(PPI) test, elevated plus maze test, Morris water maze test, passive avoidance learning test, active avoidance learning test, and forced swimming test.

Open-field test

A transparent cubic box without a top (30 × 30 × 30 cm) was used. A 40-W white lamp provided room lighting, which was approximately 110 lux on the floor of the chamber. A fan attached to the upper part of the wall at one end of the chamber presented a masking noise of 45 dB. Two infrared beams were set on each wall 2 cm above the floor at intervals of 10 cm. The total number of successive interceptions of the two adjoining beams on each bank was scored as locomotion behavior. For the rearing analysis, 12 more infrared ray beams attached 4.5 cm above the floor at 2.5-cm intervals were used. The total number of vertical beam interceptions was scored as the rearing behavior. Each mouse was allowed to explore freely in the open-field area for 20 min.

Startle response and PPI test

Using a transparent acrylic box (7 × 7 × 10 cm), the startle response was detected by an accelerometer (GH-313A, Keyence, Osaka, Japan) as vibration of the box. The acoustic startle pulses of a broadband burst (115 dB, 50 ms) and tone prepulse (85 dB, 30 ms) were presented by using a speaker located in front of the box. Prepulse using light (30 ms) was also applied by a light-emitting diode (LED); however, this was not used as data because it was found that the light prepulse did not attenuate the startle response.

At the beginning of the session, 40 startle pulses were presented to test for basal startle responsiveness and its habituation. The average values of eight blocks, consisting of five startle pulses each, were used for the statistical analysis. After that, three different types of trials were performed; that is, startle pulse alone (n = 12), startle pulse preceded by a tone prepulse (n = 12), and startle pulse preceded by a light prepulse (n = 12). Prepulses were presented 50, 100, or 200 ms before the startle pulse. In total, six types of prepulse (each n = 4) were applied. The mean interval averaged 25 s (15–45 s) throughout the session. The startle response was recorded for 200 ms with a sampling frequency of 1000 Hz. The PPI test was assessed by the ratio of the mean response of the trials with one type of prepulse (n = 4) divided by the mean response of the trials without prepulse (n = 12).

Elevated plus maze test

The maze consisted of four arms (two open arms and two closed arms), 5 cm wide and 30 cm long with a gray acrylic floor, which met at a 10 × 10 cm center zone. The two closed arms had transparent walls (15 cm in height) on both sides, and the open arms had low walls (3 mm in height) on both sides. The apparatus was mounted 75 cm above the floor of the room. The room lighting was approximately 20 lux on the maze. A video camera was placed 80 cm above the maze. A fan generated a masking noise of 45 dB. The animal was placed gently onto the center of the maze and was allowed to explore the maze freely for 10 min. The number of entries into each arm and the time spent in the open arms were videotape recorded.

Morris water maze test

A round pool, with a diameter of 95 cm and depth of 21.5 cm, was placed in the center of a 140 × 130 cm room. A platform with a diameter of 11 cm was set in one of the quadrants, 5 mm beneath the surface of black water maintained at 21 ± 1°C. On the first to fifth days, five trials per day were performed for the learning phase. Each mouse was released in one of the three quadrants of the pool without the platform, and the time to reach the platform was measured. If the mouse could not reach the platform within 60 s, the experimenter placed the mouse on the platform. On the sixth day, a probe test was performed to examine whether the mouse remembered the platform's location. The mouse was released in the quadrant on the opposite side of the platform, and its behavior for 60 s was videotaped. The time the mouse stayed in the target quadrant where the platform had been placed and the immobility time were measured.

Passive avoidance learning test

A mouse was placed in a box consisting of two rooms separated by a shutter; that is, light and dark compartments (each 10 × 10 cm). In the acquisition trial, the mouse was kept in the light compartment. Five seconds later, the door to the dark compartment was opened. When the mouse moved into the dark compartment the shutter was closed and 10 s later an electrical shock (160 V, 3 s) was delivered through the grid floor. Twenty-four hours later each mouse was placed again in the light compartment and the latency to enter the dark compartment was recorded up to a maximum of 180 s.

Active avoidance learning test

The same apparatus used for passive avoidance learning was also used in the active avoidance learning test; however, there was no shutter between the light and dark compartments. The box was set in a soundproof chamber and illuminated by a 20-W white light set on the chamber. The ceiling of the dark room was made of black acryl board and the ceiling of the light room was transparent acryl board.

The training was performed for 3 days. On each day, one session consisting of 50 trials was performed. In

each trial, a condition stimulus (CS) of a 1500-Hz sound (85 dB) was followed by an unconditioned stimulus (US) of a 140-V electrical shock. The US was given 5 s after initiation of the CS and continued until the mouse escaped to the other compartment. If the mouse did not move to the other compartment, the US lasted 15 s together with the CS. If the mouse moved within 5 s after the CS, the CS was stopped and no US was given. The time from the CS to the escape and the number of escapes were used to measure the learning performance.

Forced swimming test

The animals were placed in a square pool 24 × 24 cm in size. The water temperature was maintained at 21°C. On the first day, the mice were left in the pool for 20 min and the mobility recorded during the first 5 min by the video camera was assessed. On the second day the animals were placed in the pool for 5 min and the immobility time over 5 min was recorded. The immobility was defined by two criteria: 1) no movement of the legs and the tail and 2) a completely stationary state in the pool or movement only by inertia from previous movement. The immobility was assessed from the video by three independent raters. The median value of the three raters was used for the analysis.

Pharmacological experiments

Pharmacological experiments were performed at the RIKEN Brain Science Institute. For this analysis, *Adar2*$^{+/-}$ mice (n = 34) and the WT littermates (n = 30) were analyzed. The laboratory was air-conditioned and the temperature and humidity were maintained at approximately 22–23°C and 50–55%, respectively. Food and water were freely available except during experimentation. All of the experiments were conducted in the light phase (from 9:00 a.m. to 6:00 p.m.), and the starting times of the experiments were kept constant.

For analysis of locomotor activity, an apparatus for the open-field test equipped with a small soundproof room (185 × 185 × 225 cm) was used. Each field was made of white plastic (50 × 50 × 40 cm) and illuminated by LEDs (70 lux at the center of the field). The behavior of the mouse was monitored by a charge-coupled device camera placed on the ceiling of the rack for the open fields. The distance traveled (cm) was analyzed every 5 min. Data were collected and analyzed using an Image J OF4 (O'Hara & Company, Ltd., Tokyo, Japan), which is a modified software based on the public domain U.S. National Institutes of Health (NIH) image program (developed at the NIH and at http://rsb.info.nih.gov./nih-image/).

Amphetamine was dissolved in an aqueous 0.9% NaCl solution (1 mg/mL). The stock solution was further diluted with the 0.9% NaCl solution to achieve a dose of 1.5 mg/kg in 200 μL for injection. The locomotor activity of mice was observed for 50 min before subcutaneous administration of the amphetamine solution. In the first experiment, the fold change of the activity over 50 min after the administration of amphetamine or saline compared to the activity before treatment was used as the index.

NBQX (2,3-dioxo-6-nitro-1,2,3,4- tetrahydrobenzo[f] quinoxaline-7-sulfonamide) was dissolved in a saline solution (0.9% NaCl in water) (100 mg/mL). This stock solution was further diluted with the saline solution to achieve an dose of 10 mg/kg in 200 μL for intraperitoneal injection after a 50-min adaptation period. Fifty minutes after NBQX dosing, amphetamine or saline was administered subcutaneously. The fold change of the activity from 50 min after amphetamine treatment compared to the locomotor activity during the middle 50 min after the NBQX treatment was used as an index. These two experiments used independent mice. The first experiment consisted of 11 WT and 10 *Adar2*$^{+/-}$ mice and the second experiment consisted of 19 WT and 24 *Adar2*$^{+/-}$ mice.

Statistical analyses

For the statistical analysis, two-sample *t*-test, Pearson's coefficient of correlation, one-way ANOVA, two-way ANOVA, and RMANOVA were used. For statistical analysis of behavioral battery, at first RMANOVA with factors of time and genotype was applied. When a statistically significant interaction was detected, further post hoc analysis by *t*-test was applied. The statistical analyses were performed using IBM SPSS Statistics version 20 (IBM Corporation, New York).

Additional files

Additional file 1: Figure S1. No significant effect of NBQX on locomotor activity in both WT and *Adar2*$^{+/-}$ mice. Treatment with NBQX does not suppress a locomotor activity in both WT and *Adar2*$^{+/-}$ mice.

Additional file 2: Table S1. List of primers used in this study. Forward and reverse primer sets used in RNA editing analysis and RT-PCR for human and mouse samples.

Abbreviations

AMPA: 2-amino-3-(3-hydroxy-5-methyl-isoxazol-4-yl)-propanoic acid; *ADAR2*: Adenosine deaminase acting on RNA type 2; *HTR2C*: Serotonin 2C receptor; *GRIA2*: AMPA-type ionotropic glutamate receptor, subunit 2; *GRIA3*: AMPA-type ionotropic glutamate receptor, subunit 3; *GRIA4*: AMPA-type ionotropic glutamate receptor, subunit 4; *GRIK1*: kainate-type ionotropic glutamate receptor, subunit 1; *GRIK2*: kainate-type ionotropic glutamate receptor, subunit 2; PCR: Polymerase chain reaction.

Competing interests

The authors declare no conflict of interest regarding this work.

Authors' contribution

MKS performed behavioral and pharmacological analysis. KI and MB performed gene expression, RNA editing and behavioral analyses. TK organized the project, analyzed the data and wrote the paper. All authors read and approved the final manuscript.

Acknowledgements
The authors thank for Professor Peter H. Seeburg and Dr. Miyoko Higuchi for providing *Adar2* knockout mice. Postmortem brains were donated by the Stanley Foundation Brain Collection, courtesy of Dr. Michael B. Knable, Dr. E. Fuller Torrey, Dr. Maree J. Webster and Dr. Robert H. Yolken. We are grateful to Yoko Nakano and Taeko Miyauchi, and the staff members of the BSI Research Resource Center for technical assistance. This work was supported by grants to the Laboratory for Molecular Dynamics of Mental Disorders, RIKEN BSI, a Grant-in-Aid from the Japanese Ministry of Health and Labor and Grants-in-Aid from the Japanese Ministry of Education, Culture, Sports, Science and Technology. This study is also supported by CREST (for TK) and PRESTO (for KI) from Japan Science and Technology.

Author details
[1]Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, Wako, Saitama 351-0198, Japan. [2]Department of Molecular Psychiatry, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-8655, Japan.

Received: 15 October 2013 Accepted: 9 January 2014
Published: 21 January 2014

References

1. Charney D, Nestler S, Eric J (Eds): *Neurobiology of Mental Illness.* New York: Oxford University Press; 2011.
2. Javitt DC: **Glutamatergic theories of schizophrenia.** *Isr J Psychiatry Relat Sci* 2010, **47**:4–16.
3. Zarate CA Jr, Singh JB, Carlson PJ, Brutsche NE, Ameli R, Luckenbaugh DA, Charney DS, Manji HK: **A randomized trial of an N-methyl-D-aspartate antagonist in treatment-resistant major depression.** *Arch Gen Psychiatry* 2006, **63**:856–864.
4. Iwamoto K, Bundo M, Kato T: **Serotonin receptor 2C and mental disorders: genetic, expression and RNA editing studies.** *RNA Biol* 2009, **6**:248–253.
5. Barbon A, Barlati S: **Glutamate receptor RNA editing in health and disease.** *Biochemistry (Mosc)* 2011, **76**:882–889.
6. Silberberg G, Lundin D, Navon R, Ohman M: **Deregulation of the A-to-I RNA editing mechanism in psychiatric disorders.** *Hum Mol Genet* 2012, **21**:311–321.
7. Lomeli H, Mosbacher J, Melcher T, Hoger T, Geiger JR, Kuner T, Monyer H, Higuchi M, Bach A, Seeburg PH: **Control of kinetic properties of AMPA receptor channels by nuclear RNA editing.** *Science* 1994, **266**:1709–1713.
8. Dracheva S, Elhakem SL, Marcus SM, Siever LJ, McGurk SR, Haroutunian V: **RNA editing and alternative splicing of human serotonin 2C receptor in schizophrenia.** *J Neurochem* 2003, **87**:1402–1412.
9. Niswender CM, Herrick-Davis K, Dilley GE, Meltzer HY, Overholser JC, Stockmeier CA, Emeson RB, Sanders-Bush E: **RNA editing of the human serotonin 5-HT2C receptor. alterations in suicide and implications for serotonergic pharmacotherapy.** *Neuropsychopharmacology* 2001, **24**:478–491.
10. Dracheva S, Patel N, Woo DA, Marcus SM, Siever LJ, Haroutunian V: **Increased serotonin 2C receptor mRNA editing: a possible risk factor for suicide.** *Mol Psychiatry* 2008, **13**:1001–1010.
11. Zhu H, Urban DJ, Blashka J, McPheeters MT, Kroeze WK, Mieczkowski P, Overholser JC, Jurjus GJ, Dieter L, Mahajan GJ, Rajkowska G, Wang Z, Sullivan PF, Stockmeier CA, Roth BL: **Quantitative analysis of focused a-to-I RNA editing sites by ultra-high-throughput sequencing in psychiatric disorders.** *PLoS One* 2012, **7**:e43227.
12. Iwamoto K, Kato T: **RNA editing of serotonin 2C receptor in human postmortem brains of major mental disorders.** *Neurosci Lett* 2003, **346**:169–172.
13. Sodhi MS, Burnet PW, Makoff AJ, Kerwin RW, Harrison PJ: **RNA editing of the 5-HT(2C) receptor is reduced in schizophrenia.** *Mol Psychiatry* 2001, **6**:373–379.
14. Gurevich I, Tamir H, Arango V, Dwork AJ, Mann JJ, Schmauss C: **Altered editing of serotonin 2C receptor pre-mRNA in the prefrontal cortex of depressed suicide victims.** *Neuron* 2002, **34**:349–356.
15. Lyddon R, Dwork AJ, Keddache M, Siever LJ, Dracheva S: **Serotonin 2c receptor RNA editing in major depression and suicide.** *World J Biol Psychiatry* 2012, **14**:590–601.
16. Akbarian S, Smith MA, Jones EG: **Editing for an AMPA receptor subunit RNA in prefrontal cortex and striatum in Alzheimer's disease, Huntington's disease and schizophrenia.** *Brain Res* 1995, **699**:297–304.
17. Lyddon R, Navarrett S, Dracheva S: **Ionotropic glutamate receptor mRNA editing in the prefrontal cortex: no alterations in schizophrenia or bipolar disorder.** *J Psychiatry Neurosci* 2012, **37**:267–272.
18. Sanacora G, Treccani G, Popoli M: **Towards a glutamate hypothesis of depression: an emerging frontier of neuropsychopharmacology for mood disorders.** *Neuropharmacology* 2012, **62**:63–77.
19. Tomita H, Vawter MP, Walsh DM, Evans SJ, Choudary PV, Li J, Overman KM, Atz ME, Myers RM, Jones EG, Watson SJ, Akil H, Bunney WE Jr: **Effect of agonal and postmortem factors on gene expression profile: quality control in microarray analyses of postmortem human brain.** *Biol Psychiatry* 2004, **55**:346–352.
20. Iwamoto K, Bundo M, Kato T: **Altered expression of mitochondria-related genes in postmortem brains of patients with bipolar disorder or schizophrenia, as revealed by large-scale DNA microarray analysis.** *Hum Mol Genet* 2005, **14**:241–253.
21. Higuchi M, Maas S, Single FN, Hartner J, Rozov A, Burnashev N, Feldmeyer D, Sprengel R, Seeburg PH: **Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2.** *Nature* 2000, **406**:78–81.
22. Wolf ME: **The role of excitatory amino acids in behavioral sensitization to psychomotor stimulants.** *Prog Neurobiol* 1998, **54**:679–720.
23. Reid MS, Hsu K Jr, Berger SP: **Cocaine and amphetamine preferentially stimulate glutamate release in the limbic system: studies on the involvement of dopamine.** *Synapse* 1997, **27**:95–105.
24. Li Y, Vartanian AJ, White FJ, Xue CJ, Wolf ME: **Effects of the AMPA receptor antagonist NBQX on the development and expression of behavioral sensitization to cocaine and amphetamine.** *Psychopharmacology (Berl)* 1997, **134**:266–276.
25. Vanover KE: **Effects of AMPA receptor antagonists on dopamine-mediated behaviors in mice.** *Psychopharmacology (Berl)* 1998, **136**:123–131.
26. Simmons M, Meador-Woodruff JH, Sodhi MS: **Increased cortical expression of an RNA editing enzyme occurs in major depressive suicide victims.** *Neuroreport* 2010, **21**:993–997.
27. Kawahara Y, Ito K, Ito M, Tsuji S, Kwak S: **Novel splice variants of human ADAR2 mRNA: skipping of the exon encoding the dsRNA-binding domains, and multiple C-terminal splice sites.** *Gene* 2005, **363**:193–201.
28. Tariq A, Garncarz W, Handl C, Balik A, Pusch O, Jantsch MF: **RNA-interacting proteins act as site-specific repressors of ADAR2-mediated RNA editing and fluctuate upon neuronal stimulation.** *Nucleic Acids Res* 2013, **41**:2581–2593.
29. Kubota M, Kasahara T, Iwamoto K, Komori A, Ishiwata M, Miyauchi T, Kato T: **Therapeutic implications of down-regulation of cyclophilin D in bipolar disorder.** *Int J Neuropsychopharmacol* 2010, **13**:1355–1368.
30. Li M, Wang IX, Li Y, Bruzel A, Richards AL, Toung JM, Cheung VG: **Widespread RNA and DNA sequence differences in the human transcriptome.** *Science* 2011, **333**:53–58.
31. Ju YS, Kim JI, Kim S, Hong D, Park H, Shin JY, Lee S, Lee WC, Yu SB, Park SS, Seo SH, Yun JY, Kim HJ, Lee DS, Yavartanoo M, Kang HP, Gokcumen O, Govindaraju DR, Jung JH, Chong H, Yang KS, Kim H, Lee C, Seo JS: **Extensive genomic and transcriptional diversity identified through massively parallel DNA and RNA sequencing of eighteen Korean individuals.** *Nat Genet* 2011, **43**:745–752.
32. Bahn JH, Lee JH, Li G, Greer C, Peng G, Xiao X: **Accurate identification of A-to-I RNA editing in human by transcriptome sequencing.** *Genome Res* 2012, **22**:142–150.
33. Peng Z, Cheng Y, Tan BC, Kang L, Tian Z, Zhu Y, Zhang W, Liang Y, Hu X, Tan X, Guo J, Dong Z, Bao L, Wang J: **Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome.** *Nat Biotechnol* 2012, **30**:253–260.
34. Ramaswami G, Lin W, Piskol R, Tan MH, Davis C, Li JB: **Accurate identification of human Alu and non-Alu RNA editing sites.** *Nat Methods* 2012, **9**:579–581.
35. Schrider DR, Gout JF, Hahn MW: **Very few RNA and DNA sequence differences in the human transcriptome.** *PLoS One* 2011, **6**:e25842.
36. Picardi E, Gallo A, Galeano F, Tomaselli S, Pesole G: **A novel computational strategy to identify A-to-I RNA editing sites by RNA-Seq data: de novo detection in human spinal cord tissue.** *PLoS One* 2012, **7**:e44184.
37. Horsch M, Seeburg PH, Adler T, Aguilar-Pimentel JA, Becker L, Calzada-Wack J, Garrett L, Gotz A, Hans W, Higuchi M, Holter SM, Naton B, Prehn C, Puk O, Racz I, Rathkolb B, Rozman J, Schrewe A, Adamski J, Busch DH, Esposito I, Graw J, Ivandic B, Klingenspor M, Klopstock T, Mempel M, Ollert M, Schulz H, Wolf E, Wurst W, *et al*: **Requirement of the RNA-editing enzyme ADAR2 for normal physiology in mice.** *J Biol Chem* 2011, **286**:18614–18622.

38. Singh M, Zimmerman MB, Beltz TG, Johnson AK: **Affect-related behaviors in mice misexpressing the RNA editing enzyme ADAR2.** *Physiol Behav* 2009, **97**:446–454.
39. Singh M, Singh MM, Na E, Agassandian K, Zimmerman MB, Johnson AK: **Altered ADAR 2 equilibrium and 5HT(2C) R editing in the prefrontal cortex of ADAR 2 transgenic mice.** *Genes Brain Behav* 2011, **10**:637–647.
40. Torrey EF, Webster M, Knable M, Johnston N, Yolken RH: **The stanley foundation brain collection and neuropathology consortium.** *Schizophr Res* 2000, **44**:151–155.
41. Association AP: *Diagnostic and Statistical Manual of Mental Disorders.* 4th edition. Arlington, VA: American Psychiatric Publishing; 1994.
42. Iwamoto K, Kato T: **Effects of cocaine and reserpine administration on RNA editing of rat 5-HT2C receptor estimated by primer extension combined with denaturing high-performance liquid chromatography.** *Pharmacogenomics J* 2002, **2**:335–340.
43. Iwamoto K, Bundo M, Kato T: **Estimating RNA editing efficiency of five editing sites in the serotonin 2C receptor by pyrosequencing.** *RNA* 2005, **11**:1596–1603.

doi:10.1186/1756-6606-7-5
Cite this article as: Kubota-Sakashita *et al.*: A role of *ADAR2* and RNA editing of glutamate receptors in mood disorders and schizophrenia. *Molecular Brain* 2014 **7**:5.

Neuron
Report

Increased L1 Retrotransposition in the Neuronal Genome in Schizophrenia

Miki Bundo,[1,2] Manabu Toyoshima,[3] Yohei Okada,[4] Wado Akamatsu,[4] Junko Ueda,[2] Taeko Nemoto-Miyauchi,[2] Fumiko Sunaga,[1] Michihiro Toritsuka,[5] Daisuke Ikawa,[5] Akiyoshi Kakita,[6] Motoichiro Kato,[7] Kiyoto Kasai,[8] Toshifumi Kishimoto,[5] Hiroyuki Nawa,[9] Hideyuki Okano,[4] Takeo Yoshikawa,[3] Tadafumi Kato,[2,*] and Kazuya Iwamoto[1,10,*]
[1]Department of Molecular Psychiatry, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan
[2]Laboratory for Molecular Dynamics of Mental Disorders
[3]Laboratory for Molecular Psychiatry
RIKEN Brain Science Institute, Saitama 351-0198, Japan
[4]Department of Physiology, Keio University School of Medicine, Tokyo 160-8582, Japan
[5]Department of Psychiatry, Nara Medical University, Nara 634-8521, Japan
[6]Department of Pathology, Brain Research Institute, Niigata University, Niigata 951-8585, Japan
[7]Department of Neuropsychiatry, Keio University School of Medicine, Tokyo 160-8582, Japan
[8]Department of Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan
[9]Department of Molecular Neurobiology, Brain Research Institute, Niigata University, Niigata 951-8585, Japan
[10]PRESTO, Japan Science and Technology Agency, Saitama 332-0012, Japan
*Correspondence: kato@brain.riken.jp (T.K.), kaziwamoto-tky@umin.ac.jp (K.I.)
http://dx.doi.org/10.1016/j.neuron.2013.10.053

SUMMARY

Recent studies indicate that long interspersed nuclear element-1 (L1) are mobilized in the genome of human neural progenitor cells and enhanced in Rett syndrome and ataxia telangiectasia. However, whether aberrant L1 retrotransposition occurs in mental disorders is unknown. Here, we report high L1 copy number in schizophrenia. Increased L1 was demonstrated in neurons from prefrontal cortex of patients and in induced pluripotent stem (iPS) cell-derived neurons containing 22q11 deletions. Whole-genome sequencing revealed brain-specific L1 insertion in patients localized preferentially to synapse- and schizophrenia-related genes. To study the mechanism of L1 transposition, we examined perinatal environmental risk factors for schizophrenia in animal models and observed an increased L1 copy number after immune activation by poly-I:C or epidermal growth factor. These findings suggest that hyperactive retrotransposition of L1 in neurons triggered by environmental and/or genetic risk factors may contribute to the susceptibility and pathophysiology of schizophrenia.

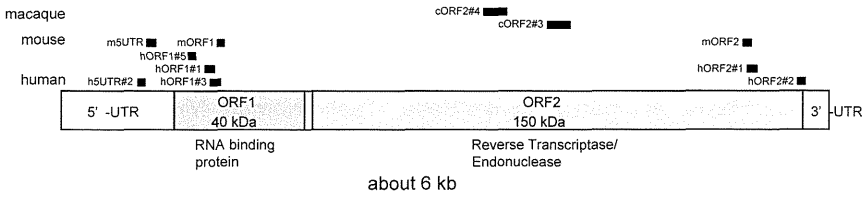
INTRODUCTION

Mental disorders including schizophrenia, bipolar disorder, and major depression affect a large proportion of the global population and have a major negative economic impact. Twin, family, and adoption studies indicate the complex involvement of both genetic and environmental factors for these diseases (Keshavan et al., 2011). Despite their apparent heritability, however, causative genetic factors are mostly unknown except for rare cases of schizophrenia associated with chromosomal abnormalities (Brandon and Sawa, 2011; Cook and Scherer, 2008; Karayiorgou et al., 2010). On the other hand, environmental risk factors including prenatal infection (Brown, 2006) and obstetric complications, such as neonatal hypoxia, embryonic ischemia, and gestational toxicosis (Lewis and Murray, 1987), are well-established risk factors for schizophrenia. However, it is not clarified how these environmental risk factors interact with genomic factors.

Accumulating evidence indicates that genomic DNA in the brain contains distinctive somatic genetic variations compared with nonbrain tissues (Poduri et al., 2013). These genetic signatures include brain-specific somatic mutations (Poduri et al., 2013), chromosomal aneuploidy (Rehen et al., 2005; Yurov et al., 2007), chromosomal microdeletion (Shibata et al., 2012), and the genome dynamics of nonlong terminal repeat (LTR) retrotransposons (Baillie et al., 2011; Evrony et al., 2012; Muotri and Gage, 2006). These observed somatic variations are hypothesized to contribute to the generation of functionally diversified brain cells (Muotri and Gage, 2006).

Among the known retrotransposons, only long interspersed nucleotide element-1 (L1) has autonomous retrotransposition activity. Full-length L1 elements include a 5′ UTR, two open reading frames (ORFs), and a 3′ UTR (Figure 1A). Encoded products from the ORFs contain activities required for retrotransposition and are employed in the insertion of new L1 copies as well as nonautonomous retrotransposons such as *Alu* and SVA (Cordaux and Batzer, 2009). Recent studies indicate that engineered L1 has retrotransposition activity in neural progenitor cells from rat hippocampus (Muotri et al., 2005), human fetal brain (Coufal et al., 2009), and human embryonic stem cells (Coufal et al., 2009). These in vitro findings were confirmed in human L1 transgenic mice in vivo (Muotri et al., 2005). Adult human brain cells also showed increased L1 copy number compared with nonbrain tissues (Coufal et al., 2009). Moreover, retrotransposition

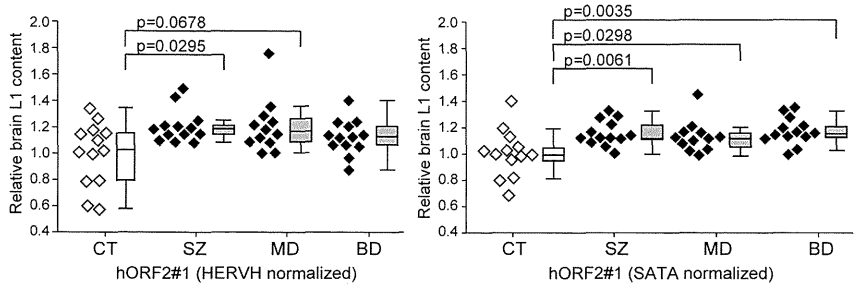
A



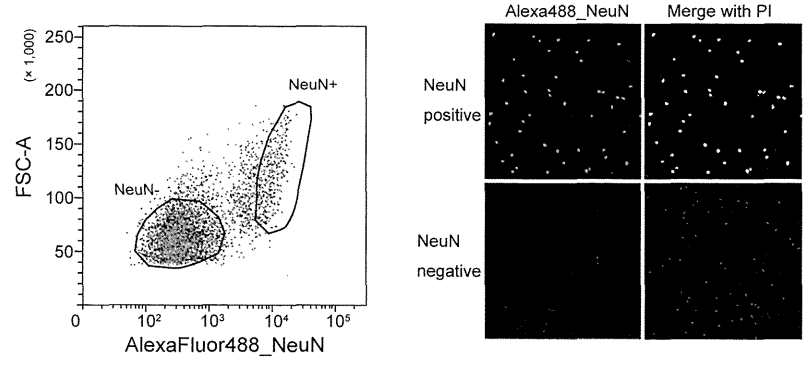
B

set	diagnosis	n	gender (F:M)	age	onset (yrs)	suicide	side (R:L)	pH	PMI (hrs)
I	CT	13	5:8	48.2±10.4	-	0	5:8	6.2±0.2	23.6±10.7
I	SZ	13	5:8	44.4±12.9	24.2±8.1	3	6:7	6.2±0.3	34.7±14.6
I	MD	12	5:7	45.2±10.0	33.5±11.8	6	3:9	6.1±0.2	27.8±10.5
I	BD	13	5:8	41.5±11.2	21.9±8.9	7	7:6	6.1±0.2	31.2±15.2
II	CT	34	8:26	44.5±7.5	-	0	18:16	6.6±0.3	29.5±13.0
II	SZ	35	9:26	42.6±8.5	21.3±6.1	7	18:17	6.5±0.2	31.4±15.3

C



D



E

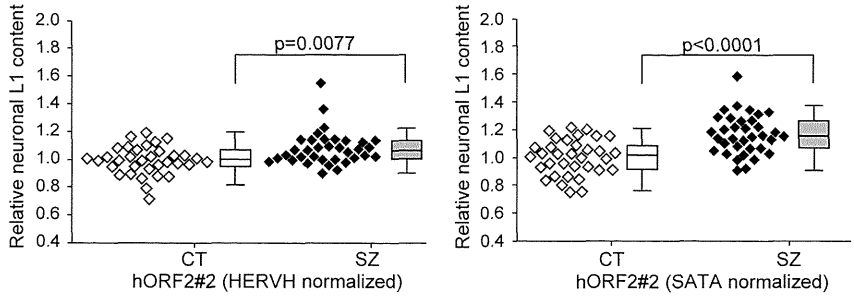


Figure 1. Increase of Brain L1 Copy Number in Schizophrenia

(A) Structure of L1 and map of the primers. Primers and probes are from previous studies (Coufal et al., 2009; Muotri et al., 2010) or designed for this study (Table S4). (B) Summary of the demographic variables of brain samples. (C) L1 copy number in set I. (D) Neuronal nuclei isolation. Left: example of NeuN-based nuclei sorting of brain cells from a patient with schizophrenia. Right: microscopic confirmation of isolated nuclei. The purity of each fraction was >95% and 99.9% for NeuN+ and NeuN−, respectively. (E) Neuronal L1 copy number in set II. In quantitative real-time PCR, L1 copy number was measured with HERVH or SATA as internal controls. The ratio of prefrontal cortex to liver (for set I) or neurons to nonneurons (for set II) was calculated and then normalized relative to the average value of control samples. Values were represented as open or closed diamonds as well as box plots. The ΔCt values of L1 and control probes were not significantly different between diagnostic groups in set I or set II. p values were determined by the Mann-Whitney U test. PMI, postmortem interval; CT, controls; SZ, schizophrenia; MD, major depression; BD, bipolar disorder; PI, propidium iodide. See also Tables S1 and S4 and Figures S1 and S2.

is active in MeCP2 mouse models and patients with Rett syndrome, indicating a role for this mechanism in this Mendelian disorder (Muotri et al., 2010). Together, these findings suggest the hypothesis that L1 retrotransposition may also be involved in the pathophysiology of mental disorders.

In this study, we quantified L1 copy number in genomic DNA derived from postmortem brains of patients with major mental disorders. We report significant increases of L1 content in the prefrontal cortex of patients with schizophrenia. To confirm this finding, we quantified L1 copy number in neurons and nonneurons from a second, independent patient cohort using NeuN-based cell sorting (Iwamoto et al., 2011; Rehen et al., 2005; Spalding et al., 2005) and found that L1 copy number in neurons was increased in patients with schizophrenia. We next quantified L1 copy number in the animal models that are known to disturb early neural development. These included maternal polyriboinosinic-polyribocytidilic acid (poly-I:C) injection in mice (Meyer and Feldon, 2012; Giovanoli et al., 2013) and chronic epidermal growth factor (EGF) injection to infant macaques (Nawa et al., 2000). We found that genomic DNA of brains from both animal models showed increased L1 copy number, addressing the importance of environmental factors during perinatal and postnatal stages. We also found that the increased L1 copy number

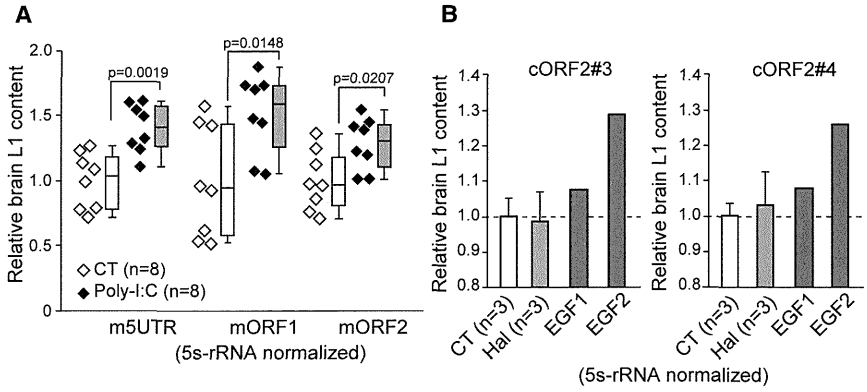


Figure 2. Increase of Brain L1 Copy Number in Animal Models

(A) Brain L1 content in the maternal poly-I:C model. p values were determined by the Mann-Whitney U test. Values were represented as open or closed diamonds as well as box plots. (B) Brain L1 content in chronic EGF or haloperidol-treated macaque models. Error bars indicate SDs. The comparative Ct method, with 5S-rRNA as an internal control, was used. The ratio of prefrontal cortex to liver (for poly-I:C model) or prefrontal gray matter to NeuN-sorted nonneurons in white matter (for macaque models) was calculated and then normalized relative to the average value of control samples. See also Table S4.

in the neurons derived from induced pluripotent stem (iPS) cells of schizophrenia patients with 22q11 deletion. The 22q11 deletion is a well-defined genetic factor and is one of the highest risk factors for schizophrenia, affecting about 1%–2% of schizophrenia patients (Karayiorgou et al., 2010). Finally, we performed whole-genome sequencing (WGS) analysis of brain and liver in controls and patients. Comparison of brain-specific L1 insertion sites revealed that brain-specific L1 insertion in patients is enriched in or near genes related to synaptic function and neuropsychiatric diseases. These results suggest that increased retrotransposition of L1 in neurons, which was triggered by genetic component and/or environmental factors at the early neural development, could contribute to the susceptibility and pathophysiology of schizophrenia.

RESULTS

Increased Brain L1 Content in Schizophrenia

We used postmortem prefrontal cortex samples of patients with schizophrenia, bipolar disorder, and major depression as well as control subjects for analysis in set I. The demographic variables are summarized in Figure 1B. We quantified L1 copy number of postmortem prefrontal cortex and liver in each subject by quantitative RT-PCR with two different internal controls, which were designed for human endogenous retrovirus (HERVH) and alpha-satellite (SATA). We found a significant increase in the brain L1ORF2 content in patients with schizophrenia (Figure 1C). A tendency toward copy number increase was also observed in mood disorders and in other L1 probes in schizophrenia (Figure S1 available online).

Somatic L1 retrotransposition was primarily found in neuronal cells (Kuwabara et al., 2009). To confirm the increased brain L1 copy number in schizophrenia and address whether this copy number increase is due to alteration of the neuronal genome, we examined an independent prefrontal cortex sample set (set II). We separated neuronal and nonneuronal nuclei from frozen brains using NeuN-based cell sorting (Figure 1D) (Iwamoto et al., 2011). NeuN is expressed in vertebrate neurons, and its antibody can be used for labeling neuronal nuclei (Mullen et al., 1992). We quantified L1ORF2 copy number of genomic DNA derived from neurons (NeuN-positive nuclei) and nonneurons (NeuN-negative nuclei) and then calculated the neuron-to-nonneuron ratio. We found a significant increase of neuronal L1ORF2 content in schizophrenia in two different internal controls (Figure 1E). The copy number of the other L1 probes tested also showed significant increase in schizophrenia compared to controls in SATA-normalized data, and similar tendency toward copy number increase was observed in HERVH-normalized data (Figure S1 and data not shown).

Assessment of Confounding Factors

We assessed the effect of confounding factors on L1 content (Table S1). Among the demographic variables tested, sample pH showed a weak correlation with L1ORF2 content in set II but not in set I. Several variables also showed weak correlations, but none showed consistency across the different internal control probes or across the two different sample sets.

To consider the possible effect of antipsychotics, we examined L1 copy number in a human neuroblastoma cell line cultured with haloperidol or risperidone for 8 days. Both antipsychotics did not modify the L1 copy number at their low or high concentrations (Figure S2). Together with the fact that the lifetime intake of antipsychotics, which was estimated as fluphenazine milligram equivalents, did not correlate with L1 copy number in both brain sets (Table S1), medication status did not affect our results.

L1 Quantification in Animal Models

To assess the potential roles of environmental factors on increased L1 copy number, we employed two different animal models that mimic environmental risk factors that affect early neural development. They included maternal poly-I:C injection in mice and chronic EGF injection to neonatal macaques. The poly-I:C, which mimics viral double-stranded RNA, injection to pregnant mice induces elevated maternal immune activation, and the offspring is known to show schizophrenia-like behavioral alterations such as impairments of prepulse inhibition and social behavior at the later stage (Meyer and Feldon, 2012). Pregnant mice received a single intraperitoneal injection of poly-I:C. L1 copy number in the prefrontal cortex of offspring was tested at postnatal day 21. We found that significant elevation of L1 copy number at all the tested probes compared to controls (Figure 2A).

We then examined the L1 copy number in macaques treated with EGF during neonatal period. Perinatal and postnatal

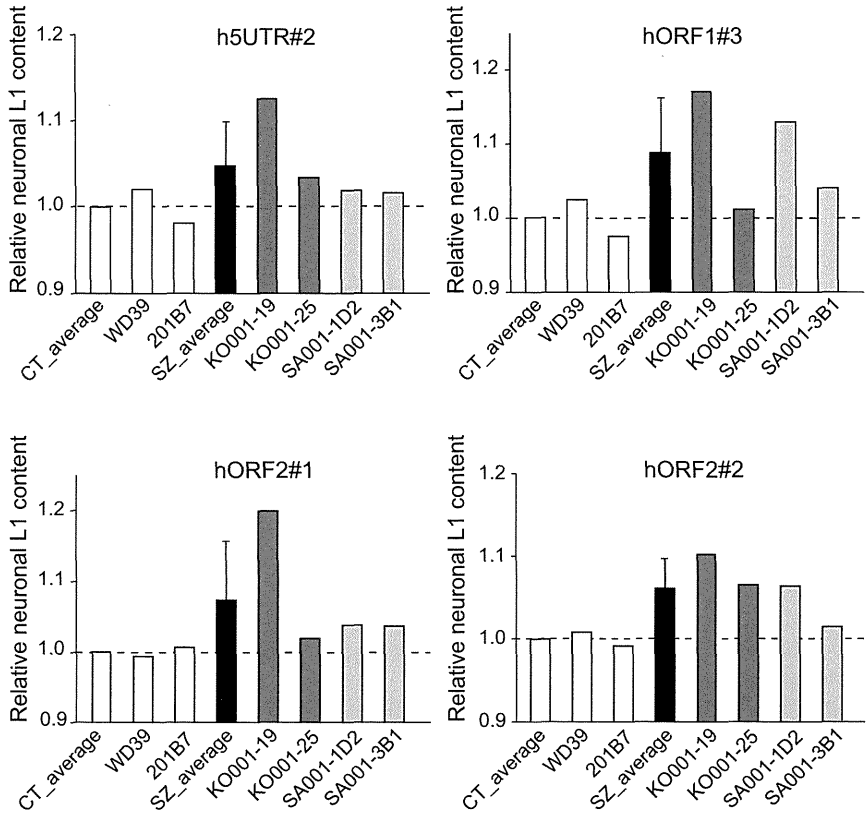


Figure 3. L1 Content in Neurons Derived from iPS Cells of Schizophrenia Patients with 22q11 Deletions
The comparative Ct method, with SATA as an internal control, was used. The ratio of NeuN-sorted neurons to nonneurons was calculated and then normalized relative to average value of control samples. Error bars indicate SDs. See also Table S4 and Figure S3.

perturbation of EGF is known to evoke schizophrenia-like phenotypes, including deficits in prepulse inhibition, latent inhibition, social interaction, and working memory, in adulthood (Nawa et al., 2009, 2000).The neonatal macaques (n = 2) subcutaneously received EGF for seven times over 11 days. After 4 and 7 years from treatment, L1 copy number in the prefrontal cortex was tested. In addition, chronic haloperidol-treated macaques (n = 3) were also tested. Due to unavailability of other tissues, we isolated nonneuronal nuclei from frozen white matter and calculated the grey matter-to-nonneuron ratio in each subject. Although statistical approach could not be applied, we observed increase of L1 copy number in EGF-treated macaques, but not in the haloperidol-treated macaques, compared to controls (Figure 2B). Taken together, these results suggest that early environmental factors play important roles in the L1 content in the brain. We further confirmed that chronic haloperidol treatment did not influence L1 copy number in this model.

L1 Quantification in the iPS Cells of Schizophrenia Patients with 22q11 Deletion

We next assessed the importance of genetic risk factor on the L1 copy number in brain. We quantified L1 copy number in the neurons derived from iPS cells of schizophrenia patients with 22q11 deletion (n = 2) as well as controls (n = 2) (Figure S3). The iPS cells were established from the fibroblasts according to the previously developed method (Imaizumi et al., 2012; Takahashi et al., 2007; M.T., unpublished data). To estimate the L1 copy number, we used two independently established iPS cell lines per patient. After induction of neuronal cells (Imaizumi et al., 2012), we isolated neuronal nuclei by NeuN-based sorting (Figure S3). We then examined L1 copy number and calculated the neuron-to-nonneuron ratio. Compared to controls, we observed consistent increase of L1 copy number in iPS cell-derived neurons of patients with schizophrenia with 22q11 deletion (Figure 3). These results suggest that the well-defined strong genetic risk factor also plays an important role in the L1 content in the brain.

Identification and Comparison of Brain-Specific L1 Transposition

We next performed WGS of brain and liver DNA from same subjects by self-assembling DNA nanoarray technology (Drmanac et al., 2010). For this experiment, schizophrenia patients (n = 3) and control subjects (n = 3) were selected to match age, PMI, gender, brain pH, and race from set I. Selected patients exhibited increased L1 content by quantitative RT-PCR assay, compared to average L1 content of the controls and selected control subjects. The WGS metrics and identified variations were summarized in Table S2. Distribution of the detected mobile elements was almost equal between the tissues and across subjects, and over the half of the identified elements was related

A

Control

insertion site	CT1	CT2	CT3	average ratio
intergenic (low)	0.64	0.76	0.69	0.69
intragenic (low)	0.36	0.24	0.31	0.31
intron (low)	0.90	1.00	0.99	0.96
exon (low)	0.10	0.00	0.01	0.04
intergenic (high)	0.65	0.81	0.69	0.72
intragenic (high)	0.35	0.19	0.31	0.28
intron (high)	0.88	1.00	0.99	0.96
exon (high)	0.12	0.00	0.01	0.04

Schizophrenia

insertion site	SZ1	SZ2	SZ3	average ratio
intergenic (low)	0.72	0.63	0.58	0.64
intragenic (low)	0.28	0.37	0.42	0.36
intron (low)	0.99	0.98	0.98	0.98
exon (low)	0.01	0.02	0.02	0.02
intergenic (high)	0.74	0.63	0.57	0.65
intragenic (high)	0.26	0.37	0.43	0.35
intron (high)	0.99	0.98	0.98	0.98
exon (high)	0.01	0.02	0.02	0.02

B

Control

Term	Count	p value
GO:0005856~cytoskeleton	74	5.92E-04
GO:0005509~calcium ion binding	56	0.0031
GO:0005930~axoneme	9	0.0095
GO:0035085~cilium axoneme	7	0.0289
GO:0003779~actin binding	26	0.0322
GO:0044425~membrane part	240	0.0387
GO:0016010~dystrophin-associated glycoprotein complex	6	0.0405

Schizophrenia

Term	Count	p value
GO:0045202~synapse	57	3.09E-09
GO:0030054~cell junction	64	8.57E-06
GO:0044459~plasma membrane part	187	1.49E-05
GO:0004674~protein serine/threonine kinase activity	58	2.27E-05
GO:0044456~synapse part	38	4.44E-05
GO:0004672~protein kinase activity	72	7.45E-05
GO:0030554~adenyl nucleotide binding	147	1.09E-04
GO:0005856~cytoskeleton	126	1.19E-04
GO:0005488~binding	820	1.25E-04
GO:0006468~protein amino acid phosphorylation	76	1.75E-04
GO:0006796~phosphate metabolic process	100	1.98E-04
GO:0006793~phosphorus metabolic process	100	1.98E-04
GO:0016773~phosphotransferase activity, alcohol group as acceptor	80	2.31E-04
GO:0001882~nucleoside binding	148	2.52E-04
GO:0005524~ATP binding	138	2.63E-04
GO:0001883~purine nucleoside binding	147	2.77E-04
GO:0032559~adenyl ribonucleotide binding	139	3.43E-04
GO:0017076~purine nucleotide binding	169	4.19E-04
GO:0000166~nucleotide binding	190	0.0011
GO:0014069~postsynaptic density	17	0.0011
GO:0032553~ribonucleotide binding	161	0.0012
GO:0032555~purine ribonucleotide binding	161	0.0012
GO:0043167~ion binding	323	0.0015
GO:0016043~cellular component organization	204	0.0018
GO:0043169~cation binding	318	0.0021
GO:0016310~phosphorylation	83	0.0021
GO:0046872~metal ion binding	315	0.0025
GO:0016301~kinase activity	85	0.0032
GO:0005737~cytoplasm	489	0.0037
GO:0008092~cytoskeletal protein binding	58	0.0045
GO:0007155~cell adhesion	74	0.0049
GO:0022610~biological adhesion	74	0.0052
GO:0019898~extrinsic to membrane	54	0.0053
GO:0043687~post-translational protein modification	108	0.0169
GO:0030030~cell projection organization	45	0.0219
GO:0005509~calcium ion binding	88	0.0234
GO:0015629~actin cytoskeleton	33	0.0439
GO:0016772~transferase activity, transferring phosphorus-containing groups	90	0.0473
GO:0045211~postsynaptic membrane	21	0.0487

C

Control

Term	Count	p value	FE
height	4	0.0132	7.7
scoliosis	3	0.0316	10.3

Schizophrenia

Term	Count	p value	FE
schizophrenia; schizoaffective disorder; bipolar disorder	5	0.0125	5.2
schizophrenia	29	0.0135	1.6
hypertension	20	0.0194	1.7
bipolar disorder	13	0.0373	1.9

Figure 4. Insertion Site, Gene Ontology, and Disease Association Analyses

(A) L1-insertion site analysis. Proportion of intergenic and intragenic L1 insertion and that of intronic and exonic L1 insertion are given. The low and high mean estimated proportions based on both less and stringent criteria are given. Note that ratios are not significantly different between patients and controls. (B) Gene ontology analysis. p values indicate Bonferroni-corrected modified Fisher's exact test p value. The terms showing $p < 0.05$ are shown for both groups. (C) Disease-association analysis. p values indicate noncorrected modified Fisher's exact test p value. FE, fold enrichment. In both analyses, gene lists generated by the stringent criteria were used. See also Tables S2 and S3 and Figure S4.

to the L1-Hs (Figure S4). Among the detected mobile element insertion sites in each sequenced sample, we first identified brain-specific L1 insertions in each subject (Tables S2 and S3). Although the total number of brain-specific L1 insertion tended to be higher in schizophrenia patients, this was not statistically significant, most likely due to the limited sample size and high interindividual variation. We then compared genomic locations of the insertion sites of brain-specific L1 between patients and controls (Figure 4A). The inter-to-intragenic L1 insertion ratio as well as exonic-to-intronic L1 insertion ratio did not differ between patients and controls. We then compared the affected genes by brain-specific L1 insertion by gene ontology approach. This analysis revealed that the number of enriched terms is higher in schizophrenia than controls, in spite that the number of brain-specific L1 insertions did not significantly differ. We found that neuronal function-related terms such as synapse and protein phosphorylation are clearly overrepresented in schizophrenia compared to controls (Figure 4B). In addition, disease-association analysis revealed that affected genes in patients are specifically enriched in terms related to schizophrenia and bipolar disorder, while those in controls are enriched in nonneuropsychiatric terms such as height and scoliosis (Figure 4C). These results were consistently confirmed when we used less stringent definition of brain-specific L1 insertion (Figure S4). In

addition, enrichment of the L1-inserted genes to the terms related to neuropsychiatric disorders in schizophrenia was also detected by the ingenuity pathway analysis (IPA) (Figure S4).

DISCUSSION

We report that the neuronal genome of schizophrenia contains higher copy number of a retrotransposon, L1. To validate this finding, we utilized iPS cells from patients with schizophrenia carrying the 22q11 deletion and observed an increase in L1 copy number in iPS cell-derived neurons. Moreover, using WGS, we found that L1 preferentially inserted into genes related to synaptic functions and schizophrenia. Animal model studies showed that environmental factors related to infection or inflammation that disturbs early neurodevelopmental processes increase L1 copy number in the brain. Collectively, these results suggest that hyperactive L1 retrotransposition into critical genes during neural development, triggered by genetic and/or environmental factors, contribute to the pathophysiology of schizophrenia. Our results significantly expand the range of neuropsychiatric illnesses linked to aberrant L1 retrotransposition, from Mendelian disease patients with *MECP2* mutations in Rett syndrome (Muotri et al., 2010) and *ATM* mutations in ataxia telangiectasia (Coufal et al., 2011) to schizophrenia, a complex mental disorder.

The observed increase of L1 content in schizophrenia was not due to, or modulated by, biological or experimental artifacts, because changes were measured in two independent patient cohorts and each result was confirmed with two different internal controls. Although the L1 region showing significant increases differed between the two brain sets, this is attributable to cohort differences amplified by the strict threshold we employed. Actually, a significant increase of L1 content was widely observed in all probes in the SATA-normalized data in set II, where neuronal L1 copy number was directly examined (Figure S1). In addition, from the data analysis utilizing lifetime intake of antipsychotics of patients, and from the cell culture and macaque experiments, we conclude that antipsychotics do not affect L1 copy number in the brain. A significant increase was also observed in patients with mood disorders in one internal control in set I (Figure 1C). Future work will clarify whether there are L1 content increases in other mental disorders using larger and/or stratified patient cohorts.

L1 retrotransposition has been detected during adult neurogenesis in the rat hippocampus, indicating that neural progenitor cells retain retrotransposition activity even in adult stages (Muotri et al., 2009). However, we analyzed potential confounding factors, including age, age of onset, and duration of illness, and did not observe any significant correlation with L1 copy number in the brain. The transcript level of L1 in adult brain sample was also increased in patients compared to controls (data not shown). However, elevated expression is unlikely to contribute to increase of L1 copy number in patients, as significant increase of L1 transcripts was detected only in the 5′ region of L1 such as 5′ UTR and ORF1. These results suggest that L1 copy number does not globally increase with aging and that the variation of L1 copy number in patients is probably confined to early neurodevelopmental stages, at least in the prefrontal cortex. This prediction would be consistent with the neurodevelopmental hypothesis of schizophrenia, where abnormalities during critical early periods of brain development may trigger the later appearance of clinical symptoms (Bloom, 1993; Murray et al., 1992; Weinberger, 1987).

In Rett syndrome, increased L1 copy number in human brain was linked to mutations in *MECP2* (Muotri et al., 2010) and *MeCP2* knockout mice also showed increased L1 content (Muotri et al., 2010). It has also been suggested that SOX2 and MECP2 regulate L1 transcription in neurons (Muotri et al., 2005; Yu et al., 2001). However, we did not observe a significant correlation between *MECP2* or *SOX2* expression and brain L1 content, by using the previously performed gene expression analyses on the same sample sets (Iwamoto et al., 2004, 2005) (data not shown). In addition, patients with high levels of L1 copy number (two schizophrenia and one major depression in set I, and two schizophrenia patients in set II) did not show altered *MECP2* or *SOX2* expression levels (data not shown). These findings suggest that the molecular mechanism of increased L1 in schizophrenia is different from Rett syndrome.

In this study, we found that both early environmental and well-defined strong genetic factors for schizophrenia are involved in the increase of L1 copy number in the brain. A recent study using the poly-I:C model indicated that the offspring of this model had exacerbated schizophrenia-like phenotypes, if they were exposed to environmental stress during puberty, suggesting that early environmental factors can lower the threshold for onset of schizophrenia (Giovanoli et al., 2013). Therefore, increased L1 insertions induced by environmental factors may increase the susceptibility to schizophrenia by disrupting synaptic and schizophrenia-related genes in neurons, rather than being a direct cause of the disease. On the other hand, the pathological consequences of increased L1 content in neurons derived from iPS cells of schizophrenia patients with 22q11 deletions remain unclear. We chose patients with 22q11 deletions to examine L1 dynamics where there is a well-defined strong genetic risk for schizophrenia. In *MeCP2*-knockout mice, Rett-like behavioral abnormalities could be rescued by the re-expression of wild-type MeCP2 at both young and adult stages (Cobb et al., 2010; Ehninger et al., 2008), suggesting that L1 content itself may not be directly causal to disease phenotypes but instead modulate phenotypic variability among patients (Muotri et al., 2010). Similarly, we speculate that the L1 increase in schizophrenia patients with 22q11 deletions is likely to modulate phenotypes of schizophrenia rather than a direct cause, because many genes related to schizophrenia, such as *TBX-1*, *SEPT5*, *COMT*, and *PRODH*, are located within the deletion (Hiroi et al., 2013; Karayiorgou et al., 2010). Nevertheless, our findings will facilitate further studies of the mechanism of increased L1 retrotransposition associated with schizophrenia.

Our WGS analysis could not detect increased brain-specific L1 insertions in schizophrenia; however, we found that L1 insertions were more frequent in genes for synaptic function and schizophrenia relative to controls. Evrony et al. cloned one L1 insertion event from 300 single neurons and showed that 2 of 83 cortical neurons from an individual had this insertion, but detection of such a low level mosaic insertion in bulk brain tissue of the same individual was difficult and needed optimization

(Evrony et al., 2012). Thus, rare L1 insertion events could be missed in our WGS analysis. Apart from L1, nonautonomous retrotransposons such as *Alu* and SVA also show an increased copy number in the brain, possibly via the aid of L1 ORF products (Baillie et al., 2011) and their copy number might also be increased in patients. Further studies on the neuronal genome of patients with mental disorders, and supporting mechanistic evidence from animal and cellular models, may establish a broader role for instability of neural genome in the pathophysiology of schizophrenia. We expect that our findings will promote the further study of genomic instability in disease etiology due to L1 retrotransposition in brain development.

EXPERIMENTAL PROCEDURES

Postmortem Samples

Postmortem brain and liver samples were obtained from the Stanley Medical Research Institute. The demographics are summarized in Figure 1B and are described at the web site (http://www.stanleyresearch.org/). Ethics committees of RIKEN and the University of Tokyo Faculty of Medicine approved the study.

Animal Models

Animal experiments were performed in accordance with the NIH Guidelines for the Care and Use of Laboratory Animals and guidelines of relevant facilities. For poly I:C model, pregnant mice (C57BL/6) received either a single intraperitoneal injection of poly-I:C (2 mg/ml, Sigma-Aldrich) dissolved in PBS (20 mg/kg) or an equivalent volume of PBS at embryonic day 9.5. At postnatal day 21, tissues were dissected from pups. For macaque models, cynomologus monkeys (*Macaca fascicularis*) (4 years old; all males) were given oral haloperidol (0.25–0.5 mg/kg; Wako Pure Chemical Industries) or vehicle for 2 months (Shibuya et al., 2010). After transiently separating two male monkey neonates (2 weeks old) from dams, neonates received subcutaneous administration of human recombinant EGF (0.3 mg/kg, Funakoshi) for seven times over 11 days and then quickly returned to their dams. Preliminary behavioral assessment of the EGF-treated monkeys was performed at ages of 4 and 6 years and reported (Nawa et al., 2009). These monkeys were sacrificed at the age of 4 and 7 years with the overdose of pentobarbital (26 mg/kg; 65 mg/ml). Experiments were subjected to review by the Ethical Committee of Shinn Nippon Biomedical Lab.

iPS Cells

All procedures for skin biopsy and iPS cell production were approved by the Keio University School of Medicine ethics committee and RIKEN ethics committee. The 201B7 iPS cells were kindly provided by Dr. Yamanaka (Takahashi et al., 2007). For the control WD39, a skin-punch biopsy from a healthy 16-year-old Japanese female obtained after written informed consent was used to generate iPS cells (Imaizumi et al., 2012). 22q11.2 deletion syndrome iPS cells (SA001 and KO001) were generated from a 37-year-old Japanese female patient (Toyosima et al., 2011) and a 30-year-old Japanese female patient, respectively, using the same method used to generate the WD39 (M.T., unpublished data). 22q11 deletion was characterized by the CGH array analysis (Figure S3). Production and maintenance of iPS cells were performed according to the previous studies (Imaizumi et al., 2012; Takahashi et al., 2007). All the iPS cells and differentiated neuronal cell lines were characterized with immunofluorescence staining and their morphologies (Figure S3).

L1 Copy Number Estimation

We performed either Taqman-based quantitative real-time PCR according to Coufal et al. (2009) with minor modifications (100 or 500 pg DNA as starting material and single amplicon analysis) or SYBR-Green-based quantitative real-time PCR according to Muotri et al. (2010). SYBR-Green assay was performed using 500 pg DNA and Power SYBR Green PCR Master Mix (Life Technologies). Primers, probe location, and reaction chemistry are listed in Figure 1A and Table S4. Quantification was performed in triplicate. A nonparametric Mann-Whitney U test was employed for two group comparison and $p < 0.05$ was considered significant.

Whole-Genome Sequencing

WGS of brain and liver samples from controls and schizophrenia patients was performed by Complete Genomics, with the paired-end library preparation and sequencing-by-ligation using self-assembling DNA nanoball (DNB) (Drmanac et al., 2010). Data process, mapping, and detection of variations were performed using the software developed by the Complete Genomics (version 2.2.0.26 and format version 2.2). Among the detected mobile insertion elements, we compared the genomic location of L1 insertion between brain and liver within an individual and identified brain-specific L1 insertions.

Further experimental details are available in the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2013.10.053.

ACKNOWLEDGMENTS

This work was supported in part by the Grant-in-Aid for Scientific Research on Innovative Areas (Unraveling the microendophenotypes of psychiatric disorders at the molecular, cellular, and circuit levels) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) to T.Y., H.N., T.K., and K.I., and a Grant-in-Aid from Ministry of Health, Labour and Welfare to T.K. This work was also supported by JST, CREST to T.K. and by JST, PRESTO to K.I. This work was also supported in part by Leading Project for Realization of Regenerative Medicine from MEXT and "Funding Program for World-Leading Innovative R&D on Science and Technology" to H.O., and by the "Development of biomarker candidates for social behavior" carried out under the Strategic Research Program for Brain Sciences from MEXT to T.Y. and K.K. This work was also supported in part by the Collaborative Research Project of the Brain Research Institute, Niigata University. Postmortem samples were donated by the Stanley Medical Research Institute, courtesy of Drs. Michael B. Knable, E. Fuller Torrey, Maree J. Webster, and Robert H. Yolken. We thank Tomoko Toyota and Atsuko Komori-Kokubo at RIKEN BSI for their technical assistance. We also thank Kenji Ohtawa at Research Resources Center at the RIKEN BSI for the cell-sorting analysis. M.B., F.S., and K.I. belong to the Department of Molecular Psychiatry, which is endowed by Dainippon Sumitomo Pharma and Yoshitomiyakuhin. H.O. is a scientific consultant for San Bio, Eisai, and Daiichi Sankyo. T.K. received a grant from Takeda Pharmaceutical. These companies had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Accepted: October 18, 2013
Published: January 2, 2014

REFERENCES

Baillie, J.K., Barnett, M.W., Upton, K.R., Gerhardt, D.J., Richmond, T.A., De Sapio, F., Brennan, P.M., Rizzu, P., Smith, S., Fell, M., et al. (2011). Somatic retrotransposition alters the genetic landscape of the human brain. Nature *479*, 534–537.

Bloom, F.E. (1993). Advancing a neurodevelopmental origin for schizophrenia. Arch. Gen. Psychiatry *50*, 224–227.

Brandon, N.J., and Sawa, A. (2011). Linking neurodevelopmental and synaptic theories of mental illness through DISC1. Nat. Rev. Neurosci. *12*, 707–722.

Brown, A.S. (2006). Prenatal infection as a risk factor for schizophrenia. Schizophr. Bull. *32*, 200–202.

Cobb, S., Guy, J., and Bird, A. (2010). Reversibility of functional deficits in experimental models of Rett syndrome. Biochem. Soc. Trans. *38*, 498–506.

Cook, E.H., Jr., and Scherer, S.W. (2008). Copy-number variations associated with neuropsychiatric conditions. Nature *455*, 919–923.

Cordaux, R., and Batzer, M.A. (2009). The impact of retrotransposons on human genome evolution. Nat. Rev. Genet. *10*, 691–703.

Coufal, N.G., Garcia-Perez, J.L., Peng, G.E., Yeo, G.W., Mu, Y., Lovci, M.T., Morell, M., O'Shea, K.S., Moran, J.V., and Gage, F.H. (2009). L1 retrotransposition in human neural progenitor cells. Nature *460*, 1127–1131.

Coufal, N.G., Garcia-Perez, J.L., Peng, G.E., Marchetto, M.C., Muotri, A.R., Mu, Y., Carson, C.T., Macia, A., Moran, J.V., and Gage, F.H. (2011). Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. Proc. Natl. Acad. Sci. USA *108*, 20382–20387.

Drmanac, R., Sparks, A.B., Callow, M.J., Halpern, A.L., Burns, N.L., Kermani, B.G., Carnevali, P., Nazarenko, I., Nilsen, G.B., Yeung, G., et al. (2010). Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. Science *327*, 78–81.

Ehninger, D., Li, W., Fox, K., Stryker, M.P., and Silva, A.J. (2008). Reversing neurodevelopmental disorders in adults. Neuron *60*, 950–960.

Evrony, G.D., Cai, X., Lee, E., Hills, L.B., Elhosary, P.C., Lehmann, H.S., Parker, J.J., Atabay, K.D., Gilmore, E.C., Poduri, A., et al. (2012). Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. Cell *151*, 483–496.

Giovanoli, S., Engler, H., Engler, A., Richetto, J., Voget, M., Willi, R., Winter, C., Riva, M.A., Mortensen, P.B., Schedlowski, M., and Meyer, U. (2013). Stress in puberty unmasks latent neuropathological consequences of prenatal immune activation in mice. Science *339*, 1095–1099.

Hiroi, N., Takahashi, T., Hishimoto, A., Izumi, T., Boku, S., and Hiramoto, T. (2013). Copy number variation at 22q11.2: from rare variants to common mechanisms of developmental neuropsychiatric disorders. Mol. Psychiatry *18*, 1153–1165.

Imaizumi, Y., Okada, Y., Akamatsu, W., Koike, M., Kuzumaki, N., Hayakawa, H., Nihira, T., Kobayashi, T., Ohyama, M., Sato, S., et al. (2012). Mitochondrial dysfunction associated with increased oxidative stress and α-synuclein accumulation in PARK2 iPSC-derived neurons and postmortem brain tissue. Mol. Brain *5*, 35.

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

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

Iwamoto, K., Bundo, M., Ueda, J., Oldham, M.C., Ukai, W., Hashimoto, E., Saito, T., Geschwind, D.H., and Kato, T. (2011). Neurons show distinctive DNA methylation profile and higher interindividual variations compared with non-neurons. Genome Res. *21*, 688–696.

Karayiorgou, M., Simon, T.J., and Gogos, J.A. (2010). 22q11.2 microdeletions: linking DNA structural variation to brain dysfunction and schizophrenia. Nat. Rev. Neurosci. *11*, 402–416.

Keshavan, M.S., Nasrallah, H.A., and Tandon, R. (2011). Schizophrenia, "Just the Facts" 6. Moving ahead with the schizophrenia concept: from the elephant to the mouse. Schizophr. Res. *127*, 3–13.

Kuwabara, T., Hsieh, J., Muotri, A., Yeo, G., Warashina, M., Lie, D.C., Moore, L., Nakashima, K., Asashima, M., and Gage, F.H. (2009). Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. Nat. Neurosci. *12*, 1097–1105.

Lewis, S.W., and Murray, R.M. (1987). Obstetric complications, neurodevelopmental deviance, and risk of schizophrenia. J. Psychiatr. Res. *21*, 413–421.

Meyer, U., and Feldon, J. (2012). To poly(I:C) or not to poly(I:C): advancing preclinical schizophrenia research through the use of prenatal immune activation models. Neuropharmacology *62*, 1308–1321.

Mullen, R.J., Buck, C.R., and Smith, A.M. (1992). NeuN, a neuronal specific nuclear protein in vertebrates. Development *116*, 201–211.

Muotri, A.R., and Gage, F.H. (2006). Generation of neuronal variability and complexity. Nature *441*, 1087–1093.

Muotri, A.R., Chu, V.T., Marchetto, M.C., Deng, W., Moran, J.V., and Gage, F.H. (2005). Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature *435*, 903–910.

Muotri, A.R., Zhao, C., Marchetto, M.C., and Gage, F.H. (2009). Environmental influence on L1 retrotransposons in the adult hippocampus. Hippocampus *19*, 1002–1007.

Muotri, A.R., Marchetto, M.C., Coufal, N.G., Oefner, R., Yeo, G., Nakashima, K., and Gage, F.H. (2010). L1 retrotransposition in neurons is modulated by MeCP2. Nature *468*, 443–446.

Murray, R.M., O'Callaghan, E., Castle, D.J., and Lewis, S.W. (1992). A neurodevelopmental approach to the classification of schizophrenia. Schizophr. Bull. *18*, 319–332.

Nawa, H., Takahashi, M., and Patterson, P.H. (2000). Cytokine and growth factor involvement in schizophrenia—support for the developmental model. Mol. Psychiatry *5*, 594–603.

Nawa, H., Someya, T., and Sakai, M. (2009). A novel schizophrenia model established by subcutaneously injecting a cytokine to a cynomolgus monkey neonate. Schizophr. Bull. *35* (*Suppl 1*), 252.

Poduri, A., Evrony, G.D., Cai, X., and Walsh, C.A. (2013). Somatic mutation, genomic variation, and neurological disease. Science *341*, 1237758.

Rehen, S.K., Yung, Y.C., McCreight, M.P., Kaushal, D., Yang, A.H., Almeida, B.S., Kingsbury, M.A., Cabral, K.M., McConnell, M.J., Anliker, B., et al. (2005). Constitutional aneuploidy in the normal human brain. J. Neurosci. *25*, 2176–2180.

Shibata, Y., Kumar, P., Layer, R., Willcox, S., Gagan, J.R., Griffith, J.D., and Dutta, A. (2012). Extrachromosomal microDNAs and chromosomal microdeletions in normal tissues. Science *336*, 82–86.

Shibuya, M., Komi, E., Wang, R., Kato, T., Watanabe, Y., Sakai, M., Ozaki, M., Someya, T., and Nawa, H. (2010). Measurement and comparison of serum neuregulin 1 immunoreactivity in control subjects and patients with schizophrenia: an influence of its genetic polymorphism. J. Neural Transm. *117*, 887–895.

Spalding, K.L., Bhardwaj, R.D., Buchholz, B.A., Druid, H., and Frisén, J. (2005). Retrospective birth dating of cells in humans. Cell *122*, 133–143.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–872.

Toyosima, M., Maekawa, M., Toyota, T., Iwayama, Y., Arai, M., Ichikawa, T., Miyashita, M., Arinami, T., Itokawa, M., and Yoshikawa, T. (2011). Schizophrenia with the 22q11.2 deletion and additional genetic defects: case history. Br. J. Psychiatry *199*, 245–246.

Weinberger, D.R. (1987). Implications of normal brain development for the pathogenesis of schizophrenia. Arch. Gen. Psychiatry *44*, 660–669.

Yu, F., Zingler, N., Schumann, G., and Strätling, W.H. (2001). Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. Nucleic Acids Res. *29*, 4493–4501.

Yurov, Y.B., Iourov, I.Y., Vorsanova, S.G., Liehr, T., Kolotii, A.D., Kutsev, S.I., Pellestor, F., Beresheva, A.K., Demidova, I.A., Kravets, V.S., et al. (2007). Aneuploidy and confined chromosomal mosaicism in the developing human brain. PLoS ONE *2*, e558.

Neuroscience Research 79 (2014) 22–33

Contents lists available at ScienceDirect

Neuroscience Research

journal homepage: www.elsevier.com/locate/neures

Comprehensive survey of CNVs influencing gene expression in the human brain and its implications for pathophysiology☆

Divya Mehta[a,b,1], Kazuya Iwamoto[c], Junko Ueda[b], Miki Bundo[c], Naoki Adati[d,2], Toshio Kojima[d,3], Tadafumi Kato[b,*]

[a] Max Planck Institute of Psychiatry, Munich 80804, Germany
[b] Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, Saitama 351-0198, Japan
[c] Department of Molecular Psychiatry, Graduate School of Medicine, University of Tokyo, Tokyo 113-8654, Japan
[d] Comparative Systems Biology Team, RIKEN Genomic Sciences Center, Yokohama 230-0045, Japan

ARTICLE INFO

Article history:
Received 8 July 2013
Received in revised form 22 October 2013
Accepted 29 October 2013
Available online 6 November 2013

Keywords:
Copy number variation (CNV)
Human post-mortem brains
Schizophrenia
Bipolar disorder
Transcriptome
Genomics

ABSTRACT

Copy number variations (CNVs) contribute to neuropsychiatric diseases, which may be partly mediated by their effects on gene expression. However, few studies have assessed the influence of CNVs on gene expression in the brain. The objective was to perform an unbiased comprehensive survey of influence of CNVs on gene expression in human brain tissues. CNV regions (CNVRs) were identified in 72 individuals (23 schizophrenia, 23 bipolar disorder and 26 controls). Significant associations between the CNVRs and gene expression levels were observed for 583 CNVR-expression probe pairs (293 unique eCNVRs and 429 unique transcripts), after corrections for multiple testing and controlling the effect of the number of subjects with CNVRs by label swapping permutations. These CNVRs affecting gene expression (eCNVRs) were significantly enriched for rare/low frequency ($p = 1.087 \times 10^{-10}$) and gene-harboring CNVRs ($p = 1.4 \times 10^{-6}$). Transcripts overlapping CNVRs were significantly enriched for glutathione metabolism and oxidative stress only for cases but not for controls. Moreover, 72 (24.6%) of eCNVRs were located within the chromosomal aberration regions implicated in psychiatric-disorders: 16p11.2, 1q21.1, 22q11.2, 3q29, 15q11.2, 17q12 and 16p13.1. These results shed light on the mechanism of how CNVs confer a risk for psychiatric disorders.

1. Introduction

A key objective in genetic research is to link genomic variation to phenotype differences to uncover normal as well as pathological variation. The influence of single nucleotide polymorphisms on phenotypic variation has been extensively studied; however, it is only recently that other DNA alterations such as copy number variations are being investigated. Copy number variations (CNVs) are DNA segments present at variable copy numbers and owing to their large size, contribute to a substantial proportion of the variation in the human genome (Ionita-Laza et al., 2009; Redon et al., 2006). Among the CNVs, rare CNVs are of more interest because they are presumably enriched in de novo events. Under the rare-variant common disease hypothesis, multiple rare variants with high effect sizes in aggregation, contribute substantially to the illness, hence these rare variants are of great interest since they have not been subject to selection as yet (Zhang et al., 2009). Rare copy number variations have been reported in individuals with neurological and psychiatric disorders such as schizophrenia (International Schizophrenia Consortium, 2008; Levinson et al., 2011; McCarthy et al., 2009), autism (Hedges et al., 2012), bipolar disorder (Zhang et al., 2009) and mental retardation (Guilmatre et al., 2009).

Although a large number of CNVs have been identified in a variety of different species and range of diseases, the functional impact of CNVs at the molecular level remains largely unexplored. One way to assess the functional impact of copy number variations is via its

Abbreviations: CNVs, copy number variations; LCLs, lymphoblastoid cell lines; CNVRs, copy number variable regions; eCNVRs, expression influencing copy number variable regions; kb, kilo base pairs; Mb, mega base pairs; fdr, false discovery rate.

* Corresponding author at: Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, Hirosawa 2-1, Wako, Saitama 351-0198, Japan. Tel.: +81 48 467 6949; fax: +81 48 467 6947.
E-mail address: kato@brain.riken.jp (T. Kato).
[1] Current address: Complex Trait Genomics Group, Queensland Brain Institute, Upland Road St Lucia, Brisbane 4067, Australia.
[2] Current address: Research Equipment Center, Hamamatsu University School of Medicine, Hamamatsu 431-3192, Japan.
[3] Current address: Research Center for Physical Fitness, Sports and Health, Toyohashi University of Technology, Toyohashi, 441-8580, Japan.

0168-0102/$ – see front matter © 2013 The Authors. Published by Elsevier Ireland Ltd and the Japan Neuroscience Society. All rights reserved.
http://dx.doi.org/10.1016/j.neures.2013.10.009

effect on different cellular processes such as gene expression levels. The initial study exploring the transcriptome-wide impact of CNVs on gene expression profiles in lymphoblastoid cell lines (LCLs) identified that approximately 20% of variation in gene expression could be attributed to copy number variations in the genome (Stranger et al., 2007). While there is a plethora of studies assessing the influence of single nucleotide polymorphisms on gene expression profiles, to the best of our knowledge, there are only four studies interrogating the influence of CNVs on gene expression in humans. Moreover, due to the limited availability of human tissues such as brain samples, three of the four studies assessing the influence of CNVs on gene expression in normal tissues till date have been performed on LCLs (Luo et al., 2012; Schlattl et al., 2011; Stranger et al., 2007) while only one recent study (Ye et al., 2012) has assessed gene expression in the human brain.

Integration of gene expression and CNV data will allow the prioritization of CNV-harboring candidate regions where the CNVs significantly alter gene expression levels of transcripts thereby providing evidence of a downstream functional consequence. The aim of this study was to perform a comprehensive and unbiased genome-wide search for functional CNVs in the human brain and to interrogate the nature of these CNVs.

2. Materials and methods

2.1. Samples

Gene expression and copy number variations (CNVs) were obtained from prefrontal cortices of postmortem brains of 105 individuals (35 controls, 35 bipolar disorder [one of which was later excluded due to the alteration of diagnosis] and 35 schizophrenia patients) from the Stanley Medical Research Institute. As described in our previous report (Iwamoto et al., 2011), to reduce confounding factors due to previously identified effects of sample pH, we interrogated 72 individuals (26 controls, 23 bipolar disorder and 23 schizophrenia patients) which were preselected for high pH levels ($pH \geq 6.4$).

2.2. Gene expression

Gene expression levels were assessed using the Affymetrix HU133A microarray which comprised of 22,283 expression probes, details of which are described elsewhere (Iwamoto et al., 2011). Briefly, the raw gene expression data was preprocessed using MAS5 (*Affymetrix*) and filtered for probes which were called present in more than 50% of the samples, allowing a total of 11,920 probes for subsequent analysis. Microarray data had been deposited to the GEO database and is available on the GEO server (GES12649) and on the Stanley Medical Research Institute database (https://www.stanleygenomics.org/).

2.3. Copy number variation

Copy number variation was measured on the Agilent 450k early access CGH array (Agilent Technologies, Inc., Santa Clara, CA, USA), which is designed based on the database of known CNVs. Sample and reference DNA (3.0 μg each) was labeled with Cy5 or Cy3 using the DNA labeling kit from Agilent. Hybridization and washes were performed following the manufacturer's recommendation. The arrays were scanned with a MicroArray Scanner G2505A (Agilent). The obtained TIFF image data were processed with Agilent Feature Extraction software (version 9.5.3.1) using the CGH-v4_95_Feb07 protocol (Agilent).

DNA from one female (NA15510, Coriell Cell Repository, Camden, NJ, USA) was used as a reference to allow detection of copy number changes. This was in accordance with previous reports which have shown that usage of a single reference increases the sensitivity to detect more CNVs and produces more consistent and reproducible data as compared to using a pooled reference (Haraksingh et al., 2011). The raw data were imported into Agilent DNA Analytics 70 software and analyzed using the Aberration Detection Method 2 (ADM-2) algorithm (Lipson et al., 2006) which uses log2 ratios weighted by log2 ratios error as calculated by Feature Extraction software to identify genomic intervals with copy number differences between the samples and the reference. The Agilent Feature Extraction software was used to compute Quality Control metrics. The Agilent protocol recommended thresholds including average signal intensity at each probe, background signal (noise) (<5) using non-hybridizing control probes and signal-to-noise ratios (>30) were used to assess the quality of DNA and the experimental workflow. The derivative log ratio spread (dLRsd) was used to calculate the robust standard deviation (spread) of the log ratio differences between consecutive probes across all chromosomes. Three samples that did not satisfy the QC metrics thresholds and had dLRsd of >0.30 were excluded from further analysis.

The following parameters were used in this analysis: threshold of ADM-2: 6.0 with a bin size of 10; fuzzy zero: on; GC correction: on, aberration filters: on (maxAberrations = 100,000 AND percentPenetrance = 0); feature level filters: on (gIsSaturated = true OR rIsSaturated = true OR gIsFeatNonUnifOL = true OR rIsFeatNonUnifOL = true). A minimum three contiguous suprathreshold probes were mandatory to define a copy number change. Data were centralized and calls with average log2 ratios of ≤0.25 were excluded from the analysis. Data were normalized using the GC correction algorithm that corrects for wavy artifacts associated with the GC content of genomic regions and fuzzy zero correction that allows correction of extended aberrant segments with low absolute mean ratios that might represent noise. In the current study we assessed only autosomal CNVs since analysis of X and Y chromosomal CNVs are difficult to interpret. After filtering, a total of 34,453 CNV probes corresponding to 6836 copy number variable regions (CNVRs) were used for the analysis.

2.4. Statistical analysis

Physical positions and annotations of the gene expression and CNV array probes were updated to the Genome build GRCh37 (hg19) using the UCSC genome browser Liftover tool (http://genome.ucsc.edu/cgi-bin/hgLiftOver). Separate analysis was performed for the continuous log2 CNV ratio and the simplified CNV state (1 = loss, 2 = normal state and 3 = gain) for each CNVR. A log2 ratio of 0 was considered the normal state, a log2 ratio of <−0.25 was considered a loss and log2 ratio of >0.25 was considered a gain. To identify the influence of CNVRs on gene expression, for each CNVR we probed a cis-window of ±1 Mb from the CNVR coordinates. For all gene expression probes located fully or partially (at least 1 bp overlap) within this window, we calculated the association between the CNVR state/CNVR log-ratio and the gene expression levels using general linear models (glm) in R, whilst covarying for age, gender, ethnicity and post-mortem interval (PMI) hours and the results were corrected for multiple testing using 5% false discovery rate (fdr). Results of the association between the CNVR state/CNVR log-ratio and the gene expression levels were very similar, therefore only the results for the CNVR state are presented.

To correct for the different numbers of gene expression probes tested for each CNVR and to account for possible inflation in association *p*-values which might result due to outliers especially for the singletons, we repeated the association analysis using label-swapping adaptive permutations with a maximum of 100,000 permutations in PLINK (Purcell et al., 2007). This method is widely accepted as the most appropriate method for multiple testing

correction and allows for outlier detection (Gibbs et al., 2010; Luo et al., 2012) since it does not assume the normal distribution of the trait and calculates the empirical p-value by label-swapping of the quantitative trait by randomly assigning each individuals quantitative trait (gene expression level) to another individuals CNVR state. Results from all permutations are used to calculate an empirical p-value of significance for each CNVR-gene expression pair (Lanktree et al., 2009).

Associations with p-values of ≤ 0.05 after 5% fdr correction and permutation empirical p-values of ≤ 0.05 were considered significant. Genomic inflation was assessed by calculation of the lambda (genomic inflation factor) for each CNVR in R. The proportion of variance in gene expression explained by the CNVR was calculated using the adjusted R^2 obtained by the glm function in R. Differences in gene expression variances across groups of transcripts were calculated using the 2-sided Kolmogorov–Smirnov test (KS-test). Enrichment of eCNVRs for low frequency and over-representation of genic eCNVRs was performed by conducting simulations. Simulations were performed by generating 1000 randomized CNVRs sets, matched for the CNVR frequency and of the same set at the eCNVR set. The randomized sets of CNVRs were sampled (without replacement) from all the tested CNVRs and based on the simulations we obtained empirical p-values for enrichment. All reported p-values were 2-sided and within 95% confidence interval.

Gene expression probes containing SNPs within their sequences were identified using the PLANdbAffy database (http://affymetrix2.bioinf.fbb.msu.ru) and are highlighted in the results table.

The functional analysis was performed using the Wikipathways and KEGG tools via the WebGestalt Gene Set Enrichment database (http://bioinfo.vanderbilt.edu/webgestalt/). The enrichment was calculated using a hypogeometric test using the human genome as the background and all results were corrected for multiple testing using the Bonferroni correction.

For comparisons of the results with previously reported eCNVRs (Luo et al., 2012; Schlattl et al., 2011; Stranger et al., 2007), we used the eCNVRs which were significant at linear regression fdr of 5% and permutation empirical p-value of ≤ 0.05 and compared these to previously reported eCNVRs to check for overlaps between the data sets.

To compare our results with previously identified brain phenotype-associated CNVRs, we performed manual data mining using NCBI PubMed to search for articles reporting significant CNVs identified in schizophrenia, major depressive disorder, bipolar disorder and/or autism. We limited our search to 7 large association studies which comprised of at least 3000 patients each (Ingason et al., 2011; International Schizophrenia Consortium, 2008; Levinson et al., 2011; McCarthy et al., 2009; Moreno-De-Luca et al., 2010; Stefansson et al., 2008; Weiss et al., 2008).

3. Results

3.1. Identification of CNVRs

A flowchart of the study design and results is provided in Fig. 1. Sample characteristics of the 72 individuals included in the study are provided in Supplementary Table 1. While results on the expression profiles of these samples have been reported previously (Iwamoto et al., 2005, 2011), in the current study we assessed the influence of CNVs on gene expression at the genome-wide scale. Using the Agilent early access 450k array, after filtering and preprocessing, a total of 34,453 autosomal CNVs were identified in the current samples, which were further grouped into 6836 unique copy number variable regions (CNVRs) using criteria of at least 3 overlapping CNV probes. Among the 6836 CNVRs, 3549 were losses, 3136 were gains and 151 were complex (gains in some and losses in others).

Of the CNVRs, 3656 (53.4%) were singletons while the remaining 3180 CNVRs were identified in at least two individuals (identical CNVR start and end coordinates). For statistical analysis purposes, the common criteria of 5% frequency was used to group the CNVRs into 5058 (74%) rare/low frequency CNVRs and 1778 common CNVRs (26%). Further, CNVRs were divided into two categories;

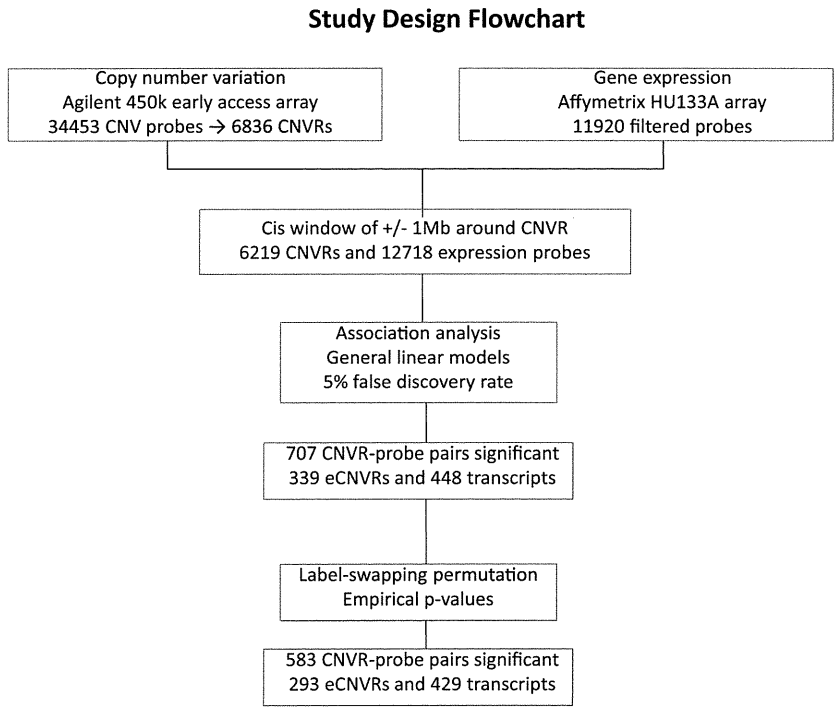


Fig. 1. Flowchart of study design and results.

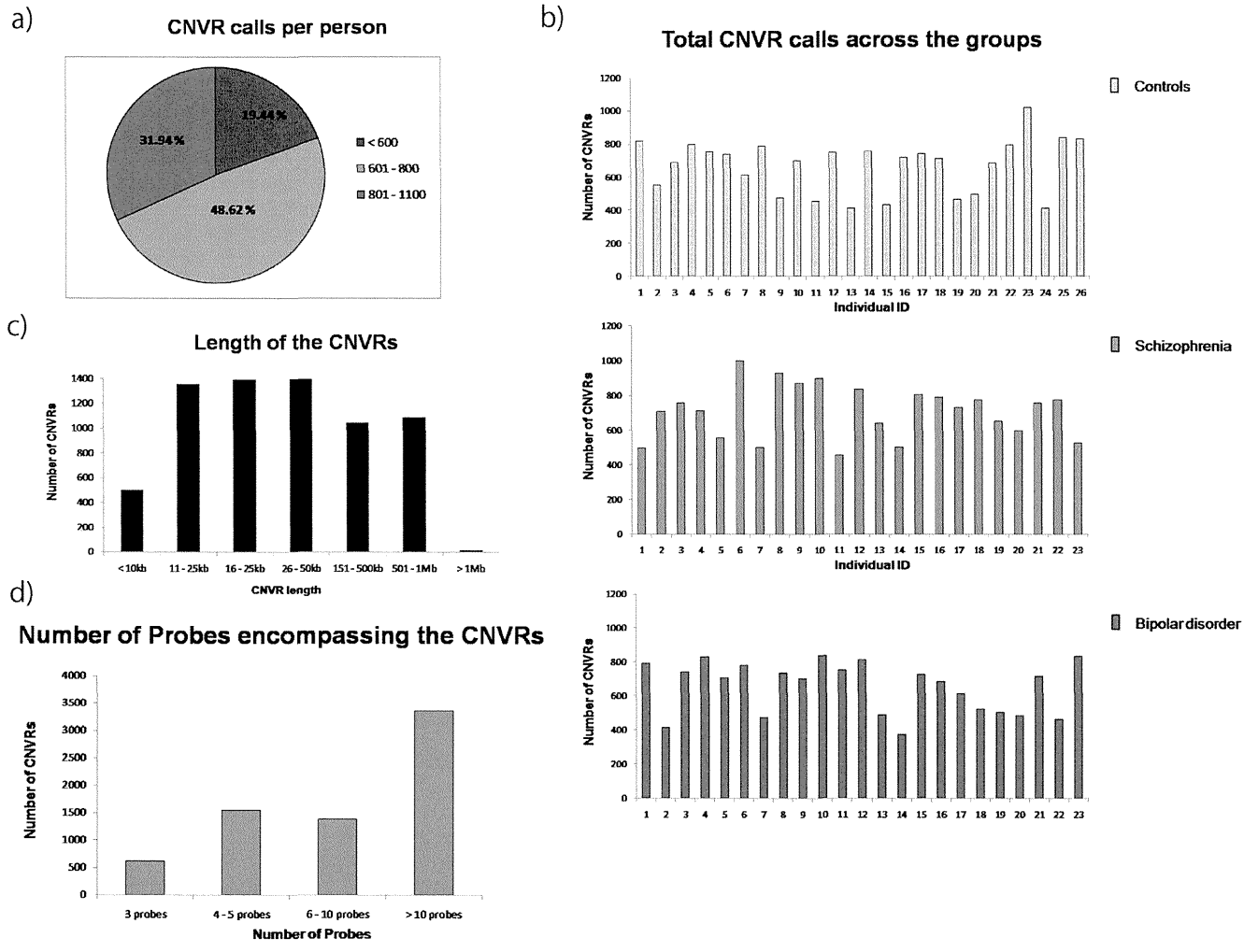


Fig. 2. Description of the copy number variable regions (CNVRs). (a) Total CNVR calls across all samples. (b) Total CNVR calls across controls, schizophrenia and bipolar disorder patients. (c) Distribution of the length of the CNVRs. (d) Number of probes encompassing CNVRs.

genic and intergenic. A total of 3051 of 6836 CNVRs (44.6%) were 'genic' CNVRs, i.e. CNVRs spanning a part of whole of the transcript gene expression probe (with at least one base pair overlap between the regions) while the remaining CNVRs did not harbor any known transcripts. The mean and median CNVR size was 66.5 kb and 11.9 kb respectively with average CNV segment numbers of 688 across all samples. The total combined CNVR burden across the individuals ranged between 11.6 and 72.4 Mb. The total number of CNVR calls and distribution of CNVR length and probes encompassing CNVRs are depicted in Fig. 2.

3.2. Influence of CNVRs on gene expression

Next, we sought to assess the influence of CNVRs on gene expression profiles in the human prefrontal cortex to identify functional CNVRs. To test the influence of CNVRs on the neighboring transcripts we defined a cis window of ±1 Mb from the CNVR to identify proximal transcripts (Fig. 3). For 6219 of the 6836 unique CNVRs, at least one transcript was located within the ±1 Mb CNVR cis coordinates. Transcripts located within the cis coordinates amounted to a total of 12,718 unique transcripts. For each CNVR, we tested the association between the CNVR and the transcripts within the cis coordinates using general linear models and adjusting for age, gender and PMI hours for the CNVR state. All results were corrected for multiple testing using 5% fdr. To avoid inflation due to outliers, we repeated the association analysis by label-swapping adaptive permutations for each CNVR-expression probe pair. Results of the permutations were compared with the linear regression results and only CNVR-expression probe pairs significant in both tests were deemed as significant. Average genomic inflation factor across all tested transcripts was 1.06, indicating no apparent inflation.

A total of 4201 CNVR-probe pairs were significant at $p \leq 0.05$ and a total of 707 CNVR-probe pairs were significant at 5% fdr threshold in the linear regression analysis. These 707 pairs corresponded to a total of 339 unique expression CNVRs (expression-influencing CNVRs or eCNVRs) and 448 unique expression probes. Of these, 583 CNVR-probe pairs (293 unique eCNVRs and 429 unique expression probes) were also significant after permutation testing (Supplementary Table 2). Representative examples of box plots of associations between CNVRs and gene expression profiles of significant eCNVRs are depicted in Fig. 4 and a list of the top 15 associations is given in Table 1. Therefore, expression levels of 429 (3.4%) of 12,718 tested transcripts were significantly influenced by CNVRs after stringent corrections for multiple testing and permutation. For 15 CNVRs overlapping a transcript (inside pairs), a positive correlation between CNV state and gene expression was obtained, as expected. On the other hand,

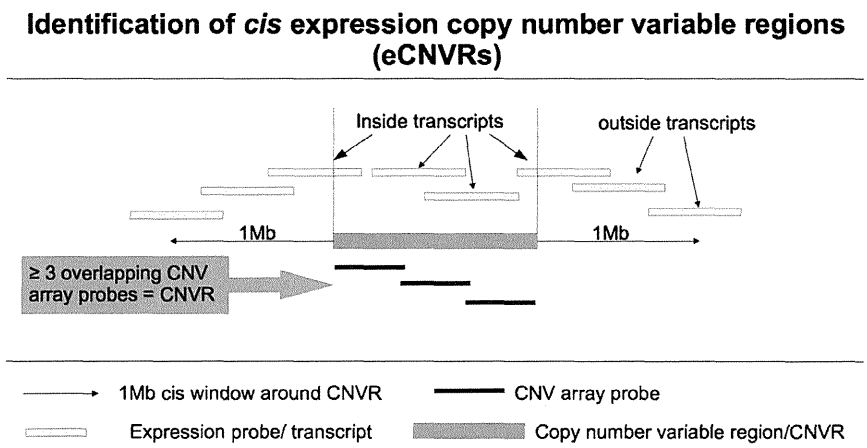


Fig. 3. Schematic figure of the cis window chosen for association testing of expression copy number variable regions (eCNVRs).

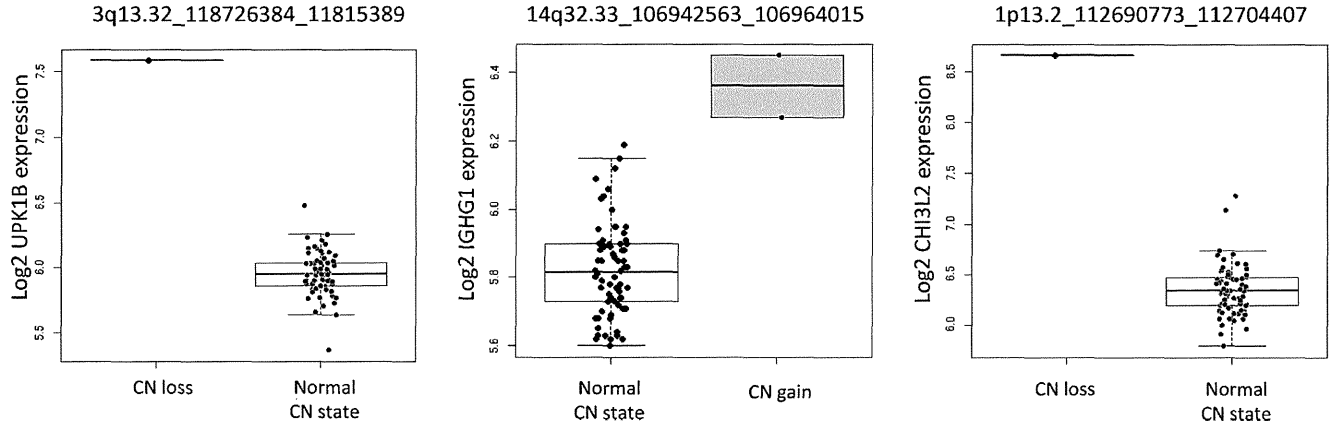


Fig. 4. Box plots of associations between CNVRs and gene expression profiles of significant eCNVR.

a certain part of CNVR-gene expression correlations were negative. Such non-conventional types of eCNVRs seen in 242 (41%) of the 293 CNVRs in this study have also been reported in previous studies (Luo et al., 2012; Schlattl et al., 2011; Stranger et al., 2007).

3.3. Nature of the significant brain eCNVRs

Of the 583 CNVR-probe pairs, we next interrogated the nature of the 293 unique CNVRs in terms of frequency, size, genes harbored, proportion of variance explained and functional annotation (see Supplementary Table 2). Of the 293 eCNVRs, 239 (81.6%) were rare/low frequency (<5% MAF) and the remaining 54 (18.4%) were common eCNVRs. There was a significant over-representation of rare/low frequency CNVRs among the significant eCNVRs ($p = 1.087 \times 10^{-10}$). The size of the CNVRs ranged between 234 bp and 1.65 Mb, with an average length of 0.73 Mb. Summary of the CNVR lengths are as follows; – ≤100 kb: 237, 101–500 kb: 48, 501 kb to 1 Mb: 7, and >1 Mb: 1. Of the significant eCNVRs, 179 (61.1%) were genic while the remaining 38.9% were non-genic, corresponding to a significant over-representation of genic CNVRs among the eCNVRs ($p = 1.4 \times 10^{-6}$). Such significant enrichment of low frequency and genic CNVRs was also observed when the analysis was restricted to non-singleton CNVRs (CNVRs robustly detected in at least 2 individuals) and when the CNVRs were restricted to CNVR intervals containing at least 5 or more probes.

Table 1
List of top 15 significant CNVR and gene expression associations.

CNVR	ProbeSetID	Chr	Cytoband	*p* value Regression	Pair type	Gene expression Transcript	CNVR_Start	CNVR_Stop
q13.32_118726384_118815389	210065_s_at	3	q13.32	2.09E-014	Outside	UPK1B	118726384	118815389
q11.21_20370979_20489827	222141_at	22	q11.21	1.38E-013	Outside	KLHL22	20370979	20489827
q11.21_21709612_21905954	222141_at	22	q11.21	1.38E-013	Outside	KLHL22	21709612	21905954
p13.2_112690773_112704407	213060_s_at	1	p13.2	1.80E-011	Outside	CHI3L2	112690773	112704407
q21.3_152555939_152586960	216701_at	1	q21.3	7.03E-011	Outside	C1orf68	152555939	152586960
q32.1_201177775_201181213	215168_at	1	q32.1	4.67E-010	Outside	TIMM17A	201177775	201181213
q34_112931419_112973293	205620_at	13	q34	2.63E-009	Outside	F10	112931419	112973293
p13.3_109749308_109757804	221874_at	1	p13.3	3.24E-009	Inside	KIAA1324	109749308	109757804
q12_34422129_34811416	218079_s_at	17	q12	3.38E-009	Outside	GGNBP2	34422129	34811416
p13.3_110228105_110254473	207464_at	1	p13.3	1.72E-008	Outside	AHCYL1	110228105	110254473
q29_195409551_195448563	217110_s_at	3	q29	3.17E-008	Outside	MUC4	195409551	195448563

Consistent with previous reports, no significant correlations between CNVR-probe distance and frequency of significant eCNVRs or the p-value of the association was observed. Such correlation was not seen even when stratifying the samples into low/high frequency CNVRs and genic/non-genic CNVRs ($p > 0.05$). The average proportion of variance in gene expression explained by the eCNVR was 26% across all significant CNVR-probe pairs.

3.4. Validation of the CNVRs and gene expression

We next compared the brain eCNVRs to previously reported eCNVRs in lymphoblast cell lines (Luo et al., 2012; Schlattl et al., 2011; Stranger et al., 2007) or in the prefrontal cortex (Ye et al., 2012). Of the 293 eCNVRs, 29 (10%) were previously reported to influence gene expression of nearby transcripts (see Supplementary Table 2).

Next, we compared the copy number data obtained by Agilent CNV array with those obtained by Affymetix GeneChip Human Mapping 500k SNP arrays for technical validation, which we have described previously (Iwamoto et al., 2011). Of the 731 autosomal CNVRs detected on the SNP array, we were able to detect 68 (9.3%) on the CNV array, these included 2 singleton eCNVRs which were significantly associated with gene expression (see Supplementary Table 2).

For validation of gene expression data, we used previously reported data from Altar et al. (2008), where gene expression profiling for the same prefrontal cortex samples was performed using the same arrays and the same normalization protocols. We assessed the correlations between the gene expression profiles from the current dataset and the published dataset. For the 429 probes significantly associated with CNVRs from Supplementary Table 2, an average correlation of 0.43 and median correlation of 0.40 was observed across all individuals. Using gene expression levels from Altar and colleagues, we were able to successfully replicate CNVR-gene expression associations for a handful of selected transcripts including PTPRN2, FHL2, SLC16A1, CH1DL and COMT genes, thereby demonstrating the technical reliability of the data (see Supplementary Table 2).

3.5. Overlap with loci associated with psychiatric disorders

While the eCNVR analysis was performed across all individuals due to limited power, the samples comprised of individuals with bipolar disorder ($n = 23$), schizophrenia ($n = 23$) and controls ($n = 26$). Of the 293 significant eCNVRs, 49 CNVRs were found only in bipolar disorder patients, 72 CNVRs were found only in schizophrenia patients, 70 CNVRs were found only in controls and the remaining 122 CNVRs were found in individuals belonging to two or more groups. The average CNVR burden and number of CNVRs were 32.23 kb and 621 regions in schizophrenia patients, 32.81 kb and 623 regions in bipolar patients and 30.12 kb and 588 regions in controls. No significant differences between the total CNVR burden, average CNVR burden and number of CNVRs were observed across the groups.

3.6. Functional annotation of CNVR-influenced genes

Next, we interrogated the functional relevance of the significant CNVR-gene expression pairs using the Web-based Gene Set Enrichment analysis (WebGestalt – Wikipathways and KEGG tools).

Among transcripts whose expression was significantly influenced by CNVRs, the corticotropin-releasing hormone pathway was significantly enriched and this enrichment was also observed when stratifying transcripts influenced by CNVRs in cases only or in controls only (Table 2).

Functional annotation of genes overlapping copy number variants identified a significant enrichment of transcripts implicated within glutathione metabolism ($p = 0.020$) and oxidative stress ($p = 0.030$) pathways in all genes ($p = 0.020$) and also only among genes overlapping CNVRs in cases only ($p = 0.0015$ for glutathione metabolism and $p = 0.0031$ for oxidative stress) but no such enrichment was observed when assessing only genes overlapping CNVRs in controls only.

Hence, the functional relevance of the CNVR-encompassed transcripts seems to be different in cases versus controls.

3.7. Comparison with loci associated with psychiatric disorders

We next compared the brain eCNVRs to previously reported loci containing CNVRs which were shown to be associated with schizophrenia, major depressive disorder, bipolar disorder and/or autism (Supplementary Table 2). We limited our search to loci from 7 large association studies (8 unique loci: 16p11.2, 1q21.1, 22q11.2, 3q29, 17q12, 16p13.1, 15q13.3 and 15q11.2) comprising of over 3000 cases per study. These 8 loci spanned approximately 44.6 Mb (~1.4% of whole genome). Of the 293 significant eCNVRs, 72 (24.6%) psychiatric-disorder associated CNVRs were found (Table 3). Of the 72 CNVRs, 19 were found only in schizophrenia patients, 21 were found only in bipolar disorder patients and 19 were found only in controls. Of the 19 CNVRs found only in controls, 7 were found in a single control individual (C_15), due to unknown or unexplained reasons, while the remaining CNVRs were found in independent individuals. The average number of CNVRs and total CNVR burden in this individual was well within the normal range (see Fig. 2b – control 15) and we technically validated two singleton CNVRs in this individual (see Supplementary Table 2), thereby reducing the possibility of sample issues such as DNA quality or hybridization problems.

The 72 eCNVRs spanned seven of the eight candidate regions including 16p11.2, 1q21.1, 22q11.2 (see Fig. 5), 3q29 (see Fig. 5), 15q11.2, 17q12 and 16p13.1 (Fig. 6 and Table 3). In a recent study, CNVRs in 1q21.1 and 22q11.2 were shown to be significantly associated with dorsolateral prefrontal cortex gene expression levels of nearby transcripts (Ye et al., 2012) while the other regions, to the best of our knowledge, has not been shown to have functional influence on gene expression regulation in the human brain.

4. Discussion

In the current study, we interrogated the influence of CNVs on gene expression in prefrontal cortex of post-mortem brain samples to identify functional CNVs. Gene expression levels of

Table 2
Pathway analysis of transcripts influenced by CNVRs.

Over-represented pathways	All transcripts	Only cases transcripts	Only controls transcripts
Transcripts whose gene expression was regulated by CNVR			
Corticotropin releasing hormone pathway	0.0430	0.0385	0.0198
Transcripts located within CNVRs			
Glutathione metabolism	0.0200	0.0015	>0.050

Table 3
List of eCNVRs within psychiatric trait-associated loci (International Schizophrenia Consortium, 2008; Levinson et al., 2011; McCarthy et al., 2009; Moreno-De-Luca et al., 2010; Stefansson et al., 2008).

CNVR	CNVR type	People	ProbeSetID	Psychiatric Trait locus	Permutations *p* value	CNVR_Start	CNVR_Stop	GeneSymbol
q21.1_144322804_144400673	Gain	bp	212392_s_at	1q21.1[4,5]	0.014808	144322804	144400673	PDE4DIP
q21.1_144672865_144709635	Gain	bp	212392_s_at	1q21.1[4,5]	0.014808	144672865	144709635	PDE4DIP
q21.1_144952007_145074984	Gain	bp	212392_s_at	1q21.1[4,5]	0.014808	144952007	145074984	PDE4DIP
q21.1_145190273_145292282	Gain	bp	212392_s_at	1q21.1[4,5]	0.014808	145190273	145292282	PDE4DIP
q21.1_145293607_145368437	Loss	bp	209207_s_at	1q21.1[4,5]	0.006770	145293607	145368437	SEC22B
q21.1_145312073_145367945	Gain	bp,bp,con*,con	206766_at	1q21.1[4,5]	0.003476	145312073	145367945	ITGA10
q21.1_145312073_145367945	Gain	bp,bp,con*,con	214113_s_at	1q21.1[4,5]	0.010387	145312073	145367945	RBM8A
q21.1_145312073_145367945	Gain	bp,bp,con*,con	209206_at	1q21.1[4,5]	0.018087	145312073	145367945	SEC22B
q21.1_145626237_145746971	Loss	bp	209207_s_at	1q21.1[4,5]	0.006718	145626237	145746971	SEC22B
q21.1_146034981_146039673	Loss	con	214113_s_at	1q21.1[4,5]	0.002943	146034981	146039673	RBM8A
q21.1_146215885_146231981	Gain	bp,bp,con*	212539_at	1q21.1[4,5]	0.009246	146215885	146231981	CHD1L
q21.1_146215885_146231981	Gain	bp,bp,con*	214474_at	1q21.1[4,5]	0.046320	146215885	146231981	PRKAB2
q21.1_146215885_146231981	Gain	bp,bp,con*	205776_at	1q21.1[4,5]	0.005915	146215885	146231981	FMO5
q21.1_146215885_146231981	Gain	bp,bp,con*	206766_at	1q21.1[4,5]	0.003218	146215885	146231981	ITGA10
q21.1_146215885_146231981	Gain	bp,bp,con*	212742_at	1q21.1[4,5]	0.012921	146215885	146231981	RNF115
q21.1_146215885_146231981	Gain	bp,bp,con*	215300_s_at	1q21.1[4,5]	0.046230	146215885	146231981	FMO5
q29_195215347_195232654	Gain	2bp,sz,2con	221536_s_at	3q29[4]	0.015525	195215347	195232654	LSG1
q29_195215347_195237188	Gain	sz,sz	217109_at	3q29[4]	0.008881	195215347	195237188	MUC4
q29_195341670_195453587	Gain	sz	212477_at	3q29[4]	0.003827	195341670	195453587	ACAP2
q29_195341670_195743252	Gain	sz	208878_s_at	3q29[4]	0.014808	195341670	195743252	PAK2
q29_195341813_195725193	Gain	con	212476_at	3q29[4]	0.004947	195341813	195725193	ACAP2
q29_195341813_195725193	Gain	con	208877_at	3q29[4]	5.22E-005	195341813	195725193	PAK2
q29_195341813_195725193	Gain	con	204210_s_at	3q29[4]	0.001016	195341813	195725193	PCYT1A
q29_195344712_195477486	Gain	sz	203838_s_at	3q29[4]	0.003567	195344712	195477486	TNK2
q29_195344712_195477486	Gain	sz	204209_at	3q29[4]	0.014120	195344712	195477486	PCYT1A
q29_195344712_195477486	Gain	sz	212476_at	3q29[4]	0.004968	195344712	195477486	ACAP2
q29_195344712_195477486	Gain	sz	221536_s_at	3q29[4]	0.001392	195344712	195477486	LSG1
q29_195393418_195452775	Gain	con	204210_s_at	3q29[4]	0.001016	195393418	195452775	PCYT1A
q29_195393418_195452775	Gain	con	212476_at	3q29[4]	0.004947	195393418	195452775	ACAP2
q29_195409551_195448563	Loss	bp	212476_at	3q29[4]	0.004287	195409551	195448563	ACAP2
q29_195409551_195448563	Loss	bp	216439_at	3q29[4]	0.005625	195409551	195448563	TNK2
q29_195409551_195448563	Loss	bp	217110_s_at	3q29[4]	9.47E-007	195409551	195448563	MUC4
q29_195411543_195448616	Complex	con,con,sz	207332_s_at	3q29[4]	0.011657	195411543	195448616	TFRC
q29_195411543_195448616	Complex	con,con,sz	208691_at	3q29[4]	0.015725	195411543	195448616	TFRC
q29_195506071_195515379	Gain	4bp,2sz,3con	207332_s_at	3q29[4]	0.054061	195506071	195515379	TFRC
q29_195506071_195516643	Gain	3sz,1bp,1con	212477_at	3q29[4]	0.005181	195506071	195516643	ACAP2
q29_195648611_195747915	Gain	sz	212477_at	3q29[4]	0.003827	195648611	195747915	ACAP2
q29_195663926_195725193	Loss	bp,bp,con	212477_at	3q29[4]	0.014808	195663926	195725193	ACAP2
q29_196555515_196559209	Loss	8bp,3sz,8con	203839_s_at	3q29[4]	0.012921	196555515	196559209	TNK2
q29_196759662_196762173	Loss	sz	207332_s_at	3q29[4]	0.000185	196759662	196762173	TFRC
q29_196759662_196762173	Loss	sz	208691_at	3q29[4]	0.004467	196759662	196762173	TFRC
q29_197347418_197394189	Loss	3sz,2bp	208877_at	3q29[4]	0.000305	197347418	197394189	PAK2
q29_197347418_197394189	Loss	3sz,2bp	208877_at	3q29[4]	0.000305	197347418	197394189	PAK2
q29_197603683_197605592	Gain	sz	211715_s_at	3q29[4]	0.010798	197603683	197605592	BDH1
q29_197603683_197605592	Gain	sz	212733_at	3q29[4]	4.15E-006	197603683	197605592	KIAA0226
q29_197603683_197605592	Gain	sz	220041_at	3q29[4]	0.004968	197603683	197605592	PIGZ
q29_197825694_197833212	Loss	sz,bp	212733_at	3q29[4]	0.006171	197825694	197833212	KIAA0226
q29_197825901_197832592	Gain	con,bp	214739_at	3q29[4]	0.006171	197825901	197832592	LRCH3
q29_197825901_197832592	Gain	con,bp	213687_s_at	3q29[4]	0.014935	197825901	197832592	RPL35A
q29_197895169_197896197	Gain	10sz,6bp,7con	212733_at	3q29[4]	0.051563	197895169	197896197	KIAA0226
q11.2_22303902_22372338	Loss	con	214876_s_at	15q11.2[27]	0.003155	22303902	22372338	TUBGCP5
q11.2_22318597_22348005	Loss	con	212133_at	15q11.2[27]	0.003062	22318597	22348005	NIPA2
p11.2_28390355_28437534	Gain	bp	221822_at	16p11.2[5,6]	0.012273	28390355	28437534	CCDC101
p11.2_28612132_28615866	Gain	sz	212808_at	16p11.2[5,6]	0.008063	28612132	28615866	NFATC2IP
p11.2_30200517_30220479	Gain	con*	221864_at	16p11.2[5,6]	0.003399	30200517	30220479	ORAI3
p11.2_30200517_30220479	Gain	con*	212275_s_at	16p11.2[5,6]	0.004775	30200517	30220479	SRCAP
p11.2_30200517_30220479	Gain	con*	45653_at	16p11.2[5,6]	0.014873	30200517	30220479	KCTD13
p11.2_30200517_30220479	Gain	con*	207684_at	16p11.2[5,6]	0.003057	30200517	30220479	TBX6
p11.2_30200517_30220479	Gain	con*	200961_at	16p11.2[5,6]	0.005915	30200517	30220479	SEPHS2
p11.2_30200517_30220479	Gain	con*	201253_s_at	16p11.2[5,6]	0.001670	30200517	30220479	CDIPT
p11.2_30200517_30220479	Gain	con*	202256_at	16p11.2[5,6]	0.005277	30200517	30220479	CD2BP2
p11.2_30200517_30220479	Gain	con*	204876_at	16p11.2[5,6]	0.007028	30200517	30220479	ZNF646
p11.2_30200517_30220479	Gain	con*	204878_s_at	16p11.2[5,6]	0.010941	30200517	30220479	TAOK2
p11.2_30200517_30220479	Gain	con*	205744_at	16p11.2[5,6]	0.000305	30200517	30220479	DOC2A
p11.2_30200517_30220479	Gain	con*	209083_at	16p11.2[5,6]	0.000158	30200517	30220479	CORO1A
p11.2_30200517_30220479	Gain	con*	214226_at	16p11.2[5,6]	0.000264	30200517	30220479	PRSS53
p11.2_30200517_30220479	Gain	con*	217949_s_at	16p11.2[5,6]	0.003144	30200517	30220479	VKORC1
p11.2_30200517_30220479	Gain	con*	218300_at	16p11.2[5,6]	0.014342	30200517	30220479	C16orf53
p11.2_30200517_30220479	Gain	con*	219072_at	16p11.2[5,6]	0.005495	30200517	30220479	BCL7C
p11.2_30200517_30220479	Gain	con*	219781_s_at	16p11.2[5,6]	0.000235	30200517	30220479	ZNF771
p11.2_30200517_30220479	Gain	con*	221968_s_at	16p11.2[5,6]	0.015791	30200517	30220479	ZNF771
p11.2_32164104_33816677	Gain	bp	219540_at	16p11.2[5,6]	0.006770	32164104	33816677	ZNF267
p13.11_15011919_15029273	Gain	con	222204_s_at	16p13.1[25]	0.014769	15011919	15029273	RRN3
p13.13_12020113_12036399	Gain	bp	205101_at	16p13.1[25]	0.014502	12020113	12036399	CIITA

Table 3 (*Continued*)

CNVR	CNVR type	People	ProbeSetID	Psychiatric Trait locus	Permutations *p* value	CNVR_Start	CNVR_Stop	GeneSymbol
p13.13_12020113_12036399	Gain	bp	210001_s_at	16p13.1[25]	1.29E-005	12020113	12036399	SOCS1
q12_34407079_34662164	Gain	con*	212186_at	17q12[26]	1.93E-005	34407079	34662164	ACACA
q12_34407079_34662164	Gain	con*	219885_at	17q12[26]	2.21E-005	34407079	34662164	SLFN12
q12_34407079_34662164	Gain	con*	1405_i_at	17q12[26]	0.000985	34407079	34662164	CCL5
q12_34407079_34662164	Gain	con*	204655_at	17q12[26]	0.005086	34407079	34662164	CCL5
q12_34407079_34662164	Gain	con*	207354_at	17q12[26]	0.005086	34407079	34662164	CCL16
q12_34407079_34662164	Gain	con*	209924_at	17q12[26]	0.006012	34407079	34662164	CCL18
q12_34407079_34662164	Gain	con*	209965_s_at	17q12[26]	0.000290	34407079	34662164	RAD51L3
q12_34407079_34662164	Gain	con*	212544_at	17q12[26]	0.000364	34407079	34662164	ZNHIT3
q12_34407079_34662164	Gain	con*	218079_s_at	17q12[26]	0.001911	34407079	34662164	GGNBP2
q12_34407079_34662164	Gain	con*	219320_at	17q12[26]	0.009997	34407079	34662164	MYO19
q12_34407079_34662164	Gain	con*	220499_at	17q12[26]	0.001219	34407079	34662164	FNDC8
q12_34407079_34662164	Gain	con*	32128_at	17q12[26]	0.006474	34407079	34662164	CCL18
q12_34408772_34646159	Gain	sz	206230_at	17q12[26]	0.014502	34408772	34646159	LHX1
q12_34422129_34811416	Gain	sz	210548_at	17q12[26]	0.003787	34422129	34811416	CCL23
q12_34422129_34811416	Gain	sz	1405_i_at	17q12[26]	0.003569	34422129	34811416	CCL5
q12_34422129_34811416	Gain	sz	200615_s_at	17q12[26]	0.018005	34422129	34811416	AP2B1
q12_34422129_34811416	Gain	sz	207343_at	17q12[26]	2.38E-005	34422129	34811416	LYZL6
q12_34422129_34811416	Gain	sz	209924_at	17q12[26]	0.000753	34422129	34811416	CCL18
q12_34422129_34811416	Gain	sz	209938_at	17q12[26]	0.004331	34422129	34811416	TADA2A
q12_34422129_34811416	Gain	sz	212186_at	17q12[26]	8.94E-005	34422129	34811416	ACACA
q12_34422129_34811416	Gain	sz	212544_at	17q12[26]	7.39E-005	34422129	34811416	ZNHIT3
q12_34422129_34811416	Gain	sz	218079_s_at	17q12[26]	1.36E-007	34422129	34811416	GGNBP2
q12_34422129_34811416	Gain	sz	218756_s_at	17q12[26]	0.004968	34422129	34811416	DHRS11
q12_34422129_34811416	Gain	sz	220499_at	17q12[26]	0.000417	34422129	34811416	FNDC8
q12_34422129_34811416	Gain	sz	32128_at	17q12[26]	0.000252	34422129	34811416	CCL18
q12_34488357_34760365	Gain	con	210548_at	17q12[26]	0.002623	34488357	34760365	CCL23
q12_34605880_34643115	Gain	con	200612_s_at	17q12[26]	0.011851	34605880	34643115	AP2B1
q12_34605880_34643115	Gain	con	212186_at	17q12[26]	0.014502	34605880	34643115	ACACA
q12_34611572_34615943	Gain	con	200612_s_at	17q12[26]	0.011851	34611572	34615943	AP2B1
q12_34611572_34615943	Gain	con	212186_at	17q12[26]	0.014502	34611572	34615943	ACACA
q12_34764374_34790180	Loss	sz	210549_s_at	17q12[26]	0.007902	34764374	34790180	CCL23
q12_34791790_34806889	Gain	sz	210549_s_at	17q12[26]	0.007902	34791790	34806889	CCL23
q12_35779149_35780902	Loss	bp	210320_s_at	17q12[26]	6.21E-006	35779149	35780902	DDX52
q12_36351950_36385101	Loss	con*	218655_s_at	17q12[26]	0.004733	36351950	36385101	CWC25
q12_36351950_36385101	Loss	con*	200618_at	17q12[26]	0.012387	36351950	36385101	LASP1
q12_36351950_36385101	Loss	con*	201080_at	17q12[26]	0.000339	36351950	36385101	PIP4K2B
q12_36351950_36385101	Loss	con*	201081_s_at	17q12[26]	0.003507	36351950	36385101	PIP4K2B
q12_36351950_36385101	Loss	con*	201400_at	17q12[26]	0.005039	36351950	36385101	PSMB3
q12_36351950_36385101	Loss	con*	210185_at	17q12[26]	0.008384	36351950	36385101	CACNB1
q12_36351950_36385101	Loss	con*	212186_at	17q12[26]	1.93E-005	36351950	36385101	ACACA
q12_36351950_36385101	Loss	con*	221937_at	17q12[26]	0.016260	36351950	36385101	SYNRG
q11.21_18618723_18621135	Loss	sz	202099_s_at	22q11.2[4,5]	0.007575	18618723	18621135	DGCR2
q11.21_18734360_18862822	Loss	bp	214371_at	22q11.2[4,5]	0.009124	18734360	18862822	TSSK2
q11.21_18734360_18862822	Loss	bp	220762_s_at	22q11.2[4,5]	0.004501	18734360	18862822	GNB1L
q11.21_19893805_19895800	Loss	bp	214371_at	22q11.2[4,5]	0.009124	19893805	19895800	TSSK2
q11.21_19893805_19895800	Loss	bp	220762_s_at	22q11.2[4,5]	0.004501	19893805	19895800	GNB1L
q11.21_20339345_20362501	Loss	bp	220762_s_at	22q11.2[4,5]	0.004501	20339345	20362501	GNB1L
q11.21_20370979_20461985	Gain	bp	214406_s_at	22q11.2[4,5]	0.003467	20370979	20461985	SLC7A4
q11.21_20370979_20461985	Gain	bp	220762_s_at	22q11.2[4,5]	0.004501	20370979	20461985	GNB1L
q11.21_20370979_20489827	Loss	con*	207081_s_at	22q11.2[4,5]	3.15E-005	20370979	20489827	PI4KA
q11.21_20370979_20489827	Loss	con*	212180_at	22q11.2[4,5]	0.003267	20370979	20489827	CRKL
q11.21_20370979_20489827	Loss	con*	204482_at	22q11.2[4,5]	0.012431	20370979	20489827	CLDN5
q11.21_20370979_20489827	Loss	con*	205576_at	22q11.2[4,5]	0.007335	20370979	20489827	SERPIND1
q11.21_20370979_20489827	Loss	con*	205881_at	22q11.2[4,5]	0.009217	20370979	20489827	ZNF74
q11.21_20370979_20489827	Loss	con*	206880_at	22q11.2[4,5]	0.005912	20370979	20489827	P2RX6
q11.21_20370979_20489827	Loss	con*	207662_at	22q11.2[4,5]	0.001885	20370979	20489827	TBX1
q11.21_20370979_20489827	Loss	con*	208818_s_at	22q11.2[4,5]	0.011028	20370979	20489827	COMT
q11.21_20370979_20489827	Loss	con*	211147_s_at	22q11.2[4,5]	0.015425	20370979	20489827	P2RX6
q11.21_20370979_20489827	Loss	con*	213981_at	22q11.2[4,5]	2.02E-005	20370979	20489827	COMT
q11.21_20370979_20489827	Loss	con*	218492_s_at	22q11.2[4,5]	0.003155	20370979	20489827	THAP7
q11.21_20370979_20489827	Loss	con*	219811_at	22q11.2[4,5]	0.010840	20370979	20489827	DGCR8
q11.21_20370979_20489827	Loss	con*	221838_at	22q11.2[4,5]	0.016453	20370979	20489827	KLHL22
q11.21_20370979_20489827	Gain	con*	222141_at	22q11.2[4,5]	1.02E-011	20370979	20489827	KLHL22
q11.21_20626904_20648019	Loss	con	205576_at	22q11.2[4,5]	0.000367	20626904	20648019	SERPIND1
q11.21_20648172_20715558	Gain	con	211177_s_at	22q11.2[4,5]	0.000771	20648172	20715558	TXNRD2
q11.21_20648172_20715558	Gain	con	218475_at	22q11.2[4,5]	0.011596	20648172	20715558	TRMT2A
q11.21_20648172_20715558	Gain	con	91617_at	22q11.2[4,5]	0.000145	20648172	20715558	DGCR8
q11.21_21455772_21667502	Loss	bp	221349_at	22q11.2[4,5]	0.002569	21455772	21667502	VPREB1
q11.21_21455772_21667502	Loss	bp	214406_s_at	22q11.2[4,5]	0.003467	21455772	21667502	SLC7A4
q11.21_21668908_21709656	Gain	sz	216301_at	22q11.2[4,5]	1.18E-005	21668908	21709656	LOC100287927
q11.21_21670977_21905954	Gain	sz	216911_s_at	22q11.2[4,5]	0.000279	21670977	21905954	HIC2
q11.21_21708235_21905954	Gain	sz	217180_at	22q11.2[4,5]	0.017219	21708235	21905954	
q11.21_21709612_21905954	Loss	con*	204086_at	22q11.2[4,5]	0.004800	21709612	21905954	PRAME
q11.21_21709612_21905954	Loss	con*	206064_s_at	22q11.2[4,5]	0.018360	21709612	21905954	PPIL2
q11.21_21709612_21905954	Loss	con*	206880_at	22q11.2[4,5]	0.005912	21709612	21905954	P2RX6