98th percentile of all the probes. Probes with no NCBI accession number were excluded. Probes under significant influence of gender ($p < 0.05$, unpaired t-test) were also excluded. This procedure reduced the number of probes included in the analysis from 54,847 to 19,121.

2.4. Overrepresentation analysis

The gene expression data were imported into ArrayAssist ver5.5.1 (Stratagene, Cedar Creek, TX). Differentially regulated probes between the two groups were analyzed by unpaired t-tests. Probes with $p < 0.01$ were considered to be differentially expressed. A correction for multiple testing was tried by calculating the false discovery rate (FDR), the rate at which significant features are truly null, using the QVALUE program (Storey and Tibshirani, 2003). For estimating q values, π_0 was determined with automated choice of λ based on the distribution of 19,121 p values using a smoothing method. FDR ≤ 0.05 was considered as significant.

Functional overrepresentation of differentially expressed genes was analyzed with the gene annotation tool DAVID, which integrates almost all major bioinformatics resources (Huang da et al., 2007). DAVID calculates the p value for the extent of overrepresentation of a given biological meaning by performing Fisher's exact test.

2.5. Artificial neural network construction and testing

For the purpose of hold-out cross-validation, subjects were subdivided into training and testing sets. The training set included 35 patients with schizophrenia (29.9 years: 14 males, 21 females) and 33 normal controls (30.9 years: 23 males, 10 females). The remaining 17 patients (32.8 years: 7 males, 10 females) and 16 controls (33.5 years: 12 males, 4 females) were held out for the testing. Mean age was similar but gender ratio was significantly different within each set ($p = 0.014$ and $p = 0.049$ in the training and testing sets, respectively).

We constructed artificial neural networks (ANNs) using the ArrayAssist program. Our ANNs were three-layer back propagation networks trained with a learning rate of 0.5, a momentum rate of 0.3, and 100 iterations (Fig. 1). The program was implemented with algorithms of stepwise forward selection to enable the determination of optimal gene features for the ANN model (Lancashire et al., 2008). Selected 127 probe datasets were used as an input for forward selection. The ideal outputs were set at 0 for controls and 1 for schizophrenia. Beginning from an empty set of features, the forward selection algorithm added a single probe that produced the best performance as a predictor at each step of 3-fold cross-validations. After determining an optimal feature size, ANNs trained 100 times to adjust weights of links. The constructed ANN model was then applied to testing samples.

2.6. Clustering analysis

To compare supervised and unsupervised methods, clustering analysis was performed on training set samples that were used for ANN construction. The ArrayAssist program was used to perform principal component analysis (PCA) clustering and agglomerative hierarchical clustering.

The method employed for hierarchical clustering was complete linkage and Euclidean distance.

3. Results

3.1. Characteristics and annotation of a blood-based gene expression signature

Among 19,121 quality filtered probes, 792 probes were identified as differentially expressed; 256 probes were down-regulated and 536 probes were up-regulated in patients with schizophrenia compared with controls. Since the gender ratio was different between the two groups, our filtering process excluded probes with significant gender differences. Down-regulated and up-regulated probes were separately profiled by DAVID. Among 256 down-regulated probes, 167 genes were annotated. Top-ranked Gene Ontology terms included cell adhesion, nervous system development, and neurotransmitter binding and receptor activity (Table 2, Table S1, S2). Unigene EST annotated 73 of these genes as expressed in brain tissue (Table S3). On the other hand, among 536 up-regulated probes, 381 genes were annotated but their association with central nervous system was not clear. The annotated genes contained 11 genes that were reported to be associated with schizophrenia (NRG1, NTF3, CHRNB2, CCKAR, DAOA, DAO, L1CAM, NEUROG1, ZNF74, PRODH, and CHGA).

As the main purpose of the study was to find the usefulness of microarray data in a clinical practice, probes suitable for laboratory testing were required. Calculation of the q value revealed that among 792 differentially expressed probes, only 17 passed a q value threshold of 0.05, which was too conservative to

Table 2
Top 10 biological process and molecular function Gene Ontology terms overrepresented by down-regulated genes.

Term	Count	%	p value
Biological process			
Biological adhesion	16	9.58%	0.0011
Cell adhesion	16	9.58%	0.0011
Nervous system development	16	9.58%	0.0012
Multicellular organismal process	45	26.95%	0.0023
Ion transport	16	9.58%	0.0024
Generation of neurons	8	4.79%	0.0079
Neurogenesis	8	4.79%	0.0118
Neuron differentiation	7	4.19%	0.0142
Sodium ion transport	5	2.99%	0.0167
Developmental process	37	22.16%	0.0169
Molecular function			
Photoreceptor activity	3	1.80%	0.0061
Alkali metal ion binding	7	4.19%	0.0063
Neurotransmitter receptor activity	5	2.99%	0.0079
Neurotransmitter binding	5	2.99%	0.0090
Calcium ion binding	16	9.58%	0.0112
Cation binding	44	26.35%	0.0154
Ion binding	47	28.14%	0.0185
Metal ion binding	46	27.54%	0.0209
Lipid binding	9	5.39%	0.0323
Gated channel activity	7	4.19%	0.0365

Count of genes involved in the term is listed, together with the percentage of involved genes in total genes. The smaller the Fisher's exact p value, the more enriched the term.

be used as a criterion for ANN input data. We therefore conducted additional filtering based on coefficient of variation (CV) and average signal intensity. CV, defined as the ratio of the standard deviation to the mean, is a measure of dispersion of probability distribution. As smaller CV is indispensable for the precise laboratory testing, we considered probes with CV less than 30% are acceptable. This resulted in an exclusion of probes with large fold change (FC) in the mean expression level, since larger FC accompanied larger CV (Fig. 2A, C). On the other hand, in spite of normal distribution of FCs (Fig. 2A), the distribution of average signal intensities was largely skewed (Fig. 2B). The large skewness of our data might lead to the construction of a less reproducible ANN model. We performed a pilot study by constructing ANNs with shifting the threshold window of average signal intensities around a median of differentially expressed probes (4.0). ANNs were most reproducible with the threshold window of 1.0 to 3.0, and therefore it was introduced as the second criterion. This procedure of additional filtering reduced the number of probes included in the further analysis from 792 to 127 (Fig. 2D).

3.2. Feature selection through ANN training and model construction

We used ANNs to examine whether a blood-based gene expression signature can distinguish patients with schizophrenia from normal controls. Forward selection (Lancashire et al., 2008) was performed using the training set of samples to identify a subset of differentially expressed probes that were important for the distinction. Among 127 probes, maximal classification accuracy was obtained with 14 probes (Fig. 3A) for 8 known genes and 6 expressed sequence tags (ESTs). Eight genes and two ESTs were expressed in the brain (UniGene), and only *DAOA* (*G72*) was associated with schizophrenia (OMIM) (Fig. 3B).

Since FC values in 14 predictors were relatively small, we carried out unsupervised standard methods to see to what extent these predictors separate patients from controls. Hierarchical clustering grouped one patient into controls and grouped 20 controls together with 34 patients (Fig. 4A), suggesting that one-dimensional Euclidean distance by 14 predictors was too close among samples to separate patients
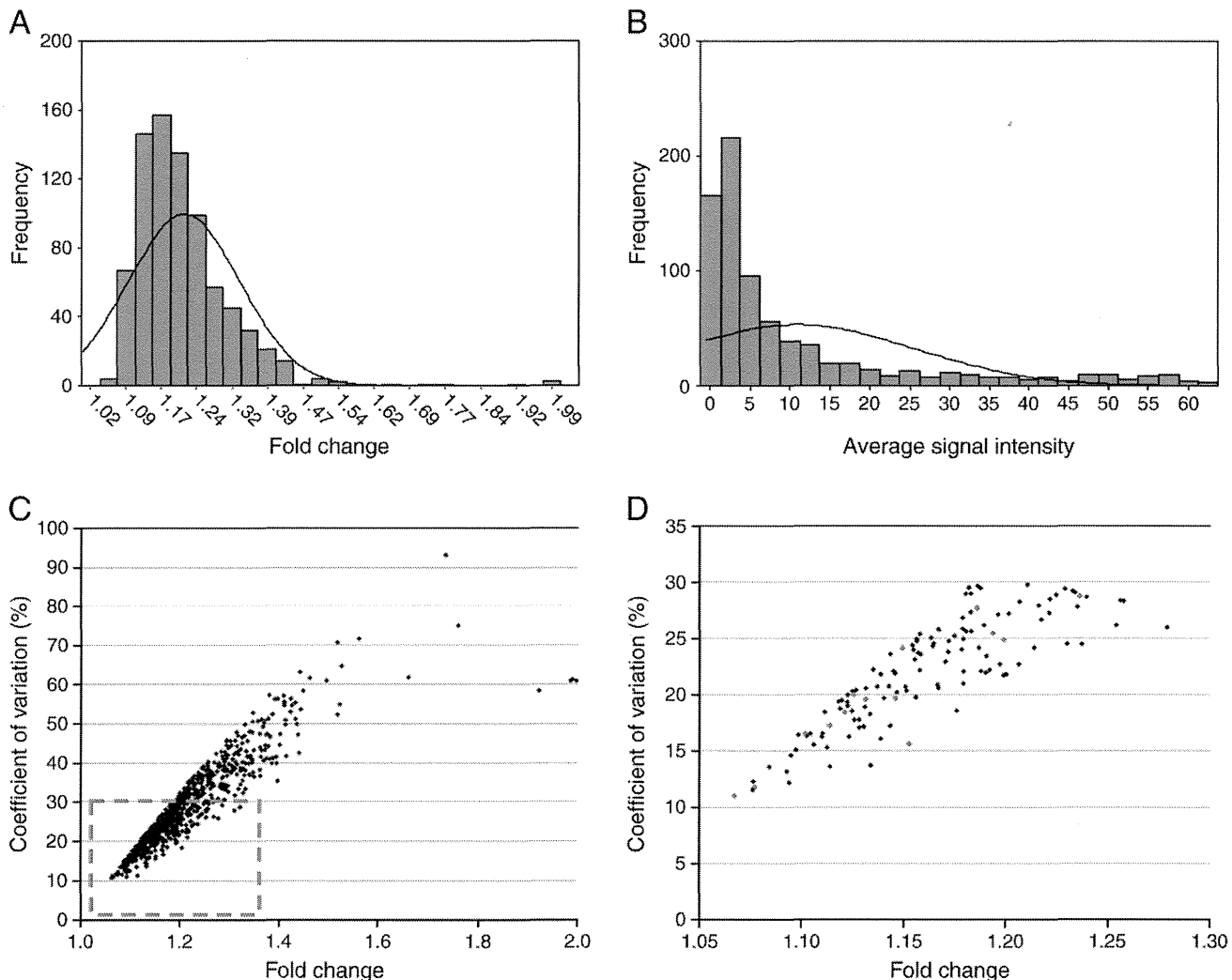

Fig. 2. Descriptive statistics of 792 differentially expressed probes identified by unpaired *t*-tests with a *p* value cut-off of 0.01. Shown are the frequencies of absolute fold change (A) and average signal intensity (B). The normal distribution curve was overlayed. (C) Values of the coefficient of variation were plotted against absolute fold change values. Probes with less than 30% CV (red dashed square) were filtered with an average signal intensity threshold of 1.0–3.0 (D). Probes selected as predictors are highlighted in red.
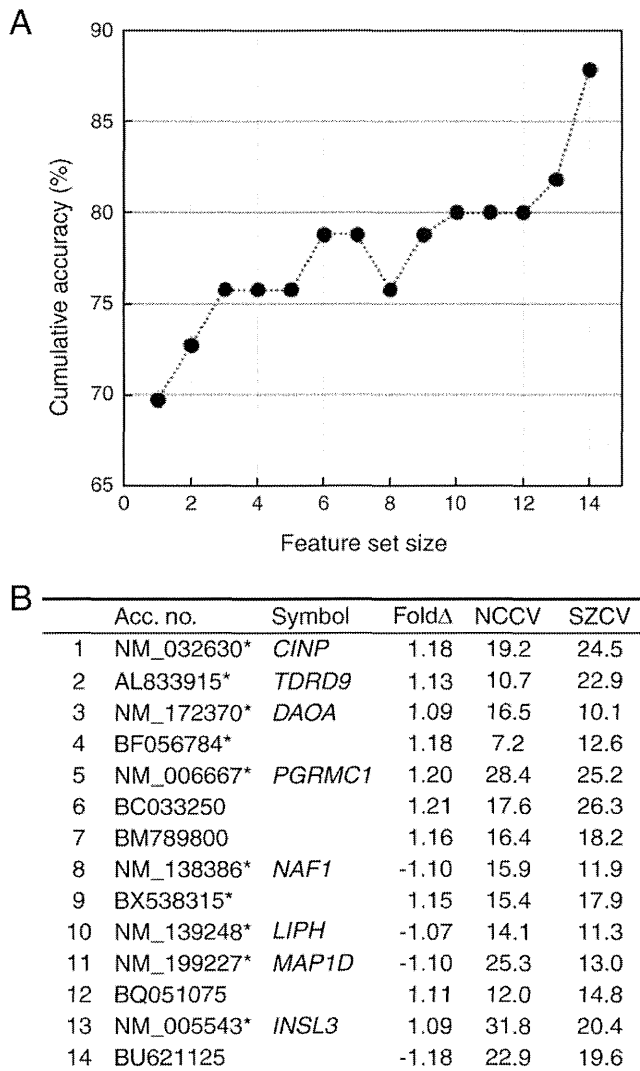
A

B

	Acc. no.	Symbol	FoldΔ	NCCV	SZCV
1	NM_032630*	CINP	1.18	19.2	24.5
2	AL833915*	TDRD9	1.13	10.7	22.9
3	NM_172370*	DAOA	1.09	16.5	10.1
4	BF056784*		1.18	7.2	12.6
5	NM_006667*	PGRMC1	1.20	28.4	25.2
6	BC033250		1.21	17.6	26.3
7	BM789800		1.16	16.4	18.2
8	NM_138386*	NAF1	-1.10	15.9	11.9
9	BX538315*		1.15	15.4	17.9
10	NM_139248*	LIPH	-1.07	14.1	11.3
11	NM_199227*	MAP1D	-1.10	25.3	13.0
12	BQ051075		1.11	12.0	14.8
13	NM_005543*	INSL3	1.09	31.8	20.4
14	BU621125		-1.18	22.9	19.6

Fig. 3. Identification of predictor genes by a stepwise forward selection ANN model. A. Each round of 3-fold cross-validation finds one probe that performs best as a predictor for classification. The cycle was repeated until selected features achieved maximal accuracy. B. Predictors in order of their appearance in the selection algorithm. Positive and negative values in fold change (Fold Δ) represents up- and down-regulation in schizophrenia patients, respectively. *Presence of expression in the brain (UniGene). CINP, cyclin-dependent kinase 2-integrating protein; TDRD9, Tudor domain containing 9; DAOA (G72), D-amino acid oxidase activator; PGRMC1, progesterone receptor membrane component 1; NAF1, nuclear assembly factor 1 homolog; LIPH, lipaze member H; MAP1D, methionine aminopeptidase 1D; INSL3, insuline-like 3 (Leidig cell); NCCV, CV in normal controls; and SZCV, CV in patients with schizophrenia.

from controls by this type of method. PCA clustering was also done to generate a three-dimensional plot from 14 dimensions (predictors). Visualization of 68 training samples demonstrated some segregation according to the clinical diagnosis (Fig. 4B), but it was difficult to find linearly separable space between the two groups.

ANNs were then trained with the training set of samples. The constructed ANN model correctly predicted 31 of 35 patients and 31 of 33 controls (Fig. 4C). Finally, performance of the model was tested using hold-out samples. Our ANN modeling was validated by correctly predicting 14 of 17 patients and 15 of 16 controls. This corresponded to 87.9% accuracy with 82.4% sensitivity and 93.8% specificity. In

addition, we subdivided all the samples again into four subsets and classified them by the ANN, respectively. Diagnosis of each set was predicted with 80.0 to 96.2% accuracy, indicating robustness of the ANN to different groupings (Table 3).

4. Discussion

Using supervised method, ANN, we found that schizophrenia can be classified by blood-based gene expression signature. Accuracy of the ANN model was 91.2%, which was not so high as similar reports in cancer research (Khan et al., 2001; Lin et al., 2007; O'Neill and Song, 2003). However, in contrast to the use of tumor samples in those studies, we are not able to obtain tissues from the brain. Considering the use of peripheral whole blood as a material, we feel that the performance of our ANN model is acceptable. In schizophrenia, Tsuang et al. (2005) presented a diagnostic equation obtained from logistic regression of eight gene expression levels in isolated peripheral blood cells. They adopted ANN to verify their finding, but their report did not describe the accuracy of it. Middleton et al. (2005) provided a list of genes expressed in peripheral blood cells that predicted diagnosis of schizophrenia with 95% or greater accuracy using two types of supervised classifiers. However, they did not cross-validated their list. The larger sample size of the present study allowed us to conduct hold-out cross-validation, which revealed that the testing set accuracy was comparable to the training set one.

Hold-out method is one of the standard cross-validation methods that are widely used to evaluate prediction by computer-based learning (Li, 2007). In ANN modeling of microarray datasets, the method is utilized by partitioning samples into training and testing sets after the identification of differentially expressed probes (Khan et al., 2001; Xu et al., 2002). Purpose of the method is to confirm that constructed ANN model does not come from overfitting of model parameters to features in the training set, which will result in poor predictive accuracy in the testing set. Therefore, validation of our ANN model does not mean to show replicated data in an independent cohort. It simply suggests that the model can be generalized to other subset of samples that are not included for the model construction.

Predictors in our ANN model were chosen by forward selection to minimize error in supervised learning. As a result, FCs in selected predictors were so small that hierarchical or PCA clustering could not discriminate between patients and controls. In postmortem brain studies of schizophrenia, top-ranked genes with 1.3-fold or greater change were shown to separate patients well from controls (Hakak et al., 2001; Iwamoto et al., 2004). We were not sure, however, whether top-ranked genes expressed in whole blood represent important biological pathways of the brain that are involved in schizophrenia. Previous research has demonstrated that the selection of top-ranked genes might lose essential information for the purpose of classification, since they are usually highly correlated (Jaeger et al., 2003; Liu et al., 2004). Low-ranked genes are sometimes necessary for classifiers to achieve greater accuracy (Li et al., 2003). We therefore consider that it is reasonable for the forward selection algorithms to identify probes with relatively small FC as predictors for classification.

In peripheral blood, variations in transcription level among individuals are attributed to genetic or epigenetic components

A

B

C

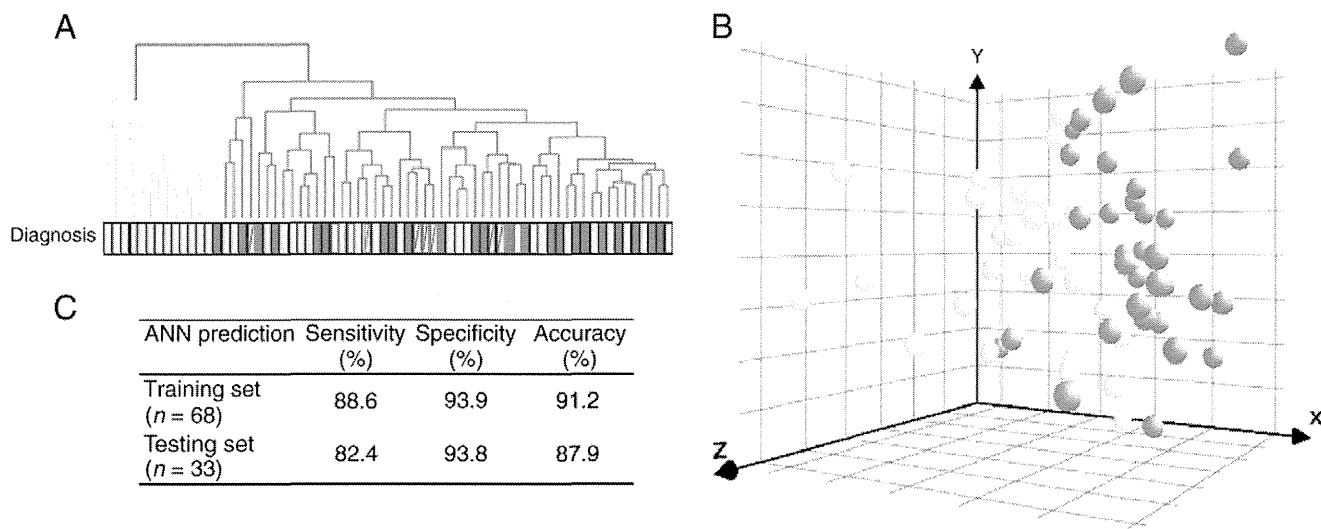| ANN prediction	Sensitivity (%)	Specificity (%)	Accuracy (%)
Training set (n = 68)	88.6	93.9	91.2
Testing set (n = 33)	82.4	93.8	87.9

Fig. 4. Classification by 14 predictors identified with stepwise forward selection. A. Hierarchical clustering resulted in two ambiguous clusters. One patient was grouped into controls, and 20 controls were grouped together with patients. Closed green square, normal controls; closed red square, drug-naïve patients with schizophrenia; open red square, drug-free patients with schizophrenia; and open red square with a slash, antipsychotic-naïve patients with schizophrenia. Note that past exposure to antipsychotics or current medication status did not affect the expression pattern of the predictors. B. Plot of the three principal components generated from 14 selected predictors revealed segregation by diagnosis but with some overlap. Green sphere, normal controls; red sphere, patients with schizophrenia. C. Performance of ANN prediction in training set and testing set samples.

and cellular composition (Cheung et al., 2003; Eady et al., 2005; Radich et al., 2004; Whitney et al., 2003). For convenience as a diagnostic tool and to minimize technical fluctuations (Debey et al., 2004; Ohmori et al., 2005), we stabilized RNA as blood was withdrawn into PAXgene Blood RNA Tubes instead of isolating a specific type of cells with time delay (Rainen et al., 2002). It has been shown that homeostasis of cellular composition in healthy individuals keeps consistent patterns of global gene expression as compared with individuals with lymphoma or bacterial infection (Whitney et al., 2003). Our forward selection algorithm did not chose genes associated with specific blood cell types. Therefore, so far as excluding samples with physiological illness, we assume that variations by cellular composition are maximally suppressed by supervised training in ANNs.

Among 14 predictors, *DAOA* (*G72*) was the only gene associated with schizophrenia (Chumakov et al., 2002). Selective degradation of D-serine by DAOA and D-amino acid oxidase results in attenuated neurotransmission via NMDA receptor (Mothet et al., 2000), which probably links the up-regulation of *DAOA* (*G72*) in the present study to the earlier finding of decreased D-serine levels in the serum of patients with schizophrenia (Hashimoto et al., 2003). However, we should investigate how 14 genes are expressed in the brain and

whether they are associated with schizophrenia before discussing the biological significance of them in peripheral blood.

We demonstrate here that by using the ANN method we can classify schizophrenia based on gene expression signature obtained from whole peripheral blood. Since we constructed the ANN model using antipsychotics-free patients without comorbidity, ANN-based classification may be applied only for a subset of patients, especially with first-episode psychosis. There are several limitations in the present study: small sample size compared to the large scale microarray, gender difference between the two groups, and absence of quantitative reverse transcription-PCR (RT-PCR) results to corroborate microarray data. Statistical power of our sample size (50) was about 22% by computation at web-based bioinformatics service (http://bioin-formatics.mdanderson.org/MicroarraySampleSize/). Therefore, our results, in terms of the predictor selection, should be interpreted cautiously. Before replicating the differential expression by RT-PCR, we are currently developing a custom array with less than several hundred probes. This will increase statistical power even with the same sample size. Moreover, high accuracy of our ANN model needs to be replicated by independent cohort with gender-matched groups before applying it to clinical practice.

Table 3
Performance of ANN prediction in regrouped sample sets.

Sample set	Group	Age (mean ± SD)	Sex (M/F)	Sensitivity (%)	Specificity (%)	Accuracy (%)
A	NC	31.3 ± 10.6	8/4	76.9	83.3	80.0
	SZ	36.8 ± 9.2	5/8			
B	NC	32.6 ± 11.0	9/3	84.6	100.0	92.0
	SZ	27.0 ± 10.9	5/8			
C	NC	32.6 ± 11.1	10/2	92.3	91.7	92.0
	SZ	28.8 ± 7.7	6/7			
D	NC	28.5 ± 4.6	8/5	92.3	100.0	96.2
	SZ	32.8 ± 14.8	5/8			

Conflict of interest

Hayashi, Iijima, Takamura and Aoshima work for R&D Department of SRL Inc. All other authors declare that they have no conflicts of interest.

Acknowledgement

We thank Dr. Hiroyuki Nawa for the helpful discussions.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.schres.2009.12.024.

References

Aston, C., Jiang, L., Sokolov, B.P., 2004. Microarray analysis of postmortem temporal cortex from patients with schizophrenia. J. Neurosci. Res. 77 (6), 858–866.

Cheung, V.G., Conlin, L.K., Weber, T.M., Arcaro, M., Jen, K.Y., Morley, M., Spielman, R.S., 2003. Natural variation in human gene expression assessed in lymphoblastoid cells. Nat. Genet. 33 (3), 422–425.

Chumakov, I., Blumenfeld, M., Guerassimenko, O., Cavarec, L., Palicio, M., Abderrahim, H., Bougueleret, L., Barry, C., Tanaka, H., La Rosa, P., Puech, A., Tahri, N., Cohen-Akenine, A., Delabrosse, S., Lissarrague, S., Picard, F.P., Maurice, K., Essioux, L., Millasseau, P., Grel, P., Debailleul, V., Simon, A.M., Caterina, D., Dufaure, I., Malekzadeh, K., Belova, M., Luan, J.J., Bouillot, M., Sambucy, J.L., Primas, G., Saumier, M., Boubkiri, N., Martin-Saumier, S., Nasroune, M., Peixoto, H., Delaye, A., Pinchot, V., Bastucci, M., Guillou, S., Chevillon, M., Sainz-Fuertes, R., Meguenni, S., Aurich-Costa, J., Cherif, D., Gimalac, A., Van Duijn, C., Gauvreau, D., Ouellette, G., Fortier, I., Raelson, J., Sherbatich, T., Riazanskaia, N., Rogaev, E., Raeymaekers, P., Aerssens, J., Konings, F., Luyten, W., Macciardi, F., Sham, P.C., Straub, R.E., Weinberger, D.R., Cohen, N., Cohen, D., 2002. Genetic and physiological data implicating the new human gene G72 and the gene for D-amino acid oxidase in schizophrenia. Proc. Natl. Acad. Sci. U. S. A. 99 (21), 13,675–13,680.

De Bruyne, V., Al-Mulla, F., Pot, B., 2007. Methods for microarray data analysis. Methods Mol. Biol. 382, 373–391.

Debey, S., Schoenbeck, U., Hellmich, M., Gathof, B.S., Pillai, R., Zander, T., Schultze, J.L., 2004. Comparison of different isolation techniques prior gene expression profiling of blood derived cells: impact on physiological responses, on overall expression and the role of different cell types. Pharmacogenomics J. 4 (3), 193–207.

Eady, J.J., Wortley, G.M., Wormstone, Y.M., Hughes, J.C., Astley, S.B., Foxall, R.J., Doleman, J.F., Elliott, R.M., 2005. Variation in gene expression profiles of peripheral blood mononuclear cells from healthy volunteers. Physiol. Genomics 22 (3), 402–411.

Hakak, Y., Walker, J.R., Li, C., Wong, W.H., Davis, K.L., Buxbaum, J.D., Haroutunian, V., Fienberg, A.A., 2001. Genome-wide expression analysis reveals dysregulation of myelination-related genes in chronic schizophrenia. Proc. Natl. Acad. Sci. U. S. A. 98 (8), 4746–4751.

Hashimoto, K., Fukushima, T., Shimizu, E., Komatsu, N., Watanabe, H., Shinoda, N., Nakazato, M., Kumakiri, C., Okada, S., Hasegawa, H., Imai, K., Iyo, M., 2003. Decreased serum levels of D-serine in patients with schizophrenia: evidence in support of the N-methyl-D-aspartate receptor hypofunction hypothesis of schizophrenia. Arch. Gen. Psychiatry 60 (6), 572–576.

Hemby, S.E., Ginsberg, S.D., Brunk, B., Arnold, S.E., Trojanowski, J.Q., Eberwine, J.H., 2002. Gene expression profile for schizophrenia: discrete neuron transcription patterns in the entorhinal cortex. Arch. Gen. Psychiatry 59 (7), 631–640.

Huang da, W., Sherman, B.T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M.W., Lane, H.C., Lempicki, R.A., 2007. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res. 35, W169–W175 Web Server issue.

Iwamoto, K., Kato, T., 2006. Gene expression profiling in schizophrenia and related mental disorders. Neuroscientist 12 (4), 349–361.

Iwamoto, K., Kakiuchi, C., Bundo, M., Ikeda, K., Kato, T., 2004. Molecular characterization of bipolar disorder by comparing gene expression profiles of postmortem brains of major mental disorders. Mol. Psychiatry 9 (4), 406–416.

Iwamoto, K., Bundo, M., Kato, T., 2005. Altered expression of mitochondria-related genes in postmortem brains of patients with bipolar disorder or schizophrenia, as revealed by large-scale DNA microarray analysis. Hum. Mol. Genet. 14 (2), 241–253.

Jaeger, J., Sengupta, R., Ruzzo, W.L., 2003. Improved gene selection for classification of microarrays. Pac. Symp. Biocomput. 53–64.

Khan, J., Wei, J.S., Ringner, M., Saal, L.H., Ladanyi, M., Westermann, F., Berthold, F., Schwab, M., Antonescu, C.R., Peterson, C., Meltzer, P.S., 2001. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nat. Med. 7 (6), 673–679.

Lancashire, L.J., Rees, R.C., Ball, G.R., 2008. Identification of gene transcript signatures predictive for estrogen receptor and lymph node status using a stepwise forward selection artificial neural network modelling approach. Artif. Intell. Med. 43 (2), 99–111.

Li, Q., 2007. Improvement of bias and generalizability for computer-aided diagnostic schemes. Comput. Med. Imaging Graph. 31 (4–5), 338–345.

Li, J., Liu, H., Ng, S.K., Wong, L., 2003. Discovery of significant rules for classifying cancer diagnosis data. Bioinformatics 19 (Suppl 2), ii93–ii102.

Lin, Y.H., Friederichs, J., Black, M.A., Mages, J., Rosenberg, R., Guilford, P.J., Phillips, V., Thompson-Fawcett, M., Kasabov, N., Toro, T., Merrie, A.E., van Rij, A., Yoon, H.S., McCall, J.L., Siewert, J.R., Holzmann, B., Reeve, A.E., 2007. Multiple gene expression classifiers from different array platforms predict poor prognosis of colorectal cancer. Clin. Cancer Res. 13 (2), 498–507.

Liu, B., Cui, Q., Jiang, T., Ma, S., 2004. A combinational feature selection and ensemble neural network method for classification of gene expression data. BMC Bioinformatics 5 (136), 1–12.

Middleton, F.A., Pato, C.N., Gentile, K.L., McGann, L., Brown, A.M., Trauzzi, M., Diab, H., Morley, C.P., Medeiros, H., Macedo, A., Azevedo, M.H., Pato, M.T., 2005. Gene expression analysis of peripheral blood leukocytes from discordant sib-pairs with schizophrenia and bipolar disorder reveals points of convergence between genetic and functional genomic approaches. Am. J. Med. Genet. B Neuropsychiatr. Genet. 136B (1), 12–25.

Mirnics, K., Middleton, F.A., Marquez, A., Lewis, D.A., Levitt, P., 2000. Molecular characterization of schizophrenia viewed by microarray analysis of gene expression in prefrontal cortex. Neuron 28 (1), 53–67.

Mirnics, K., Middleton, F.A., Stanwood, G.D., Lewis, D.A., Levitt, P., 2001. Disease-specific changes in regulator of G-protein signaling 4 (RGS4) expression in schizophrenia. Mol. Psychiatry 6 (3), 293–301.

Mothet, J.P., Parent, A.T., Wolosker, H., Brady Jr., R.O., Linden, D.J., Ferris, C.D., Rogawski, M.A., Snyder, S.H., 2000. D-serine is an endogenous ligand for the glycine site of the N-methyl-D-aspartate receptor. Proc. Natl. Acad. Sci. U. S. A. 97 (9), 4926–4931.

Ohmori, T., Morita, K., Saito, T., Ohta, M., Ueno, S., Rokutan, K., 2005. Assessment of human stress and depression by DNA microarray analysis. J. Med. Invest. 52, 266–271 Suppl.

Vawter, M.P., Ferran, E., Galke, B., Cooper, K., Bunney, W.E., Byerley, W., 2004. Microarray screening of lymphocyte gene expression differences in a multiplex schizophrenia pedigree. Schizophr. Res. 67 (1), 41–52.

O'Neill, M.C., Song, L., 2003. Neural network analysis of lymphoma microarray data: prognosis and diagnosis near-perfect. BMC Bioinformatics 4 (13), 1–12.

Radich, J.P., Mao, M., Stepaniants, S., Biery, M., Castle, J., Ward, T., Schimmack, G., Kobayashi, S., Carleton, M., Lampe, J., Linsley, P.S., 2004. Individual-specific variation of gene expression in peripheral blood leukocytes. Genomics 83 (6), 980–988.

Rainen, L., Oelmueller, U., Jurgensen, S., Wyrich, R., Ballas, C., Schram, J., Herdman, C., Bankaitis-Davis, D., Nicholls, N., Trollinger, D., Tryon, V., 2002. Stabilization of mRNA expression in whole blood samples. Clin. Chem. 48 (11), 1883–1890.

Rhodes, D.R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., Barrette, T., Pandey, A., Chinnaiyan, A.M., 2004. Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. Proc. Natl. Acad. Sci. U. S. A. 101 (25), 9309–9314.

Schwarz, E., Bahn, S., 2008. The utility of biomarker discovery approaches for the detection of disease mechanisms in psychiatric disorders. Br. J. Pharmacol. 153 (Suppl 1), S133–S136.

Storey, J.D., Tibshirani, R., 2003. Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. U. S. A. 100 (16), 9440–9445.

Sugai, T., Kawamura, M., Iritani, S., Araki, K., Makifuchi, T., Imai, C., Nakamura, R., Kakita, A., Takahashi, H., Nawa, H., 2004. Prefrontal abnormality of schizophrenia revealed by DNA microarray: impact on glial and neurotrophic gene expression. Ann. N.Y. Acad. Sci. 1025, 84–91.

Sullivan, P.F., Fan, C., Perou, C.M., 2006. Evaluating the comparability of gene expression in blood and brain. Am. J. Med. Genet. B Neuropsychiatr. Genet. 141B (3), 261–268.

Tkachev, D., Mimmack, M.L., Ryan, M.M., Wayland, M., Freeman, T., Jones, P.B., Starkey, M., Webster, M.J., Yolken, R.H., Bahn, S., 2003. Oligodendrocyte dysfunction in schizophrenia and bipolar disorder. Lancet 362 (9386), 798–805.

Tsuang, M.T., Nossova, N., Yager, T., Tsuang, M.M., Guo, S.C., Shyu, K.G., Glatt, S.J., Liew, C.C., 2005. Assessing the validity of blood-based gene expression profiles for the classification of schizophrenia and bipolar disorder: a preliminary report. Am. J. Med. Genet. B Neuropsychiatr. Genet. 133B (1), 1–5.

Whitney, A.R., Diehn, M., Popper, S.J., Alizadeh, A.A., Boldrick, J.C., Relman, D.A., Brown, P.O., 2003. Individuality and variation in gene expression patterns in human blood. Proc. Natl. Acad. Sci. U. S. A. 100 (4), 1896–1901.

Xu, Y., Selaru, F.M., Yin, J., Zou, T.T., Shustova, V., Mori, Y., Sato, F., Liu, T.C., Olaru, A., Wang, S., Kimos, M.C., Perry, K., Desai, K., Greenwald, B.D., Krasna, M.J., Shibata, D., Abraham, J.M., Meltzer, S.J., 2002. Artificial neural networks and gene filtering distinguish between global gene expression profiles of Barrett's esophagus and esophageal cancer. Cancer Res. 62 (12), 3493–3497.

Zvara, A., Szekeres, G., Janka, Z., Kelemen, J.Z., Cimmer, C., Santha, M., Puskas, L.G., 2005. Over-expression of dopamine D2 receptor and inwardly rectifying potassium channel genes in drug-naive schizophrenic peripheral blood lymphocytes as potential diagnostic markers. Dis. Markers 21 (2), 61–69.

Association of the *HSPG2* Gene with Neuroleptic-Induced Tardive Dyskinesia

Aoi Syu[1,2], Hiroki Ishiguro[1,2], Toshiya Inada[3], Yasue Horiuchi[1,2], Syunsuke Tanaka[1], Maya Ishikawa[1], Makoto Arai[4], Masanari Itokawa[4], Kazuhiro Niizato[5], Shuji Iritani[5], Norio Ozaki[6], Makoto Takahashi[7], Akiyoshi Kakita[8], Hitoshi Takahashi[8], Hiroyuki Nawa[8], Kazuko Keino-Masu[9], Eri Arikawa-Hirasawa[10] and Tadao Arinami*[1,2]

[1]Department of Medical Genetics and Department of Molecular Neurobiology, Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan; [2]CREST, Japan Science and Technology Agency, Kawaguchi-shi, Saitama, Japan; [3]Institute of Neuropsychiatry, Seiwa Hospital, Tokyo, Japan; [4]Department of Schizophrenia Research, Tokyo Institute of Psychiatry, Tokyo, Japan; [5]Department of Psychiatry, Tokyo Metropolitan Matsuzawa Hospital, Tokyo, Japan; [6]Department of Psychiatry, School of Medicine, Nagoya University, Nagoya, Aichi, Japan; [7]Department of Psychiatry, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan; [8]Brain Research Institute, Niigata University, Niigata, Japan; [9]Department of Molecular Neurobiology, Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan; [10]Research Institute for Diseases of Old Age, Department of Neurology, Juntendo University School of Medicine, Tokyo, Japan

Tardive dyskinesia (TD) is characterized by repetitive, involuntary, and purposeless movements that develop in patients treated with long-term dopaminergic antagonists, usually antipsychotics. By a genome-wide association screening of TD in 50 Japanese schizophrenia patients with treatment-resistant TD and 50 Japanese schizophrenia patients without TD (non-TD group) and subsequent confirmation in independent samples of 36 treatment-resistant TD and 136 non-TD subjects, we identified association of a single nucleotide polymorphism, rs2445142, (allelic $p = 2 \times 10^{-5}$) in the *HSPG2* (heparan sulfate proteoglycan 2, perlecan) gene with TD. The risk allele was significantly associated with higher expression of *HSPG2* in postmortem human prefrontal brain ($p < 0.01$). Administration of daily injection of haloperidol (HDL) for 50 weeks significantly reduced *Hspg2* expression in mouse brains ($p < 0.001$). Vacuous chewing movements (VCMs) induced by 7-week injection of haloperidol–reserpine were significantly infrequent in adult *Hspg2* hetero-knockout mice compared with wild-type littermates ($p < 0.001$). Treatment by the acetylcholinesterase inhibitor, physostigmine, was significantly effective for reduction of VCMs in wild-type mice but not in *Hspg2* hetero-knockout mice. These findings suggest that the *HSPG2* gene is involved in neuroleptic-induced TD and higher expression of *HSPG2*, probably even after antipsychotic treatment, and may be associated with TD susceptibility.

INTRODUCTION

Antipsychotic-induced tardive dyskinesia (TD) is an involuntary movement disorder that develops in patients who are undergoing long-term treatment with antipsychotic medications. The clinical symptoms most commonly involve orobuccal, lingual, and facial muscles, especially in older individuals. The lingual involvement in the form of fine vermicular movements of the tongue while it is sitting at the base of the oral cavity is a common early feature (Sachdev, 2000). In more severe cases, the movements may involve trunk and limbs (Tarsy and Baldessarini, 2006). Such movements lower the quality of life (QOL) of patients (Gerlach, 2002). Therefore, predicting those patients who are vulnerable to TD remains a high priority for psychiatrists in selecting the best medication for a given individual. Introduction of second-generation atypical antipsychotics has reduced the occurrence of TD to approximately 1% annually compared with the 5% frequency with typical agents (de Leon, 2007; Remington, 2007). Owing to the lack of effective treatments for TD, however, therapeutic management of TD can be problematic for schizophrenia patients receiving antipsychotic medications, especially for those patients who develop severe treatment-resistant TD. Therefore, the strategies to prevent TD are often discussed

in the context of safety and use of antipsychotic drugs (Inada *et al*, 2008).

The etiology of TD is complex and remains unclear. Age, gender, and ethnicity are all the suggested risk factors for TD. Smoking, drinking, and using street drugs may also increase the risk of TD (Menza *et al*, 1991). There is some evidence for a genetic component to TD (Muller *et al*, 2004) and molecular genetic studies of TD were conducted to identify genes related to TD (Malhotra *et al*, 2004).

The pathophysiology of TD is not completely understood. The causative role of antipsychotic and other dopamine antagonists resulted in the proposal of the dopamine supersensitivity hypothesis of TD (Klawans *et al*, 1980). However, as the hypothesis explains only some aspects of TD, many other pathophysiological models including changes in other neurotransmitter signaling systems that are affected by neuroleptics have been considered. They include gamma-aminobutyric acid (Gerlach and Casey, 1988), norepinephrine (Saito *et al*, 1986), serotonin (Haleem, 2006), and acetylcholine (Ach) (Tammenmaa *et al*, 2002).

The advent of single nucleotide polymorphism (SNP) chips for genome-wide association analysis has made screening of susceptibility genes for TD possible. We carried out a genome-wide association study of treatment-resistant TD in schizophrenia patients and reported that SNPs associated with TD were aggregated significantly in genes belonging to the gamma-aminobutyric acid receptor signaling pathway (Inada *et al*, 2008). In this study, we analyzed the *HSPG2* gene, which includes SNPs that showed the most significant association with TD in our genome-wide association study.

MATERIAL AND METHODS

Ethical Considerations

This study was initiated after approval by the ethics committee of each institution. Written informed consent was obtained from all patients after adequate explanation of the study.

Human Subjects

Human subjects in this study were 86 Japanese schizophrenia patients with TD and 186 Japanese schizophrenia patients without TD, who have been described elsewhere (Inada *et al*, 2008). Briefly, subjects were identified at psychiatric hospitals located around Tokyo and Nagoya areas of Japan. All patients satisfied the diagnostic criteria of DSM-IV (Association, 1994) for schizophrenia. All subjects and their parents were of Japanese descent. All subjects had been receiving antipsychotic therapy for at least 1 year and their TD status was monitored for at least 1 year. TD was assessed according to the Japanese version of the Abnormal Involuntary Movement Scale. TD was diagnosed according to the criteria proposed by Schooler and Kane (Schooler and Kane, 1982). Once TD was identified, the patients were followed up and received standard therapeutic regimens for TD to minimize TD symptoms. If TD persisted even after 1 year of therapy, patients were considered potential treatment-resistant TD

patients. Treatment-resistant TD patients were defined as those patients with dyskinetic movements that persisted for more than 1 year and did not improve even after 1 year of appropriate treatment after guideline-recommended therapeutic regimens for TD. We hypothesized that treatment-resistant TD, a severe form of TD, was suitable for detection of genetic association with TD. Only treatment-resistant TD patients were included as those affected with TD in this study.

Genotyping, Resequencing, and Statistics

Association screening was performed using the Illumina Sentrix Human-1 Genotyping 109k BeadChip according to the manufacturer's instructions (Illumina, San Diego CA, USA). All DNA samples were subjected to rigorous quality control to check for fragmentation and amplification. Approximately 750 ng of genomic DNA was used in each sample. Normalized bead intensity data obtained for each sample was entered into the Illumina BeadStudio 3.0 software, which converted fluorescence intensities into SNP genotypes. A GenCall Score of 0.85 was used as a minimum threshold for per-sample genotyping completeness. The mean call rate across all samples was 97%. After removing SNPs with a low genotyping rate ($p < 0.95$: $n = 3952$), SNPs deviating from the Hardy–Weinberg equilibrium ($p < 0.001$: $n = 135$), SNPs with low minor allele frequency (MAF < 0.05: $n = 2762$), and SNPs located outside exons and introns, we screened for SNPs associated with TD using 40 573 SNPs. SNPs located within 10 kb from the 5′ and 3′ ends of known genes were included. SNPs in the linkage disequilibrium (LD) of $r^2 > 0.8$ with other SNPs were excluded. The call rate was at least 99.4% for the 40 573 SNPs. The concordance rate was evaluated by comparisons of genotypes in the 100 screening samples and this gave concordance of over 98% for each sample. Genotyping using TaqMan probes (Applied Biosystems, Foster City, CA, USA) was carried out twice for each SNP, and genotype concordance was 99.7%. Genotyping completeness was > 0.99. We treated these uncalled or discrepant genotypes as missing genotypes.

To screen for novel polymorphisms, we used direct sequencing with a Big Dye Terminator Cycle Sequencing kit and ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). All exons and the exon–intron junctions of the *HSPG2* gene were amplified from the genomic DNAs of the 86 TD group patients. The sequences of primers for mutation screening are available on request.

For a more detailed analysis of the associations between SNPs in the *HSPG2* gene and TD, the tag SNPs in the gene were selected using the Haploview program (http://www.broad.mit.edu/mpg/haploview/) with the condition of an r^2 threshold of 0.8 and a minor allele frequency of 0.1, and genotyped by the TaqMan method. Allelic discrimination was performed using the ABI PRISM 7900HT Sequence Detection System using SDS 2.0 software (Applied Biosystems, Foster City, CA, USA).

Allelic associations between SNPs and TD, and departure from the Hardy–Weinberg equilibrium were evaluated by χ^2 test or Fisher's exact test. Bonferroni's correction for multiple comparisons was applied.

Human Postmortem Brains

Brain specimens were from individuals of European (Australian) and Japanese descent. The Australian sample comprised 10 schizophrenic patients and 10 age- and gender-matched controls. The diagnosis of schizophrenia was made according to the Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV criteria (American Psychiatric; Association (1994)) by a psychiatrist and a senior psychologist. Control subjects had no known history of psychiatric illness. Tissue blocks were cut from gray matter in an area of the prefrontal cortex referred to as Brodmann's area 9 (BA9). Japanese samples of BA9 gray matter from Japanese brain specimens consisted of six schizophrenic patients and 11 age- and gender-matched controls. In addition, postmortem brains of 37 deceased Japanese patients with schizophrenia were also analyzed. The Japanese subjects met the DSM-III-R criteria for schizophrenia. Details of the condition of the postmortem brains have been described elsewhere (Ishiguro et al, 2008; Koga et al, 2009).

Analysis of *Hspg2* Transcription in Human Brain Tissue

Total RNA was extracted from human brain tissues with ISOGEN Reagent (Nippon Gene, Tokyo, Japan). The RNA quality was checked using a Nanodrop ND-1000 spectrophotometer (LMS, Tokyo, Japan) to have an OD 260/280 ratio of 1.8–2 and an OD 260/230 of 1.8 or greater. Expression of the *HSPG2* genes was analyzed by the TaqMan real-time polymerase chain reaction system (Applied Biosystems, Foster City, CA). From RNA, cDNA was synthesized with Revertra Ace (Toyobo, Tokyo, Japan) and oligo dT primers. Expression of the *HSPG2* gene was analyzed with an ABI PRISM 7900 HT Sequence Detection System (Applied Biosystems), with the TaqMan gene expression assays for *HSPG2* (Hs01078535_m1), and normalized to the expression of Human GAPDH Control Reagents (Applied Biosystems).

Genotype effects on *HSPG2* expression were analyzed in Australian subjects and replicated in Japanese subjects using analysis of variance followed by Tukey's *post hoc* tests by JMP software version 7.0.1 (SAS Institute, Cary, NC, USA) was used.

Animals

Animals were same-sex housed before behavior testing. The same animals were used for all behavior tests.

Four-week-old C57BL/6J male mice (weight: 20–25 g) treated with haloperidol (HDL) or vehicle-saline and 7-week-old male mice (wild type: 8; $Hspg2^{+/-}$: 7) with orofacial dyskinesia were housed under 10 h : 14 h light/dark conditions with normal food and water *ad libitum*, with mice housed separately in groups of 4 or 5 mice.

The generation of *Hspg2* knockout mice and the phenotypes of the mice have been described elsewhere (Arikawa-Hirasawa et al, 1999). As *Hspg2* null mice are embryonic lethal, timed matings between heterozygotes were carried out to generate homozygous and wild-type mice in this study.

All animal protocols were approved by the Animal Care and Use committee of University of Tsukuba.

Drugs

Reserpine (methyl reserpate 3,4,5-trimethoxycinnamic acid ester; Wako, Osaka, Japan) and HDL (Wako, Osaka, Japan) were diluted in glacial acetic acid and then diluted in distilled water. Physostigmine (Wako), a reversible cholinesterase inhibitor, was diluted in saline. All solutions were treated subcutaneously in volumes not exceeding 10 ml/kg body weight.

HDL Treatment

To examine the effects of antipsychotic treatments on gene expression, we made two groups: an acute treatment group: 4-week-old C57BL/6J male mice were treated with intraperitoneal injection (i.p.) of 1.0 mg/kg HDL ($n = 10$) or vehicle-saline ($n = 10$) once each day for 4 weeks; and a long-term treatment group: 4-week-old C57BL/6J male mice were treated with intraperitoneal injection of 1.0 mg/kg HDL ($n = 10$) or vehicle-saline ($n = 10$) once each day for 50 weeks. Mice were killed 4 h after the last injection to obtain brain tissues.

Induction of Vacuous Chewing Movements

Mice were treated with i.p. of 2 mg/kg HDL and 0.3 mg/kg reserpine every day for 7 weeks to induce the putative TD analogue vacuous chewing movements (VCMs) (Araujo et al, 2004; Burger et al, 2005; Naidu et al, 2003). Before injection and 4 hours after the injection on the 47th day, locomotor activity test and rotarod test were carried out. On the 48th and 49th days, 1, 2, 3, 4, and 24 h after the last injection, the animals were observed for quantification of VCMs for 2 days. On the 50th day, to verify the effects of physostigmine on VCMs, mice were injected with 0.1 mg/kg physostigmine. At 1, 2, 3, 4, and 24 h after the injection of physostigmine, the animals were observed for quantification of VCMs. On the 51st and 52nd day, mice were treated with 2 mg/kg HDL and 0.3 mg/kg reserpine and then observed for quantification of VCMs. On the 53rd day, mice were treated with vehicle-saline, and 1, 2, 3, 4, and 24 h after the injection, the animals were observed for quantification of VCMs.

Analysis of *Hspg2* Transcription in Brain Tissue of Mice

The prefrontal cortex, midbrain, hippocampus, thalamus, and striatum were taken by dissection, and total RNA was extracted with an RNeasy kit (Qiagen, K.K., Tokyo, Japan). After cDNA synthesis from total RNA samples, the transcription level of cDNA samples was analyzed by a TaqMan Expression assay for *Hspg2* (Mm00464581_m1; Applied Biosystems) and normalized to that of rodent *Gapdh* with Rodent *Gapdh* Control Reagents (Applied Biosystems). The average relative expression levels of five regions were compared with the saline groups by Student's *t*-test.

1158

Table 1 Allelic p-Values of SNPs for Association with TD in Screening and Replication Samples

SNP	Chromosome	Location	Gene	Position relative to gene	Allele frequency			Uncorrected allelic p		
					TD group	Non-TD group	Screening	Replication	Combined
rs7529452	chr1	1p36.22	PLOD1	coding	0.396	0.381	**0.001**	**0.05**	
rs2445142	chr1	1p36.12	HSPG2	intron	0.579	0.380	**0.001**	**0.002**	**0.00002**
rs1934712	chr1	1p21.1	COL11A1	flanking_3UTR	0.435	0.343	**0.0007**	0.98	
rs2306444	chr1	1p12	MAN1A2	intron	0.482	0.392	**0.0005**	0.59	
rs869807	chr1	1p12	TBX15	flanking_3UTR	0.282	0.185	**0.0009**	0.97	
rs6668395	chr1	1q41	DUSP10	flanking_5UTR	0.418	0.535	**0.001**	0.59	
rs6426327	chr1	1q44	SMYD3	intron	0.412	0.324	**0.00002**	0.20	
rs4558632	chr2	2p21	EML4	flanking_5UTR	0.253	0.171	**0.0003**	0.82	
rs6714424	chr2	2p16.2	ASB3	flanking_5UTR	0.212	0.120	**0.0007**	0.76	
rs2060279	chr2	2p12	LRRTM4	flanking_5UTR	0.685	0.777	**0.0005**	0.52	
rs11694702	chr2	2q13	BUB1	flanking_5UTR	0.329	0.241	**0.001**	0.76	
rs1873201	chr2	2q24.3	KCNH7	flanking_5UTR	0.395	0.301	**0.002**	0.78	
rs11688866	chr2	2q31.3	UBE2E3	flanking_5UTR	0.452	0.581	**0.0005**	0.45	
rs3749279	chr3	3p22.3	STAC	intron	0.202	0.099	**0.001**	0.59	
rs6443468	chr3	3q26.32	TBL1XR1	flanking_5UTR	0.373	0.511	**0.001**	0.26	
rs13115988	chr4	4q22.1	LOC285513	3UTR	0.694	0.645	**0.001**	0.32	
rs700237	chr5	5p13.1	C9	flanking_5UTR	0.898	0.830	**0.0009**	0.80	
rs832582	chr5	5q11.2	MGC33648	flanking_5UTR	0.641	0.543	**0.001**	0.37	0.03
rs13153252	chr5	5q14.3	EDIL3	intron	0.375	0.486	**0.0002**	0.94	
rs6594324	chr5	5q21.3	FER	flanking_5UTR	0.692	0.578	**0.001**	0.37	
rs915125	chr6	6q14.1	FAM46A	flanking_5UTR	0.207	0.262	**0.0009**	0.41	
rs2691180	chr6	6q21	CDC2L6	intron	0.789	0.890	**0.0007**	0.41	
rs9376506	chr6	6q24.1	CITED2	flanking_5UTR	0.564	0.457	**0.001**	0.71	
rs1832445	chr6	6q24.1	FLJ39824	flanking_3UTR	0.494	0.604	**0.001**	0.94	
rs3735478	chr7	7p13	DKFZp761I2123	coding	0.058	0.145	**0.001**	0.32	0.006
rs1047053	chr7	7q36.2	DPP6	3UTR	0.657	0.758	**0.0005**	0.3	0.02
rs2583086	chr8	8q13.2	SULF1	intron	0.169	0.255	**0.001**	0.95	
rs4738269	chr8	8q13.3	KCNB2	intron	0.571	0.396	**0.0007**	**0.04**	**0.0002**
rs2927111	chr8	8q23.1	STARS	flanking_5UTR	0.369	0.487	**0.0004**	0.4	0.01
rs3019982	chr8	8q23.1	ANGPT1	flanking_3UTR	0.612	0.479	**0.0002**	0.18	0.004
rs4242345	chr8	8q24.13	ANXA13	flanking_3UTR	0.659	0.764	**0.00004**	0.56	
rs1413299	chr9	9q22.33	COL15A1	intron	0.682	0.746	**0.001**	0.73	
rs2274359	chr10	10p15.3	RBM17	intron	0.929	0.834	**0.0005**	0.51	
rs1932596	chr10	10q21.1	PCDH15	intron	0.628	0.543	**0.0007**	0.51	
rs1058198	chr10	10q22.3	DLG5	coding	0.152	0.290	**0.0006**	0.22	0.0007
rs10748816	chr10	10q24.32	ELOVL3	intron	0.494	0.634	**0.00008**	0.13	0.002
rs2246775	chr10	10q24.32	GBF1	intron	0.646	0.747	**0.0009**	0.66	
rs765934	chr10	10q26.3	MGMT	flanking_5UTR	0.732	0.642	**0.0012**	0.74	
rs886292	chr11	11p15.1	ABCC8	intron	0.825	0.696	**0.0005**	**0.02**	0.0015
rs286925	chr11	11p13	EHF	5UTR	0.542	0.611	**0.0005**	0.75	
rs568758	chr11	11q13.4	SPCS2	intron	0.738	0.818	**0.0009**	0.82	
rs624786	chr11	11q13.4	NEU3	flanking_5UTR	0.735	0.812	**0.0015**	0.83	
rs1444590	chr12	12q13.11	SLC38A1	intron	0.789	0.839	**0.0005**	0.39	
rs1154664	chr12	12q24.32	KIAA1906	flanking_3UTR	0.688	0.590	**0.0002**	0.80	
rs1924174	chr13	13q33.3	LIG4	flanking_3UTR	0.282	0.195	**0.0013**	0.19	0.04
rs1189827	chr14	14q22.3	SEC10L1	flanking_3UTR	0.741	0.663	**0.0007**	0.45	
rs11625123	chr14	14q32.12	ITPK1	intron	0.124	0.225	**0.0009**	0.45	0.007
rs10140345	chr14	14q32.2	VRK1	flanking_3UTR	0.300	0.273	**0.0011**	0.09	

128

Table 1 Continued

SNP	Chromosome	Location	Gene	Position relative to gene	Allele frequency			Uncorrected allelic *p*		
					TD group	Non-TD group	Screening	Replication	Combined
rs2061051	chr15	15q12	GABRG3	intron	0.206	0.350	**0.0014**	**0.04**	**0.0006**
rs3764211	chr15	15q13.1	APBA2	flanking_3UTR	0.726	0.815	**0.0013**	0.12	0.005
rs1036673	chr15	15q24.1	PML	3UTR	0.721	0.592	**0.0006**	0.64	
rs3809729	chr17	17p12	DNAH9	flanking_5UTR	0.867	0.869	**0.0007**	0.11	
rs4630608	chr17	17p11.2	FBXW10	intron	0.250	0.274	**0.0010**	**0.04**	
rs2287352	chr17	17q12	ACACA	flanking_5UTR	0.247	0.306	**0.0014**	0.85	
rs3744165	chr17	17q25.3	FLJ13841	5UTR	0.093	0.130	**0.0010**	0.15	
rs474122	chr18	18p11.31	DLGAP1	flanking_5UTR	0.404	0.330	**0.0002**	0.38	
rs12460403	chr19	19p13.3	HMG20B	flanking_3UTR	0.285	0.194	**0.0011**	0.55	
rs437168	chr19	19q13.12	NPHS1	coding	0.223	0.139	**0.0007**	0.9	
rs10419669	chr19	19q13.31	CBLC	intron	0.094	0.179	**0.0003**	0.84	
rs8112223	chr19	19q13.41	HAS1	flanking_5UTR	0.314	0.219	**0.0003**	0.43	
rs2328500	chr20	20p11.23	C20orf26	intron	0.376	0.324	**0.001**	0.12	
rs7281019	chr21	21q22.11	TCP10L	intron	0.924	0.862	**0.00008**	0.57	0.04
rs2056965	chr22	22q12.3	LOC91464	flanking_5UTR	0.422	0.348	**0.0002**	0.49	

Abbreviations: SNP, single nucleotide polymorphism; TD, tardive dyskinesia.
p-Values with bold emphasis indicate $p < 0.05$ in 1st *p* and 2nd *p*, and combined $p < $ 1st *p*.

Evaluation of VCMs

Mice were placed individually in observation cages ($16 \times 17 \times 19 \, cm^3$) without food. Hand-operated counters were used to quantify VCMs continuously for 5 min. VCMs were referred to as single mouth openings in the vertical plane not directed toward physical material. If VCMs occurred during a period of grooming, they were not taken into account. Mirrors were placed under the floor and behind the back wall of the cage to permit observation of oral movements when the animal faced away from the observer. The observations were made by two observers who were blind to the animal's group assignment. The observation criteria were not subjective, because an excellent inter-observer agreement was found in a previous pilot experiment (Pearson's correlation = 0.98). All behavioral experiments were conducted between 1000 and 1800 hours.

Locomotor Activity

The locomotor activity test was conducted between 1200 and 1700 hours in a dimly lit testing room. Mice were habituated to the room for at least 30 min before testing. The locomotor activity test was videotaped with a Sony Digital Video Camera (Sony, Tokyo, Japan). The behavioral testing apparatus was a black Plexiglas rectangular box (41 cm long × 22 cm wide × 20.5 cm tall) and activity was recorded for 20 min. The total distance traveled (locomotion) was scored.

Rotarod Test

The rotarod test was conducted between 1200 and 1700 hours in a dimly lit testing room. All mice were brought to the testing room in their home cages and were allowed to sit undisturbed in the testing room for at least 5 min before the start of behavioral testing. Motor performance was assessed by rotarod (Med Associates, St Albans, VT). A 1-min training session was given to each mouse on the rotarod (diameter 8 cm, 7 rpm) 5 min before the first measurement. Motor performance (time until the first fall) was registered during a 2-min session.

Statistical Analysis for Behavioral Data and Gene Expression in Animal Experiments

Effects of genotype, drug treatment, and time were analyzed using analysis of variance) followed by Tukey's *post hoc* tests or using Student's *t*-test. Individual differences of the number of VCMs between before and after injection of physostigmine and saline were tested by nonparametric test for one sample test of mean = 0.

RESULTS

Association Study

We screened for SNPs associated with TD using 40 573 tag SNPs on the Sentrix® Human-1 Genotyping BeadChip (Illumina) to identify loci associated with susceptibility to TD in 50 TD and 50 non-TD subjects (Inada *et al*, 2008). The potential impact of population structure on this association study was evaluated by using the genome-wide χ^2 inflation factor, λ, as a genomic control (Devlin and Roeder, 1999; Devlin *et al*, 2001). The estimated value of λ was 1.04, by which genome-wide association *p*-values were corrected. The lowest uncorrected allelic *p*-value for association with TD was 1×10^{-5}. Therefore, no SNP was

significantly associated with TD after Bonferroni correction. An attempt was made to replicate the association of 63 SNPs, which were allelic p-values < 0.002 and located within 10 kb from known genes with the TaqMan genotyping assay (Table 1). A potential association was found for four SNPs (allelic $p < 0.05$) (Table 1). However, no significant association was found after correction for multiple testing of 63 SNPs in the replication sample only. Among these four SNPs, an association between *GABRG3* SNP and TD has already been reported (Inada et al, 2008). The lowest allelic p-value for the association was found for rs2445142 $(p = 2 \times 10^{-5})$ when the initial genome-wide sample and replication sample were combined. The SNP is located in the *HSPG2* gene.

Next, we tested associations of 24 tag SNPs including rs2445142 in the *HSPG2* gene with TD and found a nominal significant association for five SNPs in addition to rs2445142 (Table 2). Other than rs2445142, we found a significant association of rs2124368 located in intron 43 of the *HSPG2* gene with TD even after applying Bonferroni's correction for multiple testing (uncorrected allelic $p = 0.0003$, corrected $p = 0.007$). The SNP rs2124368 was not in LD with rs2445142, which was located in intron 1 $(D' = 0.13, r^2 = 0.01)$.

Subsequently, we genotyped the SNPs of rs2501255 (intron 1), rs2501257 (intron 1), rs897474 (intron 3), rs2254357 (exon 6), rs2254358 (exon 6), and rs2497632 (intron 9), because of the expected complete LD between these SNPs and rs2445142 based on the HapMap database. We confirmed that they were in complete LD with rs2445142 $(r^2 = 1.00)$. These SNPs were located in introns 1-9 of the *HSPG2* gene. Age, sex, and age of onset were not associated with these SNPs. Acute extrapyramidal symptoms were associated with rs2445142 and the six SNPs in complete LD with rs2445142 (allelic $p = 0.00002$). Resequencing all exons of the *HSPG2* gene in patients with TD did not reveal novel SNPs. Finally, we genotyped missense SNPs of rs3736360, rs2229493, s2291827, rs2228349, rs2229491, rs2229490, rs2229489, rs2229475, rs897471, rs2229481, and rs989994, which were listed in dbSNP and did not find significant associations of these SNPs with TD (data not shown). Thus, we tested a total of 103 SNPs, including 41 SNPs in the *HSPG2* gene, in our total subjects of 86 TD and 136 non-TD patients.

Association Between *Hspg2* Expression Levels in the Postmortem Prefrontal Cortex and Rs2445142

The transcription level in the postmortem prefrontal cortex, as measured by TaqMan real-time polymerase chain reaction, was not significantly different by diagnosis, age, sex, postmortem intervals, or pH of brain samples. A significant genotype effect on *HSPG2* gene expression levels was observed in 20 Australian subjects $(F(2, 17) = 4.9, p = 0.02)$ and replicated in 54 Japanese subjects $(F(2, 51) = 3.5, p = 0.04)$. The association was significant in the combined subjects $(F(2, 71) = 7.6, p = 0.001)$. Tukey's *post hoc* tests showed that *HSPG2* expression levels were significantly higher in the subjects with the GG genotype than in those with the CC genotype (Figure 1). Unfortunately, information about TD in the brains we analyzed was not available.

Table 2 Allelic p-Values of Tag SNPs in the *HSPG2* Gene for Association with TD

Location		Allele	Allele frequency*		
			TD group	Non-TD group	Allelic p
rs3736360	exon 96 (N4331S)	A/G	0.19	0.20	0.8715
rs3767137	intron 77	A/G	0.23	0.19	0.2759
rs10917053	intron 71	A/G	1.00	0.99	0.3308
rs7355045	intron 64	G/A	0.84	0.81	0.4235
rs2290501	intron 60	C/A	0.22	0.21	0.7134
rs1563370	intron 52	A/G	0.35	0.27	0.0687
rs2229475	exon 47 (I1967V)	G/A	0.01	0.01	0.9477
rs2305562	intron 43	A/G	0.61	0.49	0.0117
rs4654991	intron 42	G/A	0.39	0.36	0.5605
rs2124368	intron 42	G/A	0.77	0.60	0.0003
rs897472	intron 36	C/A	0.09	0.05	0.1098
rs897471	exon 36 (V1503A)	A/G	0.88	0.87	0.7005
rs2229478	exon 8 (L248L)	A/G	0.53	0.42	0.0273
rs3767141	intron 6	G/A	0.66	0.60	0.1811
rs2445142	intron 1	G/C	0.58	0.38	0.00002
rs878949	intron 1	A/G	0.22	0.20	0.5867
rs1545593	intron 1	C/A	0.41	0.30	0.0122
rs1002480	intron 1	G/C	0.41	0.32	0.0368
rs6698486	intron 1	G/A	0.46	0.38	0.0754
rs10799719	intron 1	G/A	0.80	0.75	0.1789
rs9426785	intron 1	A/G	0.57	0.55	0.7389
rs4654773	intron 1	A/G	0.45	0.45	0.9165
rs11587857	intron 1	G/A	0.50	0.46	0.3465
rs4233280	5' flanking	A/G	0.07	0.03	0.0588

Abbreviations: SNP, single nucleotide polymorphism; TD, tardive dyskinesia
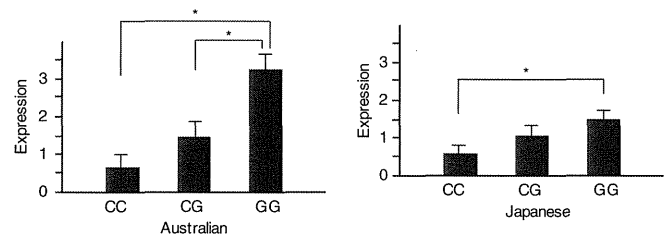*The frequency of the first allele.

Figure 1 *HSPG2* expression levels in the postmortem prefrontal region by genotype. The vertical scores show the average (SEM) of relative expression levels in each of the three genotype groups, compared with the mean gene expression in the total samples. *Indicates $p < 0.05$ by Tukey's *post hoc* tests.

Hspg2 Gene Expression in the Mouse Brains by HDL Treatment

Hspg2 expression levels were evaluated in the mouse brain after treatment with the antipsychotic drug, HDL. The expression of *Hspg2* levels did not alter after a 4-week treatment of HDL except for the striatum where *Hspg2* was expressed significantly higher than after the saline
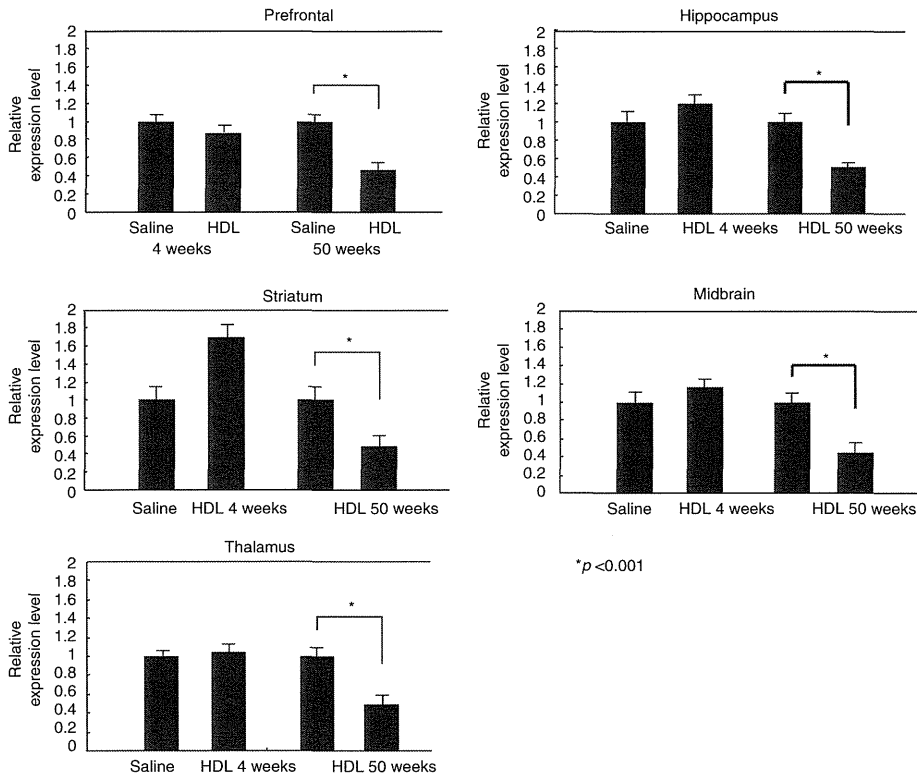
Figure 2 Effects of haloperidol (HDL) on *HSPG2* gene expression in the mouse brains. Relative expression levels of *Hspg2* from the prefrontal cortex, midbrain, hippocampus, thalamus, and striatum in the mouse brains after treatment with HDL for 4 weeks ($n = 10$) and HDL for 50 weeks ($n = 10$) were compared with the saline groups for 4 weeks ($n = 10$) and 10 weeks ($n = 10$) by Student's *t*-test.

treatment. Significantly lower expression of *Hspg2* was observed in all brain regions after a 50-week treatment with HDL than after a 50-week treatment with saline (Figure 2) ($F(1, 18) = 42.9$, $p < 0.0001$ at the prefrontal cortex; $F(1, 18) = 20.1$, $p = 0.0003$ at the hippocampus; $F(1, 18) = 15.9$, $p = 0.0009$ at the striatum; $F(1, 17) = 19.3$, $p = 0.0004$ at the midbrain; $F(1, 18) = 16.5$, $p = 0.0007$ at the thalamus).

Analysis of VCMs Induced by Haloperidol–Reserpine in *Hspg2* Knockout Mice

As we could not induce VCMs by administration of HDL only to mice, VCMs induced by long-term treatment with HDL and reserpine in female *Hspg2* hetero-knockout mice and female wild-type gene litters were measured to evaluate the relationship between expression levels of *Hspg2* and TD (Figure 3a). *Hspg2*-null knockout mice were embryonic lethal. The relative expression levels of *Hspg2* in *Hspg2* hetero-knockout mouse brains were almost half of that in the wild littermates (data not shown). Body weight, locomotor activities, and performance in the rotarod test before and after 48 days of administration of HDL and reserpine were not significantly different between *Hspg2* hetero-knockout and wild litters (data not shown). There was a significant effect of genotype ($F(1, 545) = 36.8$, $p < 0.0001$), post-treatment time ($F(4, 495) = 6.15$, $p < 0.0001$), and treatment ($F(3, 543) = 5.7$, $p = 0.0008$) for the number of VCMs for 5 min. *Post hoc* analysis showed that the number of VCMs were significantly lower

in hetero-knockout mice than in wild-type mice after the last injection of HDL and reserpine after 48 or 49 consecutive days of administration of HDL and reserpine, and subsequent injection of physostigmine on the 50th day, or saline on the 53rd day (Figure 3b). The response of VCMs to physostigmine was subsequently evaluated (Figure 3c). There was a significant effect of genotype ($F(1, 128) = 36.9$, $p < 0.0001$), but not post-treatment time ($F(4, 125) = 1.03$, $p = 0.39$) for individual differences in the number of VCMs between pre-injection and post-treatment time. As for saline treatment, there was no significant effect of genotype ($F(1, 118) = 0.13$, $p = 0.72$) and post-treatment time ($F(4, 115) = 0.31$, $p = 0.87$). The numbers of VCMs were significantly reduced by injection of physostigmine compared with those before the injection at 24 h after HDL and reserpine injection in the wild-type mice but the differences in the numbers of VCMs before and after injection of physostigmine were not significant in hetero-knockout mice. The number of VCMs did not significantly alter after injection of saline in hetero-knockout mice and wild-type mice.

DISCUSSION

From a genome-wide association analysis, this study identified the role of *HSPG2* in neuroleptic-induced TD. The association was not significant in the initial screening and second confirmation after correction for multiple testing. However, screening with the tag SNPs for *HSPG2*,
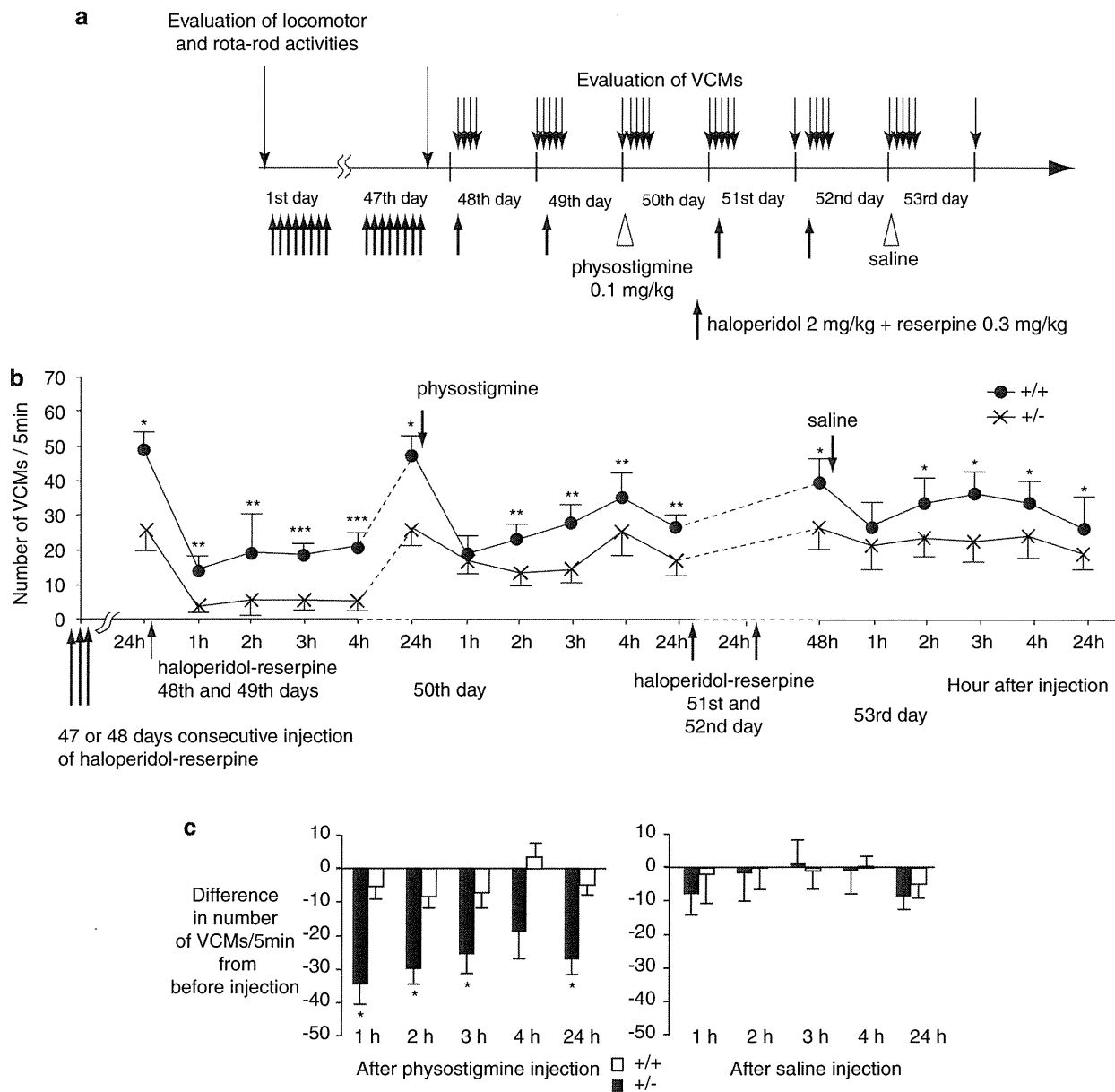
Figure 3 Analysis of vacuous chewing movements (VCMs) induced by haloperidol–reserpine in *Hspg2* knockout mice. (a) Schematic presentation of schedule of injections and measurements. (b) The average number (SEM) of VCMs for 5 min after injection. The abscissa axis shows the time after injection of HDL and reserpine, physostigmine, or saline. Significant difference between *Hspg2* hetero-knockout and wild-type mice is shown as * at $p < 0.05$, ** at $p < 0.01$, and *** at $p < 0.001$. (c) Reduction of the number of VCMs after physostigmine or saline injection. Individual differences of the number of VCMs before injection (50th day for physostigmine or 53rd day for saline) to each time after injection grouped by the genotype are shown. Significant difference from mean = 0 is shown as * at $p < 0.001$.

where the SNP (rs2445142) with the smallest p-value for association with TD in our genome-wide association study was located, identified one SNP (rs2124368) associated with TD even after correction for multiple testing. These two SNPs, which were found to be associated with TD, one identified by a genome-wide screening and another identified by screening with the tag SNPs, were not in LD. However, it is not obvious whether the finding for genetic association with TD of these SNPs in the *HSPG2* gene can be interpreted as significant, because of two steps of genome-wide association analyses before the step of screening of tag SNPs. Furthermore, the Human-1 BeadChip used in our initial screening is far from a complete genome coverage.

This may affect the credibility of the results. Confirmation of associations in other populations is necessary.

The SNP rs2445142 that showed the lowest association p-value in this study was associated with the expression levels of *HSPG2* in the human postmortem prefrontal cortex. The risk allele was associated with increased expression of *HSPG2*. The SNP rs2445142 is located in intron 1 of the *HSPG2* gene and is in complete LD with at least six SNPs located from introns 1–9. Among the SNPs associated with TD found in this study, the program TFSEARCH (http://www.cbrc.jp/research/db/TFSEARCH.html) predicts alteration of the transcription factor, LYF-1, binding affinity between the T and C alleles of rs897474 in intron 3.

Synonymous SNPs, rs2254357 (exon 6), and rs2254358 (exon 6) that were associated with TD might affect mRNA decay rates. Unfortunately, the mechanism of the association between these SNPs and HSPG2 expression levels could not be elucidated in this study.

From findings in human postmortem brain samples, we speculated that increased expression of HSPG2 is a risk factor for TD and interpreted that decreased expression of Hspg2 in mouse brains after chronic administration of HDL was a compensatory or adaptive response to neuroleptic drugs. We, therefore, hypothesized that decreased expression level of HSPG2 is protective for TD. We examined our hypothesis using hetero-knockout mice and confirmed it after finding lower numbers of VCMs in hetero-knockout mice than in the wild-type littermates after chronic administration of HDL and reserpine. We carried out the experiment using only female mice; therefore, we do not have the data on the sex difference.

The mechanism behind our hypothesis that increased expression levels of HSPG2 may induce a susceptibility to neuroleptic-induced TD is not known at present. A potential efficacy of cholinergic drugs in the treatment of TD has been reported (Caroff et al, 2001; Tammenmaa et al, 2004). AChE terminates neurotransmission at cholinergic synapses by hydrolyzing acetylcholine. At the neuromuscular junction, AChE is in the basal lamina, where AChE tetramers bind the collagen ColQ, which interacts in turn with the dystroglycan complex through perlecan (Peng et al, 1999). Perlecan is an essential component of the ColQ–AChE localization in neuromuscular junction (Rotundo et al, 2005). At central synapses, AChE tetramers bind directly to the PRiMA (Perrier et al, 2002). Although ColQ also anchors AChE in brain and heart in addition to skeletal muscle (Feng et al, 1999), the role of perlecan in acetylcholine receptor signaling in central synapses is unclear. In this study, we tested the effect of the AChE inhibitor, physostigmine, on HDL- and reserpon-induced VCMs in mice. We found significant reduction in the number of VCMs only in wild-type mice and the number of VCMs was not reduced in hetero-knockout mice. These findings indicate that perlecan may be involved in the role of AChE in TD and the genotyping and/or levels of HSPG2 may provide useful information about the effectiveness of treatment of TD with AChE.

The other important molecule to which perlecan and TD may be related is FGF2. Perlecan promotes FGF2–FGFR1 binding (Whitelock et al, 1996) and HSPGs including perlecan were upregulated by responding to injury and may have a role in intracellular trafficking of FGF2 in neurons and glia in the adult rat cerebral cortex (Leadbeater et al, 2006). Clozapine increases FGF2 expression and, on the basis of the neuroprotective activity of FGF2, a potential use of clozapine in TD was proposed (Riva et al, 1999).

Perlecan is expressed at the capillary endothelial cells in the brain and perlecan at the blood–brain barrier (BBB) may have a role in maintaining the blood–brain barrier function because of acceptance of the FGF2 secreted from astrocytes (Deguchi et al, 2002). It is reported that neuroleptics, such as HDL and chlorpromazine, alter the blood–brain barrier function and increase brain iron levels, which affect neuroleptic-induced dopamine receptor supersensitivity (Ben-Shachar et al, 1993).

Although the exact mechanisms of the association between HSPG2 and TD are unclear, this study identified the role of HSPG2 in neuroleptic-induced TD.

DISCLOSURE

The authors declare that no financial support or compensation has been received from any individual or corporate entity over the past 3 years for research or professional service and there are no personal financial holdings that could be perceived as constituting a potential conflict of interest.

REFERENCES

Araujo NP, Abilio VC, Silva RH, Pereira RC, Carvalho RC, Gonzalez C et al (2004). Effects of topiramate on oral dyskinesia induced by reserpine. Brain Res Bull 64: 331–337.

Arikawa-Hirasawa E, Watanabe H, Takami H, Hassell JR, Yamada Y (1999). Perlecan is essential for cartilage and cephalic development. Nat Genet 23: 354–358.

Association AP (1994). Diagnostic and Statistical Manual of Mental Disorders. 4th Revision. American Psychiatric Association: Washington DC.

Ben-Shachar D, Livne E, Spanier I, Zuk R, Youdim MB (1993). Iron modulates neuroleptic-induced effects related to the dopaminergic system. Isr J Med Sci 29: 587–592.

Burger ME, Fachineto R, Alves A, Callegari L, Rocha JB (2005). Acute reserpine and subchronic haloperidol treatments change synaptosomal brain glutamate uptake and elicit orofacial dyskinesia in rats. Brain Res 1031: 202–210.

Caroff SN, Campbell EC, Havey J, Sullivan KA, Mann SC, Gallop R (2001). Treatment of tardive dyskinesia with donepezil: a pilot study. J Clin Psychiatry 62: 772–775.

de Leon J (2007). The effect of atypical versus typical antipsychotics on tardive dyskinesia: A Naturalistic Study. Eur Arch Psychiatry Clin Neurosci 257: 169–172.

Deguchi Y, Okutsu H, Okura T, Yamada S, Kimura R, Yuge T et al (2002). Internalization of basic fibroblast growth factor at the mouse blood–brain barrier involves perlecan, a heparan sulfate proteoglycan. J Neurochem 83: 381–389.

Devlin B, Roeder K (1999). Genomic control for association studies. Biometrics 55: 997–1004.

Devlin B, Roeder K, Wasserman L (2001). Genomic control, a new approach to genetic-based association studies. Theor Popul Biol 60: 155–166.

Feng G, Krejci E, Molgo J, Cunningham JM, Massoulie J, Sanes JR (1999). Genetic analysis of collagen Q: roles in acetylcholinesterase and butyrylcholinesterase assembly and in synaptic structure and function. J Cell Biol 144: 1349–1360.

Gerlach J (2002). Improving outcome in schizophrenia: the potential importance of EPS and neuroleptic dysphoria. Ann Clin Psychiatry 14: 47–57.

Gerlach J, Casey DE (1988). Tardive dyskinesia. Acta Psychiatr Scand 77: 369–378.

133

Haleem DJ (2006). Serotonergic modulation of dopamine neurotransmission: a mechanism for enhancing therapeutics in schizophrenia. *J Coll Physicians Surg Pak* **16**: 556–562.

Inada T, Koga M, Ishiguro H, Horiuchi Y, Syu A, Yoshio T *et al* (2008). Pathway-based association analysis of genome-wide screening data suggest that genes associated with the gamma-aminobutyric acid receptor signaling pathway are involved in neuroleptic-induced, treatment-resistant tardive dyskinesia. *Pharmacogenet Genomics* **18**: 317–323.

Ishiguro H, Koga M, Horiuchi Y, Noguchi E, Morikawa M, Suzuki Y *et al* (2008). Supportive evidence for reduced expression of GNB1L in schizophrenia. *Schizophr Bull*. Print copy in press (originally published online 14 November, 2008, at schizophreniabulletin.oxfordjournals.org/cgi/reprint/sbn160v1).

Klawans HL, Goetz CG, Perlik S (1980). Tardive dyskinesia: review and update. *Am J Psychiatry* **137**: 900–908.

Koga M, Ishiguro H, Yazaki S, Horiuchi Y, Arai M, Niizato K *et al* (2009). Involvement of SMARCA2/BRM in the SWI/SNF chromatin-remodeling complex in schizophrenia. *Hum Mol Genet* **18**: 2483–2494.

Leadbeater WE, Gonzalez AM, Logaras N, Berry M, Turnbull JE, Logan A (2006). Intracellular trafficking in neurones and glia of fibroblast growth factor-2, fibroblast growth factor receptor 1 and heparan sulphate proteoglycans in the injured adult rat cerebral cortex. *J Neurochem* **96**: 1189–1200.

Malhotra AK, Murphy Jr GM., Kennedy JL (2004). Pharmacogenetics of psychotropic drug response. *Am J Psychiatry* **161**: 780–796.

Menza MA, Grossman N, Van Horn M, Cody R, Forman N (1991). Smoking and movement disorders in psychiatric patients. *Biol Psychiatry* **30**: 109–115.

Muller DJ, Shinkai T, De Luca V, Kennedy JL (2004). Clinical implications of pharmacogenomics for tardive dyskinesia. *Pharmacogenomics J* **4**: 77–87.

Naidu PS, Singh A, Kaur P, Sandhir R, Kulkarni SK (2003). Possible mechanism of action in melatonin attenuation of haloperidol-induced orofacial dyskinesia. *Pharmacol Biochem Behav* **74**: 641–648.

Peng HB, Xie H, Rossi SG, Rotundo RL (1999). Acetylcholinesterase clustering at the neuromuscular junction involves perlecan and dystroglycan. *J Cell Biol* **145**: 911–921.

Perrier AL, Massoulie J, Krejci E (2002). PRiMA: the membrane anchor of acetylcholinesterase in the brain. *Neuron* **33**: 275–285.

Remington G (2007). Tardive dyskinesia: eliminated, forgotten, or overshadowed? *Curr Opin Psychiatry* **20**: 131–137.

Riva MA, Molteni R, Tascedda F, Massironi A, Racagni G (1999). Selective modulation of fibroblast growth factor-2 expression in the rat brain by the atypical antipsychotic clozapine. *Neuropharmacology* **38**: 1075–1082.

Rotundo RL, Rossi SG, Kimbell LM, Ruiz C, Marrero E (2005). Targeting acetylcholinesterase to the neuromuscular synapse. *Chem Biol Interact* **157–158**: 15–21.

Sachdev PS (2000). The current status of tardive dyskinesia. *Aust N Z J Psychiatry* **34**: 355–369.

Saito T, Ishizawa H, Tsuchiya F, Ozawa H, Takahata N (1986). Neurochemical findings in the cerebrospinal fluid of schizophrenic patients with tardive dyskinesia and neuroleptic-induced parkinsonism. *Jpn J Psychiatry Neurol* **40**: 189–194.

Schooler NR, Kane JM (1982). Research diagnoses for tardive dyskinesia. *Arch Gen Psychiatry* **39**: 486–487.

Tammenmaa IA, McGrath JJ, Sailas E, Soares-Weiser K (2002). Cholinergic medication for neuroleptic-induced tardive dyskinesia. *Cochrane Database Syst Rev* **3**: CD000207.

Tammenmaa IA, Sailas E, McGrath JJ, Soares-Weiser K, Wahlbeck K (2004). Systematic review of cholinergic drugs for neuroleptic-induced tardive dyskinesia: a meta-analysis of randomized controlled trials. *Prog Neuropsychopharmacol Biol Psychiatry* **28**: 1099–1107.

Tarsy D, Baldessarini RJ (2006). Epidemiology of tardive dyskinesia: is risk declining with modern antipsychotics? *Mov Disord* **21**: 589–598.

Whitelock JM, Murdoch AD, Iozzo RV, Underwood PA (1996). The degradation of human endothelial cell-derived perlecan and release of bound basic fibroblast growth factor by stromelysin, collagenase, plasmin, and heparanases. *J Biol Chem* **271**: 10079–10086.

ORIGINAL PAPER

Translin-Associated Factor X Gene (*TSNAX*) may be Associated with Female major Depressive Disorder in the Japanese Population

Akiko Okuda · Taro Kishi · Tomo Okochi · Masashi Ikeda · Tsuyoshi Kitajima ·
Tomoko Tsunoka · Takenori Okumukura · Yasuhisa Fukuo · Yoko Kinoshita ·
Kunihiro Kawashima · Yoshio Yamanouchi · Toshiya Inada · Norio Ozaki ·
Nakao Iwata

Abstract Several investigations have reported that the translin-associated factor X gene (*TSNAX*)/disrupted-in-schizophrenia-1 gene (*DISC1*) was associated with major psychiatric disorders including schizophrenia, bipolar disorder (BP), and major depressive disorder (MDD). *TSNAX* is located immediately upstream of *DISC1*, and has been shown to undergo intergenic splicing with *DISC1*. It thus may also be influenced by translocation. To our knowledge, there are no reported gene-based association analyses between *TSNAX* and mood disorders in the Japanese population. We conducted a case–control study of Japanese samples (158 bipolar patients, 314 major depressive disorder patients, and 811 controls) with three tagging SNPs in *TSNAX*, selected using HapMap database. In addition, we

performed an association analysis between *TSNAX* and the efficacy of fluvoxamine treatment in 120 Japanese patients with MDD. The MDD patients in this study had scores of 12 or higher on the 17 items of the Structured Interview Guide for Hamilton Rating Scale for Depression (SIGH-D). We defined a clinical response as a decrease of more than 50% in baseline SIGH-D within 8 weeks, and clinical remission as an SIGH-D score of less than 7 at 8 weeks. We found an association between rs766288 in *TSNAX* and female MDD in the allele/genotype analysis. However, we did not find any association between *TSNAX* and BP or the fluvoxamine therapeutic response in MDD in the allele/genotype analysis or haplotype analysis. Our results suggest that rs766288 in *TSNAX* may play a role in the pathophysiology of female MDD in the Japanese population. A replication study using larger samples may be required for conclusive results, since our sample size was small.

Akiko Okuda, Taro Kishi participated equally in this work.

A. Okuda · T. Kishi (✉) · T. Okochi · M. Ikeda ·
T. Kitajima · T. Tsunoka · T. Okumukura · Y. Fukuo ·
Y. Kinoshita · K. Kawashima · Y. Yamanouchi · N. Iwata
Department of Psychiatry, Fujita Health University School
of Medicine, Toyoake, Aichi 470-1192, Japan
e-mail: tarok@fujita-hu.ac.jp

A. Okuda
Okehazama Hospital, Toyoake, Aichi 470-1168, Japan

M. Ikeda
Department of Psychological Medicine, School of Medicine,
Cardiff University, Heath Park, Cardiff CF14 4XN, UK

T. Inada
Neuropsychiatric Research Institute, Seiwa Hospital,
Tokyo 162-0851, Japan

N. Ozaki
Department of Psychiatry, Nagoya University Graduate
School of Medicine, Nagoya 466-8850, Japan

Introduction

The translin-associated factor X gene (*TSNAX*) and disrupted-in-schizophrenia-1 gene (*DISC1*) are located at 1q 42. These genes are associated with major psychiatric disorders, such as schizophrenia, bipolar disorder (BP), and major depressive disorder (MDD).

TSNAX (OMIM * 602964, 7 exons in this genomic region spanning 38.672 bp and 1q42), is located immediately upstream of *DISC1*, and has been shown to undergo

135

intergenic splicing with *DISC1* (Millar et al. 2000). This genomic region has been shown to be closely related to susceptibility for BP (Curtis et al. 2003; Macgregor et al. 2004). It may thus also be influenced by translocation. Hennah et al.'s (2003) haplotype transmission analysis showed that *TSNAX* was associated with schizophrenia. Palo et al. (2007) reported that *TSNAX* was associated with female psychotic disorder. Thomson et al. (2005) also showed an association between *TSNAX* and male Scottish BP patients. These studies that have found association with *TSNAX* have used Caucasian populations for which the underlying linkage disequilibrium (LD) spans *TSNAX* into the first portion of the *DISC1*. In the Japanese population the two genes are on distinct LD regions according to HapMap database (release#23a/phase II, March 2008, www.hapmap.org, population: Japanese Tokyo). However, Zhang et al. (2005) reported that *TSNAX* was not associated with schizophrenia in Japanese patients. *TSNAX* was associated with impaired spatial working memory, increased reaction time to visual targets, and reduced gray matter predominantly in the superior and middle frontal gyri (Cannon et al. 2005). There are no reported gene-based association analyses between *TSNAX* and mood disorders in the Japanese population. Therefore, we conducted a case–control study with Japanese mood disorder samples. Two recent studies reported that MDD and SSRI response in MDD have common susceptibility genes. Lekman et al. (2008) reported that *FKBP5* was associated with MDD and the citalopram therapeutic response in the White non-Hispanic population. Tsai et al. (2008) also reported significant associations between plasminogen activator inhibitor type 1 gene (*SERPINE1*) and Chinese MDD patients and the SSRI therapeutic response. We therefore performed an association analysis between *TSNAX* and the efficacy of fluvoxamine treatment in Japanese patients with MDD.

Materials and Methods

Subjects

The subjects in the association analysis were 314 MDD patients (155 males and 159 females; mean age ± standard deviation 47.3 ± 14.9 years), 158 BP patients (81 males and 77 females; 99 patients with bipolar I disorder and 59 patients with bipolar II disorder; 47.9 ± 14.2 years), and 811 healthy controls (352 males and 459 females; 37.2 ± 15.9 years). Of the 314 MDD patients, 120 (59 males and 61 females; 42.0 ± 17.2 years) were treated with fluvoxamine and diagnosed according to DSM-IV criteria with the consensus of at least two experienced

psychiatrists on the basis of a review of medical records and assessments with the Structured Interview Guide for Hamilton Rating Scale for Depression (SIGH-D). The remaining MDD patients were diagnosed according to DSM-IV criteria with the consensus of at least two experienced psychiatrists on the basis of unstructured interviews and a review of medical records. Fluvoxamine was taken two or three times a day for 8 weeks. The initial total dose was 50–100 mg per day, and the dosage was then increased gradually to a maximum of 150 mg, depending on the patients' condition. Patients with insomnia and severe anxiety were prescribed benzodiazepine drugs, but no other psychotropic drugs were permitted during the study. All subjects were unrelated to each other, ethnically Japanese, and lived in the central area of Japan.

All healthy controls were also psychiatrically screened based on unstructured interviews. None had severe medical complications such as cirrhosis, renal failure, heart failure, or other Axis-I disorders according to DSM-IV. The study was described to subjects and written informed consent was obtained from each. This study was approved by the Ethics Committees at Fujita Health University and Nagoya University School of Medicine.

Data Collection

The 120 MDD patients in this study had scores of 12 or higher on the 17 items of the SIGH-D (Peveler and Kendrick 2005). We defined a clinical response as a decrease of more than 50% in baseline SIGH-D within 8 weeks, and clinical remission as a SIGH-D score of less than 7 at 8 weeks. Detailed information on data collection was described in a previous paper (Saito et al. 2006). The clinical characteristics of patients in this study, classified according to these definitions, can be seen in Table 1.

SNP Selection and LD Evaluation

We first consulted the HapMap database (release#23a/phase II, March 2008, www.hapmap.org, population: Japanese Tokyo, minor allele frequencies (MAFs) of more than 0.05) and included 30 SNPs covering *TSNAX* (5'-flanking regions including about 55 kb from the initial exon and about 10 kb bp downstream (3') from the last exon: HapMap database contig number chr 1q42.1: 229673505.229774037). Three 'tagging SNPs' in *TSNAX* were then selected with the criteria of r^2 threshold greater than 0.8 in 'pair-wise tagging only' mode using the 'Tagger' program (Paul de Bakker, http://www/broad.mit.edu/mpg/tagger) in Haploview for the following association analysis (Barrett et al. 2005).

Table 1 Clinical characteristics of the patients in both definition groups

	N			Age (mean ± SD)	Baseline SIGH-D (avg ± SD)	Fluvoxamine dose at 8 weeks (mg/day) (avg ± SD)	Number of previous episode (avg ± SD)
	Total	Male	Female				
Overall	120	59	61	42.0 ± 17.2	20.3 ± 5.88	122 ± 3.84	1.39 ± 0.658
Clinical response group[a]							
Responders	61	31	30	42.2 ± 16.2	21.4 ± 6.14	119 ± 40.8	1.36 ± 0.570
Nonresponders	59	28	31	41.7 ± 18.5	19.1 ± 5.39	125 ± 40.7	1.44 ± 0.783
P value	0.712			0.895	0.0274	0.433	0.849
Clinical remission group[b]							
Remitters	47	22	25	40.1 ± 15.1	19.5 ± 5.01	115 ± 43.6	1.36 ± 0.110
Nonremitters	73	37	36	43.0 ± 18.3	20.7 ± 6.36	127 ± 38.2	1.42 ± 0.107
P value	0.678			0.510	0.271	0.114	0.697

[a] Clinical response was defined as a 50% or greater decrease in the baseline SIGH-D score

[b] Clinical remission was defined as a final SIGH-D score of less than 7

SNP Genotyping

We used TaqMan assays (ABI: Applied Biosystems, Inc., Foster City, CA,) for all SNPs. One allelic probe was labeled with FAM dye and the other with fluorescent VIC dye. The plates were heated at 50°C for 2 min and 95°C for 10 min, followed by 45 cycles of 95°C for 15 s and 58°C for 1 min. Please refer to ABI for the primer sequence. Detailed information is available on request.

Statistical Analysis

Genotype deviation from the Hardy–Weinberg equilibrium (HWE) was evaluated by Chi-square test (SAS/Genetics, release 8.2, SAS Japan Inc, Tokyo, Japan). Marker-trait association analysis was used to evaluate allele- and genotype association with the Chi-square test (SAS/Genetics, release 8.2, SAS Japan Inc, Tokyo, Japan), and haplotype association analysis was evaluated with a likelihood ratio test using the COCAPHASE2.403 program (Dudbridge 2003). In the haplotype analysis, we determined that the cutoff for testing haplotype frequency was 0.05. We used the permutation test option as provided in the haplotype analysis to avoid spurious results and correct for multiple testing. Permutation test correction was performed using 1000 iterations (random permutations). In addition, Bonferroni's correction was used to control inflation of the type I error rate in the single marker association analysis and in the explorative analysis. For Bonferroni correction, we employed the following numbers of multiple tests: 3 for each sample set in allele- and genotype analysis (3 tagging SNPs in TSNAX); and 6 for the explorative analysis by sex (2 × 3 tagging SNPs). We had already performed a permutation test in the haplotype analysis. Power calculation was performed using a genetic

power calculator (Purcell et al. 2003). The significance level for statistical tests was 0.05.

Results

The LD structure as determined from the HapMap database can be seen in Fig. 1. Genotype frequencies of all SNPs were in HWE. We did not detect any associations between TSNAX and mood disorders in the allele/genotype or haplotype analyses (Tables 2 and 3). It is known that there are sex differences in the pathophysiology of mood disorders (Currier et al. 2006; Faraone et al. 1987). Therefore, we performed an explorative analysis of subjects divided by sex. We found an association between rs766288 in TSNAX and female MDD in the allele/genotype analysis (Table 4). This significance remained after Bonferroni's correction. However, we did not find any association between TSNAX and BP in the allele/genotype analysis or haplotype analysis (Tables 4 and 5).

With regard to the clinical characteristics of patients, only one difference was detected between responders and nonresponders in baseline SIGH-D scores (P value = 0.0274) (Table 1). In addition to fluvoxamine treatment in this cohort, one patient each was prescribed alprazolam, loflazepate, and etizolam. Two patients each were prescribed lorazepam, brotizolm, flunitrazepam, and zopiclone. We did not find any association between TSNAX and the fluvoxamine therapeutic response in MDD patients in allele/genotype (Table 6) or haplotype analysis (Response: P value = 0.797 and Remission: P value = 0.773).

In the power analysis, we obtained more than 80% power for the detection of association when we set the genotype relative risk at 1.26–1.30 and 1.41–1.48 in MDD and BP, respectively, for TSNAX under a multiplicative model of inheritance.
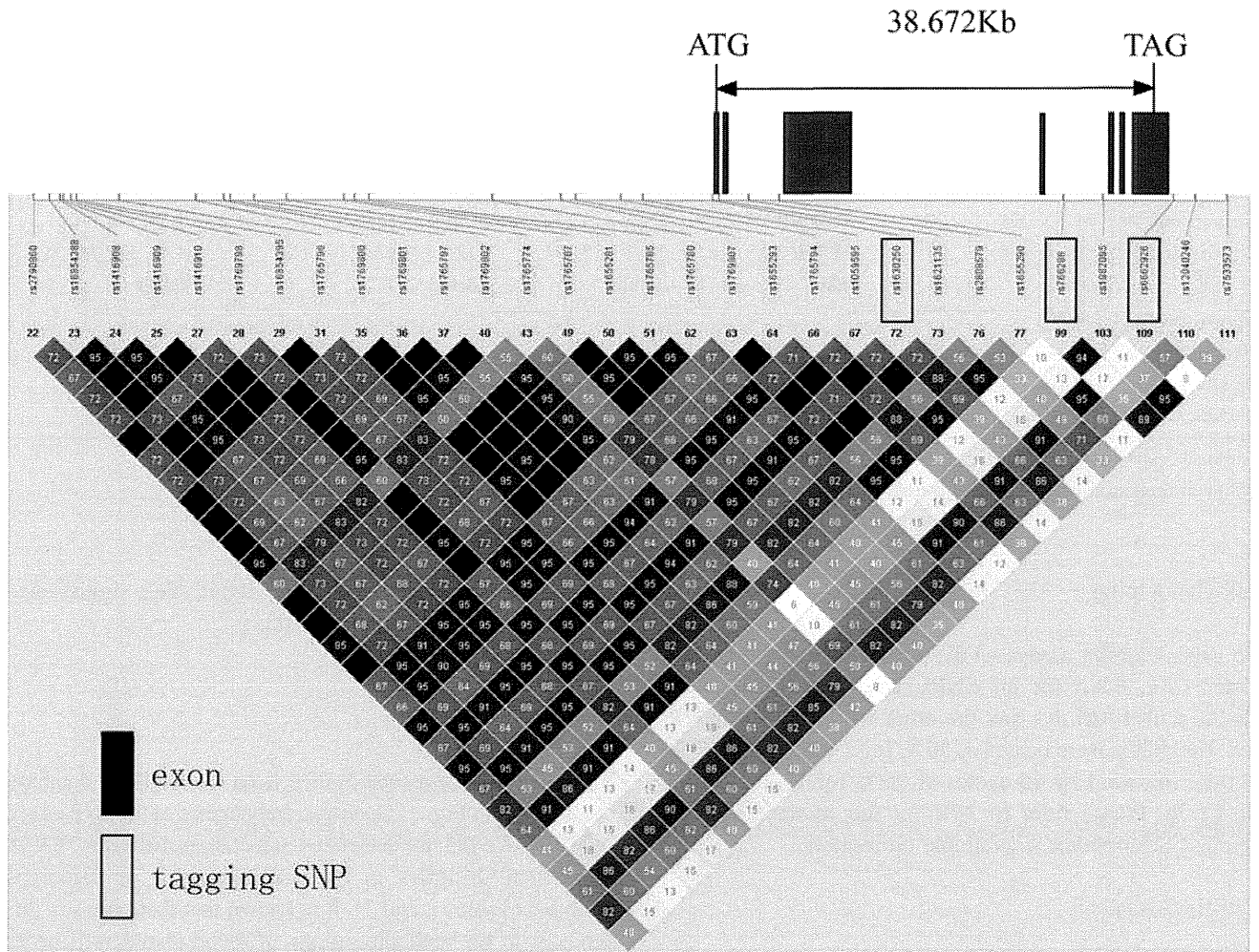
Fig. 1 LD evaluation and tagging SNPs in *TSNAX* ATG is the start codon and TAG is the stop codon. *Vertical bars* represent exons. Tagging SNPs selected from the HapMap database are represented by *black boxes*. Color scheme is based on r^2 value. Other information can be seen at the Haploview website

Table 2 Tagging SNPs and association analysis of *TSNAX*

SNP ID[a]	Phenotype[b]	MAF	N	Genotype distribution			P value[c]		
				M/M	M/m	m/m	HWE	Genotype	Allele
rs1630250	Controls	0.442	811	245	415	151	0.287		
5′ flanking region	MDD	0.475	314	85	160	69	0.700	0.356	0.165
C>G	BP	0.446	158	43	89	26	0.0789	0.493	0.892
rs766288	Controls	0.362	811	340	350	116	0.137		
Intron 4	MDD	0.322	314	141	144	29	0.367	0.0742	0.0727
C>T	BP	0.370	158	57	85	16	0.0535	0.0572	0.778
rs6662926	Controls	0.497	811	207	402	202	0.807		
3′ flanking region	MDD	0.463	314	87	163	64	0.438	0.270	0.153
C>G	BP	0.468	158	44	80	34	0.833	0.628	0.353

[a] Major allele > minor allele

[b] *MDD* Major depressive disorder, *BP* bipolar disorder, *MAF* minor allele frequency, *M* major allele, *m* minor allele

[c] Hardy–Weinberg equilibrium