

Meta-analysis of genome-wide association studies for panic disorder in the Japanese population

T Otowa^{1,2}, Y Kawamura³, N Nishida⁴, N Sugaya⁵, A Koike⁶, E Yoshida⁷, K Inoue⁸, S Yasuda^{7,9}, Y Nishimura^{1,10}, X Liu⁴, Y Konishi¹⁰, F Nishimura¹, T Shimada¹, H Kuwabara¹¹, M Tochigi¹, C Kakiuchi¹, T Umekage¹, T Miyagawa⁴, A Miyashita¹², E Shimizu¹³, J Akiyoshi¹⁴, T Someya¹⁵, T Kato¹⁶, T Yoshikawa¹⁷, R Kuwano¹², K Kasai¹, N Kato¹⁸, H Kaiya^{7,9,19}, K Tokunaga⁴, Y Okazaki²⁰, H Tani¹⁰ and T Sasaki²¹

Panic disorder (PD) is a moderately heritable anxiety disorder whose pathogenesis is not well understood. Due to the lack of power in previous association studies, genes that are truly associated with PD might not be detected. In this study, we conducted a genome-wide association study (GWAS) in two independent data sets using the Affymetrix Mapping 500K Array or Genome-Wide Human SNP Array 6.0. We obtained imputed genotypes for each GWAS and performed a meta-analysis of two GWAS data sets (718 cases and 1717 controls). For follow-up, 12 single-nucleotide polymorphisms (SNPs) were tested in 329 cases and 861 controls. Gene ontology enrichment and candidate gene analyses were conducted using the GWAS or meta-analysis results. We also applied the polygenic score analysis to our two GWAS samples to test the hypothesis of polygenic components contributing to PD. Although genome-wide significant SNPs were not detected in either of the GWAS nor the meta-analysis, suggestive associations were observed in several loci such as *BDKRB2* ($P = 1.3 \times 10^{-5}$, odds ratio = 1.31). Among previous candidate genes, supportive evidence for association of *NPY5R* with PD was obtained (gene-wise corrected $P = 6.4 \times 10^{-4}$). Polygenic scores calculated from weakly associated SNPs ($P < 0.3$ and 0.4) in the discovery sample were significantly associated with PD status in the target sample in both directions (sample I to sample II and vice versa) ($P < 0.05$). Our findings suggest that large sets of common variants of small effects collectively account for risk of PD.

Translational Psychiatry (2012) 2, e186; doi:10.1038/tp.2012.89; published online 13 November 2012

Introduction

Panic disorder (PD) is an anxiety disorder characterized by recurrent and unexpected panic attacks, subsequent worry and phobic avoidance. Lifetime prevalence of PD is 1–3%, and twice as many women as men suffer from the disorder.¹ The disorder frequently takes a chronic course, with many remissions and relapses. Genetic epidemiological studies have shown that genetic factors, as well as environmental factors, have important roles in the pathogenesis of PD. Family studies^{2,3} indicate that first-degree relatives of probands with PD have an approximately fivefold increased risk of PD, and twin studies^{4,5} estimate about 40% of the genetic liability towards PD. A large number of genetic studies of PD including linkage and association analyses have been

conducted.^{6–8} Suggestive evidence for linkage in several regions and association of candidate genes such as monoamine- and neuropeptide-related genes has been reported. However, these studies have failed to find conclusive susceptibility genes for PD.

We previously conducted the first genome-wide association study (GWAS) of PD (200 Japanese cases and 200 controls) and found several suggestively associated genes with PD.⁹ However, a follow-up study in a larger Japanese sample (558 cases and 566 controls) failed to show any significant association of these genes with PD.¹⁰ Erhardt *et al.*¹¹ conducted a three-stage GWAS of PD in German samples (216 cases and 222 controls in the GWAS stage, and 909 cases and 915 controls in total). They found two

¹Department of Neuropsychiatry, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; ²Department of Psychiatry, Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA; ³Department of Psychiatry, Sakae Seijinkai Hospital, Kanagawa, Japan; ⁴Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; ⁵Research Project for Addictive Substances, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan; ⁶Central Research Laboratory, Hitachi, Tokyo, Japan; ⁷Outpatient Clinic for Anxiety Disorders, Akasaka Mental Clinic, Tokyo, Japan; ⁸Department of Public Health, Fujita Health University School of Medicine, Aichi, Japan; ⁹Outpatient Clinic for Anxiety Disorders, Yokohama Clinic, Kanagawa, Japan; ¹⁰Department of Neuropsychiatry, Graduate School of Medicine, Mie University, Mie, Japan; ¹¹Department of Child Neuropsychiatry, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; ¹²Department of Molecular Genetics, Center for Bioresources, Brain Research Institute, Niigata University, Niigata, Japan; ¹³Department of Cognitive Behavioral Physiology, Research Center for Child Mental Development, Graduate School of Medicine, Chiba University, Chiba, Japan; ¹⁴Department of Neuropsychiatry, Oita University Faculty of Medicine, Oita, Japan; ¹⁵Department of Psychiatry, Graduate School of Medical and Dental Sciences, Niigata University, Niigata, Japan; ¹⁶Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, Saitama, Japan; ¹⁷Laboratory for Molecular Psychiatry, RIKEN Brain Science Institute, Saitama, Japan; ¹⁸Department of Psychiatry, Karasuyama Hospital, Showa University School of Medicine, Tokyo, Japan; ¹⁹Research Center for Panic Disorder, Nagoya Mental Clinic, Aichi, Japan; ²⁰Department of Neurology, Tokyo Metropolitan Matsuzawa Hospital, Tokyo, Japan and ²¹Laboratory of Health Education, Graduate School of Education, University of Tokyo, Tokyo, Japan

Correspondence: Professor T Sasaki, Laboratory of Health Education, Graduate School of Education, University of Tokyo, 7-3-1 Hongo Bunkyo, Tokyo 113-8625, Japan.

E-mail: psytokyo@yahoo.co.jp

Keywords: *BDKRB2*; gene ontology; GWAS; *NPY5R*; panic disorder; polygenic score

Received 16 July 2012; accepted 25 July 2012

single-nucleotide polymorphisms (SNPs) (rs7309727 and rs11060369), located in *TMEM132D* on 12q24.3, to be associated with PD. Gregersen *et al.*¹² conducted a genome-wide scan using microsatellite markers in the isolated population of the Faroe Islands (13 distantly related PD cases and 43 controls). They found an association of *ACCN1* with PD in an extended Faroese sample (31 cases and 162 controls), whereas they could not replicate the association in an outbred Danish sample (243 cases and 645 controls). Kawamura *et al.*¹³ recently performed a genome-wide copy number variation association study of PD in 2055 Japanese subjects (535 cases and 1520 controls) and reported that common duplications in 16p11.2 were associated with PD. This region includes several genes such as *IGH*, *HERC2P4*, *TP53TG3*, *SLC6A8* and *SLC6A10P* and small RNAs. However, none of them overlapped with genetic markers associated in earlier candidate gene studies. Due to the lack of power in previous association studies, genes that are truly associated with PD might not be detected.

To overcome the limitations of previous association studies targeting candidate genes with small sample sizes, we conducted a new GWAS of PD in a larger sample and combined our first and second GWAS (GWAS I and II: 718 cases and 1717 controls in the meta-analysis stage). For a follow-up study, 12 suggestively significant SNPs were tested in an independent sample set (329 cases and 861 controls). Considering evidence for a substantial polygenic contribution to psychiatric disorders, we also undertook a polygenic score analysis to test whether large sets of common variants of small effects accounted for risk of PD.

Materials and methods

Subjects. The two samples included in the meta-analysis consisted of unrelated Japanese subjects living in Tokyo, Nagoya and Niigata, all of which are located in the mainland of Japan. Sample I consisted of 200 PD cases and 200 controls described in our previous GWAS.⁹ In the present study, a new GWAS (II) of PD was conducted using sample II (579 PD cases and 1568 controls), in which part of the sample (545 cases and 414 controls) overlapped with those reported in our previous paper.¹⁰ The diagnosis was confirmed according to the Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (DSM-IV) criteria,¹⁴ using the Mini International Neuropsychiatric Interview (MINI)¹⁵ and reviewing clinical records. Control subjects were interviewed by one of the authors and screened for history of major psychiatric disorders by using either a short questionnaire or the structured interview (MINI).

Replication sample comprised 329 cases and 863 controls, all recruited in Japan through multiple institutions. Cases were diagnosed according to the DSM-IV criteria by at least two experienced psychiatrists on the basis of all available sources of information, including unstructured interviews, clinical observations and medical records. Controls were screened for history of psychiatric disorders including PD based upon self report. After complete description of the study to the subjects, written informed consent was obtained. This study was approved by the ethical committees of relevant institutions (University of Tokyo, Chiba University, Niigata University, Mie University, Oita University Faculty of Medicine and RIKEN). Details of sample recruitment are provided in Supplementary Materials and Methods.

Genotyping and quality control (QC). Samples I and II were genotyped on the GeneChip Human Mapping 500K Array Set and the Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara, CA, USA), respectively. For replication, genotyping was performed by the Taqman assay with the ABI PRISM 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA, USA). Genotyping for GWAS and replication was conducted at Human SNP Typing Center and Department of Psychiatry at the University of Tokyo, respectively.

For the GWAS stage, stringent QC procedures were applied to individual subjects and SNP data (for example, sample-wise call rate ≥ 0.95 , SNP call rate ≥ 0.95 , Hardy–Weinberg equilibrium P -value ≥ 0.001 and minor allele frequency ≥ 0.05). The subjects QC left 177 cases and 178 controls for sample I, and 541 cases and 1539 controls for sample II (Table 1). The final GWAS consisted of 282 173 SNPs for sample I and 633 001 SNPs for sample II on autosomal chromosomes. In addition, 6043 SNPs for sample I and 23 131 SNPs for sample II on X chromosome in females were also analyzed.

For replication, we investigated independent autosomal SNPs ($P < 5 \times 10^{-5}$ in the meta-analysis) genotyped on both of the two genotyping platforms. The QC (sample-wise and SNP call rate ≥ 0.95 and Hardy–Weinberg equilibrium P -value ≥ 0.001) left 12 SNPs for 329 cases and 861 controls for the final replication data set. Details of QC procedure are provided in Supplementary Materials and Methods.

Demographic variables of each sample set after QC filtering are summarized in Table 1.

SNP-based association analysis. GWAS in each study was conducted independently using logistic regression under an additive model using PLINK ver. 1.07.¹⁶ As it is known

Table 1 Demographic characteristics of samples totaling 1047 cases and 2578 controls

	Sample I		Sample II		Replication	
	Cases	Controls	Cases	Controls	Cases	Controls
Number	177	178	541	1539	329	861
Female/male ratio	2.1	2.1	2.1	1.2	1.5	1.0
Age \pm s.d.	38.9 \pm 10.9	38.2 \pm 9.3	38.0 \pm 12.2	39.9 \pm 10.8	38.3 \pm 12.1	37.9 \pm 14.6

Abbreviation: s.d., standard deviation.

that there is a population structure in the Japanese population,¹⁷ multidimensional scaling (MDS) analysis was performed using PLINK, and 10 MDS components were added as covariates in each GWAS to control for population stratification. SNPs with P -value $< 5 \times 10^{-8}$ were considered to be genome-wide significant in individual GWAS and meta-analysis. A $P < 0.05$ was taken as significant evidence of replication.

A power calculation was performed using the method as described by Ohashi *et al.*¹⁸ The meta-analysis sample (718 cases and 1717 controls) was tested to detect $\alpha = 5 \times 10^{-8}$, assuming disease prevalence of 0.03. For a power of 80%, the minimum detectable relative risks were 1.75, 1.52 and 1.47 for risk allele frequency of 0.1, 0.25 and 0.4, respectively.

The quantile–quantile plot was used to evaluate overall significance of the GWA results and the potential impact of population stratification. The inflation factor λ was calculated on the basis of the median χ^2 . Haploview 4.2¹⁹ was used to create Manhattan plots of P -values from the GWAS and meta-analysis. SNAP 2.0²⁰ was used to plot association results in each gene region in the meta-analysis stage of the GWAS data.

Meta-analysis. A meta-analysis of the GWAS I and II was conducted using PLINK ver. 1.07. The combined sample comprised 718 cases and 1717 controls. To detect strong PD-associated SNPs that were not genotyped on the platforms, imputation was performed with MACH ver. 1.0 software,²¹ using the HapMap CHB + JPT phase II data as the reference sample. Association analyses in each GWAS were conducted by logistic regression with 10 MDS components as covariates using ProbABEL.²² The location of SNPs reported is taken from the NCBI build 36 (UCSC hg18). Poorly imputed SNPs ($r^2 < 0.30$) and SNPs with low minor allele frequency (< 0.01) were excluded, resulting in a final meta-analysis data set of 1.9M SNPs. For each SNP, we combined the odds ratios (ORs) for a given reference allele weighted by the confidence intervals using a fixed-effects or random-effects model. To assess the homogeneity of ORs, the Cochran's Q statistic was used.²³ If the Cochran's Q statistic showed P -value < 0.01 , a random-effects model (DerSimonian and Laird method) was employed.²⁴ If the Cochran's Q statistic showed P -value ≥ 0.01 , a fixed-effects model (Mantel–Haenszel method) was employed.²⁵ We also combined GWAS and replication data using the same approaches (fixed-effects or random-effects model).

Gene ontology (GO) term enrichment analysis. For enrichment analyses, we used the public domain tool provided by the Database for Annotation, Visualization and Integrated Discovery (DAVID) bioinformatics platform.²⁶ We examined GO because it is used widely in functional annotation and enrichment analysis. As DAVID accepts only a gene list, markers were converted to representative genes using the ProxyGeneLD software²⁷ (see details in Supplementary Materials and Methods). We used 2.0 and 1.9M imputed markers in individual GWAS and the meta-analysis as an input file, respectively. After conversion, top 1% of genes (172 genes in each GWAS and 166 in the meta-analysis) were tested by enrichment analysis using DAVID with the total genes (17 180 in GWAS I, 17 173 in GWAS II

and 16 645 in the meta-analysis) as the base set. Considering the redundant nature of annotations, the groups of similar annotations were grouped together using 'Functional Annotation Clustering' (kappa value > 0.5). We selected the best significantly enriched terms of individual groups. The enrichment P -value was calculated from the number of genes in the list that hit a given biology class compared to pure random chance. To control the family-wide false-positive rate in the result list, the multiple test correction of enrichment P -values must be performed on the functional annotation categories being tested at the same time. To correct this multiple testing, false discovery rate was calculated, which is the expected proportion of false discoveries among the discoveries (%).²⁸ The false discovery rate less than 25% was taken as significance. We attempted to replicate the significant GO terms found in one GWAS using the other GWAS, and also analyzed enriched GO terms based on the meta-analysis of the two GWAS data sets.

Association analysis of prior candidate genes. To find supportive evidence for associations of previously implicated candidate genes, we additionally tested for the association with a set of 356 prior candidate genes summarized by Maron *et al.*⁸ We used ProxyGeneLD software to convert SNPs from the meta-analysis of two GWAS data to specific genes as described above. After conversion, 38 341 SNPs were assigned to 327 candidate genes.

Polygenic score analysis. Polygenic scores were calculated and tested for their effect following the method described by International Schizophrenia Consortium.²⁹ Five sets of LD-pruned SNPs (r^2 threshold at 0.25 and window size 200 SNPs) were selected from the discovery data. This selection of SNPs was based on their nominal P -value (significance threshold $P_T < 0.1, 0.2, 0.3, 0.4$ and 0.5). Polygenic score was calculated by weighting scores of risk allele count by the logOR observed in the discovery data. We calculated individual scores for each set of SNPs using PLINK ver. 1.07. Nagelkerke's pseudo R^2 was calculated by logistic regression analysis using R software (<http://www.r-project.org>), with the number of non-missing SNPs and 10 MDS components as covariates in the target sample.

Data sharing. Portions of our GWAS results for PD are publicly available in the Genome-Wide Association Database (GWAS DB; <http://gwas.lifesciencedb.jp/>).³⁰

Results

GWAS results. In this study, we conducted a new GWAS with a larger sample size (sample II) than that in our previous reports (sample I). The quantile–quantile and Manhattan plots of the GWAS I and II are illustrated in Supplementary Figures S1 and S2, respectively. The genomic inflation factor (λ) of the GWAS I and II was 1.06 and 1.05, respectively, suggesting no substantial effects of population structure. The Manhattan plots of the GWAS I and II, respectively, showed several genomic regions as potential risk loci (Supplementary Figures S1 and S2, respectively), although none was genome-wide significant. The results of genotyped SNPs showing P -value

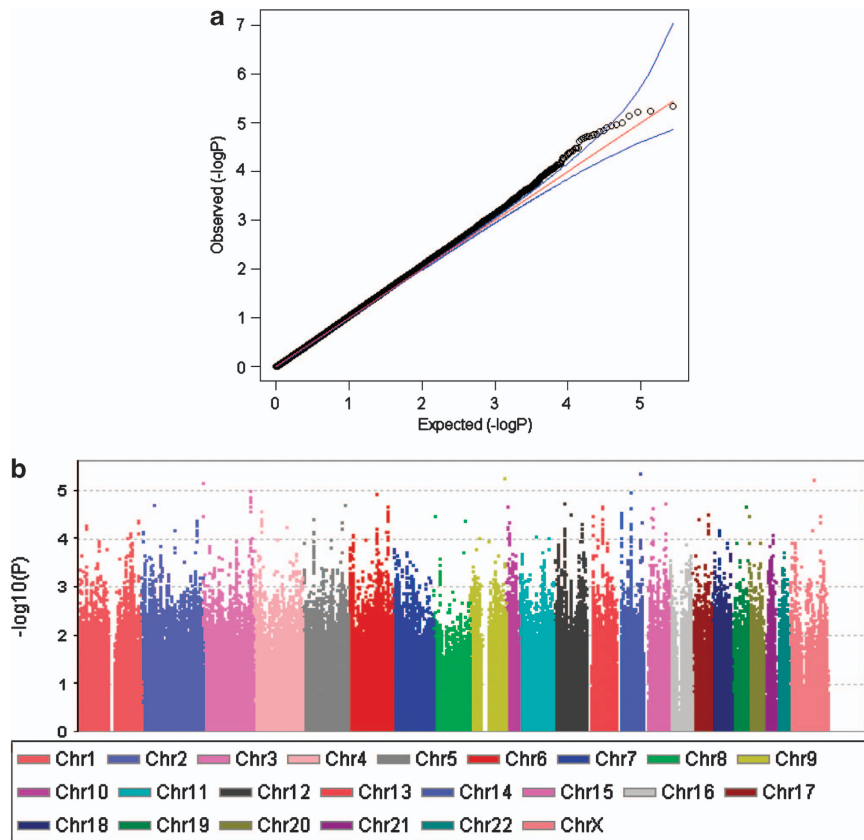


Figure 1 Results of the meta-analysis of the two genome-wide association studies (GWAS I and II). (a) Quantile–quantile plots of the association results. Observed association results ($-\log_{10} P$) are plotted against the expected distribution under the null hypothesis of no association. (b) Manhattan plots present the P -values across the genome. The association results ($-\log_{10} P$) are plotted in chromosomal order. Single-nucleotide polymorphisms in X chromosome were analyzed only in females.

$< 10^{-4}$ in the GWAS I and II are provided in Supplementary Tables S1 and S2, respectively.

Meta-analysis of the GWAS data. The quantile–quantile and Manhattan plots in the meta-analysis of the two GWAS are illustrated in Figure 1. The genomic inflation factor was 1.04. Among 1.9M imputed SNPs in the meta-analysis, no genome-wide significance was observed. Four SNPs showed significance at the level of $P < 10^{-5}$ (Supplementary Table S3). Three of the top four SNPs (rs10144552, rs944805 and rs4129976) were genotyped on both of the platforms. The strongest association was observed at rs10144552 located 36 kb upstream of *BDKRB2* ($P = 4.4 \times 10^{-6}$) on 14q32.2 (Figure 2). SNPs showing P -value $< 5 \times 10^{-5}$ in the meta-analysis were listed in Supplementary Table S3. The tests of heterogeneity of effects in each sample were non-significant ($P \geq 0.01$), showing that estimates were similar between the two samples and justifying the fixed effects of the meta-analysis (data not shown).

Follow-up study. We investigated the associations of 12 SNPs selected from the top meta-analysis findings in an independent Japanese sample (329 cases and 861 controls). Among 12 SNPs, 9 SNPs showed the same direction of effects between the meta-analysis and replication samples,

the number of which could be expected by chance ($P = 0.64$). None of these SNPs was significant (Table 2). When these SNPs were analyzed by sex, no significant association was identified (Supplementary Tables S4 and S5).

In the combined analysis across the two GWAS and replication samples, no genome-wide significant association was found (Table 2). The ORs ranged between 1.15 and 1.49. Rs10144552 in *BDKRB2* showed the most significant P -value ($P = 1.3 \times 10^{-5}$, OR = 1.31). The second strongest association was found at rs2911968 in *MCPH1* on 8p23.1 ($P = 2.6 \times 10^{-5}$, OR = 1.30). The third significant SNP was found at rs4129976 located in *CNTN4* on 3p26.3 ($P = 5.0 \times 10^{-5}$, OR = 1.29).

GO term enrichment analysis. To further explore the GWAS data, we took a GO-based approach, which provides complementary information to single-marker analysis. We found that five terms in GWAS I and four terms in GWAS II were associated with PD at a nominal P -value < 0.05 (Supplementary Tables S6 and S7). Among these significantly enriched terms, three in GWAS I (phosphoinositide binding, axon guidance and transcription factor activity) and one in GWAS II (negative regulation of secretion) had false discovery rate $< 25\%$. However, these significantly enriched terms in one GWAS were not replicated using the other GWAS. When the analysis was conducted based on the

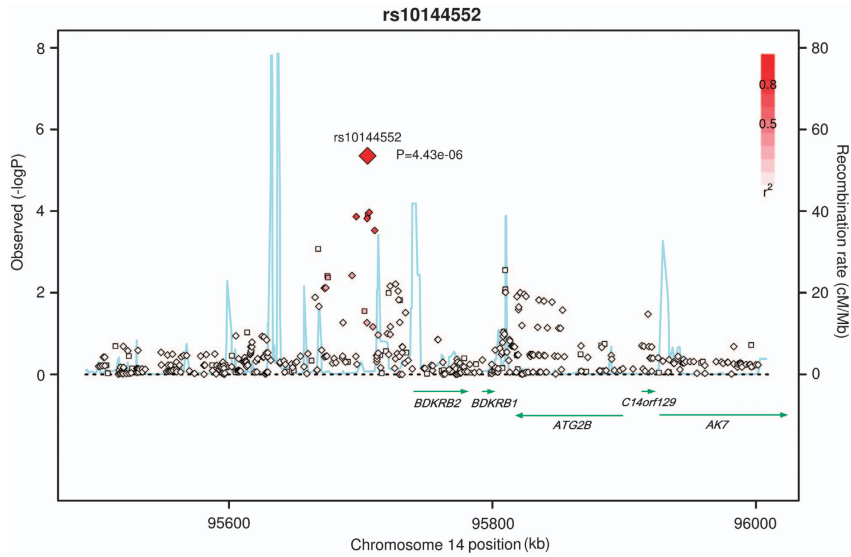


Figure 2 Plots of association results ($-\log_{10} P$) in *BDKRB2* region in the meta-analysis of the genome-wide association studies (GWAS). Chromosome position is plotted according to its physical position with the reference to the NCBI build 36. Recombination rate as estimated from the JPT and CHB HapMap data is plotted in light blue. Large red diamond (genotyped on both the platforms): single-nucleotide polymorphism (SNP) with strongest evidence for association (rs10144552). Small diamond and square represent imputed SNPs and SNPs genotyped on both the GWAS platforms, respectively. Strengths of linkage disequilibrium (LD) (r^2) with SNP rs10144552 in the plots are shown (darker red indicates stronger LD).

Table 2 Top findings based on the meta-analysis of two GWAS data sets and their follow-up study

Chr	SNP	BP	M ^a	m ^a	MAF ^b (control)	GWAS I		GWAS II		GWAS _{Meta}		Replication		Total		Gene ^d
						P	OR ^c	P	OR ^c	P	OR ^c	P	OR ^c	P	OR ^c	
14	rs10144552	95 705 200	G	A	0.213	0.035	1.51	3.62E-05	1.42	4.43E-06	1.43	0.29	1.12	1.32E-05	1.31	<i>BDKRB2</i> ^e
8	rs2911968	6 475 913	C	T	0.226	0.98	1.00	5.05E-06	1.46	3.33E-05	1.17	0.15	1.16	2.58E-05	1.30	<i>MCPH1</i>
3	rs4129976	2 144 803	C	T	0.201	0.021	1.60	2.39E-04	1.36	7.12E-06	1.43	0.21	1.16	4.96E-05	1.29	<i>CNTN4</i>
9	rs944805	134 387 909	T	C	0.196	0.29	1.23	6.79E-06	1.50	5.78E-06	1.34	0.82	1.03	1.41E-04	1.29	<i>C9orf171</i>
13	rs936336	59 060 504	A	G	0.176	0.12	0.71	6.22E-05	0.65	2.27E-05	0.67	0.56	0.93	1.70E-04	0.75	
13	rs7324557	23 194 862	G	A	0.343	0.20	0.80	7.14E-05	0.72	3.55E-05	0.75	0.51	0.94	2.16E-04	0.80	<i>MIPEP</i>
13	rs9563740	59 038 977	A	G	0.176	0.11	0.70	8.58E-05	0.66	2.22E-05	0.67	0.58	0.93	2.18E-04	0.75	
15	rs8040090	84 970 275	G	A	0.052	0.035	2.07	1.32E-04	1.74	1.85E-05	1.79	0.87	1.03	2.82E-04	1.49	<i>AGBL1</i>
6	rs802669	111 278 575	G	A	0.309	0.27	1.20	1.74E-05	1.38	1.22E-05	1.32	0.92	0.99	4.07E-04	1.22	
5	rs4976549	167 764 211	T	G	0.220	0.030	1.50	2.02E-04	1.36	1.96E-05	1.38	0.83	0.98	6.41E-04	1.24	<i>WWC1</i>
8	rs6470380	126 864 763	A	G	0.078	0.15	1.51	1.23E-04	1.60	4.14E-05	1.29	0.96	0.99	7.24E-04	1.37	
20	rs12480756	4 697 032	C	T	0.177	0.57	0.89	2.27E-06	1.55	3.45E-05	1.43	0.97	1.00	4.58E-01	1.15	<i>RASSF2</i>

Abbreviations: BP, base position; Chr, chromosome; GWAS, genome-wide association study; OR, odds ratio; SNP, single-nucleotide polymorphism.

^aM, major allele; m: minor allele. ^bMAF: minor allele frequency, the frequency was calculated based on all combined controls. ^cOR was calculated for the minor allele. ^dGenes with SNPs located up to 20 kb down- or upstream were shown. ^eRs10144552 is located 36 kb upstream of *BDKRB2*.

meta-analysis of the two GWAS data sets, none of the GO-enriched terms had false discovery rate <25% (Supplementary Table 8).

Candidate gene association analysis. We tested the 327 prior candidate genes using a gene list from the results of meta-analysis of the GWAS data. Although none of these genes achieved significance after experiment-wise correction for multiple testing ($P < 1.53 \times 10^{-4} = 0.05/327$), gene-wise evaluation produced 20 genes yielding significance at the level of $P < 0.05$ (Supplementary Table S9). Among them, the strongest evidence for association was found at rs12501691 located in *NPY5R* on 4q31.3–q32 (gene-wise corrected $P = 6.4 \times 10^{-4}$). In addition to the prior candidate gene list, we examined the associations of *TMEM132D*¹¹ and

*ACCN1*¹² with PD in our results. Neither of SNPs rs7309727 nor rs11060369 in *TMEM132D* was significant (nominal $P = 0.38$ and 0.55, respectively). The most significantly associated SNP (rs1397504) in *TMEM132D* in our results did not reach the gene-wide significance (gene-wise corrected $P = 0.37$). On the other hand, although the most strongly associated SNP within *ACCN1* showed non-significance (rs880305, gene-wise corrected $P = 0.13$), rs280039 located 350 kb upstream of *ACCN1* showed strong association with PD (nominal $P = 4.0 \times 10^{-5}$). The regional association plots of *TMEM132D* and *ACCN1* are presented in Supplementary Figures S3 and S4. Furthermore, when we investigated the region of 16p11.2 where common duplications were associated with PD,¹³ no significant SNP was found.

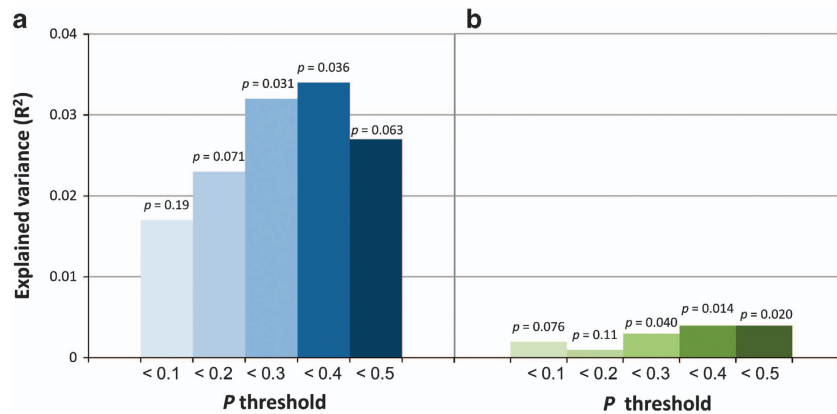


Figure 3 Pseudo R^2 explained by the PD polygenic scores in the genome-wide association study samples. X axis represents sets of SNPs with P -values threshold. Y axis represents Nagelkerke's F^2 . P -values in each set of single-nucleotide polymorphisms represent the significance of correlation between the PD status in the target sample and polygenic scores based on the discovery sample. (a) Sample I (target) predicted by sample II (discovery). (b) Sample II (target) predicted by sample I (discovery).

Polygenic score analysis. Figure 3 and Supplementary Table S10 present the P -values and Nagelkerke's F^2 statistics for the logistic regression analysis based on the two Japanese samples. Significant associations were found between case–control status in the target sample (sample I) and polygenic scores based on the sets of SNPs with $P_T < 0.3$ and < 0.4 in the discovery sample (sample II) ($P = 0.031$ and 0.036 , $F^2 = 3.2\%$ and 3.4% , respectively Figure 3). The polygenic scores based on the sets of SNPs with $P_T < 0.3$, 0.4 and 0.5 in the discovery sample (sample I) also showed significant associations with PD status in the target sample (sample II) ($P = 0.040$, 0.014 and 0.020 , respectively). However, pseudo F^2 was lower than that based on a larger discovery sample (sample II) (maximum $F^2 < 0.5\%$).

Discussion

To date, most of candidate gene association studies and GWAS have been underpowered to detect the small effect sizes of susceptibility loci for PD. In this study, we performed a new GWAS and a meta-analysis of our two GWAS data sets followed by a replication study in Japanese (1047 PD cases and 2578 controls in total). To our knowledge, the sample size of the present study was the largest in the GWAS of PD. However, the meta-analysis of the GWAS data produced no genome-wide significance and none of the top SNPs selected from the meta-analysis was associated with PD in the replication sample. This may have been attributable to chance findings in the meta-analysis and inadequate replication sample size. Recent large-scale GWAS of schizophrenia³¹ and bipolar disorder³² suggest that the effect sizes of common risk alleles are small ($OR < 1.2$). The power of our meta-analysis sample (718 cases and 1717 controls) was 0.3% to detect $\alpha = 5 \times 10^{-8}$ conferred by an allele with a frequency of 0.25 and an OR of 1.2. Therefore, it is very unlikely to detect any locus at the genome-wide significance with this power. Moreover, the power of replication sample (329 cases and 861 controls) was 45% to detect $\alpha = 0.05$ conferred by an allele with a frequency of 0.25 and an OR of 1.2. Besides the

possibility that the meta-analysis findings were false positives, the lack of replication may be due to the limited power of the replication sample.

Despite the limited power of the samples, the regions with statistically significant association might have some relevance with PD. Among these loci, the most strongly associated SNP was rs10144552 at upstream of *BDKRB2* on 14q32 (all combined $P = 1.3 \times 10^{-5}$). The direction of the effects was the same across the three (two GWAS and replication) samples. Gratacòs *et al.*³³ reported that a SNP rs945032, located in the promoter region of *BDKRB2*, was associated with PD, substance abuse and bipolar disorder. Our imputation results did not include this SNP. As rs945032 is in the middle of a recombination hotspot, even nearest imputed SNPs in the present study presented low LD with this SNP ($r^2 < 0.01$). *BDKRB2* encodes bradykinin receptor that activates second messenger regulating blood pressure, pain perception and neuronal differentiation.³³ Several studies have suggested that mortality due to cardiovascular diseases is high in PD patients.³⁴ These reports highlight *BDKRB2* as a promising candidate for PD that is worthy of additional follow-up.

In the analysis of the prior PD candidate genes, the strongest evidence for association was found at rs12501691 in *NPY5R* on 4q31.3–q32 (gene-wise corrected $P = 6.4 \times 10^{-4}$). *NPY5R* was reported to be associated with PD, especially in females.³⁵ In the present study, sex-specific analyses showed that rs12501691 in *NPY5R* were associated with PD in only females (female, $P = 2.7 \times 10^{-3}$; male, $P = 0.085$). *Neuropeptide Y (NPY)* is a particularly plausible candidate for modulating effects of environmental stress exposure on susceptibility to anxiety disorders as seen in animal models.³⁶ Therefore, changes in *NPY* expression due to genetic variations in the *NPY*-related genes affect stress response and emotion,³⁷ and may contribute to anxiety disorders such as PD. We also investigated PD candidate genes *TMEM132D*¹¹ and *ACCN1*¹² reported in previous GWAS. Erhardt *et al.*¹¹ reported that two SNPs rs7309727 and rs11060369 in *TMEM132D* were strongly associated with PD in the German population. Risk allele of rs11060369 was

associated with higher *TMEM132D* mRNA expression levels in the frontal cortex. In our study, however, neither of the SNPs showed significant association with PD. Gregersen *et al.*¹² conducted a genome-wide scan using the isolated population of the Faroe Islands, and suggested *ACCN1* as a candidate gene for PD. Our meta-analysis of the GWAS provided supportive evidence for association of rs280039 at upstream of *ACCN1* with PD in the Japanese population ($P=4.0 \times 10^{-5}$), whereas no association of SNPs within the gene was found at the gene-wide levels of significance. The inconsistency in the results between Caucasian and Japanese samples may be due to differences in clinical and genetic heterogeneity.

The lack of replication in GWAS- and GO-based enrichment analyses across samples is not surprising if the effect of each locus is too small to surpass the threshold specified for significance. Therefore, we performed the polygenic score analysis using our GWAS data. The power of the discovery sample (sample II, 541 cases and 1539 controls; sample I, 177 cases and 178 controls), respectively, was 65% and 39% to detect $\alpha=0.4$ conferred by an allele with a frequency of 0.25 and an OR of 1.1. Despite the low power in the discovery sample, the sets of SNPs with $P_T < 0.3$ and 0.4 based on the discovery sample (also $P_T < 0.5$ in sample I) were significantly associated with PD status in the target sample. The largest variance was explained by the set of SNPs at the threshold of $P_T < 0.4$ in the discovery sample (sample II, $P=0.036$ and $R^2=3.4\%$; sample I, $P=0.014$ and $R^2=0.4\%$). These results are compatible with previous studies in which around 1–2% of the variance for anxiety and depression could be explained by the polygenic score approach.³⁸ Our findings suggest that ‘true’ risk variants of PD may be included in the sets of the more liberal thresholds.

The present study has limitations. First, although controls were screened to exclude history of major psychiatric disorders, some subjects in the control group were still at risk for developing PD. This may have decreased sensitivity. However, given the relatively low prevalence of PD (<2–3%), it was less likely that inclusion of controls with potential PD had major effect on statistical power. Second, as our study was designed to identify common variants conferring risk for PD, we did not assess the impact of rare variants. Recently, multiple rare variants have been reported in genes shown to harbor common variants associated with common diseases.^{39,40} Therefore, resequencing genomic regions suggested by GWAS might be powerful approach to identify rare coding (that is, missense or nonsense) variants associated with PD. Third, we did not investigate gene–gene and gene–environment interactions. It has been recognized that anxiety disorders including PD are the results of multiple, complex interactions between genes and environmental influences.⁴¹ The candidate genes suggested in the present study are needed for further research on interactions with environmental factors such as childhood experience.

In conclusion, the present study suggest that *BDKRB2* and several other genes are worthy of follow-up as candidates for PD. Our findings also suggest that a large number of common alleles of small effects collectively account for risk of PD.

GWAS with multiple large samples and their meta/mega-analyses are warranted to confirm our findings.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements. This work was supported by Grants-in-Aid for Scientific Research on Priority Areas ‘Comprehensive Genomics’ and ‘Applied Genomics’ (no. 17019029; no. 21300242) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (Principal Investigator: Professor T Sasaki). TO was supported by research fellowship from the Japan Society for the Promotion of Science (no. 21-8373). We are grateful to Ms Yoshimi Ishibashi, Ms Yuko Ogasawara and Ms. Kayoko Kato for their excellent technical assistance. We also thank Drs Kenneth S Kendler and John M Hettema at Virginia Commonwealth University and Dr Kazuya Iwamoto at the University of Tokyo for their helpful comments. Most importantly, we thank the individuals who have participated in and contributed to this study.

- Eaton WW, Kessler RC, Wittchen HU, Magee WJ. Panic and panic disorder in the United States. *Am J Psychiatry* 1994; **151**: 413–420.
- Crowe RR, Noyes R, Pauls DL, Slymen D. A family study of panic disorder. *Arch Gen Psychiatry* 1983; **40**: 1065–1069.
- Weissman MM, Bland RC, Canino GJ, Faravelli C, Greenwald S, Hwu HG *et al.* The cross-national epidemiology of panic disorder. *Arch Gen Psychiatry* 1997; **54**: 305–309.
- Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ. Panic disorder in women: a population-based twin study. *Psychol Med* 1993; **23**: 397–406.
- Hettema JM, Neale MC, Kendler KS. A review and meta-analysis of the genetic epidemiology of anxiety disorders. *Am J Psychiatry* 2001; **158**: 1568–1578.
- Hovatta I, Barlow C. Molecular genetics of anxiety in mice and men. *Ann Med* 2008; **40**: 92–109.
- Hamilton SP. Linkage and association studies of anxiety disorders. *Depress Anxiety* 2009; **26**: 976–983.
- Maron E, Hettema JM, Shlik J. Advances in molecular genetics of panic disorder. *Mol Psychiatry* 2010; **15**: 681–701.
- Otowa T, Yoshida E, Sugaya N, Yasuda S, Nishimura Y, Inoue K *et al.* Genome-wide association study of panic disorder in the Japanese population. *J Hum Genet* 2009; **54**: 122–126.
- Otowa T, Tani H, Sugaya N, Yoshida E, Inoue K, Yasuda S *et al.* Replication of a genome-wide association study of panic disorder in a Japanese population. *J Hum Genet* 2010; **55**: 91–96.
- Erhardt A, Czibere L, Roeske D, Lucae S, Unschuld PG, Ripke S *et al.* *TMEM132D*, a new candidate for anxiety phenotypes: evidence from human and mouse studies. *Mol Psychiatry* 2011; **16**: 647–663.
- Gregersen N, Dahl HA, Buttenschøn HN, Nyegaard M, Hedemand A, Als TD *et al.* A genome-wide study of panic disorder suggests the amiloride-sensitive cation channel 1 as a candidate gene. *Eur J Hum Genet* 2011; **20**: 84–90.
- Kawamura Y, Otowa T, Koike A, Sugaya N, Yoshida E, Yasuda S *et al.* A genome-wide CNV association study on panic disorder in a Japanese population. *J Hum Genet* 2011; **56**: 852–856.
- American Psychiatric Association *Diagnostic and Statistical Manual of Mental Disorders*. 4th edn. American Psychiatric Press; Washington, DC, 1994.
- Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E *et al.* The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998; **59**(Suppl 20): 22–33.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
- Yamaguchi-Kabata Y, Nakazono K, Takahashi A, Saito S, Hosono N, Kubo M *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**: 1–12.
- Ohashi J, Yamamoto S, Tsuchiya N, Hatta Y, Komata T, Matsushita M *et al.* Comparison of statistical power between 2×2 allele frequency and allele positivity tables in case-control studies of complex disease genes. *Ann Hum Genet* 2001; **65**: 197–206.
- Barrett JC, Fry B, Maller J, Daly M. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263–265.
- Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O’Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 2008; **24**: 2938–2939.

21. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* 2010; **34**: 816–834.
22. Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics* 2010; **11**: 134.
23. Nakaoka H, Inoue I. Meta-analysis of genetic association studies: methodologies, between-study heterogeneity and winner's curse. *J Hum Genet* 2009; **54**: 615–623.
24. Ioannidis JP, Patsopoulos NA, Evangelou E. Heterogeneity in meta-analyses of genome-wide association investigations. *PLoS One* 2007; **2**: e841.
25. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959; **22**: 719–748.
26. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nat Protoc* 2009; **4**: 44–57.
27. Hong MG, Pawitan Y, Magnusson PK, Prince JA. Strategies and issues in the detection of pathway enrichment in genome-wide association studies. *Hum Genet* 2009; **126**: 289–301.
28. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc B* 1995; **57**: 289–300.
29. International Schizophrenia Consortium Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC *et al*. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
30. Koike A, Nishida N, Inoue I, Tsuji S, Tokunaga K. Genome-wide association database developed in the Japanese Integrated Database Project. *J Hum Genet* 2009; **54**: 543–546.
31. The Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P *et al*. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 2011; **43**: 969–976.
32. Sklar P, Ripke S, Scott LJ, Andreassen OA, Cichon S, Craddock N *et al*. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet* 2011; **43**: 977–983.
33. Gratacòs M, Costas J, de Cid R, Bayés M, González JR, Baca-García E *et al*. Identification of new putative susceptibility genes for several psychiatric disorders by association analysis of regulatory and non-synonymous SNPs of 306 genes involved in neurotransmission and neurodevelopment. *Am J Med Genet B Neuropsychiatr Genet* 2009; **150B**: 808–816.
34. Chen YH, Tsai SY, Lee HC, Lin HC. Increased risk of acute myocardial infarction for patients with panic disorder: a nationwide population-based study. *Psychosom Med* 2009; **71**: 798–804.
35. Domschke K, Hohoff C, Jacob C, Maier W, Fritze J, Bandelow B *et al*. Chromosome 4q31-34 panic disorder risk locus: association of neuropeptide Y Y5 receptor variants. *Am J Med Genet B Neuropsychiatr Genet* 2008; **147B**: 510–516.
36. Thorsell A. Brain neuropeptide Y and corticotropin-releasing hormone in mediating stress and anxiety. *Exp Biol Med* 2010; **235**: 1163–1167.
37. Zhou Z, Zhu G, Hariri AR, Enoch MA, Scott D, Sinha R *et al*. Genetic variation in human NPY expression affects stress response and emotion. *Nature* 2008; **452**: 997–1001.
38. Demirkan A, Penninx BW, Hek K, Wray NR, Amin N, Aulchenko YS. Genetic risk profiles for depression and anxiety in adult and elderly cohorts. *Mol Psychiatry* 2011; **16**: 773–783.
39. Johansen CT, Wang J, Lanktree MB, Cao H, McIntyre AD, Ban MR *et al*. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. *Nat Genet* 2010; **42**: 684–687.
40. Momozawa Y, Mni M, Nakamura K, Coppieters W, Almer S, Amininejad L *et al*. Resequencing of positional candidates identifies low frequency IL23R coding variants protecting against inflammatory bowel disease. *Nat Genet* 2011; **43**: 43–47.
41. Smoller JW, Block SR, Young MM. Genetics of anxiety disorders: the complex road from DSM to DNA. *Depress Anxiety* 2009; **26**: 965–975.



Translational Psychiatry is an open-access journal published by **Nature Publishing Group**. This work is licensed under the **Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 Unported License**. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

Supplementary Information accompanies the paper on the Translational Psychiatry website (<http://www.nature.com/tp>)