15) Jaferi A, Bhatnagar S : Corticosterone can act at the posterior paraventricular thalamus to inhibit hypothalamic-pituitary-adrenal activity in animals that habituate to repeated stress. *Endocrinology* **147** : 4917-4930, 2006
16) Teves D, Videen TO, Cryer PE *et al* : Activation of human medial prefrontal cortex during autonomic responses to hypoglycemia. *Proc Natl Acad Sci U S A* **101** : 6217-6221, 2004
17) Arbelaez AM, Powers WJ, Videen TO *et al* : Attenuation of counterregulatory responses to recurrent hypoglycemia by active thalamic inhibition : a mechanism for hypoglycemia-associated autonomic failure. *Diabetes* **57** : 470-475, 2008
18) Moore RY, Weis R, Moga MM : Efferent projections of the intergeniculate leaflet and the ventral lateral geniculate nucleus in the rat. *J Comp Neurol* **420** : 398-418, 2000
19) Nascimento ES Jr, Duarte RB, Silva SF *et al* : Retinal projections to the thalamic paraventricular nucleus in the rock cavy (Kerodon rupestris). *Brain Res* **1241** : 56-61, 2008
20) Peng ZC, Grassi-Zucconi G, Bentivoglio M : Fos-related protein expression in the midline paraventricular nucleus of the rat thalamus : basal oscillation and relationship with limbic efferents. *Exp Brain Res* **104** : 21-29, 1995
21) Novak CM, Harris JA, Smale L *et al* : Suprachiasmatic nucleus projections to the paraventricular thalamic nucleus in nocturnal rats (Rattus norvegicus) and diurnal nile grass rats (Arviacanthis niloticus). *Brain Res* **874** : 147-157, 2000
22) Salazar-Juárez A, Escobar C, Aguilar-Roblero R : Anterior paraventricular thalamus modulates light-induced phase shifts in circadian rhythmicity in rats. *Am J Physiol Regul Integr Comp Physiol* **283** : R897-R904, 2002
23) Nakahara K, Fukui K, Murakami N : Involvement of thalamic paraventricular nucleus in the anticipatory reaction under food restriction in the rat. *J Vet Med Sci* **66** : 1297-1300, 2004
24) Landry GJ, Yamakawa GR, Mistlberger RE : Robust food anticipatory circadian rhythms in rats with complete ablation of the thalamic paraventricular nucleus. *Brain Res* **1141** : 108-118, 2007
25) Ji RR, Schlaepfer TE, Aizenman CD *et al* : Repetitive transcranial magnetic stimulation activates specific regions in rat brain. *Proc Natl Acad Sci U S A* **95** : 15635-15640, 1998
26) Price JL, Drevets WC : Neurocircuitry of mood disorders. *Neuropsychopharmacology* **35** : 192-216, 2010
27) Deutch AY, Ongür D, Duman RS : Antipsychotic drugs induce Fos protein in the thalamic paraventricular nucleus : a novel locus of antipsychotic drug action. *Neuroscience* **66** : 337-346, 1995
28) Allen Mouse Brain Atlas [internet]. Seattle (WA) : Allen Institute for Brain Science ©2009. Available from : http://mouse.brain-map.org.
29) Mai JK, Paxinos G, Voss T : *Atlas of the human brain*, third edition. Academic Press, San Diego, 2007

平成23年度　　資　　料

Citation: Transl Psychiatry (2011) 1, e24, doi:10.1038/tp.2011.26

www.nature.com/tp

Hypermethylation of serotonin transporter gene in bipolar disorder detected by epigenome analysis of discordant monozygotic twins

H Sugawara[1,2], K Iwamoto[1,3], M Bundo[1,3], J Ueda[1], T Miyauchi[1], A Komori[1], A Kazuno[1], N Adati[4], I Kusumi[5], Y Okazaki[6], J Ishigooka[2], T Kojima[4] and T Kato[1]

Bipolar disorder (BD) is a severe mental disorder characterized by recurrent episodes of mania and depression. Serotonin transporter (HTT) is a target of antidepressants and is one of the strongest candidate molecules of mood disorder, however, genetic study showed equivocal results. Here, we performed promoter-wide DNA methylation analysis of lymphoblastoid cell lines (LCLs) derived from two pairs of monozygotic twins discordant for BD. To rule out the possible discordance of copy number variation (CNV) between twins, we performed CNV analysis and found the copy number profiles were nearly identical between the twin pairs except for immunoglobulin-related regions. Among the three genes we obtained as candidate regions showing distinct difference of DNA methylation between one of the two pairs, hypermethylation of *SLC6A4*, encoding HTT, in the bipolar twin was only confirmed by bisulfite sequencing. Then, promoter hypermethylation of *SLC6A4* in LCLs of BD patients was confirmed in a case–control analysis. DNA methylation of *SLC6A4* was significantly correlated with its mRNA expression level in individuals with the S/S genotype of HTTLPR, and mRNA expression level was lower in BD patients carrying the S/S genotype. Finally, DNA methylation of the same site was also higher in the postmortem brains of BD patients. This is the first study to report the role of epigenetic modification of *SLC6A4* in BD using an unbiased approach, which provides an insight for its pathophysiology.
Translational Psychiatry (2011) **1**, e24; doi:10.1038/tp.2011.26; published online 26 July 2011

Introduction

Mood disorders such as major depression and bipolar disorder (BD) cause severe social burdens and there is an urgent need to elucidate the pathophysiology of these disorders. Studying the mechanisms of action of psychotropic drugs has been a major strategy to understand the pathophysiology of mental disorders, and research has focused on serotonin transporter (HTT), the target molecule of antidepressants. Altered density of HTT in the brains of patients with mood disorders has been reported in postmortem samples[1] and by *in vivo* molecular imaging.[2–6] A serotonin-transporter-linked promoter region (HTTLPR) includes the functional polymorphism, in which the short (S) allele has lower promoter activity than the long (L) allele,[7] and it has been implicated in mood disorders.[1,8–10] The S allele was reported to be a risk factor for mood disorders, interacting with stress and/or early environment.[8] However, the results of subsequent studies are conflicting[11–13] and none of the recent genome-wide association studies supported the role of *SLC6A4* in mood disorders.[14]

We have been searching for gene expression and epigenetic differences between monozygotic (MZ) twins discordant for BD to elucidate the molecular pathways relevant to its pathophysiology.[15,16] Although high concordance rate of BD in MZ twins supports the contribution of genetic factor in BD, importantly, it is not 100%. Because MZ twins have been regarded as having identical genomes, these facts suggest the importance of environmental or epigenetic factor for the onset of mental disorders. In fact, considerable epigenomic differences between MZ twins have been reported by several groups.[17–22] On the other hand, there are also reports of genomic differences between MZ twins such as point mutations,[23–25] microsatellite repeat length[26] and copy number variations (CNVs).[27] Therefore, both genetic and epigenetic factors may have a role in the discordance of the onset of diseases between MZ twins.

Here, we performed promoter-wide DNA methylation analysis of lymphoblastoid cell lines (LCLs) derived from two pairs of MZ twins discordant for BD. To rule out the possible discordance of CNV between twins, we performed CNV analysis.

Materials and methods

Samples. We enrolled two pairs of MZ twins discordant with respect to BD (Supplementary Table S1). The twins

[1]Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, Saitama, Japan; [2]Department of Psychiatry, Tokyo Women's Medical University, Tokyo, Japan; [3]Department of Molecular Psychiatry, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; [4]Research Equipment Center, Hamamatsu University School of Medicine, Shizuoka, Japan; [5]Department of Psychiatry, Hokkaido University, Sapporo, Japan and [6]Tokyo Metropolitan Matsuzawa Hospital, Tokyo, Japan
Correspondence: Dr T Kato, Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. E-mail: kato@brain.riken.jp
Keywords: bipolar disorder; DNA methylation; monozygotic twins; serotonin transporter
Received 17 June 2011; accepted 21 July 2011

were 49-year-old males (pair 1, previously reported in Kusumi *et al.*[28]) and 42-year-old males (pair 2). Detailed clinical information was described in previous reports and their zygosities were determined by micro-satellite markers and single nucleotide polymorphisms.[15,16] We used their genomic DNA derived from LCLs for all analyses.

In the CNV analysis, we used a female sample (NA15500, Coriell Cell Repository, Camden, NJ, USA) as a reference and an unrelated healthy female sample for quality control of real-time PCR analysis. Methods of CNV analysis are shown in Supplementary Information.

To exclude the genomic regions in which DNA methylation is different between peripheral blood leukocytes (PBLs) and LCLs, we used four sets of PBLs and LCLs from the same individuals who were unrelated healthy Japanese males and females.[29] For the collection of PBLs, blood was collected into a 7 ml tube containing EDTA. Red blood cells were lysed by osmotic shock by adding 40 ml of red blood cell lysis buffer containing 10 mM Tris/HCl, pH 7.6, 6 mM $MgCl_2$ and 10 mM NaCl. After centrifugation, the pellet was used for DNA extraction.

For the case–control analysis, we examined 20 unrelated BD patients (12 males and 8 females, 38.8 ± 13.1 years old (mean ± s.d.)) and 20 unrelated control subjects (16 males and 4 females, 38.5 ± 8.9 years old) using their genomic DNA derived from LCLs. They were subjects from our previous report,[30] and all of them were Japanese. Written informed consent was obtained from all the subjects.

Postmortem prefrontal cortices (BA10) derived from BD patients ($N=35$) and controls ($N=35$) were provided by the Stanley Foundation Brain Collection (The Stanley Medical Research Institute, Chevy Chase, MD, USA). We were able to use 31 BD samples and 32 control samples for DNA methylation analysis.

This study was approved by the Research Ethics Committee of RIKEN.

Cell cultures. LCLs were previously established by transforming the lymphocytes with Epstein–Barr (EB) virus following standard protocol.[31] Briefly, lymphocytes were separated from peripheral blood and cultured with RPMI 1640 medium containing 20% fetal bovine serum (Gibco, Carlsbad, CA, USA), penicillin and streptomycin (50 μg ml^{-1} each) and supernatant of the B95-8 cell culture infected by the EB virus. After the cells showed stable growth, they were passaged two or three times a week using similar medium except for 10% fetal bovine serum. The cells were then kept frozen and recultured for the experiment. We used low-passage LCLs for case–control studies. Genomic DNA was extracted using Gentra Puregene Cell kit (Qiagen, Hilden, Germany).

DNA methylation analysis

Enrichment of methylated DNA. We used a MethylCollector Ultra Kit (Active Motif, Carlsbad, CA, USA), which can enrich methylated DNA fragments with as few as five methylated CpG sites. DNA fragments (100 ng) were sonicated to produce 100- to 300-bp fragments using a sonicator (Covaris, Woburn, MA, USA) and they were incubated with a His-tagged recombinant MBD2b/MBD3L1 protein complex. After stringent washing, methylated DNA was eluted and purified using the MiniElute PCR purification kit (Qiagen).

Promoter tiling array. We used a GeneChip Human Promoter 1.0R array (Affymetrix, Santa Clara, CA, USA), which contains 4.6 million probes tiled to cover more than 25 500 human promoter regions. Sample preparation for the tiling array was performed according to the Affymetrix chromatin immunoprecipitation assay protocol provided by the manufacturer. Briefly, after amplification, DNA samples were labeled using a GeneChip WT Double-Stranded DNA Terminal Labeling Kit (Affymetrix), and they were then hybridized with Affymetrix GeneChip Human Promoter 1.0R arrays. The arrays were stained and washed using GeneChip Fluidics Station 450 and scanned with the GeneChip 3000 7G Scanner. We used fully CpG unmethylated DNA obtained by whole genome amplification as a reference sample.

Tiling array data analysis. For the detection of methylated regions (MRs), model-based analysis of tiling arrays (MAT) software was used.[32] By using MAT, the signals of each probe in each array can be standardized individually. Parameters used in the MAT analysis were as follows: bandwidth, 300; maximum gap, 300; minimum number of probes, 10; *P* value, 10^{-4}.

The filtering process was performed with R script (http://www.r-project.org) using the following steps (Supplementary Figure S4). Scripts are available upon request.

(1) Data of each twin pair were directly compared. The regions showing significant differences between the BD twin and the healthy co-twin were selected. Only those regions containing six or more CpG sites were further selected because the MethylCollector can collect only the DNA fragments that contain five or more CpG sites. The selected MRs were named BD (bipolar twin)-dominant MRs and C (healthy co-twin)-dominant MRs.
(2) The data of each twin were compared with those of a reference sample (that is, unmethylated DNA). Among the BD-dominant MRs, the regions showing the significantly methylated signal compared with a reference sample in BD twin but not in healthy co-twin were selected. The selected regions were named BD-specific MRs. C-specific MRs were determined vice versa. This step highlighted the methylation difference between twins and eliminates significant but subtle differences to maximize the chance for successful confirmation by other methods.
(3) The MRs overlapping with a CpG islands were selected. For validation of tiling array data, we arbitrarily selected 13 representative MRs for bisulfite sequencing.

Bisulfite conversion. Either 2 μg (for LCL samples) or 500 ng (for postmortem brain samples) of genomic DNA was converted for methylation sequencing using Epitect bisulfite kits (Qiagen), according to the manufacturer's standard protocol. Completion of sodium bisulfite conversion was confirmed during sequencing by ensuring that known lone cytosines were read as thymines.

Bisulfite sequencing. Primer pairs were determined using Meth Primer software.[33] PCR products were purified by MinElute PCR Purification Kit (Qiagen) or digested with thermostable β-agarase (NIPPON GENE, Tokyo, Japan) and cloned using a TOPO TA cloning kit (Invitrogen, Carlsbad, CA, USA). Single bacterial colonies were subjected to sequencing analysis. Each bisulfite sequencing data was quantified by quantification tool for methyaltion analysis.[34]

Filtering of tiling array data. To further narrow the number of MRs by setting more stringent criteria based on the results of bisulfite sequencing, we reset the threshold for filtering as a P value $<10^{-6}$ of direct comparison between each twin pair (Supplementary Figure S5, step 4).

Next, we excluded the regions that showed alteration of DNA methylation status before and after the transformation by EB virus. We previously performed the DNA methylation analysis using the same platform in four sets of PBLs and LCLs established from the same individuals and identified the regions that showed methylation only in LCLs or only in PBLs.[30] MRs that overlapped with those regions were excluded (Supplementary Figure S5, step 5).

Finally, from the candidate regions for BD-specific MRs, the regions that were detected as MRs in at least one of LCL were excluded (Supplementary Figure S4), because the regions methylated in LCLs of healthy individuals cannot be related to disease. With regard to the C-specific MRs, only the regions overlapping with common MRs in four LCLs were selected. Loss of methylation in patients can be regarded as related to disease only when methylation in that region is consistently seen in healthy individuals (Supplementary Figure S5, step 6).

Case–control analysis

Pyrosequencing. The PCR product of bisulfite-modified DNA was used for pyrosequencing analysis according to the manufacturer's standard protocol (Qiagen). Briefly, 4 μl of streptavidin-sepharose beads (Amersham Biosciences, Piscataway, NJ, USA) and 54 μl of binding buffer (10 mM Tris-HCl, 1 mM EDTA, 2 M NaCl, 0.1% Tween-20 at pH 7.6) were mixed with 50 μl of PCR product for 10 min at room temperature. The reaction mixture was placed onto a MultiScreen-HV, Clear Plate (Millipore, Billerica, MA, USA). After applying the vacuum, the beads were treated with a denaturation solution (0.2 N NaOH) for 1 min and washed twice with washing buffer (10 mM Tris-acetate at pH 7.6). The beads were then suspended with 50 μl of annealing buffer (20 mM Tris-acetate, 2 mM Mg-acetate at pH 7.6) containing 10 pmol of sequencing primer. The template-sequencing primer mixture was transferred onto a PSQ 96 Plate (Biotage, Uppsala, Sweden), heated to 90 °C for 2 min and cooled to room temperature. Sequencing reactions were performed with a PyroMark Gold Q96 Reagents Kit (Qiagen) according to the manufacturer's instructions. The percentages of methylation were calculated from the raw data using the allele quantification algorithm of the software provided by the manufacturer (PSQ96MA2.1.1 software, Qiagen).[35] The PCR primers used for bisulfite sequencing of the promoter region of *SLC6A4* were 5′-TTTTTAGTTGTTTG GGTATTTGTGTTA-3′ (forward) and 5′-AAAACTTTACAAC CTCTTAAAAACCC-3′ (reverse). The sequencing primers were 5′-TTTTGTATAAAGTTATTTGT-3′, 5′-AATATAAATTA TGGGTTGAA-3′ and 5′-ATTTTTTTTAAGGGGTTTTT-3′, and the reading sequences were AT/CGT/CAT, AT/CGA AAGTAAGTAATTTTTTTTAAAGT/CG and TAT/CGG.

Genotyping. Using genomic DNA derived from LCLs, we investigated the length polymorphism repeat (HTTLPR) in the promoter region of *SLC6A4*, which comprises a 44-bp insertion (long, L allele) or deletion (short, S allele), and one single nucleotide polymorphism in L allele (LA/LG), which was reported to be related to gene expression as well as HTTLPR genotype.[36]

For HTTLPR genotype analysis, amplification of genomic DNA was performed as initially described by Kaiser *et al.*,[37,38] with slight modifications. In a total volume of 10 μl, we used 40 ng DNA; 200 μM each of dATP, dCTP, and dTTP; 150 μM dGTP, 50 μM; 7-deaza-GTP (Roche, Mannheim, Germany); 1 μM of each primer; 10* buffer 1; and 2.8 units of mixed polymerase (Taq- and Tgo-polymerase) from the Expand Long-Range PCR System (Roche). Applied primer sequences were 5′-GCAACCTCCCAGCAACTCCCTGTA-3′ (forward) and 5′-GAGGTGCAGGGGGATGCTGGAA-3′ (reverse). The PCR program of both reactions consisted of an initial cycle of 2 min at 94 °C followed by 35 cycles each consisting of 10 s at 98 °C, 30 s at 68 °C and 30 s at 72 °C. After amplification, HTTLPR amplicons were separated by electrophoresis in 3% agarose gels after staining with ethidium bromide.

For LA/LG single nucleotide polymorphism analysis, each product of L allele in L/S genotype subjects was digested with thermostable β-agarase (NIPPON GENE) and subjected to sequencing analysis.

Real-time quantitative RT-PCR. Total RNA was extracted using Trizol (Invitrogen) and cleaned following the standard protocol. After DNase I treatment, 5 μg of total RNA was used for complementary DNA synthesis by oligo (dT) 12–18 primer and SuperScript II reverse transcriptase (Invitrogen). Real-time quantitative RT-PCR using Taqman Universal PCR master mix, No AmpErase UNG (Applied Biosystems, a part of Life Technologies, Carlsbad, CA, USA), was performed with an ABI Prism 7900HT (Applied Biosystems). The PCR conditions were denaturation at 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 60 s. The comparative Ct method was used to quantify transcripts according to the manufacturer's protocol (User Bulletin 2, Applied Biosystems). Each sample was quantified in triplicate. Amplification of single products was confirmed by monitoring the dissociation curve. We used the expression level of actin beta (*ACTB*) as an internal control. The standard ABI primer probe sets were used (*SLC6A4*, Hs00984356; *ACTB*, Hs99999903). We performed statistical analysis using SPSS (SPSS Japan, Tokyo, Japan).

Results

Copy number differences between the twins. The copy number profiles were seemingly identical between both twin pairs (Supplementary Figure S1). To find subtle copy number differences between the twins, we detected their respective

copy number changes compared with a reference and calculated the fold change of each probe. In both twin pairs, the differences with fold changes of more than 1.2 were mostly found on chromosomes 2, 14 and 22, which contain immunoglobulin (Ig)-related regions (Supplementary Figure S2). We examined 40 regions including one Ig-related region by real time PCR analysis. Although the copy number change of Ig-related region was confirmed, none of other regions with copy number differences between the twins were confirmed by real-time PCR analysis (Supplementary Figure S3). This suggested that most of the apparent copy number differences detected by the array were false positives, except for Ig-related regions. The differences of the copy number in Ig-related regions between the twins are reasonable because a single B cell having a specific genome rearrangement of Ig-related gene[39] is amplified by EB virus transformation in LCLs.

DNA methylation differences between the twins. Promoter-wide DNA methylation profiles of the twins were examined by Affymetrix GeneChip Human Promoter 1.0R tilling arrays after methylated DNA was enriched using MBD2b and MBD3L1 conjugated beads. Fully unmethylated DNA obtained by whole genome amplification was used as a reference. Using this method, the DNA methylated status within the 10 kb range for each of more than 25 500 promoters could be analyzed. MAT[32] was used to identify MRs. By using MAT, the signals of each probe in each array can be standardized individually.

DNA methylation analysis by the promoter tiling array showed MRs specific to the bipolar twin (80 or 4, respectively) or those to the healthy co-twin (71 or 81, respectively) in each twin pair (Supplementary Figure S4). Among the 236 regions containing six or more CpG sites, 13 representative regions overlapping with CpG islands were analyzed by bisulfite sequencing to validate the tiling array data. Although the DNA methylation differences between the twins were confirmed in most regions, some differences with P values more than 10^{-6} could not be confirmed (Supplementary Figures S5 and S6). To narrow the number of candidate MRs, we applied more stringent criteria (P value $\leqslant 10^{-6}$) based on the results of bisulfite sequencing. We compared the DNA methylation status between four sets of PBLs and LCLs obtained from the same individuals and found that MRs in PBLs were largely conserved in LCLs but a part of the regions were methylated only in LCLs or only in PBLs.[30] Thus, we eliminated those 'variable' MRs during establishment of LCLs. Furthermore, the regions detected as MRs in at least one of the LCL were excluded from the BD-specific MRs, and only the regions overlapping with common MRs in four LCLs were selected with regard to the C-specific MRs (Supplementary Figure S4). We finally obtained three candidate MRs: *KIAA1530* on chromosome 4 as a hypermethylated region in the healthy co-twin, and *FANK1* (fibronectin type III and ankyrin repeat domains 1) on chromosome 10 and *SLC6A4*, encoding serotonin transporter on chromosome 17 (Supplementary Figure S7), as a hypermethylated region in the bipolar twin, only from pair 1.

These three regions were examined by bisulfite sequencing, and we could confirm the differential DNA methylation status between the twins only in the promoter region of *SLC6A4* (Figure 1b and Supplementary Figures S8 and S9).

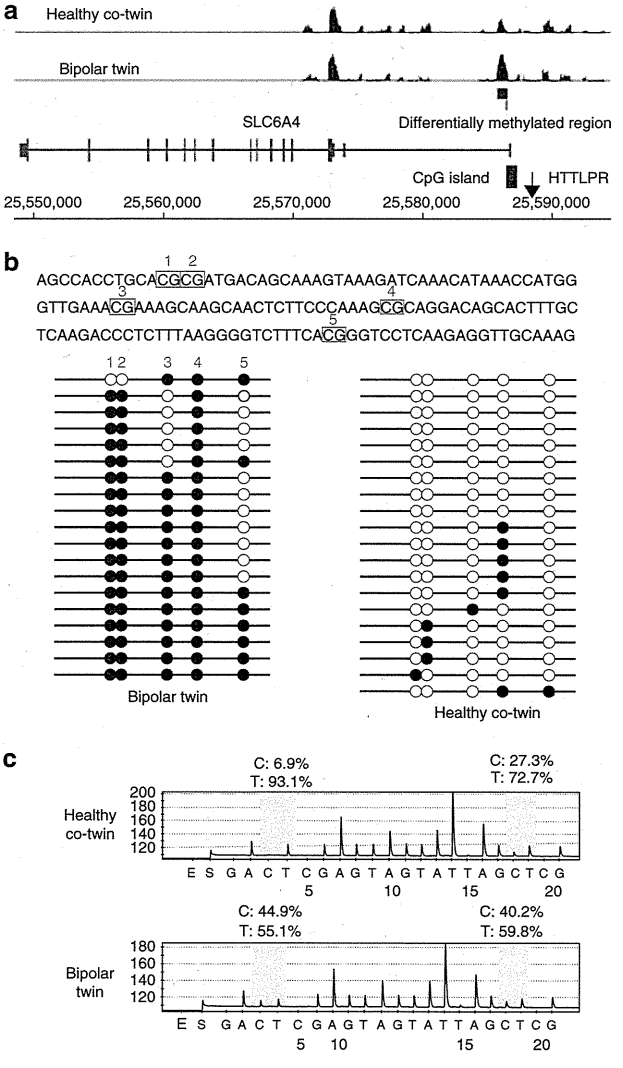

Figure 1 DNA methylation difference between twins in *SLC6A4*. (**a**) Results of comprehensive DNA methylation analysis of lymphoblastoid cell lines (LCLs) of a pair of monozygotic twins discordant for bipolar disorder using tiling arrays. The vertical axis represents the signal intensity, and the horizontal axis represents the base number on the chromosome 17 (NCBI36/hg18). Exon–intron structure of the *SLC6A4* is shown below the data of tiling arrays. The CpG island and the regions of HTTLPR are shown by a green square and an arrow, respectively. The region showing statistically significant methylation difference between the bipolar twin and the healthy co-twin, and the region examined by bisulfate sequencing are shown by a blue square and a red bar, respectively. (**b**) Results of bisulfite sequencing. The genomic region examined by bisulfite sequencing, which corresponds to the base numbers from 25 586 333 and 25 586 482, is shown above. The five CpG sites are surrounded by red squares. Black and white circles represent the methylated and unmethylated CpGs, respectively. Each raw shows the data of one clone. Five circles in one raw represent the five CpG sites shown above. This region is methylated in the bipolar twin but not in the healthy co-twin. (**c**) Representative results of pyrosequencing using independent samples. The results of two CpG sites (3 and 4) were shown by yellow shadows. Percentages of C and T mean fractions of methylation and unmethylation on each CpG site, respectively. Each fraction of methylation on other CpG sites of the bipolar twin and the healthy co-twin was 24.2 and 0 on CpG1, 28.8 and 0 on CpG2, and 9.3 and 4.3 on CpG5, respectively. The differences of DNA methylation in *SLC6A4* between the twins were confirmed using the other LCL samples of the twins recultured for this experiment.

This region located on the CpG island shore at the 2–3 kb downstream of HTTLPR was hypermethylated in the bipolar twin. We also examined the DNA methylation status of the twins using independent cell culture by pyrosequencing to rule out a possibility of cell culture-induced artificial or stochastic change (Figure 1c).

Case–control analysis for DNA methylation of LCLs in *SLC6A4*. Because there are DNA methylation changes even between healthy MZ twins, the existence of DNA methylation difference does not warrants the relationship with the disease. Thus, we next performed a case–control analysis using the LCLs of 20 patients with BD and 20 controls. The DNA methylation of the two CpG sites (CpG3 and CpG4) was significantly higher in BD patients than in controls ($P<0.05$, Mann–Whitney *U*-test; Figure 2a). We also analyzed the expression level of *SLC6A4* in LCLs by real-time quantitative RT-PCR. HTTLPR, including a single nucleotide polymorphism in L allele (LA/LG) that reportedly affects gene expression, was genotyped.[36] Because all subjects with L allele, but one control subject with S/LG genotype, had LA allele, LA and LG alleles were not separately analyzed. Among the subjects with the S/S genotype, the mRNA expression level was significantly lower ($P<0.05$, Mann–Whitney *U*-test; Figure 2b) and the DNA methylation level of CpG3 was significantly higher ($P<0.05$; Mann–Whitney *U*-test; Figure 2c) in the BD patients compared with controls. The same trend was seen in CpG4 ($P=0.059$, Mann–Whitney *U*-test). The DNA methylation level of CpG3 was significantly correlated with its mRNA expression level only in subjects with the S/S genotype ($r=-0.425$, $P=0.043$, $n=23$; Figure 2d) and not in those with the S/L genotype ($r=-0.262$, $P=0.346$, $n=15$; Figure 2e).

DNA methylation status of postmortem brain samples in *SLC6A4*. DNA methylation patterns in blood cells may not directly reflect the status of the brain. To test whether hypermethylation of *SLC6A4* is also observed in the brain, we examined the DNA methylation status of these

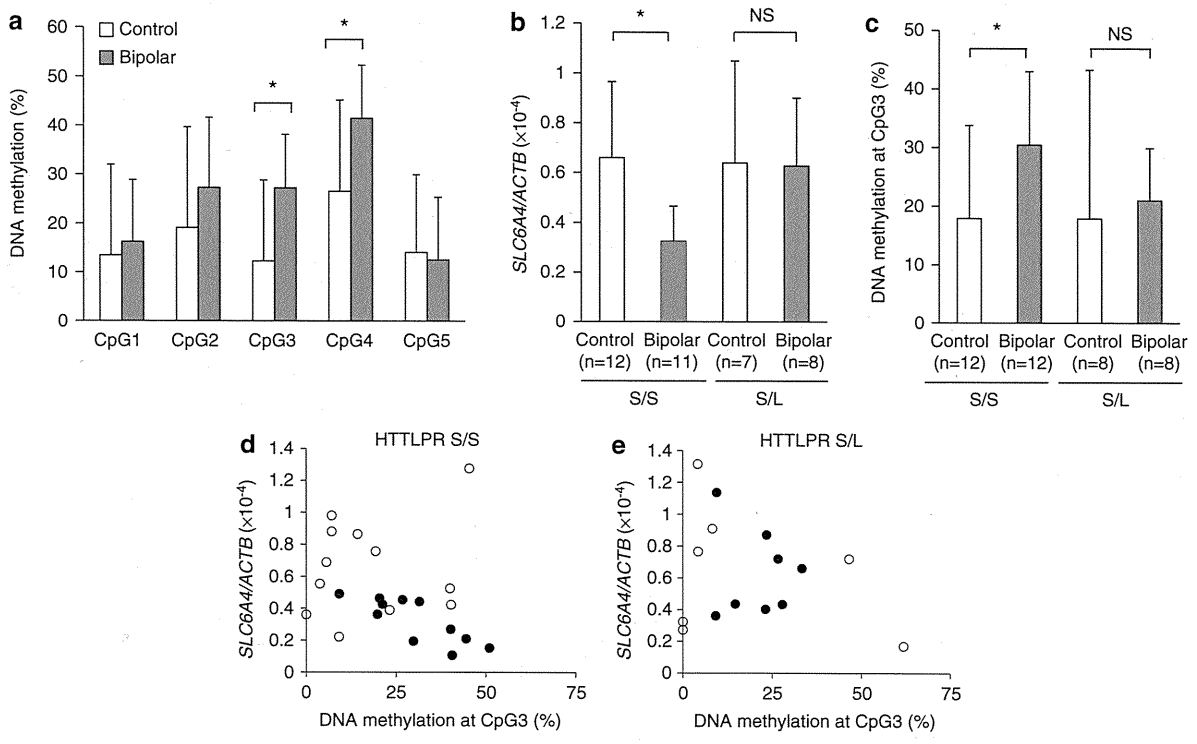

Figure 2 Case–control analysis of DNA methylation and gene expression in LCLs. (**a**) DNA methylation levels of five CpG sites shown in the Figure 1 in LCLs (bipolar disorder (BD), $n=20$, controls, $n=20$) using pyrosequencing. Two CpG sites (3 and 4) showed significantly higher DNA methylation in BD patients. Values indicate mean ± s.d. Asterisk indicates significance; $P=0.008$ and $P=0.035$ for CpG3 and CpG4, respectively, by Mann–Whitney *U* test. (**b**) Relationship between expression level of *SLC6A4* and HTTLPR genotype (BD, $n=19$, controls, $n=19$). One datum in each diagnostic group was turned out to be at outlier by Smirnov–Grubbs test and removed from data analysis. Among the all samples ($n=38$), the mRNA level of *SLC6A4* tended to be lower in BD patients compared with controls ($P=0.08$; Mann–Whitney *U*-test). The mRNA expression level was significantly lower in BD patients compared with controls only in subjects with the S/S genotype ($P=0.007$; Mann–Whitney *U*-test) and not in those with the S/L genotype ($P=0.867$; Mann–Whitney *U*-test). Values indicate mean ± s.d. An asterisk indicates significant difference ($P<0.05$). NS, not statistically significant. (**c**) Relationship between the DNA methylation of CpG3 and HTTLPR genotype (BD, $n=20$, controls, $n=20$). The DNA methylation of the CpG3 site was significantly higher in BD patients compared with controls only in the subjects with the S/S genotype ($P=0.020$, Mann–Whitney *U*-test) but not in those with the S/L genotype ($P=0.105$, Mann–Whitney *U*-test). The DNA methylation of the CpG4 site also tended to be higher in BD patients compared with controls only in the subjects with the S/S genotype ($P=0.052$, Mann–Whitney *U*-test) and not in those with the S/L genotype ($P=0.279$, Mann–Whitney *U*-test). Values indicate mean ± s.d. (**d**, **e**) Correlation between the DNA methylation at the CpG3 site and mRNA expression level of *SLC6A4* (BD, $n=19$, controls, $n=19$). There was a significant correlation in the subjects with the S/S genotype ($r=-0.425$, $P=0.043$, $n=23$); (**d**), whereas no significant correlation was found in the subjects with the S/L genotype ($r=-0.262$, $P=0.346$, $n=15$); (**e**). Closed circles represent patients with BD and open circles represent control subjects.

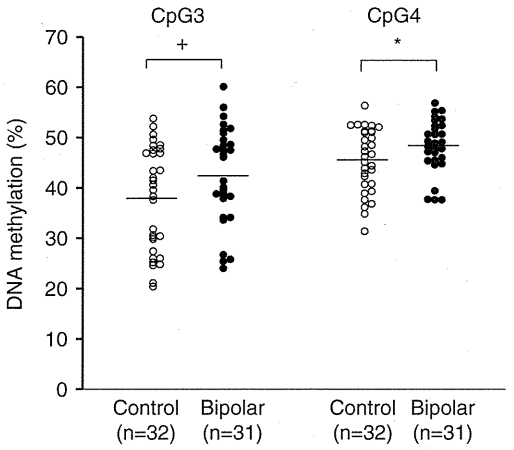

Figure 3 DNA methylation analysis of *SLC6A4* in the prefrontal cortices of bipolar disorder (BD) patients and control subjects. The CpG3 and CpG4, shown in the Figure 1, were analyzed by pyrosequencing. Closed circles represent patients with BD and open circles represent control subjects. Horizontal bars indicate averages. $^{*}P<0.05$, $^{+}P<0.10$ (Mann–Whitney *U*-test, one-tailed). There was no significant difference of gender between the BD patients and controls. Age, postmortem interval (hours) and brain pH were not significantly correlated with the DNA methylation level of both CpG3 and CpG4. In the BD patients, there was no significant difference of the DNA methylation level of the two CpG sites between the subjects with and without antidepressant treatment.

two CpG sites in postmortem brains of 31 BD patients and 32 controls obtained from the Stanley Foundation Brain Collection. CpG4 showed a significantly higher DNA methylation level in BD patients compared with controls ($P<0.05$, Mann–Whitney *U*-test, one-tailed), and CpG3 also showed the same trend ($P=0.059$, Mann–Whitney *U*-test, one-tailed; Figure 3). There was no significant effect of confounding factors, such as gender, age, postmortem interval (hours), brain pH and administration of antidepressants, on the DNA methylation level of the two CpG sites.

Discussion

Multiple lines of evidence suggest the role of *SLC6A4* in BD. Although the results of serotonin transporter-binding potential in BD patients are conflicting,[2–6] one of them reported the low potential in the depressive state,[40] and less gene expression of serotonin transporter in postmortem brains of depressed suicide victims was also reported.[41] HTTLPR was initially reported to be associated with BD, though subsequent studies show conflicting results.[9,42,43] Gene-environment (G × E) interaction[14] and resulting epigenetic changes[44] might explain the inconsistent results of genetic studies in mood disorder.[14] In this study, we found a possible interaction between HTTLPR genotype and BD diagnosis on the DNA methylation and mRNA expression level of *SLC6A4*. DNA methylation level of the CpG island upstream of *SLC6A4* was reported to be associated with environmental factors[45,46] and possibly with HTTLPR genotype.[47] Weaver *et al.*[48] hypothesized that DNA methylation plays a role as an epigenetic mark of G × E interaction, and they suggested that hypermethylation of the promoter region of the glucocorticoid receptor gene associated with poor maternal care causes stress vulnerability. Hypermethylation of *SLC6A4* in BD may reflect epigenetic effect in G × E interaction and contribute to the patho physiology of BD.

In this study, we performed promoter-wide DNA methylation analysis and detected hypermethylation of *SLC6A4* in the bipolar twin. The differences of methylation patterns between MZ twins may depend on age and difference of nurturing environment,[18] and a twin study showed that the variation of DNA methylation in *SLC6A4* is attributable to unique environmental factors rather than heritable factors.[49] It would be important to elucidate the origin and timing of the methylation change in *SLC6A4*[50] in the future study. In this study, only one of the two twins showed marked difference of DNA methylation of *SLC6A4*. Considering the heterogeneity of BD, it is not realistic to postulate that all patients with BD are caused by the same genetic or epigenetic factor. Thus, a close investigation in a large series of well-characterized discordant twins may provide a clue to understand the pathophysiology of BD.[51] In this study, the difference between the groups was relatively modest in case–control studies. There was a large inter-individual difference of DNA methylation at CpG3 in LCLs of patients with BD (9.2 to 51%), possibly reflecting the heterogeneity of BD. In these analyses, we used genomic DNA derived from LCLs cultured with drug-free medium to minimize the effect of medication on DNA methylation. Although none of BD patients were taking antidepressants at the blood sampling, it is difficult to know what the effects of medications and co-morbid illnesses are on the methylation signatures observed in these samples. Moreover, we only examined a portion of the *SLC6A4*. Although DNA methylation in LCLs is not the same as PBLs, MRs in PBLs are largely maintained in LCLs.[30] In this study, we excluded the MRs affected by EB virus transformation and finally detected the hypermethylation of *SLC6A4* in the brains of BD patients. We could only examine the methylation of *SLC6A4* in prefrontal cortices, but further analysis of other regions expressing high levels of HTT, such as dorsal raphe nucleus, would be interesting.

In summary, we detected hypermethylation of *SLC6A4* in BD, using an unbiased approach. DNA methylation of *SLC6A4* may be an epigenetic mark resulting from a G × E interaction that leads to the development of BD. Our findings add a new insight to elucidate the pathophysiology of mood disorder.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements. Postmortem brain samples were donated by The Stanley Medical Research Institute Brain Collection courtesy of Drs Michael B Knable, E Fuller Torrey, Maree J Webster and Robert H Yolken. We are grateful to Dr Kakiuchi and Ms Ishiwata for the establishment of LCLs of twin pairs, and the Research Resource Center of RIKEN Brain Science Institute for performing the microarray analysis.

1. Mann JJ, Huang YY, Underwood MD, Kassir SA, Oppenheim S, Kelly TM *et al.* A serotonin transporter gene promoter polymorphism (5-HTTLPR) and prefrontal cortical binding in major depression and suicide. *Arch Gen Psychiatry* 2000; **57**: 729–738.
2. Cannon DM, Ichise M, Fromm SJ, Nugent AC, Rollis D, Gandhi SK *et al.* Serotonin transporter binding in bipolar disorder assessed using [11C]DASB and positron emission tomography. *Biol Psychiatry* 2006; **60**: 207–217.

3. Cannon DM, Ichise M, Rollis D, Klaver JM, Gandhi SK, Charney DS *et al.* Elevated serotonin transporter binding in major depressive disorder assessed using positron emission tomography and [11C]DASB; comparison with bipolar disorder. *Biol Psychiatry* 2007; **62**: 870–877.
4. Chou YH, Wang SJ, Lin CL, Mao WC, Lee SM, Liao MH. Decreased brain serotonin transporter binding in the euthymic state of bipolar I but not bipolar II disorder: a SPECT study. *Bipolar Disord* 2010; **12**: 312–318.
5. Ichimiya T, Suhara T, Sudo Y, Okubo Y, Nakayama K, Nankai M *et al.* Serotonin transporter binding in patients with mood disorders: a PET study with [11C](+)McN5652. *Biol Psychiatry* 2002; **51**: 715–722.
6. Parsey RV, Hastings RS, Oquendo MA, Hu X, Goldman D, Huang YY *et al.* Effect of a triallelic functional polymorphism of the serotonin-transporter-linked promoter region on expression of serotonin transporter in the human brain. *Am J Psychiatry* 2006; **163**: 48–51.
7. Heils A, Teufel A, Petri S, Seemann M, Bengel D, Balling U *et al.* Functional promoter and polyadenylation site mapping of the human serotonin (5-HT) transporter gene. *J Neural Transm Gen Sect* 1995; **102**: 247–254.
8. Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H *et al.* Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 2003; **301**: 386–389.
9. Collier DA, Stober G, Li T, Heils A, Catalano M, Di Bella D *et al.* A novel functional polymorphism within the promoter of the serotonin transporter gene: possible role in susceptibility to affective disorders. *Mol Psychiatry* 1996; **1**: 453–460.
10. Pezawas L, Meyer-Lindenberg A, Drabant EM, Verchinski BA, Munoz KE, Kolachana BS *et al.* 5-HTTLPR polymorphism impacts human cingulate-amygdala interactions: a genetic susceptibility mechanism for depression. *Nat Neurosci* 2005; **8**: 828–834.
11. Kendler KS, Kuhn JW, Vittum J, Prescott CA, Riley B. The interaction of stressful life events and a serotonin transporter polymorphism in the prediction of episodes of major depression: a replication. *Arch Gen Psychiatry* 2005; **62**: 529–535.
12. Surtees PG, Wainwright NW, Willis-Owen SA, Luben R, Day NE, Flint J. Social adversity, the serotonin transporter (5-HTTLPR) polymorphism and major depressive disorder. *Biol Psychiatry* 2006; **59**: 224–229.
13. Uher R, McGuffin P. The moderation by the serotonin transporter gene of environmental adversity in the aetiology of mental illness: review and methodological analysis. *Mol Psychiatry* 2008; **13**: 131–146.
14. Caspi A, Hariri AR, Holmes A, Uher R, Moffitt TE. Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *Am J Psychiatry* 2010; **167**: 509–527.
15. Kakiuchi C, Iwamoto K, Ishiwata M, Bundo M, Kasahara T, Kusumi I *et al.* Impaired feedback regulation of XBP1 as a genetic risk factor for bipolar disorder. *Nat Genet* 2003; **35**: 171–175.
16. Kuratomi G, Iwamoto K, Bundo M, Kusumi I, Kato N, Iwata N *et al.* Aberrant DNA methylation associated with bipolar disorder identified from discordant monozygotic twins. *Mol Psychiatry* 2008; **13**: 429–441.
17. Baranzini SE, Mudge J, van Velkinburgh JC, Khankhanian P, Khrebtukova I, Miller NA *et al.* Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. *Nature* 2010; **464**: 1351–1356.
18. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML *et al.* Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci USA* 2005; **102**: 10604–10609.
19. Kaminsky ZA, Tang T, Wang SC, Ptak C, Oh GH, Wong AH *et al.* DNA methylation profiles in monozygotic and dizygotic twins. *Nat Genet* 2009; **41**: 240–245.
20. Mill J, Dempster E, Caspi A, Williams B, Moffitt T, Craig I. Evidence for monozygotic twin (MZ) discordance in methylation level at two CpG sites in the promoter region of the catechol-O-methyltransferase (COMT) gene. *Am J Med Genet B Neuropsychiatr Genet* 2006; **141B**: 421–425.
21. Petronis A, Gottesman II, Kan P, Kennedy JL, Basile VS, Paterson AD *et al.* Monozygotic twins exhibit numerous epigenetic differences: clues to twin discordance? *Schizophr Bull* 2003; **29**: 169–178.
22. Rosa A, Picchioni MM, Kalidindi S, Loat CS, Knight J, Toulopoulou T *et al.* Differential methylation of the X-chromosome is a possible source of discordance for bipolar disorder female monozygotic twins. *Am J Med Genet B Neuropsychiatr Genet* 2008; **147B**: 459–462.
23. Kaplan L, Foster R, Shen Y, Parry DM, McMaster ML, O'Leary MC *et al.* Monozygotic twins discordant for neurofibromatosis 1. *Am J Med Genet A* 2010; **152A**: 601–606.
24. Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y *et al.* Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. *Nat Genet* 2002; **32**: 285–289.
25. Sakuntabhai A, Ruiz-Perez V, Carter S, Jacobsen N, Burge S, Monk S *et al.* Mutations in ATP2A2, encoding a Ca2+ pump, cause Darier disease. *Nat Genet* 1999; **21**: 271–277.
26. Helderman-van den Enden AT, Maaswinkel-Mooij PD, Hoogendoorn E, Willemsen R, Maat-Kievit JA, Losekoot M *et al.* Monozygotic twin brothers with the fragile X syndrome: different CGG repeats and different mental capacities. *J Med Genet* 1999; **36**: 253–257.
27. Bruder CE, Piotrowski A, Gijsbers AA, Andersson R, Erickson S, Diaz de Stahl T *et al.* Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. *Am J Hum Genet* 2008; **82**: 763–771.
28. Kusumi I, Ohmori T, Kohsaka M, Ito M, Honma H, Koyama T. Chronobiological approach for treatment-resistant rapid cycling affective disorders. *Biol Psychiatry* 1995; **37**: 553–559.
29. Sugawara H, Iwamoto K, Bundo M, Ueda J, Ishigooka J, Kato T. Comprehensive DNA methylation analysis of human peripheral blood leukocytes and lymphoblastoid cell lines. *Epigenetics* 2011; **6**: 508–515.
30. Hayashi A, Kasahara T, Kametani M, Toyota T, Yoshikawa T, Kato T. Aberrant endoplasmic reticulum stress response in lymphoblastoid cells from patients with bipolar disorder. *Int J Neuropsychopharmacol* 2009; **12**: 33–43.
31. Kato T, Ishiwata M, Mori K, Washizuka S, Tajima O, Akiyama T *et al.* Mechanisms of altered Ca2+ signalling in transformed lymphoblastoid cells from patients with bipolar disorder. *Int J Neuropsychopharmacol* 2003; **6**: 379–389.
32. Johnson WE, Li W, Meyer CA, Gottardo R, Carroll JS, Brown M *et al.* Model-based analysis of tiling-arrays for ChIP-chip. *Proc Natl Acad Sci USA* 2006; **103**: 12457–12462.
33. Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. *Bioinformatics* 2002; **18**: 1427–1431.
34. Kumaki Y, Oda M, Okano M. QUMA: quantification tool for methylation analysis. *Nucleic Acids Res* 2008; **36**: W170–W175.
35. Colella S, Shen L, Baggerly KA, Issa JP, Krahe R. Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites. *Biotechniques* 2003; **35**: 146–150.
36. Hu XZ, Lipsky RH, Zhu G, Akhtar LA, Taubman J, Greenberg BD *et al.* Serotonin transporter promoter gain-of-function genotypes are linked to obsessive-compulsive disorder. *Am J Hum Genet* 2006; **78**: 815–826.
37. Kaiser R, Muller-Oerlinghausen B, Filler D, Tremblay PB, Berghofer A, Roots I *et al.* Correlation between serotonin uptake in human blood platelets with the 44-bp polymorphism and the 17-bp variable number of tandem repeat of the serotonin transporter. *Am J Med Genet* 2002; **114**: 323–328.
38. Kaiser R, Tremblay PB, Schmider J, Henneken M, Dettling M, Muller-Oerlinghausen B *et al.* Serotonin transporter polymorphisms: no association with response to antipsychotic treatment, but associations with the schizoparanoid and residual subtypes of schizophrenia. *Mol Psychiatry* 2001; **6**: 179–185.
39. Hozumi N, Tonegawa S. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc Natl Acad Sci USA* 1976; **73**: 3628–3632.
40. Oquendo MA, Hastings RS, Huang YY, Simpson N, Ogden RT, Hu XZ *et al.* Brain serotonin transporter binding in depressed patients with bipolar disorder using positron emission tomography. *Arch Gen Psychiatry* 2007; **64**: 201–208.
41. Arango V, Underwood MD, Boldrini M, Tamir H, Kassir SA, Hsiung S *et al.* Serotonin 1A receptors, serotonin transporter binding and serotonin transporter mRNA expression in the brainstem of depressed suicide victims. *Neuropsychopharmacology* 2001; **25**: 892–903.
42. Cho HJ, Meira-Lima I, Cordeiro Q, Michelon L, Sham P, Vallada H *et al.* Population-based and family-based studies on the serotonin transporter gene polymorphisms and bipolar disorder: a systematic review and meta-analysis. *Mol Psychiatry* 2005; **10**: 771–781.
43. Kunugi H, Hattori M, Kato T, Tatsumi M, Sakai T, Sasaki T *et al.* Serotonin transporter gene polymorphisms: ethnic difference and possible association with bipolar affective disorder. *Mol Psychiatry* 1997; **2**: 457–462.
44. Mill J, Petronis A. Molecular studies of major depressive disorder: the epigenetic perspective. *Mol Psychiatry* 2007; **12**: 799–814.
45. Beach SR, Brody GH, Todorov AA, Gunter TD, Philibert RA. Methylation at SLC6A4 is linked to family history of child abuse: an examination of the Iowa Adoptee sample. *Am J Med Genet B Neuropsychiatr Genet* 2010; **153B**: 710–713.
46. van IMH, Caspers K, Bakermans-Kranenburg MJ, Beach SR, Philibert R. Methylation matters: interaction between methylation density and serotonin transporter genotype predicts unresolved loss or trauma. *Biol Psychiatry* 2010; **68**: 405–407.
47. Philibert R, Madan A, Andersen A, Cadoret R, Packer H, Sandhu H. Serotonin transporter mRNA levels are associated with the methylation of an upstream CpG island. *Am J Med Genet B Neuropsychiatr Genet* 2007; **144B**: 101–105.
48. Weaver IC, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR *et al.* Epigenetic programming by maternal behavior. *Nat Neurosci* 2004; **7**: 847–854.
49. Wong CC, Caspi A, Williams B, Craig IW, Houts R, Ambler A *et al.* A longitudinal study of epigenetic variation in twins. *Epigenetics* 2010; **5**: 516–526.
50. Connor CM, Akbarian S. DNA methylation changes in schizophrenia and bipolar disorder. *Epigenetics* 2008; **3**: 55–58.
51. Kato T, Iwamoto K, Kakiuchi C, Kuratomi G, Okazaki Y. Genetic or epigenetic difference causing discordance between monozygotic twins as a clue to molecular basis of mental disorders. *Mol Psychiatry* 2005; **10**: 622–630.

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

RESEARCH ARTICLE

AMERICAN JOURNAL OF medical genetics PART B
Neuropsychiatric Genetics

Association of *ANK3* With Bipolar Disorder Confirmed in East Asia

Atsushi Takata,[1,2] Se Hyun Kim,[3] Norio Ozaki,[4] Nakao Iwata,[5] Hiroshi Kunugi,[6] Toshiya Inada,[7] Hiroshi Ujike,[8] Kazuhiko Nakamura,[9] Norio Mori,[9] Yong Min Ahn,[3] Eun-Jeong Joo,[10] Joo Yun Song,[3] Shigenobu Kanba,[2] Takeo Yoshikawa,[11] Yong Sik Kim,[3] and Tadafumi Kato[1]*

[1]Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, Saitama, Japan
[2]Department of Psychiatry, Kyushu University, School of Medicine, Fukuoka, Japan
[3]Department of Psychiatry, Seoul National University College of Medicine, Seoul, Korea
[4]Department of Psychiatry, Graduate School of Medicine, Nagoya University, Aichi, Japan
[5]Department of Psychiatry, Fujita Health University School of Medicine, Aichi, Japan
[6]Department of Mental Disorder Research, National Institute of Neuroscience, Tokyo, Japan
[7]Seiwa Hospital, Institute of Neuropsychiatry, Tokyo, Japan
[8]Ujike Nishiguchi Clinic, Okayama, Japan
[9]Department of Psychiatry and Neurology, Hamamatsu University School of Medicine, Hamamatsu, Japan
[10]Department of Neuropsychiatry, Eulji University School of Medicine, Eulji General Hospital, Seoul, Korea
[11]Laboratory for Molecular Psychiatry, RIKEN Brain Science Institute, Saitama, Japan

Received 7 September 2010; Accepted 1 December 2010

Results of genome-wide association studies (GWASs) for bipolar disorder (BD) have indicated *ANK3* as one of the most promising candidates for a susceptibility gene. In this study, we performed genetic association analysis of two single-nucleotide polymorphisms (SNPs) in *ANK3* (rs1938526 and rs10994336), whose genome-wide significant associations were reported in a previous meta-analysis of GWASs, using genotyping data of Korean and Japanese case–control samples and a part of data from a GWAS in Han-Chinese from Taiwan. The total number of participants was 2,212 cases (352 from Korea, 860 from Japan, and 1,000 from Taiwan) and 2,244 controls (349 from Korea, 895 from Japan, and 1,000 from Taiwan). We could not detect any significant difference of allele frequency in individual analyses using each of the three populations. However, when we combined the three data sets and performed a meta-analysis, rs1938526 showed nominally significant association ($P = 0.048$, odds ratio = 1.09). The over-represented allele in BD was same as that reported in Caucasian GWASs. On the other hand, any significant association was not detected in rs10994336. This discrepancy between two SNPs may be explained by the different degree of linkage disequilibrium between Asian and Caucasian. These findings further supported the association between *ANK3* and BD, and also suggested the genomic region around rs1938526 as a common risk locus across ethnicities.

Key words: GWAS; mood disorder; Asian; association study; replication

How to Cite this Article:
Takata A, Kim SH, Ozaki N, Iwata N, Kunugi H, Inada T, Ujike H, Nakamura K, Mori N, Ahn YM, Joo EJ, Song JY, Kanba S, Yoshikawa T, Kim YS, Kato T. 2011. Association of *ANK3* With Bipolar Disorder Confirmed in East Asia.
Am J Med Genet Part B 156:312–315.

INTRODUCTION

Recent genome-wide association studies (GWASs) for bipolar disorder (BD) have identified a number of new candidate genes. The *ANK3* gene, encoding ankyrin-G, is one of these newly found candidates. The ankyrin family proteins link the spectrin cytoskeleton to various membrane proteins, including ion channels [Bennett and Baines, 2001]. Increased content of a minor form of

Atsushi Takata and Se Hyun Kim contributed equally to this work.
*Correspondence to:
Tadafumi Kato, M.D., Ph.D., Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, Hirosawa 2-1, Wako, Saitama 351-0198, Japan. E-mail: kato@brain.riken.jp
Published online 13 January 2011 in Wiley Online Library (wileyonlinelibrary.com)
DOI 10.1002/ajmg.b.31164

ankyrin in erythrocyte membranes of BD patients treated with lithium was reported [Zhang and Meltzer, 1989]. Ankyrin-G was originally detected at the axonal initial segment and nodes of Ranvier [Kordeli et al., 1995]. In Purkinje cells derived from mutant mice with cerebellum-specific knockout of ankyrin-G, deficient ability to initiate action potentials and to support rapid and repetitive firing was observed [Zhou et al., 1998]. Thus, ankyrin-G plays a certain role in neuronal excitability.

Association between *ANK3* and BD was initially reported in the first GWAS using DNA pooling [Baum et al., 2008]. Afterwards, genome-wide significant associations in two single-nucleotide polymorphisms (SNPs) in *ANK3* (rs1938526 and rs10994336) were detected in a meta-analysis of GWASs for BD with 4,387 cases and 6,209 controls [Ferreira et al., 2008]. Schulze et al. [2009] confirmed the association reported in the first pooling-GWAS by individual genotyping and replicated the association of rs10994336 in an independent sample. Other GWASs also reported suggestive association signals in *ANK3* region [Scott et al., 2009; Smith et al., 2009; Lee et al., 2010]. According to these results, *ANK3* has attracted huge attentions in the field of genetic research for BD. However, except for the suggestive association reported in a GWAS of Han-Chinese in Taiwan [Lee et al., 2010], all of the previously reported associations were of Caucasian samples. Therefore, further analysis of *ANK3* in BD using samples with an ethnicity different from Caucasian is warranted to test the robustness of this association. In this study, we focused on the two SNPs in *ANK3*, whose genome-wide significant associations were observed in Ferreira et al. and evaluated their association with BD, using genotyping data of these two SNPs in case–control samples from Korea and Japan and a part of data from a GWAS in Taiwan. The total number of participants was 2,212 cases (352 from Korea, 860 from Japan and 1,000 from Taiwan) and 2,244 controls (Korea, 349, Japan, 895, and Taiwan, 1,000).

MATERIALS AND METHODS

Samples and Data

Korean case–control samples. A total 701 individuals of unrelated patient and control samples were included in this study, which consisted of 349 controls (173 men, 176 women; mean age = 25.9 ± 6.6 years) and 352 patients with BD (149 men, 203 women; mean age = 33.5 ± 12.0 years; 254 bipolar I disorder [BDI], 98 bipolar II disorder [BDII]). Participants were all ethnically Korean. All the patients met the diagnostic criteria of DSM-IV for BD. The patients were individually interviewed by trained researchers using the Korean-translated version of the Diagnostic Interview for Genetic Studies [DIGS, Joo et al., 2004]. Consensus diagnostic meetings with more than two psychiatrists were held regularly to evaluate the participants' final diagnoses and to rule out subjects with organic brain disease, alcohol or drug problems, or other general medical conditions possibly manifesting as psychiatric symptoms. Healthy controls were free from present, past, and family history (first-degree relatives) of psychiatric illness or substance abuse diagnoses. The study protocol was approved by the ethics committee of Seoul National University Hospital. Written informed consent was obtained from each patient prior to enrollment.

Japanese COSMO case–control samples. These samples were recruited through the COSMO (Collaborative Study of Mood Disorders) consortium [Iwayama et al., 2010], which comprises 860 unrelated bipolar patients (424 men, 436 women; mean age 50.2 ± 14.4 years; 584 BDI, 276 BDII) and 895 age- and sex-matched controls (445 men, 450 women; mean age 49.9 ± 13.5 years). All samples are of Japanese origin. The patients were diagnosed by at least two experienced psychiatrists according to the DSM-IV criteria for BDI or BDII on the basis of unstructured interviews and reviews of their medical records. All healthy control subjects were also psychiatrically screened on the basis of unstructured interviews. The present study was approved by the ethics committees of all participating institutes. All controls and patients gave informed written consent to participate in the study, after provision and explanation of study protocols and objectives.

GWAS data in Taiwan. Genotype data of two *ANK3* SNPs were obtained by a GWAS in Taiwan [Lee et al., 2010] and were kindly provided by Dr. Andrew Cheng (Institute of Biomedical Sciences, Academia Sinica, Taiwan). The samples consisted of 1,000 BDI patients (464 men, 536 women; mean age 42.8 ± 12.9 years) and 1,000 normal controls (502 men, 498 women; mean age 51.7 ± 17.8 years). All samples are of Han Chinese origin. Detailed clinical information and method of diagnosis is described in the previous paper [Lee et al., 2010].

Markers, genotyping, and data analysis. In this study, we genotyped following two SNP markers in *ANK3*: rs1938526 (A/G) and rs10994336 (C/T). In Ferreira et al., rs1938526, which was genotyped, was associated with BD with odds ratio (OR) of 1.395 and *P*-value of 1.3×10^{-8} and rs10994336, whose genotype was imputed, was associated with OR of 1.450 and *P*-value of 9.1×10^{-9}.

In Korean and Japanese samples, TaqMan assay (Applied Biosystems, Foster City, CA) was used for genotyping. In GWAS in Taiwan, genotypes were determined using the Illumina HumanHap550-Duo BeadChip (Illumina, San Diego, CA).

Deviation from Hardy–Weinberg equilibrium and linkage disequilibrium between the two SNPs were calculated using Haploview software v 4.2 [Barrett et al., 2005]. Association between markers and BD in each individual population was assessed using the chi-square test. A meta-analysis using Mantel–Haenszel model and evaluation of sample heterogeneity were performed on Review Manager computer software [RevMan, 2008]. The I^2 statistics were used for the assessment of heterogeneity between the samples [Higgins and Thompson 2004; Trikalinos et al., 2008]. Statistical power for detection was calculated by Genetic Power Calculator [http://pngu.mgh.harvard.edu/~purcell/gpc/, Purcell et al., 2003] with a multiplicative model and following parameters; risk allele frequency, 0.3642 for rs1938526 and 0.2699 for rs10994336, these values were calculated as the frequencies of the previously reported risk alleles in our control samples; number of cases, 2,212; number of controls, 2,244; prevalence, 0.01; OR, 1.1, or 1.2; α-level, 0.05.

RESULTS

The results of the case–control association analyses in each population and the meta-analysis were summarized in Table I. When genetic association between rs1938526 or rs10994336 and BD were

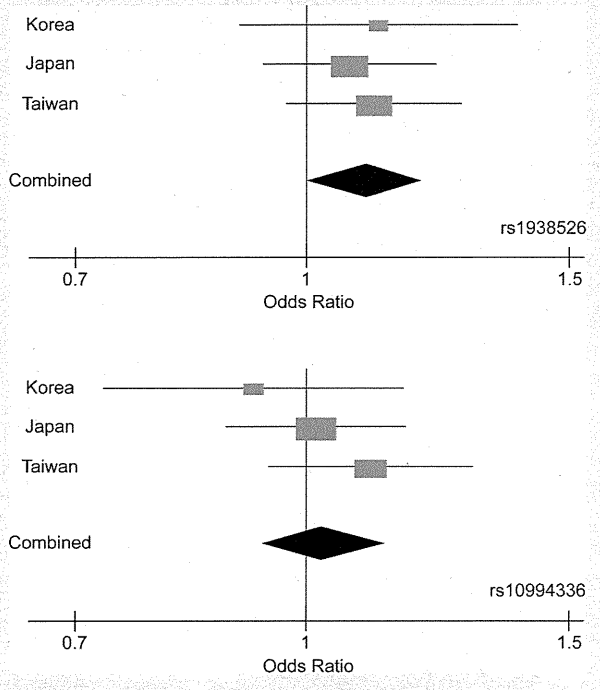

FIG. 1. Forest plots of odds ratios and their 95% confidence intervals for the individual case–control analyses and the meta-analysis.

individually analyzed using each of the three populations, we could not detect any significant difference of allele frequency. However, in the meta-analysis combining the three data sets, rs1938526 showed nominally significant association ($P = 0.048$, OR = 1.09) (Fig. 1). Except for rs1938526 in Korean control population ($P = 0.045$), genotype frequencies of both SNPs in each diagnostic group and each recruited country were within Hardy–Weinberg equilibrium ($P > 0.05$). Although the minor allele frequencies (MAF) were different among three populations, the I^2 statistics were 0% in both SNPs, indicating there was no variation in OR attributable to heterogeneity. Linkage disequilibrium between the two SNPs assessed by D′ was 0.76 in Korean, 0.71 in Japanese and 0.84 in Han-Chinese in Taiwan. The statistical power for detection obtained from overall sample size of our populations were estimated as follows: rs1938526, 98.8% for OR 1.2 and 59.5% for OR 1.1; rs10994336, 97.6% for OR 1.2 and 53.2% for OR1.1.

DISCUSSION

In this study, we detected a nominal association between rs1938526 in *ANK3* and BD, as a result of the meta-analysis using data sets from Korea, Japan, and Taiwan. The G allele of rs1938526 was more frequent in BD cases, in both of our study, using East Asian samples, and Ferreira et al., using Caucasian samples. The statistical power obtained from the overall sample size in this study was substantial for a replication study analyzing a single marker, owing to the relatively high MAF in our populations (MAF in controls of present

TABLE I. Results of the Case–Control Association Analyses in Each Population and the Meta-Analysis

	Korea				Japan				Taiwan				Meta-analysis	
	MAF				MAF				MAF					
SNP, minor allele	Cases (N = 352)	Controls (N = 349)	*P*-value	OR (95% CI)	Cases (N = 860)	Controls (N = 895)	*P*-value	OR (95% CI)	Cases (N = 1,000)	Controls (N = 1,000)	*P*-value	OR (95% CI)	*P*-value	OR (95% CI)
rs1938526, G	0.411	0.384	0.310	1.12 [0.90–1.38]	0.453	0.438	0.359	1.06 [0.93–1.22]	0.314	0.292	0.130	1.11 [0.97–1.27]	0.048*	1.09 [1.00–1.19]
rs10994336, T	0.280	0.297	0.489	0.92 [0.73–1.16]	0.357	0.355	0.918	1.01 [0.88–1.16]	0.201	0.185	0.214	1.10 [0.94–1.29]	0.591*	1.03 [0.93–1.13]

CI, confidential interval; MAF, minor allele frequency; OR, odds ratio; SNP, single-nucleotide polymorphism.
*Test for heterogeneity; $I^2 = 0\%$.

study, 0.3642 for rs1938526 and 0.2699 for rs10994336; MAF in controls of Ferreira et al., 0.056 for rs1938526 and 0.053 for rs10994336). Although significant difference in allele frequencies was not detected in the analyses using each individual population, they showed similar ORs, indicating insufficient statistical power rather than truly negative association as a source of the lack of significance. These results collectively suggest this SNP as a risk marker common across ethnicities.

On the other hand, we could not observe association between rs10994336 and BD. This discrepancy between two SNPs in our populations, both of which showed genome-wide significant association in Ferreira et al., may be explained by the different degree of LD among populations, particularly low LD in Korean and Japanese. The D′ value between two SNPs was 0.86 in Caucasian (from HapMap data), 0.84 in Han-Chinese in Taiwan, 0.76 in Korean, and 0.71 in Japanese (from our data). In accordance with different degree of LD, ORs of rs1938526 and rs10994336 in our data were similar in Han-Chinese in Taiwan (1.11 and 1.10) and slightly different in Korean (1.12 and 0.92) and Japanese (1.06 and 1.01).

The OR of rs1938526 estimated in this study was quite lower than that detected in previous Caucasian studies. Previous studies for *ANK3* in BD suggested the presence of allelic heterogeneity in this gene [Schulze et al., 2009; Smith et al., 2009]. In the case of Parkinson's disease, association signals in GWASs were detected in genes whose rare functional mutations have already known to be disease causing [Satake et al., 2009; Simon-Sanchez et al., 2009; Edwards et al., 2010]. In the case of hypertriglyceridemia, excess of rare mutations in patients was observed in genes identified by GWASs was reported [Johansen et al., 2010]. Therefore, it should be worthwhile to resequence the locus around rs1938526 seeking for rare functional mutations.

In conclusion, the results of this study as well as the findings in previous GWASs and its replication study supported the association of *ANK3* with BD. Especially, the genomic region around rs1938526 is a valuable target for a rare mutation search by resequencing.

ACKNOWLEDGMENTS

We thank for all the patients and control subjects who volunteered to the study. We are grateful to Dr. Andrew Cheng (Institute of Biomedical Sciences, Academia Sinica, Taiwan) who generously provided the raw data of GWAS and D′ of two SNPs in Taiwan.

REFERENCES

Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21(2):263–265.

Baum AE, Akula N, Cabanero M, Cardona I, Corona W, Klemens B, Schulze TG, Cichon S, Rietschel M, Nothen MM, et al. 2008. A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. Mol Psychiatry 13(2):197–207.

Bennett V, Baines AJ. 2001. Spectrin and ankyrin-based pathways: Metazoan inventions for integrating cells into tissues. Physiol Rev 81(3): 1353–1392.

Edwards TL, Scott WK, Almonte C, Burt A, Powell EH, Beecham GW, Wang L, Zuchner S, Konidari I, Wang G, et al. 2010. Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. Ann Hum Genet 74(2): 97–109.

Ferreira MA, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L, Fan J, Kirov G, Perlis RH, Green EK, et al. 2008. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. Nat Genet 40(9):1056–1058.

Higgins JP, Thompson SG. 2004. Controlling the risk of spurious findings from meta-regression. Stat Med 23(11):1663–1682.

Iwayama Y, Hattori E, Maekawa M, Yamada K, Toyota T, Ohnishi T, Iwata Y, Tsuchiya KJ, Sugihara G, Kikuchi M, et al. 2010. Association analyses between brain-expressed fatty-acid binding protein (FABP) genes and schizophrenia and bipolar disorder. Am J Med Genet Part B 153B(2): 484–493.

Johansen CT, Wang J, Lanktree MB, Cao H, McIntyre AD, Ban MR, Martins RA, Kennedy BA, Hassell RG, Visser ME, et al. 2010. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. Nat Genet 42(8):684–687.

Joo EJ, Joo YH, Hong JP, Hwang S, Maeng SJ, Han JH, Yang BH, Lee YS, Kim YS. 2004. Korean version of the diagnostic interview for genetic studies: Validity and reliability. Compr Psychiatry 45(3):225–229.

Kordeli E, Lambert S, Bennett V. 1995. AnkyrinG. A new ankyrin gene with neural-specific isoforms localized at the axonal initial segment and node of Ranvier. J Biol Chem 270(5):2352–2359.

Lee MT, Chen CH, Lee CS, Chen CC, Chong MY, Ouyang WC, Chiu NY, Chuo LJ, Chen CY, Tan HK, et al. 2010. Genome-wide association study of bipolar I disorder in the Han Chinese population. Mol Psychiatry [Epub ahead of print].

Purcell S, Cherny SS, Sham PC. 2003. Genetic Power Calculator: Design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19(1):149–150.

The Cochrane Collaboration. Review Manager (RevMan) [Computer program]. Version 5.0. Copenhagen: The Nordic Cochrane Centre. 2008.

Satake W, Nakabayashi Y, Mizuta I, Hirota Y, Ito C, Kubo M, Kawaguchi T, Tsunoda T, Watanabe M, Takeda A, et al. 2009. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. Nat Genet 41(12):1303–1307.

Schulze TG, Detera-Wadleigh SD, Akula N, Gupta A, Kassem L, Steele J, Pearl J, Strohmaier J, Breuer R, Schwarz M, et al. 2009. Two variants in Ankyrin 3 (ANK3) are independent genetic risk factors for bipolar disorder. Mol Psychiatry 14(5):487–491.

Scott LJ, Muglia P, Kong XQ, Guan W, Flickinger M, Upmanyu R, Tozzi F, Li JZ, Burmeister M, Absher D, et al. 2009. Genome-wide association and meta-analysis of bipolar disorder in individuals of European ancestry. Proc Natl Acad Sci USA 106(18):7501–7506.

Simon-Sanchez J, Schulte C, Bras JM, Sharma M, Gibbs JR, Berg D, Paisan-Ruiz C, Lichtner P, Scholz SW, Hernandez DG, et al. 2009. Genome-wide association study reveals genetic risk underlying Parkinson's disease. Nat Genet 41(12):1308–1312.

Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, Berrettini W, Byerley W, Coryell W, Craig D, Edenberg HJ, et al. 2009. Genome-wide association study of bipolar disorder in European American and African American individuals. Mol Psychiatry 14(8):755–763.

Trikalinos TA, Salanti G, Zintzaras E, Ioannidis JP. 2008. Meta-analysis methods. Adv Genet 60:311–334.

Zhang Y, Meltzer HL. 1989. Increased content of a minor ankyrin in erythrocyte membranes of bipolar subjects. Psychiatry Res 27(3): 267–275.

Zhou D, Lambert S, Malen PL, Carpenter S, Boland LM, Bennett V. 1998. AnkyrinG is required for clustering of voltage-gated Na channels at axon initial segments and for normal action potential firing. J Cell Biol 143(5):1295–1304.

Takata et al. Genome Biology 2011, 12:R92
http://genomebiology.com/2011/12/9/R92

Genome Biology

RESEARCH Open Access

Exome sequencing identifies a novel missense variant in *RRM2B* associated with autosomal recessive progressive external ophthalmoplegia

Atsushi Takata[1,2,3], Maiko Kato[4], Masayuki Nakamura[4], Takeo Yoshikawa[3], Shigenobu Kanba[2], Akira Sano[4] and Tadafumi Kato[1*]

Abstract

Background: Whole-exome sequencing using next-generation technologies has been previously demonstrated to be able to detect rare disease-causing variants. Progressive external ophthalmoplegia (PEO) is an inherited mitochondrial disease that follows either autosomal dominant or recessive forms of inheritance (adPEO or arPEO). AdPEO is a genetically heterogeneous disease and several genes, including *POLG1* and *C10orf2*/Twinkle, have been identified as responsible genes. On the other hand, *POLG1* was the only established gene causing arPEO with mitochondrial DNA deletions. We previously reported a case of PEO with unidentified genetic etiology. The patient was born of a first-cousin marriage. Therefore, the recessive form of inheritance was suspected.

Results: To identify the disease-causing variant in this patient, we subjected the patient's DNA to whole-exome sequencing and narrowed down the candidate variants using public data and runs of homozygosity analysis. A total of 35 novel, putatively functional variants were detected in the homozygous segments. When we sorted these variants by the conservation score, a novel missense variant in *RRM2B*, whose heterozygous rare variant had been known to cause adPEO, was ranked at the top. The list of novel, putatively functional variants did not contain any other variant in genes encoding mitochondrial proteins registered in MitoCarta.

Conclusions: Exome sequencing efficiently and effectively identified a novel, homozygous missense variant in *RRM2B*, which was strongly suggested to be causative for arPEO. The findings in this study indicate arPEO to be a genetically heterogeneous disorder, as is the case for adPEO.

Background

Massively parallel sequencing, also known as next generation-sequencing, is a revolutionary technology that enables us to obtain large amounts of genomic sequence information in an incomparably more rapid and less expensive manner than before [1]. This technology is applicable for various investigations, including resequencing of full genomes or more targeted parts thereof for discovery of genomic variations, genome-wide mapping of structural rearrangements, transcriptome sequencing, genome-wide epigenetic analysis, metagenomic sequencing, and so on [2]. Whole-genome and whole-exome (sequences of all protein-coding regions) resequencing aiming at identification of causative variants for rare, inherited diseases is one of these applications, and have demonstrated their efficiency and effectiveness (reviewed in [3]).

Previously, we reported a patient who had been born of a first-cousin marriage and was suspected to be affected by inherited progressive external ophthalmoplegia (PEO) [4]. Inherited PEO is a form of mitochondrial disease that follows either autosomal dominant or recessive forms of inheritance (adPEO (MIM 157640; 609283; 609286; 610131, 613077) or arPEO (MIM 258450)). The characteristic findings of inherited PEOs are multiple mitochondrial DNA (mtDNA) deletions and ragged red fibers in the muscle biopsy [5]. Typical clinical symptoms are bilateral ptosis and paralysis of the extraocular muscle. Other symptoms include exercise intolerance, cataracts, hearing loss, sensory axonal neuropathy, optic atrophy,

* Correspondence: kato@brain.riken.jp
[1]Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan
Full list of author information is available at the end of the article

ataxia, depression, hypogonadism, and Parkinsonism [6-10].

In the present case, the recessive form of inheritance was suspected because of the patient's family history. However, no pathogenic variant in *POLG1* (MIM 174763), which encodes a mitochondrial DNA polymerase and was the only established gene whose variants were known to cause arPEO so far, was identified [4].

The proband in this study was the only child and the available genetic information from family members was limited. Therefore, it was almost impossible to identify the causative variant using linkage analysis. On the other hand, exome sequencing using a next-generation sequencer has demonstrated its utility to detect causative variants of rare disease using a small number of samples, especially in the case of consanguineous family. Here, we performed exome sequencing in combination with runs of homozygosity (ROH) analysis in order to identify the causative variant in this patient.

Results

Exome sequencing identifies a novel, homozygous missense variant in *RRM2B*

A total of 3.2 Gb of sequence was generated from one lane of sequencing using the Illumina Genome Analyzer II (Illumina, San Diego, CA, USA). The proportion of the targeted exome covered at 1×, 5× and 10× was 96.3%, 88.0% and 78.3%, respectively. The mean coverage was 37.2×. A total of 19, 215 variants were detected in the coding regions defined by RefSeq Gene [11] and their flanking splice sites. The number of detected coding variants does not deviated greatly from that in previous reports [3,12]. After removing variants registered on the public database of sequence variants (dbSNP, build 130) or found in eight exomes of HapMap individuals [12] or the exome of a single, healthy, unrelated Japanese individual, which was analyzed in the same run of Illumina Genome Analyzer II sequencing, 1, 336 variants remained. Among these, 592 variants, including 141 homozygous ones, were functional (nonsense, missense, frameshift or splice site). Next, we performed ROH analysis to narrow down the candidate regions, using the base calling data on single nucleotide variants in this patient. To enhance the accuracy of the variant calling used for this analysis, 1) only the data of single nucleotide variants were used and insertion/deletion variants were excluded because of lower reliability of the detection of insertion/deletion variants [13], 2) variants called with coverage less than 8× were excluded, 3) variants called with a coverage of more than 100× were excluded because genomic regions that are known to be duplicated or have similar sequences such as pseudogenes tend to be read with high coverage. Because the primary aim of this analysis was not to evaluate ROH segments precisely, but to narrow down the list of candidate variants without overlooking the causative variant, we used relaxed criteria of ROH segments. The total size of ROH regions was 992 Mb (about 32% of the genome), which was significantly larger than the expected total size of ROH segments in an offspring born from a first cousin marriage (one-eighth of the genome). A total of 35 novel and functional variants in 33 genes were identified in ROH segments. A summary of the filtering strategy is given in Table 1.

When we sorted these listed variants by a conservation score (phyloP score) to identify those that were most likely to be functional, a novel missense variant in *RRM2B* (g.341G > A, p.P33S), whose rare, heterozygous variant had been known to cause adPEO, was ranked at the top (Table 2).

The existence of the *RRM2B* variant in the patient's DNA was confirmed by Sanger sequencing (Figure 1a). As expected, each of the parents had this variant in the heterozygous state. This variant changes an amino acid residue that is highly conserved across 44 vertebrates (Figure 1b). Among 359 control subjects (718 chromosomes) of Japanese origin, one subject carried this variant in the heterozygous state.

Exclusion of other variants that could cause PEO

In the list of 35 novel and functional variants in the ROH segments, no other variants in genes encoding mitochondrial proteins were registered in Human MitoCarta [14]. We could not find any pathogenic mutations in other genes known to cause mitochondrial diseases with multiple mtDNA deletions (*POLG1*, *POLG2* (MIM 604983), *C10orf2* (MIM 606075), *SLC25A4* (MIM 103220), *OPA1* (MIM 605290), *TYMP* (MIM 131222) and *WFS1* (MIM 606201)) in exome analysis, as was observed in a previous study using Sanger sequencing [4]. Although the mtDNA sequence was not targeted by the SureSelect Human All Exon Kit (Agilent, Santa Clara, CA, USA), 16, 558 of 16, 568 (99.9%) bases in mtDNA were read four or more times due to its higher copy number than nuclear DNA, and no known pathogenic variant was found. Because of the family history of the patient, we suspected that his disease was caused by a recessive mutation. However, there was another possibility that *de novo* variants affect him in a dominant manner. To test this possibility, we investigated whether he had *de novo* variants that could explain his symptoms. In the list of 592 novel and putatively functional variants, there were 26 heterozygous variants in genes registered in MitoCarta. Among them, five variants were not found in dbSNP132 or 1000 Genome Project data [15] (SNP calls released in June 2011), and were located at conserved base positions (phyloP score > 2). By performing Sanger sequencing, we confirmed that all of these variants were not *de novo*, but inherited from either of his healthy parents or found as a false positive (Table 3).

Table 1 Summary of the filtering to narrow down the candidates for the causal variant

Criteria for the filtering	Number of remaining variants
Coding variants	19, 215
Not in dbSNP130	2, 015
Not in eight HapMap exomes [12]	1, 833
Not in in-house data of a healthy Japanese individual	1, 336
Functional (missense, nonsense, frameshift and splice site)	592
In run-of-homozygosity regions	35 (in 33 genes)

The filtering was performed using the listed criteria in descending order.

Evaluation of the amount of mtDNA

The mtDNA copy number relative to nuclear DNA in the patient's skeletal muscle was not decreased, but rather increased (Figure 2). As expected, the *ND4/RNaseP* ratio was lower than the *ND1/RNaseP* ratio in the patient, which suggests increased levels of mtDNA deletions that include the *ND4* region, such as the 4, 977-bp common mtDNA deletion [16]. This result indicated that the clinical manifestation in the present patient was not due to mtDNA depletion.

Discussion

In this study, we subjected DNA from a PEO patient with unidentified genetic etiology to exome sequencing and detected a novel, homozygous missense variant in *RRM2B*. *RRM2B* encodes p53-inducible ribonucleotide reductase small subunit 2-like protein (p53R2) and this protein plays an essential role in the maintenance of mtDNA by reducing ribonucleotides in the cytosol [17], as is indicated by the fact that rare variants in this gene cause various forms of mitochondrial diseases characterized by mtDNA depletion and deletions. To our knowledge, 15 cases of mitochondrial depletion syndrome (MIM 612075) from 11 families [18-22] and one sporadic case of mitochondrial neurogastrointestinal encephalopathy [23] (MIM 603041) associated with homozygous or compound heterozygous rare variants in *RRM2B* have been reported. More recently, two families with adPEO due to a heterozygous nonsense variant were described [24]. In the screening of *RRM2B* variants in 50 mitochondrial disease patients without causative variants in *POLG1* and *C10orf2*, one Kearns-Sayre syndrome (MIM 530000) patient who carried two different novel missense variants and one PEO patient who carried an in-frame deletion were identified [25].

The clinical symptoms and findings in the muscle biopsy of our case were typical for Mendelian-inherited PEO. No members of his maternal family have shown any neuromuscular symptoms, suggesting that the mtDNA deletions of the patient were not maternally inherited. Real-time quantitative PCR analysis revealed that there was no mtDNA depletion. We did not observe gastrointestinal dysmotility, cardiac conduction abnormalities, pancreatic dysfunction and sensory ataxic neuropathy, which are characteristic symptoms for other mitochondrial diseases associated with mtDNA deletions, namely mitochondrial neurogastrointestinal encephalopathy, Kearns-Sayre syndrome, Pearson syndrome, and

Table 2 List of novel and functional variants in run-of-homozygosity regions

Chromosome	Position	Reference allele	Variant allele	Variant calling/coverage	Gene	Amino acid change	PhyloP score
8	103313660	G	A	58/58	*RRM2B*	Pro33Ser	6.741
1	39620317	G	A	5/7*	*MACF1*	Arg2523Gln; Arg3025Gln	5.329
4	107449465	A	C	63/63	*MGC16169*	Asn34Lys	5.199
22	15980313	C	T	5/5*	*LOC100287323*	Val569Ile	4.997
11	64117795	G	A	4/4*	*SLC22A12*	Trp37Stp; Trp258Stp	4.945
10	29010439	G	C	24/24	*BAMBI*	Gly108Ala	4.878
20	49482400	G	A	4/4*	*NFATC2*	Ala778Val	4.437
1	238437608	C	T	10/12	*FMN2*	Pro1101Leu	3.804
1	85362528	T	-	65/69	*WDR63*	Splice site	3.503
3	99094433	A	G	24/34	*DKFZp667G2110*	Lys546Glu	3.299
3	336547	T	G	23/23	*CHL1*	Ser30Ala	3.014
3	46595758	C	G	27/40*	*LRRC2*	Arg41Gly	2.522
4	169335658	A	C	9/13*	*ANXA10*	Thr193Pro	2.257
5	140538797	C	T	127/127	*PCDHB8*	Thr333Ile	2.011

Variants with PhyloP score > 2 are listed. Asterisks indicate variants with coverage < 8× or a variant calling/coverage ratio < 0.7; the reliability of these variant calls is generally lower than that of the others.

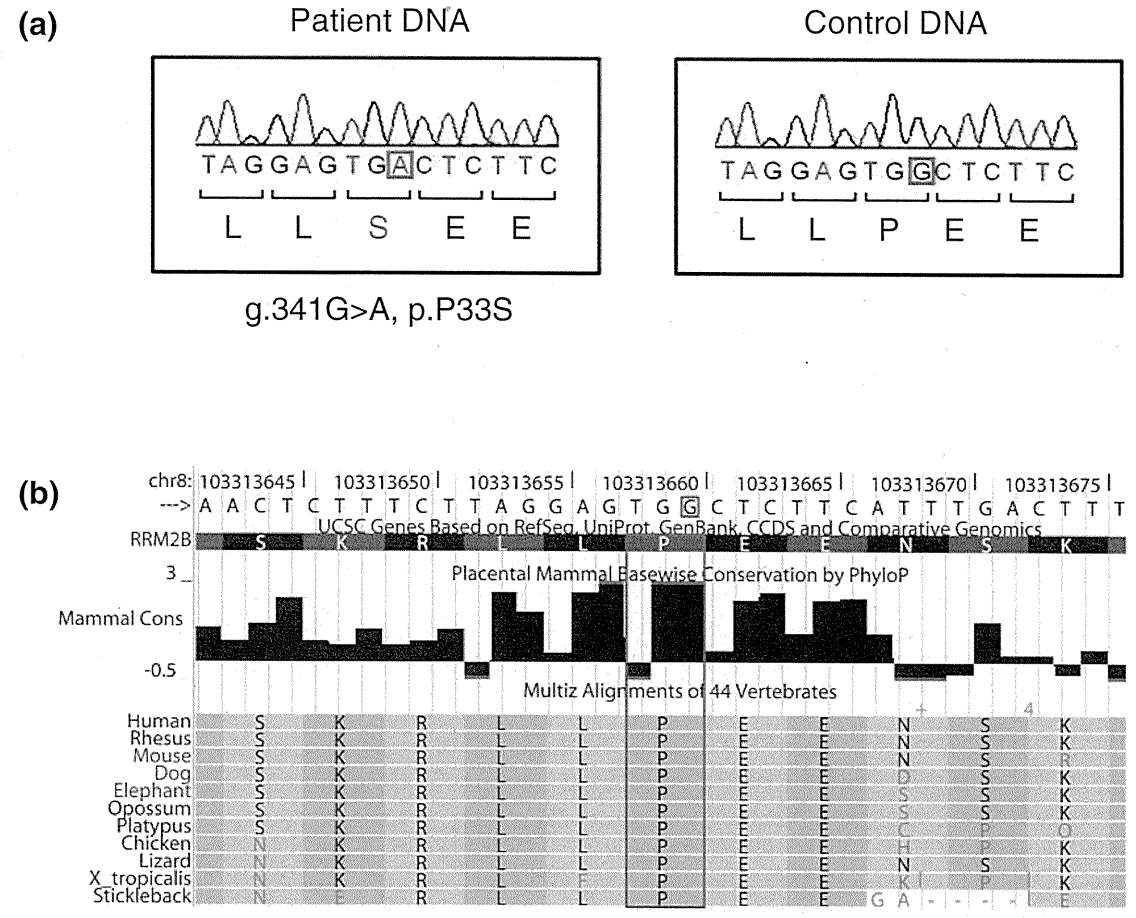

Figure 1 The identified disease-associated variant in *RRM2B*. **(a)** Partial sequence of *RRM2B* in the patient's DNA (left) and control DNA (right). Red squares indicate the base position of the g.341G > A, p.P33S variant. **(b)** The substituted amino acid residue (red box) is highly conserved across 44 vertebrate species (from the UCSC genome browser [31]).

sensory ataxic neuropathy, dysarthria, and ophthalmoparesis (MIM 607459), respectively. Therefore, this patient was diagnosed as having arPEO caused by a homozygous missense variant of *RRM2B*.

Before this study, *POLG1* had been the only established gene responsible for arPEO, while adPEO is a genetically heterogeneous disease, caused by rare variants in *POLG1, POLG2, C10orf2, SLC25A4, OPA1* and *RRM2B*. The results of this study identifying the second responsible gene for arPEO indicate that arPEO is also a genetically heterogeneous disease, as is the case for adPEO.

The symptoms observed in this patient included major depressive episodes. Frequent comorbidity of mood disorders in patients of mitochondrial disease has been generally recognized [26] and several lines of evidences have supported the possible involvement of mitochondrial dysfunctions in the pathophysiology of mood

Table 3 List of novel, putatively functional and heterozygous variants in mitochondrial genes

Chromosome	Position	Reference allele	Variant allele	Variant calling/ coverage	Gene	Amino acid change	PhyloP score	Inheritance
7	30615756	G	C	36/69	*GARS*	Asp256His	6.494	Paternally inherited
10	104476790	T	T	14/30	*SFXN2*	Leu73Pro	4.906	Maternally inherited
7	100670236	C	C	20/51	*FIS1*	Ala90Pro	3.824	Maternally inherited
11	47620527	A	A	3/8	*MTCH2*	Tyr23His	3.680	Not confirmed in Sanger sequencing
1	10286026	C	G	22/46	*KIF1B*	Ile732Met	3.092	Maternally inherited

Variants with PhyloP score > 2 are listed.

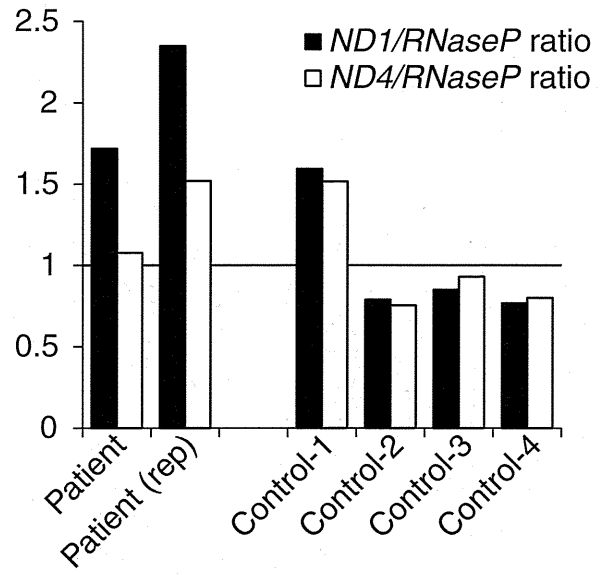

Figure 2 Relative amounts of mtDNA in skeletal muscle tissues from the patient and four control subjects. *ND1/RNaseP* and *ND4/RNaseP* ratios calculated by real-time quantitative PCR were used to evaluate mtDNA levels. The mtDNA level in the patient was comparable to those of controls. Values are relative to the average of all four controls.

disorders [27]. So far, rare variants of *POLG1*, *C10orf2* and *SLC25A4* have been reported in inherited PEO pedigrees with frequent comorbidity of mood disorders [28]. Given the typical symptoms of major depressive disorder in the present case, *RRM2B* should be added to the list of genes causal for PEO associated with mood disorders.

The identified P33S variant changes an amino acid residue highly conserved among vertebrates. The amino-terminal region of p53R2, in which this altered amino acid is located, is suggested to be crucial for interaction with p21 protein. p53R2 may contribute to DNA repair in cooperation with p21 [29]. In its amino-terminal region, the homozygous p.R41P variant was detected in a mitochondrial depletion syndrome case [21]. On the other hand, other pathogenic missense variants have been located in various sites of p53R2, including those involved in iron-binding [18,20], those putatively crucial for homodimerization of p53R2 [21,23] or heterotetramerization with the RRM1 (ribonucleoside-diphosphate reductase large subunit) homodimer [18,22], and so on. The relationships between clinical phenotypes and the properties of variants, as well as their underlying mechanisms, should be the subject of further investigations.

Conclusions

In this study, we describe a homozygous missense variant in *RRM2B* that is strongly suggested to cause arPEO. We were not only able to identify the disease-associated variant, but could also exclude other candidates (that is, variants in known PEO-related genes such as *POLG1*, other mitochondrial genes in nucleic DNA and mtDNA) using data from single exome sequencing. This result further demonstrates the efficiency and effectiveness of exome sequencing to detect causative variants of rare, inherited, and genetically heterogeneous diseases.

Materials and methods

Clinical information of the patient

The detailed clinical history, family history and laboratory data of the studied subject are described elsewhere [4]. Briefly, a 43-year-old man presented with hearing loss, bilateral ptosis, external ophthalmoplegia and muscle weakness. Examinations revealed the existence of pigmentary degeneration of the retina and gonadal atrophy. The initial symptom of progressive hearing loss began at age 16 years. Depressive mood, anxiety and hypochondriacal complaints were observed in his clinical course. His parents were first cousins, he had no siblings, and no other member of his family has a known history of neurological illness. In the muscle biopsy, marked variation of muscle fiber size, ragged red fibers, COX-negative fibers and multiple mtDNA deletions were detected. According to his clinical history, family history and laboratory data, arPEO was suspected.

The present study conformed to the Declaration of Helsinki, and was approved by the RIKEN Wako Institute Ethics Committee I, as well as the ethics committees of Kagoshima University Graduate School of Medical and Dental Sciences and other participating institutes. Written informed consent was obtained from every subject.

Exome sequencing and data analysis

Total DNA was obtained from peripheral blood of the patient using standard protocols. Total DNA (3 μg) was sheared into approximately 300-bp fragments using a Covaris sonicator (Covaris, Woburn, MA, USA). A paired-end exome library for Illumina sequencing was prepared using the SureSelect Human All Exon Kit (Agilent) following the manufacturer's instructions. Massively parallel sequencing was performed using one lane of the Genome Analyzer II (Illumina) at RIKEN Omics Science Center by the Life Science Accelerator system. Base calling was performed by the Illumina pipeline with default parameters. Obtained reads were mapped against the human reference genome (UCSC hg18/GRCh36) using CLC Genomics Workbench v4.0.2 software (CLC Bio, Aarhus, Denmark) with default parameters. Variant calling was performed using the SNP and DIP detection tools in CLC Genomics Workbench v4.0.2 with default parameters. Analysis of ROH was performed using PLINK software v1.0.7 [30]. The primary aim of this

analysis was not to evaluate ROH segments precisely, but to narrow down the list of candidate variants without overlooking the causative variant. Therefore, we used relatively small (1, 000 kb) sliding windows for ROH segments, did not consider local blocks of linkage disequilibrium in the Japanese population, and did not exclude the data of variants whose frequency was not registered in dbSNP; those variants might not be polymorphic in the Japanese population and possibly contributed to extend the length of ROH. Conservation information for the variants among 44 vertebrate species (phyloP score) was collected from the UCSC genome browser [31].

Sanger sequencing

Sanger sequencing of PCR amplicons was performed to confirm the detected disease-associated variant using a 3730 × L DNA Analyser (Applied Biosystems, Foster City, CA, USA). The primers used were: forward, 5'-AGGCAGACAGGCTCTCAAAC-3'; reverse, 5'-GGCAGAATTAGATGCCATTG-3'.

Real-time quantitative PCR

The amount of nuclear DNA and mtDNA in the skeletal muscle of the patient and four age- and sex-matched controls (all males aged 39 to 48 years) was evaluated by real-time quantitative PCR analysis according to the previously validated methods [32]. Briefly, copy numbers of *RNaseP* (for nuclear DNA), *ND1* and *ND4* (for mtDNA) were evaluated using the TaqMan method (Applied Biosystems). Analysis of the patient's tissue was performed in two independent reactions, and each experiment was triplicated. *ND1*/*RNaseP* and *ND4*/*RNaseP* ratios were calculated as $2^{[Ct(RNaseP)-Ct(\text{each gene})]}$.

Data accessibility

The sequence data from this study have been submitted to dbGaP [33] (study accession [phs000392.v1.p1]).

Abbreviations

adPEO: autosomal dominant progressive external ophthalmoplegia; arPEO: autosomal recessive progressive external ophthalmoplegia; mtDNA: mitochondrial DNA; PEO: progressive external ophthalmoplegia; ROH: runs of homozygosity.

Acknowledgements

We would like to thank Dr Yu-ichi Goto (Department of Clinical Laboratory, National Center Hospital for Mental, Nervous and Muscular Disorders, National Center of Neurology and Psychiatry, Tokyo, Japan), who kindly provided us with control muscle DNA samples from subjects without muscle disease. We are grateful to the Support Unit for Bio-material Analysis (Research Resources Center, RIKEN BSI) and LSA System Development Unit (Omics Science Center, RIKEN Yokohama Institute) for performing exome library construction and massively parallel sequencing.

Author details

[1]Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan. [2]Department of Neuropsychiatry, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan. [3]Laboratory for Molecular Psychiatry, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan. [4]Department of Psychiatry, Kagoshima University Graduate School of Medical and Dental Sciences, 8-35-1 Sakuragaoka, Kagoshima-shi, Kagoshima 890-8520, Japan.

Authors' contributions

AT and TK designed the study and drafted the manuscript. AT performed data analysis and molecular experiments. MK, MN and AS performed clinical assessment. MK, MN, TY and AS provided materials for experiments. TY, SK, AS and TK coordinated the study and performed critical revision of the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 25 April 2011 Revised: 28 June 2011
Accepted: 28 September 2011 Published: 28 September 2011

References

1. Metzker ML: **Sequencing technologies - the next generation.** *Nat Rev* 2010, **11**:31-46.
2. Shendure J, Ji H: **Next-generation DNA sequencing.** *Nat Biotechnol* 2008, **26**:1135-1145.
3. Ng SB, Nickerson DA, Bamshad MJ, Shendure J: **Massively parallel sequencing and rare disease.** *Hum Mol Genet* 2010, **19**:R119-124.
4. Kato M, Nakamura M, Ichiba M, Tomiyasu A, Shimo H, Higuchi I, Ueno S, Sano A: **Mitochondrial DNA deletion mutations in patients with neuropsychiatric symptoms.** *Neurosci Res* 2010, **69**:331-336.
5. Zeviani M, Servidei S, Gellera C, Bertini E, DiMauro S, DiDonato S: **An autosomal dominant disorder with multiple deletions of mitochondrial DNA starting at the D-loop region.** *Nature* 1989, **339**:309-311.
6. Servidei S, Zeviani M, Manfredi G, Ricci E, Silvestri G, Bertini E, Gellera C, Di Mauro S, Di Donato S, Tonali P: **Dominantly inherited mitochondrial myopathy with multiple deletions of mitochondrial DNA: clinical, morphologic, and biochemical studies.** *Neurology* 1991, **41**:1053-1059.
7. Suomalainen A, Majander A, Haltia M, Somer H, Lonnqvist J, Savontaus ML, Peltonen L: **Multiple deletions of mitochondrial DNA in several tissues of a patient with severe retarded depression and familial progressive external ophthalmoplegia.** *J Clin Invest* 1992, **90**:61-66.
8. Melberg A, Lundberg PO, Henriksson KG, Olsson Y, Stalberg E: **Muscle-nerve involvement in autosomal dominant progressive external ophthalmoplegia with hypogonadism.** *Muscle Nerve* 1996, **19**:751-757.
9. Suomalainen A, Majander A, Wallin M, Setala K, Kontula K, Leinonen H, Salmi T, Paetau A, Haltia M, Valanne L, Lonnqvist J, Peltonen L, Somer H: **Autosomal dominant progressive external ophthalmoplegia with multiple deletions of mtDNA: clinical, biochemical, and molecular genetic features of the 10q-linked disease.** *Neurology* 1997, **48**:1244-1253.
10. Luoma P, Melberg A, Rinne JO, Kaukonen JA, Nupponen NN, Chalmers RM, Oldfors A, Rautakorpi I, Peltonen L, Majamaa K, Somer H, Suomalainen A: **Parkinsonism, premature menopause, and mitochondrial DNA polymerase gamma mutations: clinical and molecular genetic study.** *Lancet* 2004, **364**:875-882.
11. **The RefSeqGene Project..** [http://www.ncbi.nlm.nih.gov/projects/RefSeq/RSG/].
12. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, Shaffer T, Wong M, Bhattacharjee A, Eichler EE, Bamshad M, Nickerson DA, Shendure J: **Targeted capture and massively parallel sequencing of 12 human exomes.** *Nature* 2009, **461**:272-276.
13. Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, Gildersleeve HI, Beck AE, Tabor HK, Cooper GM, Mefford HC, Lee C, Turner EH, Smith JD, Rieder MJ, Yoshiura K, Matsumoto N, Ohta T, Niikawa N, Nickerson DA, Bamshad MJ, Shendure J: **Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome.** *Nat Genet* 2010, **42**:790-793.
14. Pagliarini DJ, Calvo SE, Chang B, Sheth SA, Vafai SB, Ong SE, Walford GA, Sugiana C, Boneh A, Chen WK, Hill DE, Vidal M, Evans JG, Thorburn DR, Carr SA, Mootha VK: **A mitochondrial protein compendium elucidates complex I disease biology.** *Cell* 2008, **134**:112-123.
15. **1000 Genomes Project..** [http://www.1000genomes.org/page.php].
16. He L, Chinnery PF, Durham SE, Blakely EL, Wardell TM, Borthwick GM, Taylor RW, Turnbull DM: **Detection and quantification of mitochondrial**

DNA deletions in individual cells by real-time PCR. *Nucleic Acids Res* 2002, **30**:e68.
17. Pontarin G, Fijolek A, Pizzo P, Ferraro P, Rampazzo C, Pozzan T, Thelander L, Reichard PA, Bianchi V: **Ribonucleotide reduction is a cytosolic process in mammalian cells independently of DNA damage.** *Proc Natl Acad Sci USA* 2008, **105**:17801-17806.
18. Bourdon A, Minai L, Serre V, Jais JP, Sarzi E, Aubert S, Chretien D, de Lonlay P, Paquis-Flucklinger V, Arakawa H, Nakamura Y, Munnich A, Rotig A: **Mutation of RRM2B, encoding p53-controlled ribonucleotide reductase (p53R2), causes severe mitochondrial DNA depletion.** *Nat Genet* 2007, **39**:776-780.
19. Acham-Roschitz B, Plecko B, Lindbichler F, Bittner R, Mache CJ, Sperl W, Mayr JA: **A novel mutation of the RRM2B gene in an infant with early fatal encephalomyopathy, central hypomyelination, and tubulopathy.** *Mol Genet Metab* 2009, **98**:300-304.
20. Kollberg G, Darin N, Benan K, Moslemi AR, Lindal S, Tulinius M, Oldfors A, Holme E: **A novel homozygous RRM2B missense mutation in association with severe mtDNA depletion.** *Neuromuscul Disord* 2009, **19**:147-150.
21. Spinazzola A, Invernizzi F, Carrara F, Lamantea E, Donati A, Dirocco M, Giordano I, Meznaric-Petrusa M, Baruffini E, Ferrero I, Zeviani M: **Clinical and molecular features of mitochondrial DNA depletion syndromes.** *J Inherit Metab Dis* 2009, **32**:143-158.
22. Bornstein B, Area E, Flanigan KM, Ganesh J, Jayakar P, Swoboda KJ, Coku J, Naini A, Shanske S, Tanji K, Hirano M, DiMauro S: **Mitochondrial DNA depletion syndrome due to mutations in the RRM2B gene.** *Neuromuscul Disord* 2008, **18**:453-459.
23. Shaibani A, Shchelochkov OA, Zhang S, Katsonis P, Lichtarge O, Wong LJ, Shinawi M: **Mitochondrial neurogastrointestinal encephalopathy due to mutations in RRM2B.** *Arch Neurol* 2009, **66**:1028-1032.
24. Tyynismaa H, Ylikallio E, Patel M, Molnar MJ, Haller RG, Suomalainen A: **A heterozygous truncating mutation in RRM2B causes autosomal-dominant progressive external ophthalmoplegia with multiple mtDNA deletions.** *Am J Hum Genet* 2009, **85**:290-295.
25. Pitceathly RD, Fassone E, Taanman JW, Sadowski M, Fratter C, Mudanohwo EE, Woodward CE, Sweeney MG, Holton JL, Hanna MG, Rahman S: **Kearns-Sayre syndrome caused by defective R1/p53R2 assembly.** *J Med Genet* 2011, **48**:610-617.
26. Fattal O, Link J, Quinn K, Cohen BH, Franco K: **Psychiatric comorbidity in 36 adults with mitochondrial cytopathies.** *CNS Spectrums* 2007, **12**:429-438.
27. Kato T: **Molecular neurobiology of bipolar disorder: a disease of 'mood-stabilizing neurons'?** *Trends Neurosci* 2008, **31**:495-503.
28. Kasahara T, Kubota M, Miyauchi T, Noda Y, Mouri A, Nabeshima T, Kato T: **Mice with neuron-specific accumulation of mitochondrial DNA mutations show mood disorder-like phenotypes.** *Mol Psychiatry* 2006, **11**:577-593, 523.
29. Xue L, Zhou B, Liu X, Heung Y, Chau J, Chu E, Li S, Jiang C, Un F, Yen Y: **Ribonucleotide reductase small subunit p53R2 facilitates p21 induction of G1 arrest under UV irradiation.** *Cancer Res* 2007, **67**:16-21.
30. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: **PLINK: a tool set for whole-genome association and population-based linkage analyses.** *Am J Hum Genet* 2007, **81**:559-575.
31. **Genome browser at UCSC..** [http://genome.ucsc.edu].
32. Kakiuchi C, Ishiwata M, Kametani M, Nelson C, Iwamoto K, Kato T: **Quantitative analysis of mitochondrial DNA deletions in the brains of patients with bipolar disorder and schizophrenia.** *Int J Neuropsychopharmacol* 2005, **8**:515-522.
33. **The Database of Genotypes and Phenotypes (dbGaP)..** [http://www.ncbi.nlm.nih.gov/gap].

doi:10.1186/gb-2011-12-9-r92
Cite this article as: Takata *et al.*: Exome sequencing identifies a novel missense variant in *RRM2B* associated with autosomal recessive progressive external ophthalmoplegia. *Genome Biology* 2011 **12**:R92.