospital and hospitals related to Department of Psychiatry, Nagoya University Graduate School of Medicine or Department of Psychiatry, Showa University School of Medicine. Healthy controls, including hospital and institutional staff, were recruited from local advertisements in Tokyo and Aichi. Magnetic resonance (MR) measurements and neurocognitive tests were performed only on some subjects (MR measurements: 81 patients with schizophrenia and 201 healthy controls; neurocognitive tests: 62 patients with schizophrenia and 139 healthy controls), all of whom were recruited at National Center of Neurology and Psychiatry. Demographic information for the subjects receiving MR measurements and neurocognitive tests is shown in detail in Supplementary Table 1 and Figure 1b. Consensus diagnosis was made for each patient by at least two trained psychiatrists, according to the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) criteria, based on clinical interview and other available information including medical records and other research assessments. No patient was diagnosed by medical records alone. Controls were healthy volunteers who had no current or past contact to psychiatric services. After a description of the study, written informed consent was obtained from every subject. The study protocol was approved by institutional ethics committees.

Genetic analysis

Venous blood was drawn from subjects and genomic DNA was extracted from whole blood according to standard procedures. Seven single nucleotide polymorphisms (SNPs) in the PACAP gene and three SNPs in the PAC1, VPAC1 and VPAC2 genes were genotyped using the TaqMan 5′-exonuclease allelic discrimination assay, as described previously.[11,12] Primers and probes for the detection of the SNPs are available on request. Statistical analysis of genetic association studies was performed using SNPAlyse (DYNACOM, Yokohama, Japan). The presence of Hardy–Weinberg equilibrium was examined by using the χ^2 test for goodness of fit. Allele distributions between patients and controls were analyzed by the χ^2 test for independence. All P-values reported are two-tailed. Statistical significance was defined as $P<0.05$.

Neuroimaging analysis

All MR studies were performed on a 1.5 T Siemens Magnetom Vision plus system (Siemens, Erlangen, Germany). A three-dimensional volumetric acquisition of a T1-weighted gradient echo sequence produced a gapless series of 144 sagittal sections using an MPRage sequence (TE/TR, 4.4/11.4 ms; flip angle, 15°; acquisition matrix, 256 × 256; 1NEX, field of view, 31.5 cm; slice thickness, 1.23 mm).

Data were analyzed with Statistical Parametric Mapping 2 (SPM2) running on MATLAB 6.5. MR images were processed using optimized voxel-based morphometry (VBM) in SPM2 as described in detail previously.[13,14] Normalized segmented images were modulated by multiplication with Jacobian determinants of the spatial normalization to encode the deformation field for each subject as tissue density changes in normal space. Following modulation, images were smoothed using a 12 mm full-width half-maximum of isotropic Gaussian kernel, because previous studies had proved that this should be a reasonable filter.[13,15,16] In addition, we confirmed that the results of statistical analyses with three different smoothing filters (6, 8 and 12 mm Gaussian kernels) were essentially the same.

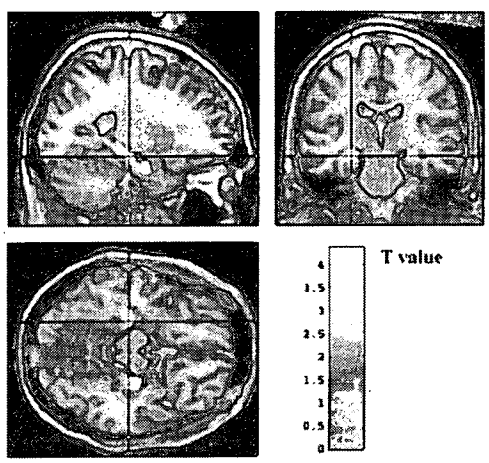
Statistical analyses were performed with SPM2, which implemented a general linear model. A hypothesis-driven regions of interest (ROIs) approach was used to investigate the hippocampus, using an ROI from the Wake Forest University PickAtlas.[17] Our hypothesis is that the PACAP genotype related to the risk of developing schizophrenia is associated with hippocampal volume, because PACAP is associated with hippocampal function in rodents, and hippocampal volume is reported to be reduced in schizophrenia. The genotype and diagnostic effects on hippocampal gray matter volume change were assessed statistically using a single-subject condition and covariate model with a significance level set to 0.05 (corrected for multiple comparisons within the ROI). Age and gender were included in the model to control for confounds. Anatomic localization was according to both MNI coordinates and Talairach coordinates, obtained from M. Brett's transformations (http://www.mrccbu.cam.ac.uk/Imaging/Common/mnispace.shtml) and presented as Talairach coordinates.

Neurocognitive tests

Several memory tests, subscales of the Wechsler Memory Scale revised version (logical memory I, logical memory II, visual reproduction I, visual reproduction II, verbal paired associates I (VPAI),

verbal paired associates II, visual paired associates I and visual paired associates II) and the general intelligence IQ (from full scale of the Wechsler Adult Intelligence Scale, revised edition, WAIS-R), were performed by some of the subjects recruited at National Center of Neurology and Psychiatry. In association analysis between SNP3 of the PACAP gene and VPAI, group comparisons of demographic data were performed by using unpaired *t*-tests or χ^2, as appropriate. There were no differences between genotype groups and demographic variables, for example, age, gender, education years and full-scale IQ, except for gender distribution in patients with schizophrenia ($P=0.026$) (Figure 1b). The effects of the SNP3 genotype of the PACAP gene and diagnosis on scores of memory tests were analyzed by a two-way analysis of covariance (ANCOVA), with age, gender and education years as covariates using SPSS 11.0J for Windows (SPSS Japan Inc., Tokyo, Japan).

a



b

	Schizophrenia (n=62)		Control (n=139)	
	A/A, A/G	G/G	A/A, A/G	G/G
Variables	(n=10)	(n=52)	(n=27)	(n=112)
Age (SD)	39.2 (10.5)	47.1(13.7)	34.6 (11.0)	37.3 (11.8)
Gender (M/F)	3/7	35/17	13/14	34/78
Education years (SD)	14.0 (2.6)	12.8 (3.3)	16.7 (2.8)	16.5 (3.1)
Full scale IQ (SD)	86.2 (21.5)	83.8 (18.8)	109.9 (13.4)	109.6 (11.6)
VPA1 (SD)	13.3 (3.9)	9.0 (4.6)*	16.1 (2.7)	16.1 (2.3)

Figure 1 Genetic variation of PACAP is associated with hippocampal morphology and memory in humans. (**a**) Statistical maps of *t*-transformed hippocampal volume differences derived by optimized VBM of individuals homozygous for the G allele in SNP3 of the PACAP gene, relative to A-carriers, in all subjects, thresholded at $P<0.05$ (corrected) in coronal, sagittal and axial views. These data show bilateral significant hippocampal volume reduction in individuals homozygous for the G allele. (**b**) Lower visual associate memory I score in individuals homozygous for the G allele in SNP3 of the PACAP gene, compared to A-carriers, in the schizophrenia group. Means±s.d. are shown. VPAI, visual paired associates I. *$P<0.05$, compared with A-carriers.

When genotype effects on VPAI in controls or patients with schizophrenia were examined separately, a Mann–Whitney *U*-test and ANCOVA with gender as a covariate were used.

Animal study

All animal experiments were carried out in accordance with protocols approved by the Animal Research Committee of Osaka University and by the Ethics Review Committee for Animal Experimentation of the National Institute of Neuroscience. Generation of PACAP$^{-/-}$ mice by a gene targeting technique has been reported previously.[7] The null mutation was backcrossed onto the genetic background of Crlj:CD1 (Institute of Cancer Research) mice purchased from Charles River (Tokyo, Japan). All wild-type control mice and PACAP$^{-/-}$ mice (homozygous for the mutant PACAP gene) used in locomotor activity and PPI experiments were obtained from the intercross of heterozygous animals. C57BL/6J mice were purchased from Charles River and were allowed to acclimate in our animal facility for at least 5 days before initiation of experiments. Mice were housed in a temperature- (23±1°C) and light-controlled room with a 12 h light–dark cycle (lights on from 0800 to 2000) and allowed free access to water and food, except during behavioral testing.

Locomotor activity was quantified using an infrared photocell beam detection system, Acti-Track (Panlab, Barcelona, Spain). Following intraperitoneal injection of risperidone (0.1 mg/kg) or an equivalent amount of saline, mice were placed in plastic activity monitoring boxes (30 × 30 × 30 cm) and tracked for 60 min, with data being stored permanently; parameters indicative of locomotor activity, such as distance traveled, were assessed. Each mouse was tested individually and had no contact with the other mice. The PACAP mutant cohort used in locomotor activity testing consisted of 12 wild-type mice and 12 PACAP$^{-/-}$ mice ($n=6$ each for saline control and risperidone groups).

Acoustic startle responses for PPI were measured in a startle chamber (SR-LAB; San Diego Instruments, CA, USA) as described.[18] Mice were placed in the startle chamber for 30 min after intraperitoneal injection of risperidone (0.1 mg/kg) or an equal amount of saline. The testing session started with 5 min of acclimatization to the startle chamber in the presence of 65 dB background broadband (white) noise. Testing consisted of forty 120 dB pulses alone and 10 pulses preceded (100 ms) by a prepulse of 66, 68, 71 or 77 dB. Pulses were randomly presented with an average of 15 s between pulses. Twelve no-stimulus trials were included to assess spontaneous activity during testing. PPI was calculated as a percentage score: PPI (%) = (1−((startle response for pulse with prepulse)/(startle response for pulse alone))) × 100. The PACAP mutant cohort used in PPI testing consisted of 35 wild-type mice (saline control group = 22; risperidone group = 13) and 33 PACAP$^{-/-}$ mice (saline control group = 17; risperidone group = 16).

Male C57BL/6J mice weighing 20–25 g received once-daily injections intraperitoneally for 14 days with phencyclidine (PCP) (5 mg/kg; $n=13$) or saline for control ($n=12$). PACAP and PAC1 mRNA levels were measured by a real-time quantitative RT–PCR method (TaqMan assay, Applied Biosystems, Tokyo, Japan), using total RNA extracted from the frontal cortex or hippocampus of mice treated with PCP or saline, as described previously.[19] Statistically significant differences were assessed by the Mann–Whitney *U*-test.

Results

Genetic analysis

We examined the possible association between schizophrenia and genetic variations in the PACAP gene. Seven SNPs in the PACAP gene, selected from public databases, were genotyped, and the genotype distributions of all seven SNPs in the PACAP gene were in Hardy–Weinberg equilibrium in both controls and patients with schizophrenia (data not shown). The allele frequencies of the seven SNPs in patients and controls are shown in Table 1. The major allele of SNP3 and the minor allele of SNP5 were in excess in patients with schizophrenia when compared to controls (SNP3: $\chi^2=7.6$, $P=0.0059$, odds ratio = 0.74, 95% confidence interval (CI) 0.59–0.92; SNP5: $\chi^2=4.2$, $P=0.041$, odds ratio = 1.38, 95% CI 1.01–1.84), whereas no significant association of the other five SNPs with schizophrenia was observed (Table 1). SNP3 was significantly associated with schizophrenia after Bonferroni correction (corrected $P=0.041$). We next examined the possible association between schizophrenia and genes encoding the receptors for PACAP, such as the PAC1, VPAC1 and VPAC2 receptor genes. The genotype distributions of all three SNPs in the PAC1, VPAC1 and VPAC2 genes were in Hardy–Weinberg equilibrium in both controls and patients with schizophrenia, except for that of SNP3 of the VPAC1 gene in controls (data not shown). The allele frequencies of the three SNPs in each receptor gene in the patients and controls are shown in Table 2. There was significant evidence for an association between a genetic variant of the PAC1 gene and schizophrenia (SNP2: $\chi^2=6.0$, $P=0.014$, odds ratio = 1.18, 95% CI 1.03–1.35, corrected $P=0.042$), whereas none of the SNPs in the genes encoding VPAC1 or VPAC2 was associated with schizophrenia (Table 2). The evidence that the genes encoding PACAP and its receptor PAC1 are associated with schizophrenia suggests that signaling through PACAP and PAC1 might be associated with the pathophysiology of schizophrenia.

Intermediate phenotype

As the PACAP gene has been reported to play a role in learning and memory and hippocampal long-term potentiation in rodents,[20,21] we next examined the possible impact of SNP3 of the PACAP gene, which was associated with schizophrenia, on hippocampal volume in patients with schizophrenia and controls. A genotype effect was found as bilateral reductions of hippocampal volumes (right: $P=0.04$, $t=3.2$; left: $P=0.002$, $t=4.1$) in homozygous G subjects compared with A-carriers (Figure 1a). There was also a diagnostic effect, a significant reduction in left hippocampal volume in patients with schizophrenia compared with controls ($P=0.033$, $t=3.3$). Genotype–diagnosis interaction effects on brain morphology were not found, even at a lenient threshold (uncorrected $P=0.05$). We next estimated the effects of genotypes on hippocampal volume in the control groups and schizophrenic groups, separately. Schizophrenic patients homozygous for the G allele showed a significant reduction in bilateral hippocampal volumes (right: $P=0.013$, $t=3.5$; left: $P=0.005$, $t=3.9$). On the other hand, we found significantly decreased volumes of the bilateral hippocampi in homozygous G subjects compared with the A-carriers, at a lenient threshold (uncorrected $P=0.05$) in controls; however, no voxels could survive after the correction for multiple comparisons. These data

Table 1 Allele frequencies of seven SNPs in the PACAP gene between the patients with schizophrenia and controls

SNP-ID	*dbSNP*	*Distance from SNP1*	*Major/minor polymorphism*	*Location*	*Number of subjects*		*Minor allele frequency*		P-value	*Odds ratio (95% CI)*
					Controls	*Patients*	*Controls*	*Patients*		
SNP1	rs2846584	—	C/T	5′-region	967	804	0.362	0.373	0.54	
SNP2	rs2231181	712	G/C	5′-UTR	960	795	0.336	0.330	0.69	
SNP3	rs1893154	1071	G/A	Intron1	951	797	0.126	0.097	0.0059	0.74 (0.59–0.92)
SNP4	rs1893153	1149	T/A	Intron1	953	793	0.174	0.163	0.37	
SNP5	rs2856966	3656	A/G	Exon3 (D54G)	953	786	0.047	0.063	0.041	1.38 (1.01–1.84)
SNP6	rs928978	4481	C/A	Intron4	958	798	0.475	0.485	0.58	
SNP7	rs1610037	6581	A/G	3′-region	962	794	0.216	0.211	0.73	

Abbreviations: CI, confidence interval; PACAP, pituitary adenylate cyclase-activating polypeptide; SNPs, single nucleotide polymorphisms.
Minor allele frequencies in controls are shown. Significant results ($P<0.05$) are indicated with underline.

Table 2 Allele frequencies of SNPs in the PAC1, VPAC1 and VPAC2 gene between the patients with schizophrenia and controls

Gene name	*SNP-ID*	*dbSNP*	*Distance from SNP1*	*Major/minor polymorphism*	*Location*	*Number of subjects*		*Minor allele frequency*		*P-value*	*Odds ratio (95% CI)*
						Controls	*Patients*	*Controls*	*Patients*		
PAC1	SNP1	rs1468687	—	T/C	Intron2	950	796	0.287	0.264	0.12	
	SNP2	rs2302475	15553	C/T	Intron5	958	797	0.479	0.520	0.014	1.18 (1.03–1.35)
	SNP3	rs2267742	34598	A/G	Intron12	936	786	0.127	0.133	0.58	
VPAC1	SNP1	rs735773	—	C/G	Intron1	937	784	0.357	0.38	0.16	
	SNP2	rs406360	12972	A/G	Intron4	948	789	0.431	0.433	0.91	
	SNP3	rs3733055	22942	G/T	Exon13 (R445L)	958	801	0.041	0.035	0.33	
VPAC2	SNP1	rs885861	—	C/T	3′-UTR	963	802	0.208	0.232	0.090	1.15 (0.98–1.36)
	SNP2	rs3793224	55026	C/T	Intron4	944	791	0.247	0.232	0.29	
	SNP3	rs3812312	109228	C/T	Intron2	923	781	0.221	0.218	0.85	

Abbreviations: CI, confidence interval; SNPs, single nucleotide polymorphisms.
Minor allele frequencies in controls are shown. Significant results ($P<0.05$) are indicated with underline.

suggest that SNP3 in the PACAP gene could have an impact on hippocampal morphology.

As the human hippocampus is related to memory function, we also examined the association between SNP3 of the PACAP gene and several subscales of the Wechsler memory scale revised version in patients with schizophrenia and controls (Figure 1b). Two-way ANCOVA on VPAI revealed significant effects of diagnosis ($F=33.8$, $P<0.0001$) and genotype of SNP3 ($F=5.2$, $P=0.024$), and an interaction between diagnosis and genotype ($F=6.6$, $P=0.011$), whereas an effect of genotype was not found in other memory subscales (data not shown). Individuals homozygous for the G allele of SNP3, which was enriched in schizophrenia, had lower scores of VPAI than schizophrenic patients carrying the A allele (Mann–Whitney *U*-test: $P=0.015$); however, there was no difference between the two genotypes in the control group ($P>0.8$). ANCOVA with gender as a covariate did not alter the statistical significance of these results in patients with schizophrenia ($P=0.029$). These data suggest that the risk SNP of the PACAP gene could be associated with reduced hippocampal volume and poorer memory performance, which are neurobiological traits related to risk for schizophrenia.

Animal study

As our data indicate that PACAP might be associated with schizophrenia, PACAP knockout mice ($PACAP^{-/-}$ mice) could be a possible animal model for schizophrenia. Several schizophrenia-related behaviors in rodents, such as hyperactivity, deficits in PPI, locomotor response to antipsychotics, disturbance in social interaction and cognitive deficits, have been commonly observed in previous pharmacological and genetic animal models for schizophrenia.[22] Therefore, we examined the impact of an atypical antipsychotic, risperidone, on hyperactivity and deficits in PPI in $PACAP^{-/-}$ mice. $PACAP^{-/-}$ mice maintained high initial levels of locomotor activity during the open field test (Figure 2a and b), as reported previously.[7] When treated with risperidone, hyperlocomotion in $PACAP^{-/-}$ mice was attenuated almost to the normal levels seen in wild-type mice; however, treatment with risperidone had no significant effect on locomotor activity in wild-type mice (Figure 2a and b). Risperidone also reversed the diminished PPI in $PACAP^{-/-}$ mice[8] to the control level seen in wild-type mice (Figure 2c). Risperidone had no significant effect on PPI levels in wild-type mice (Figure 2c) and startle amplitudes in both $PACAP^{-/-}$ and wild-type mice (data not shown). These results suggest that the abnormal behaviors in $PACAP^{-/-}$ mice, which are believed to be schizophrenia-like phenotypes in rodents, can be rescued by an atypical antipsychotic, risperidone.

The abuse of PCP, an *N*-methyl-D-aspartic acid receptor antagonist, results in positive symptoms, negative symptoms and cognitive impairments, similar to those seen in patients with schizophrenia. Thus, mice chronically treated with PCP have been used as a potential animal model for schizophrenia.[23] To assess a possible change in the expression of PACAP and PAC1 receptor in the pathological state, we performed mRNA expression analysis for PACAP and PAC1 in the frontal cortex and hippocampus of mice chronically treated with PCP. The expression level of PACAP mRNA was significantly reduced in the frontal cortex, but not in the hippocampus (Supplementary Figure 1). On the other hand, increased expression of PAC1 mRNA was observed in both frontal cortex and hippocampus (Supplementary Figure 1). Although the altered expression of PACAP and PAC1 in mouse brains treated with PCP was subtle, these data are considered to be in line with the behavioral abnormalities in $PACAP^{-/-}$ mice, a possible animal model for schizophrenia.

These results using animal models support the notion that PACAP is associated with the pathophysiology of schizophrenia.

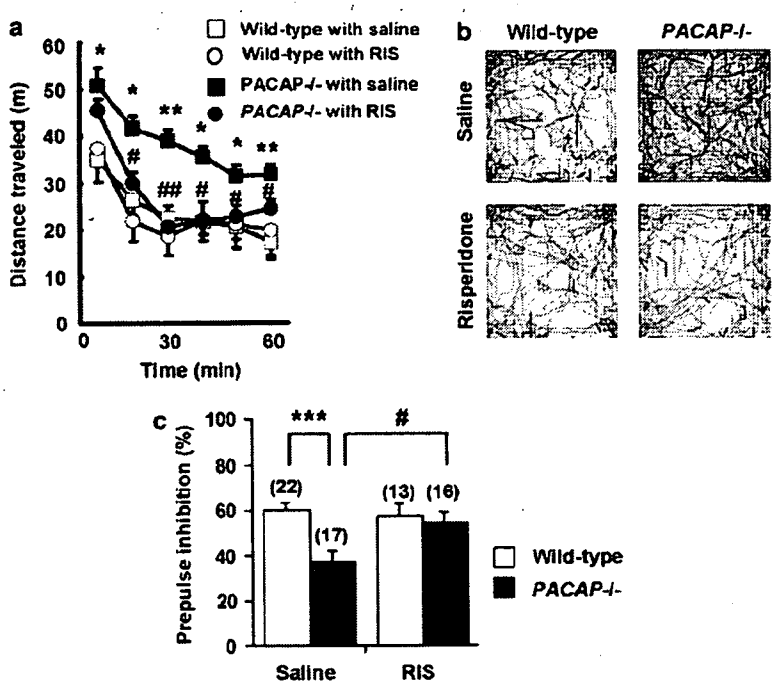


Figure 2 Hyperlocomotion and deficits in the PPI of PACAP$^{-/-}$ mice were normalized by risperidone treatment. (**a**) Locomotor activity in wild-type and PACAP$^{-/-}$ mice that received 0.1 mg/kg risperidone (RIS) or saline. $n=6$ per group. (**b**) Representative locomotor patterns of saline- or 0.1 mg/kg risperidone-treated wild-type and PACAP$^{-/-}$ mice during 25–30 min of a 60 min recording in an open field test. (**c**) PPI levels induced by a 77 dB prepulse in wild-type and PACAP$^{-/-}$ mice after pretreatment with risperidone (0.1 mg/kg) or saline. Numbers of animals for experiments are shown in parentheses. Data are given as means ± s.e.m. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, compared to wild-type. $^{\#}P<0.05$, $^{\#\#}P<0.01$, compared with saline in PACAP$^{-/-}$ mice.

Discussion

Our findings support the possibility that PACAP is a potential schizophrenia susceptibility gene. Clinical association between schizophrenia and the genes encoding PACAP and PAC1 and an association between intermediate phenotypes, hippocampal volume and visual associate memory performance and a risk SNP in the PACAP gene have been demonstrated in our study. There are several limitations in our results. We screened control subjects with no past or current visits to psychiatric services; however, we could not exclude the possibility that they have an undiagnosed or untreated psychiatric disorder. The obtained evidence for association was not very strong, especially in the association between the genotype and visual associate memory performance ($P<0.05$ level). When we applied corrections for multiple testing for several memory tests, this positive association became negative. This association is not conclusive, although the association between the risk allele for schizophrenia and poorer memory performance might be attractive. Thus, replication studies should be conducted to confirm our findings. We do not know whether SNP3 alters the expression/function of the PACAP gene. Accordingly, there remains the possibility that other polymorphisms, which are in linkage disequilibrium to this polymorphism, are truly responsible for giving susceptibility.

Studies aiming to identify susceptibility genes for schizophrenia are faced with the confounds of subjective clinical criteria and the likelihood of allelic and locus heterogeneity. Although schizophrenia is substantially heritable, the mode of inheritance is complex, involving numerous genes of small effect and a nontrivial environmental component. The concept of intermediate phenotype (endophenotype) assumes that neurobiological deficits occur across the schizophrenia spectrum in schizophrenia patients, schizotypal patients and clinically unaffected relatives of schizophrenia patients. The intermediate phenotype approach is an alternative method for measuring phenotypic variation that may facilitate the identification of susceptibility genes in the context of complexly inherited traits. Using this approach, we showed an association between the PACAP gene and two intermediate phenotypes, hippocampal volume and visual associate memory, in addition to the genetic association with schizophrenia. Our study could be a successful example of using this strategy to find susceptibility genes for complex diseases.

The hyperactivity and deficits in PPI observed in PACAP$^{-/-}$ mice[7,8] are believed to be schizophrenia-like behaviors in rodents. PAC1 knockout mice also show abnormal behaviors, including elevated locomotor activity and abnormal social behavior.[24,25] Our genetic findings, which demonstrate an association

between schizophrenia and two genes, PACAP and PAC1, are supported by the abnormal behaviors in knockout mice of PACAP and PAC1. Risperidone, an atypical antipsychotic, has the advantage of better extrapyramidal tolerability than conventional antipsychotics, but also has advantages in cognitive disturbances and the treatment of negative and depressive symptoms.[26] Our previous study showed that haloperidol, a representative conventional antipsychotic, rescued hyperactivity,[7] but did not rescue deficits in PPI.[8] As risperidone treatment rescued both of these abnormalities in PACAP$^{-/-}$ mice, and as risperidone is a combined D2 and 5-HT$_{2A}$ receptor antagonist, either dopamine or serotonin signaling, or both, could be relevant to the abnormal behaviors in PACAP$^{-/-}$ mice.

Our convergent evidence suggests that investigation of PACAP-PAC1 signaling in the brain could provide a clue to elucidating the possible mechanisms of pathophysiology in schizophrenia.

Acknowledgments

We thank Ms Tomoko Shizuno, Keiko Okada and Akiko Murakami for technical assistance and staff of the National Center of Neurology and Psychiatry for recruiting patients and healthy subjects. This work was supported in part by Grants-in-Aid from the Japanese Ministry of Health, Labor and Welfare (H18-kokoro-005, H17-kokoro-001, H17-kokoro-007 and H16-kokoro-002); the Japanese Ministry of Education, Culture, Sports, Science and Technology; Japan Society for the Promotion of Science; CREST (Core Research for Evolutional Science and Technology) of JST (Japan Science and Technology Agency); Japan Foundation for Neuroscience and Mental Health; the Sankyo Foundation of Life Science; and Taisho Pharmaceutical Co Ltd.

References

1 Owen MJ, Williams NM, O'Donovan MC. The molecular genetics of schizophrenia: new findings promise new insights. *Mol Psychiatry* 2004; **9**: 14–27.
2 Preston GA, Weinberger DR. Intermediate phenotypes in schizophrenia: a selective review. *Dialog Clin Neurosci* 2005; **7**: 165–179.
3 Braff DL, Light GA. The use of neurophysiological endophenotypes to understand the genetic basis of schizophrenia. *Dialog Clin Neurosci* 2005; **7**: 125–135.
4 Hashimoto H, Shintani N, Baba A. Higher brain functions of PACAP and a homologous Drosophila memory gene amnesiac: insights from knockouts and mutants. *Biochem Biophys Res Commun* 2002; **297**: 427–431.
5 Vaudry D, Gonzalez BJ, Basille M, Yon L, Fournier A, Vaudry H. Pituitary adenylate cyclase-activating polypeptide and its receptors: from structure to functions. *Pharmacol Rev* 2000; **52**: 269–324.
6 Arimura A. Perspectives on pituitary adenylate cyclase activating polypeptide (PACAP) in the neuroendocrine, endocrine, and nervous systems. *Jpn J Physiol* 1998; **48**: 301–331.
7 Hashimoto H, Shintani N, Tanaka K, Mori W, Hirose M, Matsuda T *et al*. Altered psychomotor behaviors in mice lacking pituitary adenylate cyclase-activating polypeptide (PACAP). *Proc Natl Acad Sci USA* 2001; **98**: 13355–13360.
8 Tanaka K, Shintani N, Hashimoto H, Kawagishi N, Ago Y, Matsuda T *et al*. Psychostimulant-induced attenuation of hyperactivity and prepulse inhibition deficits in Adcyap1-deficient mice. *J Neurosci* 2006; **26**: 5091–5097.
9 Nurnberger Jr JI, Foroud T. Genetics of bipolar affective disorder. *Curr Psychiatry Rep* 2000; **2**: 147–157.
10 Ishiguro H, Ohtsuki T, Okubo Y, Kurumaji A, Arinami T. Association analysis of the pituitary adenyl cyclase activating peptide gene (PACAP) on chromosome 18p11 with schizophrenia and bipolar disorders. *J Neural Transm* 2001; **108**: 849–854.
11 Hashimoto R, Suzuki T, Iwata N, Yamanouchi Y, Kitajima T, Kosuga A *et al*. Association study of the frizzled-3 (FZD3) gene with schizophrenia and mood disorders. *J Neural Transm* 2005; **112**: 303–307.
12 Hashimoto R, Okada T, Kato T, Kosuga A, Tatsumi M, Kamijima K *et al*. The breakpoint cluster region gene on chromosome 22q11 is associated with bipolar disorder. *Biol Psychiatry* 2005; **57**: 1097–1102.
13 Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, Frackowiak RS. A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage* 2001; **14**: 21–36.
14 Ashburner J, Friston KJ. Voxel-based morphometry – the methods. *Neuroimage* 2000; **11**: 805–821.
15 Pezawas L, Verchinski BA, Mattay VS, Callicott JH, Kolachana BS, Straub RE *et al*. The brain-derived neurotrophic factor val66met polymorphism and variation in human cortical morphology. *J Neurosci* 2004; **24**: 10099–10102.
16 Mechelli A, Friston KJ, Frackowiak RS, Price CJ. Structural covariance in the human cortex. *J Neurosci* 2005; **25**: 8303–8310.
17 Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 2003; **19**: 1233–1239.
18 Sakaue M, Ago Y, Baba A, Matsuda T. The 5-HT1A receptor agonist MKC-242 reverses isolation rearing-induced deficits of prepulse inhibition in mice. *Psychopharmacology (Berl)* 2003; **170**: 73–79.
19 Chiba S, Hashimoto R, Hattori S, Yohda M, Lipska B, Weinberger DR *et al*. Effect of antipsychotic drugs on DISC1 and dysbindin expression in mouse frontal cortex and hippocampus. *J Neural Transm* 2006; **113**: 1337–1346.
20 Matsuyama S, Matsumoto A, Hashimoto H, Shintani N, Baba A. Impaired long-term potentiation *in vivo* in the dentate gyrus of pituitary adenylate cyclase-activating polypeptide (PACAP) or PACAP type 1 receptor-mutant mice. *Neuroreport* 2003; **14**: 2095–2098.
21 Sacchetti B, Lorenzini CA, Baldi E, Bucherelli C, Roberto M, Tassoni G *et al*. Pituitary adenylate cyclase-activating polypeptide hormone (PACAP) at very low dosages improves memory in the rat. *Neurobiol Learn Mem* 2001; **76**: 1–6.
22 Gainetdinov RR, Mohn AR, Caron MG. Genetic animal models: focus on schizophrenia. *Trends Neurosci* 2001; **24**: 527–533.
23 Jentsch JD, Roth RH. The neuropsychopharmacology of phencyclidine: from NMDA receptor hypofunction to the dopamine hypothesis of schizophrenia. *Neuropsychopharmacology* 1999; **20**: 201–225.
24 Otto C, Martin M, Wolfer DP, Lipp HP, Maldonado R, Schutz G. Altered emotional behavior in PACAP-type-I-receptor-deficient mice. *Brain Res Mol Brain Res* 2001; **92**: 78–84.
25 Nicot A, Otto T, Brabet P, Dicicco-Bloom EM. Altered social behavior in pituitary adenylate cyclase-activating polypeptide type I receptor-deficient mice. *J Neurosci* 2004; **24**: 8786–8795.
26 Moller HJ. Risperidone: a review. *Expert Opin Pharmacother* 2005; **6**: 803–818.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)

Available online at www.sciencedirect.com
ScienceDirect
Psychiatry Research 153 (2007) 61 – 67
PSYCHIATRY RESEARCH
www.elsevier.com/locate/psychres

The new GRID Hamilton Rating Scale for Depression demonstrates excellent inter-rater reliability for inexperienced and experienced raters before and after training

Hideaki Tabuse [a,b], Amir Kalali [c], Hideki Azuma [b], Norio Ozaki [d], Nakao Iwata [e], Hiroshi Naitoh [e], Teruhiko Higuchi [f], Shigenobu Kanba [g], Kunihiko Shioe [h], Tatsuo Akechi [b], Toshi A. Furukawa [b,*]

[a] *Holy Cross Hospital, Kujiri 2431-160, Izumi-cho, Toki, 509-5142 Japan*
[b] *Department of Psychiatry and Cognitive-Behavioral Medicine, Nagoya City University Graduate School of Medical Sciences, Mizuho-cho, Mizuho-ku, Nagoya, 467-8601 Japan*
[c] *International Society for CNS Drug Development, San Diego, CA, USA*
[d] *Department of Psychiatry, Nagoya University Graduate School of Medicine, Nagoya, Japan*
[e] *Department of Psychiatry, Fujita Health University, Toyoake, Japan*
[f] *Musashi Hospital, National Center of Neurology and Psychiatry, Kodaira, Japan*
[g] *Department of Neuropsychiatry, Kyushu University Graduate School of Medical Sciences, Fukuoka, Japan*
[h] *Department of Neuropsychiatry, Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Yamanashi, Japan*

Received 16 January 2006; received in revised form 15 May 2006; accepted 23 July 2006

Abstract

The Hamilton Rating Scale for Depression (HAMD) is the *de facto* international gold standard for the assessment of depression. There are some criticisms, however, especially with regard to its inter-rater reliability, due to the lack of standardized questions or explicit scoring procedures. The GRID-HAMD was developed to provide standardized explicit scoring conventions and a structured interview guide for administration and scoring of the HAMD. We developed the Japanese version of the GRID-HAMD and examined its inter-rater reliability among experienced and inexperienced clinicians ($n=70$), how rater characteristics may affect it, and how training can improve it in the course of a model training program using videotaped interviews. The results showed that the inter-rater reliability of the GRID-HAMD total score was excellent to almost perfect and those of most individual items were also satisfactory to excellent, both with experienced and inexperienced raters, and both before and after the training. With its standardized definitions, questions and detailed scoring conventions, the GRID-HAMD appears to be the best achievable set of interview guides for the HAMD and can provide a solid tool for highly reliable assessment of depression severity.

Keywords: Depressive disorder; Psychiatric status rating scales; Mental status schedule; Reproducibility of results

1. Introduction

The Japanese Society of Clinical Psychopharmacology has long realized the need to standardize the administration of the Hamilton Rating Scale for Depression

* Corresponding author. Tel.: +81 52 853 8271; fax: +81 52 852 0837.
E-mail address: furukawa@med.nagoya-cu.ac.jp (T.A. Furukawa).

0165-1781/$ - see front matter © 2006 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.psychres.2006.07.004

(HAMD) (Hamilton, 1960), the *de facto* international standard for the assessment of depression (Furukawa et al., 2005), within Japan and appointed a team headed by Dr. Higuchi to develop a model training program in 2000. In the course of these efforts, we learned that a group of researchers had met in the USA in 1999 and proposed to establish a common set of standards for scoring and administering the HAMD that would be acceptable to the Food and Drug Administration and be used by pharmaceutical, academic and clinical researchers. This proposal led to the formation of the Depression Rating Scale Standardization Team (DRSST), a group of individuals representing clinicians, academia, government and the pharmaceutical industry. The goal of this group was to standardize the administration and scoring of the HAMD without significantly altering the original intent of Hamilton's items or the scoring profile rather than to develop a new instrument (Kalali et al., 2002; Bech et al., 2005).

The product of their efforts is the GRID-HAMD, which has three components: the GRID scoring system, (scoring intensity and frequency separately to obtain the severity score), the manual of scoring conventions with detailed anchor descriptions and more behavioral exemplars, and a semi-structured interview guide. The DRSST clarified and operationalized ambiguous anchor descriptions and incorporated the new definitions into the individual items. The GRID-HAMD can be downloaded free of charge at the International Society for CNS Drug Development homepage (http://www.iscdd.org). Given the many versions of the scale in use, the DRSST concluded that standardization would improve the current scale and lay the groundwork for development of a new scale.

The Japanese team felt that the GRID-HAMD would set a new standard in depression rating and decided to develop the Japanese training program around it. We developed the Japanese version of the GRID-HAMD (see Section 2) and then conducted a model training course for the GRID-HAMD in March 2004. The primary purpose of the study is to examine the inter-rater reliability of the Japanese version of the GRID-HAMD among experienced and inexperienced Japanese psychiatrists and psychologists, how rater characteristics may affect it, and how training can enhance it.

2. Methods

2.1. Participants

Psychiatrists ($n=52$), clinical psychologists ($n=12$) and medical students ($n=6$) from three university medical schools in Japan (Nagoya City University, Nagoya University and Fujita Health University) took part in a full day training course for the newly developed Japanese version of the GRID-HAMD. Of the 70 participants, 20 had no previous experience with any version of the HAMD, whereas 17 had administered it between one and five times and 33 had administered it six or more times. However, only 16 of the last group had ever received formal training in the administration of the instrument. The mean (S.D.) of clinical experience was 6.3 (6.1) years for the psychiatrists and 3.5 (3.0) years for the clinical psychologists.

2.2. Instrument

The Japanese version of the GRID-HAMD was developed in collaboration with the DRSST. The original English version of the GRID-HAMD was translated into Japanese by TAF. A team of seven psychiatrists, all of them experts in depression treatment and research, checked the translation and amended it where necessary, based upon the consensus of the team. Two research assistants, both proficient in English and one with a Bachelor's degree in psychology, and both blind to the original English version, then back-translated the Japanese translation of the probe questions into English. AK checked the backtranslation and pointed out possible discrepancies, based upon which TAF retranslated the questioned sentences into Japanese. This process was repeated three times, until AK was able to ascertain semantic equivalence between the original and back-translated versions.

2.3. Procedure

We used three pairs of videotapes of pre- and post-treatment administration of the HAMD. Two pairs used simulated Japanese patients (one man and one woman) and the other pair used a simulated English patient. The Japanese man, woman and their interviewers were played by professional actors and actresses, based on rough scenarios but including a substantial amount of ad lib interactions. The participants' general impression was that the patients were very well played and appeared natural, but that the interviewers appeared rather stiff. The English patient's interviews had Japanese subtitles. Each interview lasted between 15 and 40 min. The experts' consensus total scores for the six videotapes were 26 for the Japanese man pre-treatment, 10 post-treatment, 37 for the Japanese woman pre-treatment, 19 post-treatment, 21 for the English woman pre-treatment and 0 post-treatment. These videotapes were prepared independently of and before our training workshop for the GRID-

HAMD. The interviewers in these videotapes by and large followed the conventions of the Structured Interview Guide for the Hamilton Depression Rating Scale (SIGH-D) (Williams, 1988), which sometimes did not probe specifically enough into the frequency of the symptoms during the last week.

The participants in the workshop used GRID-HAMD to rate each interview. When the videotape failed to ask for the frequency, the participants were instructed to assume that the frequency was 50% of the time. This was the case for items 2, 3, 7, 10, 11, 12, and 13 of the pre-treatment videotape of the Japanese man, for items 2, 11, and 13 of the post-treatment videotape of the Japanese man, for items 2, 3, 6, 7, 10, 12, 13, and 15 of the pre-treatment video of the Japanese woman, and for items 2, 5, 7, 10, 12, and 13 of the post-treatment video of the Japanese woman. In other words, 24 out of the 68 items (35%) required participants to rely on this rating convention.

Because the rating difficulty might differ between the videotapes of the Japanese man and woman, the participants were randomly divided into two groups, and each group saw either the man's videos or the woman's videos first. There was no discussion immediately following the two videos. The videos therefore served as pre-training and post-training assessments of the raters' reliability. After this pre-training assessment in the reliability of the GRID-HAMD, their training began with a lecture on the history of the Hamilton Rating Scale for Depression and a general discussion of assessment in psychiatry. The training of the GRID-HAMD formed the core of the workshop and used the English woman's videotapes. After scoring each English woman's videotapes, possible discrepancies and questions were discussed among the participants and the trainers. The three pairs of videotapes were therefore presented during this 1-day course as shown in Fig. 1.

2.4. Analyses

The inter-rater reliability for each item of the GRID-HAMD and for its total score was estimated by way of the ANOVA intraclass correlation coefficient (ICC) (one-way random effects model, single rater) of the SPSS (SPSS Inc., 2002). Because of its intrinsic paradoxical characteristic whereby we obtain low ICC despite high agreement (Feinstein and Cicchetti, 1990), we did not calculate ICC

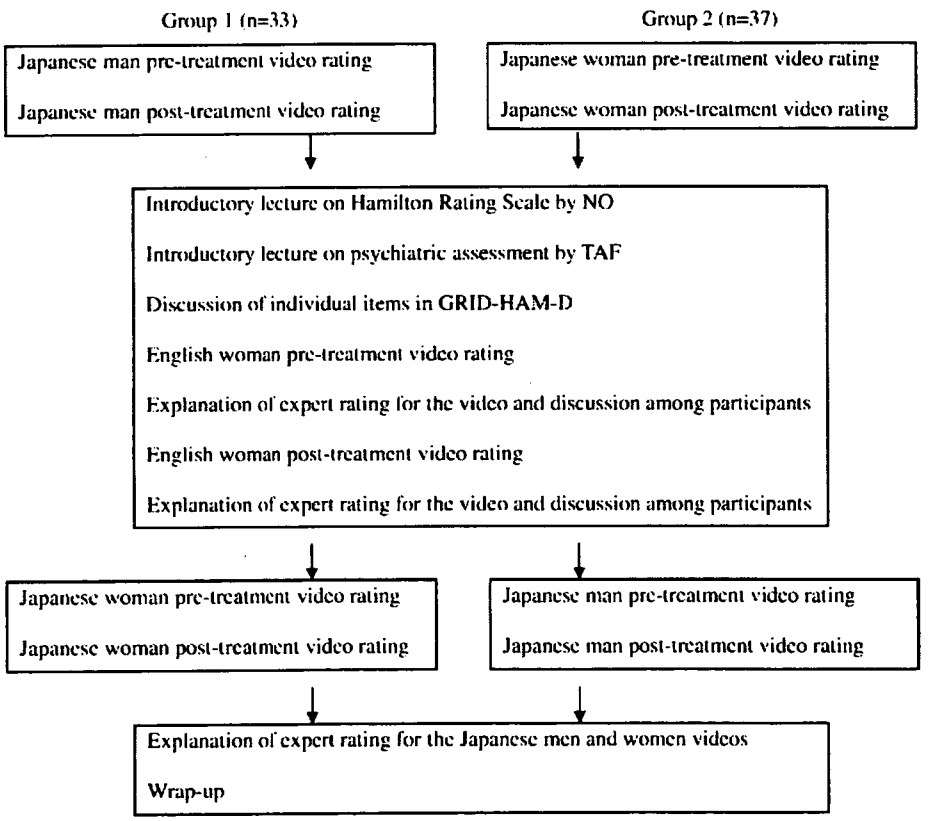


Fig. 1. Procedure of the model training program.

when one rating predominated (≥0.90 of all the ratings) for a particular item. It has been suggested that ICCs between 0.61 and 0.80 are "substantial" and those greater than 0.80 are "almost perfect" (Landis and Koch, 1977).

Because we were interested in the effects of experience and training, we subgrouped our participants based on their previous experience with the HAMD as follows.

Group A (n=20) No previous experience with the HAMD
Group B (n=17) Have administered the HAMD between one and five times
Group C (n=17) Have administered the HAMD six or more times, but have never had formal training in its administration
Group D (n=16) Have administered the HAMD six or more times, and have received formal training in its administration.

To examine the influence of the rating convention of assigning 50% frequency to such items as where the interviewer failed to ask for frequency in the videotape, we ran a supplementary sensitivity analysis by comparing

Table 1
ANOVA ICC for each item and the total score of the GRID-HAMD for the four subgroups of participants before and after training

Item		Group A		Group B		Group C		Group D	
		Before training	After training	Before training	After training	Before training	After training	Before training	After training
1	Depressed mood	0.83 (0.64–0.95)	0.89 (0.76–0.97)	0.78 (0.50–0.96)	0.84 (0.61–0.97)	0.91 (0.74–0.98)	0.84 (0.62–0.97)	0.84 (0.57–0.97)	0.92 (0.55–0.99)
2	Guilt	0.69 (0.41–0.89)	0.60 (0.30–0.86)	0.41 (0.07–0.84)	0.71 (0.40–0.94)	0.67 (0.34–0.93)	0.59 (0.27–0.91)	0.70 (0.34–0.94)	0.58 (0.19–0.91)
3	Suicide	0.89 (0.75–0.97)	0.87 (0.70–0.96)	0.92 (0.79–0.99)	0.90 (0.73–0.98)	0.92 (0.78–0.99)	0.89 (0.72–0.98)	0.94 (0.83–0.99)	0.97 (0.91–1.00)
4	Insomnia, early	0.90 (0.76–0.97)	0.90 (0.77–0.97)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	0.72 (0.41–0.95)	0.85 (0.63–0.97)	0.86 (0.62–0.98)	1.00 (1.00–1.00)
5	Insomnia, middle	0.78 (0.55–0.93)	0.75 (0.50–0.92)	0.66 (0.32–0.93)	0.79 (0.51–0.96)	0.78 (0.50–0.96)	0.70 (0.39–0.94)	0.87 (0.64–0.98)	0.78 (0.47–0.96)
6	Insomnia, late	0.86 (0.69–0.96)	0.93 (0.83–0.98)	0.91 (0.74–0.98)	0.97 (0.91–1.00)	0.87 (0.66–0.98)	0.95 (0.86–0.99)	0.92 (0.77–0.99)	0.96 (0.8–0.99)
7	Work and activities	0.63 (0.34–0.87)	0.69 (0.42–0.90)	0.70 (0.37–0.94)	0.72 (0.40–0.94)	0.73 (0.42–0.95)	0.75 (0.47–0.95)	0.73 (0.39–0.95)	0.71 (0.35–0.94)
8	Psychomotor retardation	0.67 (0.39–0.89)	0.80 (0.68–0.94)	0.67 (0.34–0.93)	0.78 (0.50–0.96)	0.85 (0.63–0.97)	0.83 (0.61–0.97)	0.59 (0.20–0.91)	0.61 (0.22–0.92)
9	Psychomotor agitation	na	na	−0.05 (−0.18–0.43)	−0.05 (−0.18–0.43)	0.13 (−0.10–0.67)	na	na	0.00 (−0.21–0.57)
10	Anxiety, psychic	0.75 (0.50–0.92)	0.38 (0.08–0.74)	0.64 (0.30–0.92)	0.75 (0.45–0.95)	0.73 (0.41–0.95)	0.74 (0.45–0.95)	0.82 (0.54–0.97)	0.96 (0.88–0.99)
11	Anxiety, somatic	0.88 (0.73–0.96)	0.82 (0.61–0.94)	0.80 (0.53–0.96)	0.87 (0.66–0.98)	0.89 (0.70–0.98)	0.86 (0.65–0.97)	0.93 (0.80–0.99)	0.89 (0.69–0.98)
12	Loss of appetite	0.87 (0.72–0.96)	0.91 (0.78–0.97)	0.89 (0.71–0.98)	0.95 (0.84–0.99)	0.88 (0.69–0.98)	0.82 (0.58–0.97)	0.91 (0.74–0.99)	0.87 (0.64–0.98)
13	Somatic symptoms, general	0.53 (0.21–0.82)	0.50 (0.20–0.81)	0.64 (0.30–0.93)	0.48 (0.13–0.87)	0.57 (0.21–0.90)	0.36 (0.07–0.81)	0.63 (0.25–0.93)	0.44 (0.05–0.86)
14	Sexual interest	na	na	na	na	na	na	na	na
15	Hypochondriasis	0.60 (0.30–0.86)	0.62 (0.33–0.87)	0.53 (0.18–0.89)	0.73 (0.42–0.95)	0.69 (0.37–0.94)	0.83 (0.60–0.97)	0.85 (0.59–0.97)	0.73 (0.39–0.95)
16	Loss of weight	0.63 (0.34–0.87)	0.65 (0.36–0.88)	0.64 (0.30–0.92)	0.79 (0.51–0.96)	0.71 (0.39–0.94)	0.84 (0.73–0.98)	0.77 (0.46–0.96)	0.71 (0.35–0.94)
17	Loss of insight	na	na	na	na	0.03 (−0.15–0.55)	na	na	na
Total		0.95 (0.87–0.98)	0.95 (0.87–0.91)	0.93 (0.82–0.99)	0.95 (0.86–0.99)	0.97 (0.91–1.00)	0.95 (0.85–0.99)	0.97 (0.90–1.00)	0.99 (0.96–1.00)

Figures in parentheses indicate the 95% confidence intervals.
na = not applicable due to too little variation because the particular score predominated and more than 90% of the obtained ratings were the same.
Group A (n=20): No previous experience with the HAMD.
Group B (n=17): Have administered the HAMD between one to five times.
Group C (n=17): Have administered the HAMD six or more times, but have never had formal training in its administration.
Group D (n=16): Have administered the HAMD six or more times, and have received formal training in its administration.

the average ANOVA ICCs between items for which the interviewers did not ask about frequency in more than half of the videotapes (items 2, 7, 10, 12, and 13) and those for which the interviewers asked (items 1, 4, 5, 6, and 15).

3. Results

Table 1 shows the ANOVA ICCs and their 95% confidence intervals for each item and the total score of the GRID-HAMD as applied to the Japanese man and woman's videotapes, for Groups A through D, both before and after training with lectures and practice with the English woman's videotapes. Excluding items 9, 14, and 17 (Psychomotor agitation, Sexual interest, and Loss of insight), which showed too little variation among raters to calculate meaningful chance-corrected agreement coefficients, and item 13 (Somatic symptoms, general), which often had ANOVA ICCs below 0.60, the inter-rater reliability of individual items was already largely in the substantial to excellent range before the training and did not show much increase after the training. Thus the respective averages of the ANOVA ICCs for individual items were 0.75 and 0.74 for Group A before and after training, 0.73 and 0.81 for Group B, 0.78 and 0.79 for Group C, and 0.81 and 0.79 for Group D. The ANOVA ICCs for the total score were almost perfect for all groups both before and after the training (range: 0.93 to 0.99). The average ICC for the items where the interviewers asked for frequency was 0.83 (range: 0.70 to 0.92) and that for the items where they failed to ask and where therefore the subjects were instructed to assume 50% frequency was 0.69 (range: 0.52 to 0.89).

4. Discussion

Our results suggest that when we relied on the GRID-HAMD scoring conventions, the inter-rater reliability of the total score was excellent to almost perfect and that satisfactory inter-rater reliability for individual items was also achievable, even with inexperienced raters and even without training. These findings are at variance with some previous studies on inter-rater reliability for HAMD items, which often reported poor reliability at the individual item levels. Cicchetti and Prusoff (1983) assessed reliability before treatment initiation and 16 weeks later at trial end. Before treatment, only one item was sufficiently reliable and 13 items had coefficients below 0.50. After treatment, again only one item was sufficiently reliable and 11 items had coefficients below 0.50. Craig et al. (1985) also found that only one item had adequate inter-rater reliability. On the other hand, Moberg et al. (2001) reported that nine items showed adequate reliability when the standard HAMD depression scale was administered, but all items showed adequate reliability when the scale was administered with the SIGH-D interview guidelines of Williams (1988). Our findings appear to extend theirs. Narita et al. (2002) pointed out specific weaknesses/ambiguities in the rating instructions in the SIGH-D, especially with regard to items for middle insomnia, somatic anxiety, loss of weight, depersonalization/derealization, and loss of insight; all of these are well anchored in the GRID-HAMD.

On the other hand, our results suggested that inter-rater reliability for general somatic symptoms may be low. However, we suspect that this was due to the difference in emphasis between SIGH-D item 13 and GRID-HAMD item 13, the former following the traditional HAM-D interpretation and focusing on heaviness and aches and the latter emphasizing fatigue and anergia in accordance with DSM-IV criterion symptoms.

With regard to the total score of the HAMD, most of the previous studies reported substantial to satisfactory inter-rater reliability, with ICCs ranging from 0.46 to 0.99 (Bagby et al., 2004). Some investigators provided evidence that the skill level or expertise of the interviewer and the provision of structured queries and scoring guidelines affect reliability (O'Hara and Rehm, 1983; Hooijer et al., 1991). Our findings suggest that with the use of explicit scoring conventions as outlined in the GRID-HAMD, even inexperienced raters can achieve satisfactory inter-rater reliability. We failed to show a significant effect of expertise or training, possibly because of the ceiling effect of these already high baseline reliability coefficients, although the raw scores do hint at even higher reliability coefficients after training and for more experienced users.

Weaknesses of the present study may be as follows: Firstly, the present study is based on videotaped interviews with simulated patients. Although the actor and actress played their roles naturally, with much ad lib interaction, the generalizability of the present findings to bona fide patients cannot be taken for granted and warrant another study. However, it should be pointed out that experienced physicians have been reported to be unable to differentiate standardized patients from real patients when they were sent unannounced into a physician's office, even when the physician was told in advance that this would be occurring (Kobak et al., 2003). The videotaped reliability study with simulated patients may also have inflated reliability estimates in comparison with test–retest design with real patients, which would more accurately reproduce clinical realities. Secondly, we used videotapes that had been made prior to and independently of our workshop for the GRID-HAMD.

The interviewers in the videotapes therefore did not abide by the GRID-HAMD conventions but roughly followed the SIGH-D questions. They therefore did not probe specifically enough about the frequency of some symptoms. The rating convention of assigning a 50% frequency to such items may have inflated the reliability estimates, but our sensitivity analysis did not support this possibility. Had the interviewers in the videotapes followed the GRID-HAMD interview guides, it is safe to assume that reliability could have been even higher. Thirdly, the videotaped interviews in the present study were such that there was little variation for three out of 17 items of the GRID-HAMD. We could therefore not ascertain satisfactory reliability for these items. In future studies we need to prepare videotapes that allow more variation in ratings for these items. Fourthly, although the ICCs did not change materially before and after the training, it must be pointed out that the present findings do not obviate the need for clinical expertise in depression assessment, as almost all the participants had substantial clinical experience already. In order to assure satisfactory rater performance, the raters' ability to conduct assessments on real patients is important in itself, in addition to the reliability of the instrument (Lipsitz et al., 2004). Lastly, the present study was conducted in Japanese with the Japanese version of the GRID-HAMD. The Japanese version was developed in strict adherence to the established back-translation procedure to ensure its linguistic equivalence with the English original, and we believe the present findings can be replicated with the original version as well, as it is thanks to the well-structured, adequately explained nature of the GRID-HAMD and not to any particularities of its Japanese version that we could achieve satisfactory reliability. Strictly speaking, however, the cross-cultural generalizability of the present findings must await independent replication studies in English and other languages and cultures.

Recently, a comprehensive review of the HAMD by Bagby et al. (2004, 2005) concurred that the GRID-HAMD is a major improvement over the previous versions in developing clear structured interview prompts and scoring guidelines, and in standardizing the scoring system. However, the retention of "loss of insight" that makes neither a conceptual nor an empirical contribution to the severity of depression or the lack of such DSM-IV criterion symptoms as "loss of concentration" remain major difficulties with the GRID-HAMD. Also the report from a 2002 National Institute of Health sponsored conference in the US on the assessment of depression and anxiety in clinical trials recommended the GRID-HAMD as the optimal way to administer the HAMD. A recent National Institutes of Health sponsored conference on assessment of suicidality also recommended the GRID-HAMD as the preferred version of the HAMD for assessing suicidality.

We feel that the GRID-HAMD is the best achievable set of semi-structured guides for the HAMD, the *de facto* standard in depression rating for over four decades, and this fact was corroborated in the present study by its robust reliability findings. In conclusion, the GRID-HAMD appears to provide a solid tool for highly reliable assessment of depression severity for both experienced and inexperienced mental health professionals.

References

Bagby, R.M., Ryder, A.G., Schuller, D.R., Marshall, M.B., 2004. The Hamilton Depression Rating Scale: has the gold standard become a lead weight? American Journal of Psychiatry 161, 2163–2177.

Bagby, R.M., Schuller, D.R., Ryder, A.G., Marshall, M.B., 2005. Dr. Bagby and colleagues reply. American Journal of Psychiatry 162, 2397–2398.

Bech, P., Engelhardt, N., Evans, K.R., Gibertini, M., Kalali, A.H., Kobak, K.A., Lipsitz, J.D., Williams, J.B., Pearson, J.D., Rothman, M., 2005. Why the Hamilton Depression Rating Scale endures. American Journal of Psychiatry 162, 2396.

Cicchetti, D.V., Prusoff, B.A., 1983. Reliability of depression and associated clinical symptoms. Archives of General Psychiatry 40, 987–990.

Craig, T.J., Richardson, M.A., Pass, R., Bregman, Z., 1985. Measurement of mood and affect in schizophrenic inpatients. American Journal of Psychiatry 142, 1272–1277.

Feinstein, A.R., Cicchetti, D.V., 1990. High agreement but low kappa: I. The problems of two paradoxes. Journal of Clinical Epidemiology 43, 543–549.

Furukawa, T.A., Streiner, D.L., Azuma, H., Higuchi, T., Kamijima, K., Kanba, S., Ozaki, N., Aoba, A., Murasaki, M., Miura, S., 2005. Cross-cultural equivalence in depression assessment: Japan–Europe–North American study. Acta Psychiatrica Scandinavica 112, 279–285.

Hamilton, M., 1960. A rating scale for depression. Journal of Neurology, Neurosurgery and Psychiatry 23, 56–62.

Hooijer, C., Zitman, F.G., Griez, E., van Tilburg, W., Willemse, A., Dinkgreve, M.A., 1991. The Hamilton Depression Rating Scale (HDRS); changes in scores as a function of training and version used. Journal of Affective Disorders 22, 21–29.

Kalali, A., Williams, J., Kobak, K.A., Lipsitz, J., Engelhardt, N., Evans, K., Olin, J., Pearson, J., Rothman, M.L., Bech, P., 2002. The new GRID HAM-D: pilot testing and international field trials. International Journal of Neuropsychopharmacology 5, S147–S148.

Kobak, K.A., Lipsitz, J.D., Feiger, A., 2003. Development of a standardized training program for the Hamilton Depression Scale using internet-based technologies: results from a pilot study. Journal of Psychiatric Research 37, 509–515.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159–174.

Lipsitz, J., Kobak, K., Feiger, A., Sikich, D., Moroz, G., Engelhard, A., 2004. The Rater Applied Performance Scale: development and reliability. Psychiatry Research 127, 147–155.

Moberg, P.J., Lazarus, L.W., Mesholam, R.I., Bilker, W., Chuy, I.L., Neyman, I., Markvart, V., 2001. Comparison of the standard and structured interview guide for the Hamilton Depression Rating Scale

in depressed geriatric inpatients. American Journal of Geriatric Psychiatry 9, 35–40.

Narita, T., Kim, N., Nakane, Y., Ozaki, N., 2002. Reliability and validity of the Structured Interview Guide of the Hamilton Depression Rating Scale (SIGH-D). Japanese Journal of Clinical Psychopharmacology 6, 77–82 [in Japanese].

O'Hara, M.W., Rehm, L.P., 1983. Hamilton Rating Scale for Depression: reliability and validity of judgments of novice raters. Journal of Consulting and Clinical Psychology 51, 318–319.

SPSS Inc., 2002. SPSS for Windows Version 11.5. SPSS Inc., Chicago.

Williams, J.B., 1988. A structured interview guide for the Hamilton Depression Rating Scale. Archives of General Psychiatry 45, 742–747.

Available online at www.sciencedirect.com
ScienceDirect
Schizophrenia Research 93 (2007) 58–65
SCHIZOPHRENIA RESEARCH
www.elsevier.com/locate/schres

Association analysis of AKT1 and schizophrenia in a UK case control sample

Nadine Norton [a,*], Hywel J. Williams [a], Sarah Dwyer [a], Liam Carroll [a], Tim Peirce [a], Valentina Moskvina [a], Ricardo Segurado [a], Ivan Nikolov [a], Nigel M. Williams [a], Masashi Ikeda [b], Nakao Iwata [b], Michael J. Owen [a], Michael C. O'Donovan [a]

[a] *Department of Psychological Medicine, Wales School of Medicine, Henry Wellcome Building, Cardiff University, Heath Park, Cardiff CF14 4XN, UK*
[b] *Department of Psychiatry, Fujita Health University School of Medicine, Toyoake, Aichi 470-1192, Japan*

Received 8 January 2007; received in revised form 13 February 2007; accepted 16 February 2007
Available online 26 March 2007

Abstract

AKT1 (V-akt murine thyoma viral oncogene homolog 1) is involved in intracellular signalling pathways postulated as of aetiological importance in schizophrenia. Markers in the *AKT1* gene have also recently been associated with schizophrenia in two samples of European origin and in Japanese and Iranian samples. Aiming to replicate these findings, we examined ten SNPs spanning *AKT1* in a UK case-control sample (schizophrenia cases $n=673$, controls $n=716$). These included all SNPs previously reported to be associated in European, Japanese and Iranian samples, alone or in haplotypes, as well as additional markers defined by the Haploview Tagger program (pair-wise tagging, minimum $r^2=0.8$, minor allele frequency=0.02). We found no association with single markers (min $p=0.17$). We found weak evidence for association ($p=0.04$) with a four marker haplotype reported as significant in the original positive European sample of Emamian et al. [Emamian, E.S., Hall, D., Birnbaum, M.J., Karayiorgou, M., Gogos, J.A., 2004. Convergent evidence for impaired AKT1-GSK3beta signaling in schizophrenia. Nat. Genet. 36, 131–137] and also an overlapping three marker haplotype ($p=0.016$) that had previously been reported as significant in a Japanese sample. Nominal p-values for these haplotypes did not survive correction for multiple testing. Our study provides at best weak support for the hypothesis that *AKT1* is a susceptibility gene for schizophrenia. Examination of our own data and those of other groups leads us to conclude that overall, the evidence for association of *AKT1* as a susceptibility gene for schizophrenia is weakly positive, but not yet convincing.

Keywords: AKT1; Schizophrenia; Association; Candidate gene

1. Introduction

Emamian et al. (2004) proposed that alterations in brain protein kinase activity contribute to the aetiology of schizophrenia. In pursuit of this hypothesis, they examined the abundance of seven protein kinases in lymphoblast cell lines. Reduced AKT1 expression was found in cell lines derived from schizophrenic patients compared to controls, a finding subsequently confirmed in *post-mortem* frontal cortex and hippocampus. Moreover, they also found reduced phosphorylation of GSKβ3, a substrate of AKT1. These data provide a

* Corresponding author. Tel.: +44 2920744212; fax: +44 2020746554.
E-mail address: nortonn@cf.ac.uk (N. Norton).

0920-9964/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.schres.2007.02.006

plausible case that AKT1 might be involved in the pathophysiology of schizophrenia, a hypothesis whose aetiological relevance they explored by genetic association using 5 SNPs spanning the gene (see Tables 1–3 for SNP nomenclature) in 268 US families of North European origin containing one or more individuals with schizophrenia. The initial evidence for association was weak but one marker, (SNP3), yielded evidence for association ($p=0.05$, uncorrected) as did a number of haplotypes (minimum $p=0.04$, corrected) each of which shared alleles T and C at SNPs 2 and 3 (Emamian et al., 2004), (referenced hereafter as the core haplotype). The core haplotype was also associated with reduced AKT1 protein expression in 20 control lymphoblast cell lines. Follow-up studies in three independent Japanese samples gave mixed results. Two studies consisted of over 500 cases and over 400 controls. The first (Ohtsuki et al., 2004) found no association (allelic or haplotypic) while the second (Ikeda et al., 2004) reported weak evidence for association with a different variant and different haplotype to that of Emamian et al. (2004) (Table 2), with allele C of the core being carried in haplotypes that were both over and underrepresented in cases. A third study in a Japanese sample of 124 families found no association (allelic or haplotypic) (Ide et al., 2006). Schwab et al. (2005) found significant association with 3 of 7 SNPs tested in *AKT1* in 79 sib pair families of German origin. The associated SNPs included SNP3, $p=0.027$, which was nominally significant in the Emamian study as well as two other SNPs, with the strongest result (SNP2a, rs10149779, $p=0.002$) remaining significant after correction for multiple testing ($p=0.014$). The most significant haplotype from the Emamian study (SNP2/SNP3/SNP4, TCG, Table 2) was also significantly over-represented in cases as was the TTA haplotype (formed by the same SNPs), which had been under-transmitted to cases in the study of Emamian et al. (2004) and which does not carry the core TC haplotype. Several other haplotypes created by various permutations of markers were also significantly over-transmitted with illness, with the strongest evidence coming from a haplotype derived from SNP1/SNP2a/SNP3 ($p=0.0013$ corrected for multiple testing), Table 2. For all haplotypes in which SNP2 was included, the over-transmitted haplotype carried Emamian's core T allele at SNP2 but the finding of the earlier study was not precisely recapitulated since haplotypes carrying either C or T at core SNP3 were significantly over-transmitted.

Further studies have been less supportive (Bajestan et al., 2006; Liu et al., 2006). The 5 SNPs genotyped by Emamian et al. (2004) were genotyped in 218 families from Taiwan (Liu et al., 2006) with no significant association from either single markers or haplotypes. The same SNPs were also typed in an Iranian case control sample, (schizophrenia cases $n=321$, controls $n=383$) (Bajestan et al., 2006). Again, neither the SNPs nor the

Table 1
Results: single markers

SNP ID	Dist to next SNP (bp)	Base change allele 1/2	Sample sized	Allele 1 count (freq)	Allele 2 count (freq)	*p*-value (1df)
rs3803300 (SNP1)	4816	G/A	Case ($N=660$)	1194 (0.90)	126 (0.10)	0.40
			Control ($N=707$)	1292 (0.91)	122 (0.09)	
rs2498784 (SNP1a)	5229	T/C	Case ($N=658$)	102 (0.08)	1214 (0.92)	0.93
			Control ($N=712$)	109 (0.08)	1315 (0.92)	
rs1130214 (SNP2)	8648	T/G	Case ($N=586$)	361 (0.31)	811 (0.69)	0.73
			Control ($N=660$)	415 (0.31)	905 (0.69)	
rs10149779 (SNP2a)	4400	A/G	Case ($N=658$)	398 (0.30)	918 (0.70)	0.69
			Control ($N=711$)	440 (0.31)	982 (0.69)	
rs2494738	279	A/G	Case ($N=662$)	109 (0.08)	1215 (0.92)	0.68
			Control ($N=705$)	110 (0.08)	1300 (0.92)	
rs3730358 (SNP3)	6513	C/T	Case ($N=608$)	1043 (0.86)	173 (0.14)	0.30
			Control ($N=679$)	1145 (0.84)	213 (0.16)	
rs2498799 (SNP4)	702	G/A	Case ($N=592$)	918 (0.78)	266 (0.22)	0.17
			Control ($N=659$)	991 (0.75)	327 (0.25)	
rs2494732 (SNP5)	46	T/C	Case ($N=588$)	652 (0.55)	524 (0.45)	0.80
			Control ($N=663$)	742 (0.56)	584 (0.44)	
rs3803304	6051	C/G	Case ($N=653$)	327 (0.25)	979 (0.75)	0.56
			Control ($N=707$)	368 (0.26)	1046 (0.74)	
rs2498804 (SNP A)	–	G/T	Case ($N=660$)	911 (0.69)	409 (0.31)	0.46
			Control ($N=709$)	960 (0.68)	458 (0.32)	

Allele counts, frequencies and *p*-values across *AKT1* locus. SNP ID includes both rs no. and ID used in Emamian et al. (2004), Ikeda et al. (2004) and Schwab et al. (2005).

Table 2
Comparison of associated haplotypes

Study	Population	rs3803300 (SNP1)	rs2498784 (SNP1a)	rs1130214 (SNP2)	rs10149779 (SNP2a)	rs2494738	rs3730358 (SNP3)	rs2498799 (SNP4)	rs2494732 (SNP5)	rs3803304	rs2498804 (SNPA)	SCZ	CON	*P*-value
Emamian et al.	US			T			C*	G*				–	0.15	0.0006
Schwab et al.	German			T			C*	G*				0.17	0.10	0.023
This study	UK			T			C	G				0.19	0.17	0.37
Schwab et al.	German	G*			T		C*					0.17	0.09	0.0013
This study	UK	G			T		C					0.18	0.16	0.51
Emamian et al.	US			T			C*	G*	G			–	–	0.004
This study	UK			T			C	G	G			0.13	0.10	0.04
Schwab et al.	German			T			C	G	G			0.10	0.07	0.11[a]
Ikeda et al.	Japanese			T			C	G	G			0.02	0.01	0.18[a]
Ikeda et al.	Japanese						C*	G*	G			0.32	0.27	0.014
This study	UK						C	G	G			0.22	0.17	0.016
Bajestan et al.	Iranian	A		G*			C*	A	G			0.07	0.03	0.004
This study	UK	A		G			C	A	G			0.05	0.05	0.73
This study	UK	A					C*	G*				0.04	0.02	0.006
Ikeda et al.	Japanese						C*	G*	A			0.17	0.24	0.0001
This study	UK						C	G	A			0.55	0.56	0.55

Comparison of the most significant *p*-values from current studies reporting positive association with *AKT1* and schizophrenia. Significant haplotypes are marked in grey.
[a]Personal communications, *Ancestral allele in NCBI entrez SNP.

haplotypes from the Emamian study were associated. However, a novel five marker haplotype comprised of SNPs1-5, showed some evidence for association (global $p=0.05$ uncorrected) with haplotype AGCAG being more frequent in cases compared to controls (uncorrected $p=0.004$, Bonferroni corrected, $p=0.03$, case freq 0.068, control freq 0.034). Given the diverse range of ethnicities studied so far, lack of consistency of the patterns of association between studies is potentially explicable in terms of population differences in LD and modest power to detect weak genetic effects. Moreover, in the light of partial replication of the original findings at the level of a specific haplotype in the only other European origin sample so far reported, *AKT1* is clearly worth further investigation in other samples of broadly similar ethnicity.

We set out to investigate *AKT1* in schizophrenia using a moderately large UK based case control sample under the following strategies. We genotyped SNPs 1–5 from Emamian et al. (2004), and additional markers

Table 3
LD data for control sample

	rs3803300 (SNP1)	rs2498784 (SNP1a)	rs1130214 (SNP2)	rs10149779 (SNP2a)	rs2494738	rs3730358 (SNP3)	rs2498799 (SNP4)	rs2494732 (SNP5)	rs3803304	rs2498804 (SNPA)
rs3803300 (SNP1)	x	0.86	0.01	0.01	0.46	0.01	0.11	0.07	0.00	0.12
rs2498784 (SNP1a)	0.98	x	0.02	0.02	0.38	0.00	0.11	0.06	0.00	0.10
rs1130214 (SNP2)	0.36	0.78	x	0.95	0.00	0.25	0.06	0.21	0.11	0.06
rs10149779 (SNP2a)	0.41	0.81	0.99	x	0.01	0.26	0.06	0.20	0.12	0.07
rs2494738	0.71	0.62	0.30	0.36	x	0.02	0.12	0.09	0.00	0.16
rs3730358 (SNP3)	0.63	0.55	0.80	0.80	1	x	0.40	0.22	0.49	0.37
rs2498799 (SNP4)	0.63	0.70	0.29	0.30	0.71	0.85	x	0.43	0.40	0.69
rs2494732 (SNP5)	0.77	0.79	0.60	0.59	0.92	0.98	1	x	0.44	0.59
rs3803304	0.05	0.18	0.37	0.38	0.14	0.96	0.65	1	x	0.69
rs2498804 (SNPA)	0.79	0.76	0.25	0.27	0.95	0.96	0.99	1	0.97	x

LD data for control sample. D' is below diagonal and r^2 is above diagonal.

reported by others, SNP1a, SNP2a, (Schwab et al., 2005) and SNPA (Ikeda et al., 2004). We specifically tested all significant associated haplotypes reported by Emamian et al. (2004), (n=7), Ikeda et al. (2004), (n=9), Schwab et al. (2005), (n=23) and the Iranian 5 marker haplotype (Bajestan et al., 2006), (a total of 30 tests), although our primary hypothesis concerned the European origin haplotypes (n=28). Additionally, we derived tagged SNPs across the *AKT1* locus after genotyping all the above markers in the CEU panel used by the HapMap project and combining those data with all additional markers available in the HapMap (version 1.65) and performed two and three marker haplotype analyses for all marker combinations.

2. Materials and methods

2.1. Subjects

All case-control subjects used in this study were unrelated Caucasians born in the UK or Ireland. All cases met DSM-IV criteria for schizophrenia. Consensus diagnoses were made by two raters from all available information following a semi-structured interview, SCAN or PSE (Wing et al., 1974, 1990), and examination of case notes. The cases consisted of 456 males and 217 females, average age at collection 44.5 years ±14.6, whilst the controls consisted of 482 males and 234 females, average age at collection 41.5 years ±11.5 years. Control individuals were group matched to cases for age, sex, and ethnicity from more than 1400 blood donors recruited from the National Blood Transfusion Service. Individuals on medication are not allowed to donate blood in the UK nor are they remunerated even for expenses. Thus unlike in some countries, donating blood in the UK is entirely an altruistic process that does not tend to enrich for indigents, or people with substance abuse or psychosis. Donors were not screened for the absence of psychiatric illness, as this does not affect the power when a disease has the population prevalence of schizophrenia (Owen et al., 1997). Multicentre and Local Research Ethics Committee approval was obtained, and all subjects, both cases and controls, gave written informed consent to participate. We previously found no evidence for population stratification within the samples based on the distribution of p-values obtained from genotyping pooled samples for >300 SNPs (Williams et al., 2005a). We also tested for evidence of substructure in approximately one-third of our sample with STRUCTURE (Pritchard et al., 2000) by using 97 SNPs scattered across the genome and 1000 SNPs targeted to regions on chromosomes 10 and 22 (Georgieva et al., 2006).

Sample power was estimated to be 80% for the "core TC haplotype" given our observed frequency, an OR of 1.3, alpha=0.05 and 79% for the associated TCG haplotype in Table 2, under the same parameters. For rs3730358 (associated in both Emamian et al., 2004 and Schwab et al., 2005), we estimated power to be 73% given an OR of 1.3, alpha=0.05.

2.2. SNP selection

We initially selected for genotyping, SNPs 1–5 from Emamian et al. (2004), (rs3803300, rs1130214, rs3730358, rs2498799, rs2494732 respectively), two additional SNPs from Schwab et al. (2005), (rs2498784 and rs10149779, SNP1a and SNP2a respectively), and 1 additional SNP from Ikeda et al. (2004), (rs2498804, SNPA), in order to be able to test the specific marker and haplotype hypotheses generated by those studies. All SNPs were optimised on the same CEPH DNA samples used in the international HapMap project for purposes of both error checking (all genotypes were checked against HapMap data for concordance) and also for tag SNP selection. We used our CEPH data and all other available CEPH data from the HapMap release 16C.1 June 2005 (Generic genome browser version 1.65) across the *AKT1 locus* from UCSC May2004 chr14:104304140–104341530 (including 8.4 kb sequence upstream and 2.6 kb sequence downstream of *AKT1*) and performed pairwise analysis with TAGGER as implemented in Haploview (Barrett et al., 2005) using settings $r^2>0.8$, minimum MAF 2%. This suggested as additional tagging SNPs rs2494738 and rs3803304, none of which have been genotyped in previous *AKT1* association studies (Emamian et al., 2004; Ohtsuki et al., 2004; Ikeda et al., 2004; Schwab et al., 2005; Liu et al., 2006; Bajestan et al., 2006; Ide et al., 2006).

2.3. Genotyping

8/10 SNPs were genotyped using the Sequenom MassARRAY™ system as per the manufacturer's instructions with either hME or iplex chemistries. SNPs 1a and 2a were genotyped with allele-specific PCR using the Amplifluor system (Myakishev et al., 2001; Hawskins et al., 2002).

Assay design and PCR conditions are available on request. All assays used to type the full association sample were optimised initially by genotyping DNA from 30 CEPH parent–offspring trios from 21 families (Utah residents with ancestry from northern and western Europe), as detailed in the international HapMap project

(The International HapMap Consortium, 2005). We re-genotyped 46 of these samples along with the case control sample to provide a measure of genotyping accuracy. All genotypes were called blind to sample identity and affected status.

2.4. Statistical analysis

Tests of genotypic and allelic association were performed using contingency tables. Haplotype analyses were performed using the EM algorithm and a permutation test as implemented in program EH plus (Zhao et al., 2000) for global significance. Association of specific haplotypes was estimated with Cocaphase (Dudbridge, 2003). LD values were calculated using the ldmax program within the GOLD software (Abecasis and Cookson, 2000).

3. Results

Genotype data for SNP2, rs2494738, rs3803304 and SNPA from our assays in the same 90 CEPH DNA samples used in the International HapMap Project were 100% concordant with HapMap data. 100% concordance was also achieved between genotype data of 46 CEPH DNA samples typed in our initial assay optimisation stage and the same samples contained within our case control sample set for all 10 SNPs.

Genotype data were in Hardy Weinberg equilibrium for both cases and controls for all SNPs. No significant differences between cases and controls were observed for any single markers by allele (Table 1) or genotype (data not shown).

We specifically tested those marker combinations reported to yield the most significantly associated haplotypes by Emamian et al. (2004) including the core haplotype, (SNP2/SNP3, TC), those of Schwab et al. (2005), (SNP1/SNP2a/SNP3, GTC) and Ikeda et al. (2004), (SNP3/SNP4/SNP5, CGG and CGA as well as seven other overlapping haplotypes also significantly associated in the Japanese sample). Table 2 summarizes the most significant haplotypes with the ancestral alleles marked as *. Marker combination SNP2/SNP3/SNP4 which gave the most significant results in the Emamian study, gave a global p-value of 0.08 in our sample. However, the TCG haplotype for this marker combination which gave nominal significance in the initial Emamian study ($p=0.0006$) and also in that of Schwab et al. (2005), ($p=0.02$) was not significantly associated in our sample, $p=0.37$, although a non significant trend was observed in the previously reported direction (Table 2). Our case and control frequencies for this haplotype were 0.19 and 0.17 respectively, compared to 0.17 and 0.10 in the German sample (Schwab et al., 2005) and 0.15 in the US sample of European origin of Emamian et al. (2004).

Global haplotype analysis of markers (SNP2/SNP3) forming the core two-marker haplotype of Emamian et al. (2004) revealed no significant evidence for association (global $p=0.09$) nor did specific analysis of the TC phased core haplotype ($p=0.41$). However, a specific phased 4 marker haplotype (SNP2/SNP3/SNP4/SNP5, TCGG, Table 2) which was also significant in the Emamian study, $p=0.004$, was associated in our sample, $p=0.04$ (case freq=0.13, control freq=0.10). The same haplotype was not significantly associated in either the German (Schwab et al., 2005) or in the Japanese (Ikeda et al., 2004) samples, ($p=0.11$ and 0.18 respectively), (personal communications).

The most significant haplotype of (SNP1/SNP2a/SNP3 GTC, Table 2) reported by Schwab et al. (2005) was not associated in our sample, ($p=0.51$), although the trend with frequencies of 0.18 and 0.16 in our cases and controls respectively, compared to 0.17 and 0.09 in the German sample was in the same direction.

Global haplotype analysis of SNP3/SNP4/SNP5 which was the most significant haplotype in the Japanese study of Ikeda et al. (2004) was significantly associated (uncorrected $p=0.04$) in our sample. The largest difference in haplotype frequency for this combination was 5% for the haplotype CGG (Table 2, $p=0.016$). The same haplotype was also significantly associated in the same direction in the Japanese sample (Ikeda et al., 2004), $p=0.014$, although frequencies in the samples are different (CGG=0.27 in Japanese controls vs. 0.17 in Cardiff controls). The most significant haplotype associated in the Japanese sample was SNP3/SNP4/SNP5 CGA, ($p=0.0001$), (Ikeda et al., 2004). This was not significantly associated in our sample. The haplotype frequencies in our sample and in that of Ikeda et al. (2004) are substantially different (Table 2). Ikeda et al. (2004) also reported 7 overlapping haplotypes with individual p-values less than 0.05. None of these was significantly associated in our sample.

Although we selected the markers predicated on single locus (i.e. pairwise tagging), in order to try to capture the effect of unknown variants that are not represented in HapMap, we undertook 2 and 3 marker haplotype analysis across all the markers including those additional SNPs recommended by our tagging procedure. We obtained evidence for association for haplotypes of SNP1/SNP3/SNP4, global $p=0.04$, which overlaps physically with the most significant haplotypes reported by Schwab et al. (2005) and Emamian et al. (2004). On closer inspection, the effect came from two haplotypes with frequencies of

less than 5%, (GTG case=0.009, control=0.018, $p=0.06$ and ACG, case=0.043, control=0.023, $p=0.006$). Allele C of SNP3 is common to our significant haplotype (SNP1/SNP3/SNP4, ACG) and the most significant haplotypes of Emamian et al. (2004) and Schwab et al. (2005), whilst allele G of SNP4 is common to our SNP1/SNP3/SNP4, ACG haplotype and the SNP2/SNP3/SNP4, TCG haplotype of Emamian et al. (2004) and Schwab et al. (2005).

4. Discussion

Following the initial report (Emamian et al., 2004) and mixed replication evidence (Ohtsuki et al., 2004; Ikeda et al., 2004; Ide et al., 2006; Schwab et al., 2005; Liu et al., 2006; Bajestan et al., 2006) we sought to provide further evidence for association between schizophrenia and polymorphisms in *AKT1*. The question of when the evidence for association between disease and gene is convincing is a vexing one for several reasons. Ideally, such evidence would come from repeated demonstration of a directional association (even if not significant) between a disorder and a specified allele such that pooled or meta-analyses demonstrates a clear highly significant directional effect. However, when based upon indirect association, replication of specified alleles may not be easily obtained due to a mixture of population differences in allelic heterogeneity at the locus, patterns of LD, allele frequencies, phenotypic variation relevant to the associated allele, or exposure to environmental variables with which a risk allele interacts (O'Donovan and Owen, 1999). Moreover, mathematical analyses show that where the true effect size of a susceptibility allele is weak, opposite alleles may be genuinely associated with disease even in populations with similar LD measures, allele frequencies, and identical effect sizes at the functional locus (Moskvina and O'Donovan, in press). Given the above, while association to the same allele across studies should at least be sought, it cannot be a prerequisite for considering a study as supportive of association between disease and a gene. Instead, we believe it is legitimate to view association to any allele or haplotype that both survives honest appropriate correction for multiple testing and is based on a well-designed quality-controlled study as significant evidence for replication at the gene-level. When multiple studies meet this criterion, as we consider to be the case for example for dysbindin (Williams et al., 2005b), then in our view, the evidence can be considered convincing.

Our single marker data for AKT1 provide no evidence for association with schizophrenia, but haplotype analysis showed some trends similar to the existing data, albeit, none that remain significantly associated in the context of multiple testing. When associated haplotypes from all of these studies are aligned (Table 2), it is apparent that the most significant risk haplotypes across studies overlap, making the correction for multiple testing over conservative. (Of the p-values reported in Table 2, the most significant haplotype of Schwab et al. (2005) is already corrected for multiple testing by simulation (Becker and Knapp, 2004). The most significant haplotype in the Emamian study ($p=0.006$) is uncorrected, but remains significant after adjustment for the comparisons made in the study, ($p<0.04$).

In Table 2 where we show the most significant haplotype reported from each study and compare these specific haplotypes across the published data and with our own data, the SNP3/4, alleles CG combination occurs in 6/7 of the significant haplotypes. This trend also extends to the Japanese sample (Ikeda et al., 2004), but not in the Iranian sample (Bajestan et al., 2006) where the only significant haplotype contains the A allele at SNP 4. Also, when all ethnicities are considered together, a trend for a longer overlap is observed with SNPs 3/4/5, CGG in the significant risk haplotypes of both this, the initial positive and the Japanese study, (4/10 significant haplotypes). However, given that alleles C and G are respectively the major alleles at SNP3 and SNP4, and are present on numerically more haplotypes than SNP3 allele T, this may simply be chance rather than a reflection of a genuine pattern in the data. Nevertheless, similar directions of effect were observed in our sample for the haplotypes most significantly over-represented in samples of European origin. Further confidence could be achieved if the four marker haplotype (SNP2/SNP3/SNP4/SNP5, TCGG) which was associated in the original positive study and in our own data was also significant in the German study and Japanese samples, particularly since (SNP3/SNP4/SNP5, CGG) was also significant in the Japanese sample. However, the TCGG haplotype was not significantly associated in the German sample, $p=0.11$ (personal communication) although the trend was in the same direction, and the same haplotype was of low frequency in the Japanese sample of Ikeda et al. (2004), (personal communication) and was not significant in the Japanese sample of Ohtsuki et al. (2004), (personal communication).

Our haplotype frequencies were broadly similar to those other samples of European, origin (Emamian et al., 2004; Schwab et al., 2005), although when compared to the German sample (Schwab et al., 2005), both our case and control frequencies were more similar to the German cases than to the German controls (Table 2). Comparison of LD patterns across studies showed D'