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

Please cite this article in press as: Mehta, D., et al., Comprehensive survey of CNVs influencing gene expression in the human brain and its implications for pathophysiology. Neurosci. Res. (2013), http://dx.doi.org/10.1016/j.neures.2013.10.009

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;

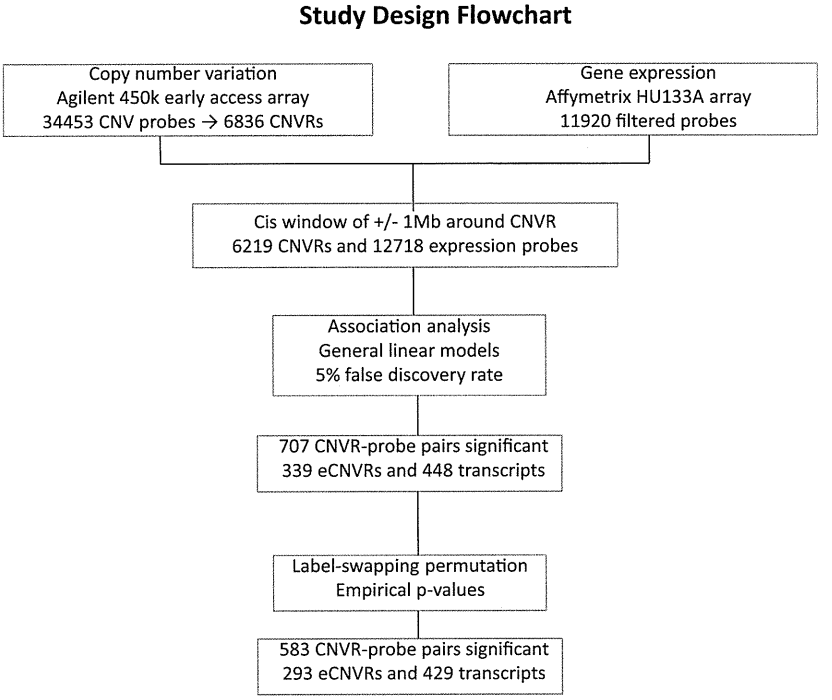

Fig. 1. Flowchart of study design and results.

Please cite this article in press as: Mehta, D., et al., Comprehensive survey of CNVs influencing gene expression in the human brain and its implications for pathophysiology. Neurosci. Res. (2013), http://dx.doi.org/10.1016/j.neures.2013.10.009

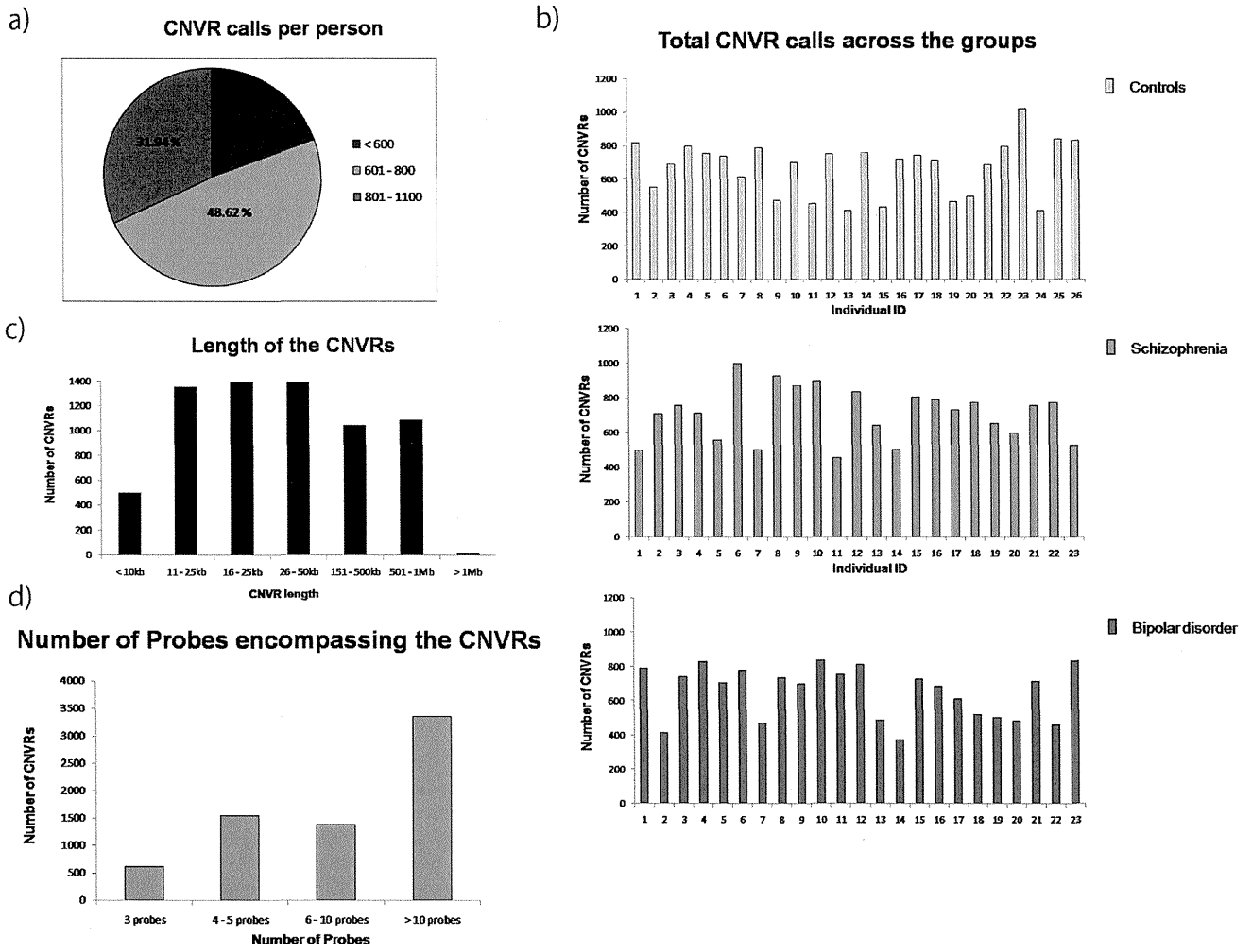

Fig. 2. Description of the copy number variable regions (CNVRs). (a) Total CNVR calls across all samples. (b) Total CNVR calls across controls, schizophrenia and bipolar disorder patients. (c) Distribution of the length of the CNVRs. (d) Number of probes encompassing CNVRs.

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 ±1Mb 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,

Please cite this article in press as: Mehta, D., et al., Comprehensive survey of CNVs influencing gene expression in the human brain and its implications for pathophysiology. Neurosci. Res. (2013), http://dx.doi.org/10.1016/j.neures.2013.10.009

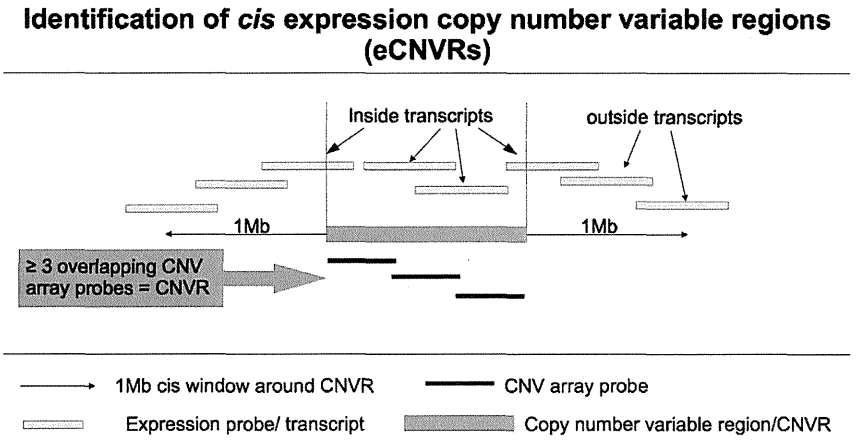

Fig. 3. Schematic figure of the cis window chosen for association testing of expression copy number variable regions (eCNVRs).

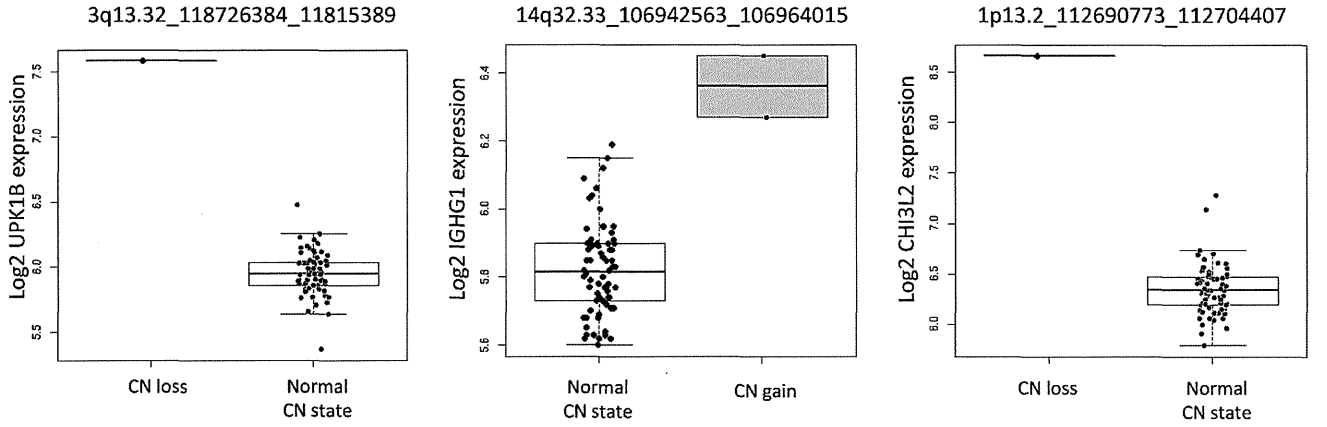

Fig. 4. *Box plots of associations between CNVRs and gene expression profiles of significant eCNVR.*

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

Please cite this article in press as: Mehta, D., et al., Comprehensive survey of CNVs influencing gene expression in the human brain and its implications for pathophysiology. Neurosci. Res. (2013), http://dx.doi.org/10.1016/j.neures.2013.10.009

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

Please cite this article in press as: Mehta, D., et al., Comprehensive survey of CNVs influencing gene expression in the human brain and its implications for pathophysiology. Neurosci. Res. (2013), http://dx.doi.org/10.1016/j.neures.2013.10.009

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

Please cite this article in press as: Mehta, D., et al., Comprehensive survey of CNVs influencing gene expression in the human brain and its implications for pathophysiology. Neurosci. Res. (2013), http://dx.doi.org/10.1016/j.neures.2013.10.009

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

Please cite this article in press as: Mehta, D., et al., Comprehensive survey of CNVs influencing gene expression in the human brain and its implications for pathophysiology. Neurosci. Res. (2013), http://dx.doi.org/10.1016/j.neures.2013.10.009

Table 3 (*Continued*)

CNVR	CNVR type	People	ProbeSetID	Psychiatric Trait locus	Permutations *p* value	CNVR_Start	CNVR_Stop	GeneSymbol
q11.21_21709612_21905954	Loss	con*	212180_at	22q11.2[4,5]	0.003267	21709612	21905954	CRKL
q11.21_21709612_21905954	Loss	con*	218492_s_at	22q11.2[4,5]	0.003155	21709612	21905954	THAP7
q11.21_21709612_21905954	Loss	con*	205881_at	22q11.2[4,5]	0.009217	21709612	21905954	ZNF74
q11.21_21709612_21905954	Loss	con*	207081_s_at	22q11.2[4,5]	3.15E-005	21709612	21905954	PI4KA
q11.21_21709612_21905954	Loss	con*	211147_s_at	22q11.2[4,5]	0.015425	21709612	21905954	P2RX6
q11.21_21709612_21905954	Loss	con*	221838_at	22q11.2[4,5]	0.016453	21709612	21905954	KLHL22
q11.21_21709612_21905954	Loss	con*	222141_at	22q11.2[4,5]	1.02E-011	21709612	21905954	KLHL22
q11.21_21709612_21905954	Loss	con*	200684_s_at	22q11.2[4,5]	0.007571	21709612	21905954	UBE2L3
q11.21_21709612_21905954	Loss	con*	205576_at	22q11.2[4,5]	0.007335	21709612	21905954	SERPIND1
q11.21_21709612_21905954	Loss	con*	212271_at	22q11.2[4,5]	0.000261	21709612	21905954	MAPK1
q11.21_21711906_21905954	Gain	bp	214406_s_at	22q11.2[4,5]	0.003467	21711906	21905954	SLC7A4
q11.21_21711906_21905954	Gain	bp	215048_at	22q11.2[4,5]	0.018037	21711906	21905954	ZNF280B
q11.21_21711906_21905954	Gain	bp	221349_at	22q11.2[4,5]	0.002569	21711906	21905954	VPREB1
q11.22_22605295_22630082	Gain	bp	211655_at	22q11.2[4,5]	0.018087	22605295	22630082	LOC100287927
q11.22_23241489_23252126	Gain	sz	215036_at	22q11.2[4,5]	0.016135	23241489	23252126	
q11.22_23242646_23248046	Gain	sz	217180_at	22q11.2[4,5]	0.017219	23242646	23248046	
q11.23_23805014_23825653	Gain	sz	203815_at	22q11.2[4,5]	0.007688	23805014	23825653	GSTT1
q11.23_24271987_24343125	Loss	sz	203815_at	22q11.2[4,5]	0.007688	24271987	24343125	GSTT1
q11.23_24278085_24341961	Loss	sz	202624_s_at	22q11.2[4,5]	0.006683	24278085	24341961	CABIN1
q11.23_24278085_24341961	Loss	sz	214623_at	22q11.2[4,5]	0.008420	24278085	24341961	FBXW4P1
q11.23_24291835_24345621	Gain	con*	204993_at	22q11.2[4,5]	5.26E-005	24291835	24345621	GNAZ
q11.23_24291835_24345621	Gain	con*	217668_at	22q11.2[4,5]	0.000836	24291835	24345621	C22orf36
q11.23_24291835_24345621	Gain	con*	215202_at	22q11.2[4,5]	0.001560	24291835	24345621	LOC91316
q11.23_24291835_24345621	Gain	con*	203878_s_at	22q11.2[4,5]	0.017484	24291835	24345621	MMP11
q11.23_24291835_24345621	Gain	con*	205582_s_at	22q11.2[4,5]	4.46E-005	24291835	24345621	GGT5
q11.23_24291835_24345621	Gain	con*	202929_s_at	22q11.2[4,5]	0.003827	24291835	24345621	DDT
q11.23_24291835_24345621	Gain	con*	211471_s_at	22q11.2[4,5]	0.000417	24291835	24345621	RAB36
q11.23_24291835_24345621	Gain	con*	217871_s_at	22q11.2[4,5]	3.15E-005	24291835	24345621	MIF
q11.23_24329367_24398674	Loss	bp	212167_s_at	22q11.2[4,5]	0.018093	24329367	24398674	SMARCB1
q11.23_24341917_24400174	Loss	bp	207215_at	22q11.2[4,5]	0.015425	24341917	24400174	GSTTP1
q11.23_24341917_24400174	Loss	bp	203877_at	22q11.2[4,5]	0.002009	24341917	24400174	MMP11
q11.23_24341917_24400174	Loss	bp	221108_at	22q11.2[4,5]	0.000987	24341917	24400174	C22orf43
q11.23_24341917_24400174	Loss	bp	206532_at	22q11.2[4,5]	2.32E-006	24341917	24400174	
q11.23_24341917_24400174	Loss	bp	215816_at	22q11.2[4,5]	3.15E-005	24341917	24400174	LOC91316
q11.23_24341917_24400174	Loss	bp	220507_s_at	22q11.2[4,5]	0.017079	24341917	24400174	UPB1
q11.23_24344364_24398674	Loss	con	211471_s_at	22q11.2[4,5]	0.016695	24344364	24398674	RAB36
q11.23_24344364_24398674	Loss	con	204993_at	22q11.2[4,5]	0.003924	24344364	24398674	GNAZ
q11.23_24344364_24398674	Loss	con	217871_s_at	22q11.2[4,5]	0.013295	24344364	24398674	MIF
q11.23_24344364_24398674	Loss	con	202315_s_at	22q11.2[4,5]	0.001560	24344364	24398674	BCR
q11.23_24344364_24398674	Loss	con	217223_s_at	22q11.2[4,5]	0.000426	24344364	24398674	BCR
q11.23_24344364_24398674	Loss	con	37652_at	22q11.2[4,5]	0.009212	24344364	24398674	CABIN1
q11.23_24356690_24369021	Gain	con*	215202_at	22q11.2[4,5]	0.001560	24356690	24369021	LOC91316
q11.23_24356690_24369021	Gain	con*	217668_at	22q11.2[4,5]	0.000836	24356690	24369021	C22orf36
q11.23_24356690_24369021	Gain	con*	204993_at	22q11.2[4,5]	5.26E-005	24356690	24369021	GNAZ
q11.23_24356690_24369021	Gain	con*	211471_s_at	22q11.2[4,5]	0.000417	24356690	24369021	RAB36
q11.23_24356690_24369021	Gain	con*	205582_s_at	22q11.2[4,5]	4.46E-005	24356690	24369021	GGT5
q11.23_24356690_24369021	Gain	con*	217871_s_at	22q11.2[4,5]	3.15E-005	24356690	24369021	MIF
q11.23_24356690_24369021	Gain	con*	202929_s_at	22q11.2[4,5]	0.003827	24356690	24369021	DDT
q11.23_24356690_24369021	Gain	con*	203878_s_at	22q11.2[4,5]	0.017484	24356690	24369021	MMP11
q11.23_25756694_25775816	Gain	bp	220507_s_at	22q11.2[4,5]	0.017079	25756694	25775816	UPB1
q11.23_25756694_25775816	Gain	bp	204183_s_at	22q11.2[4,5]	0.001594	25756694	25775816	ADRBK2
q11.23_25756694_25775816	Gain	bp	204184_s_at	22q11.2[4,5]	0.002150	25756694	25775816	ADRBK2

x = in LCLs and xxx = in brain tissue; sz = schizophrenia, bp = bipolar, con = control, con* = control outlier C_15.

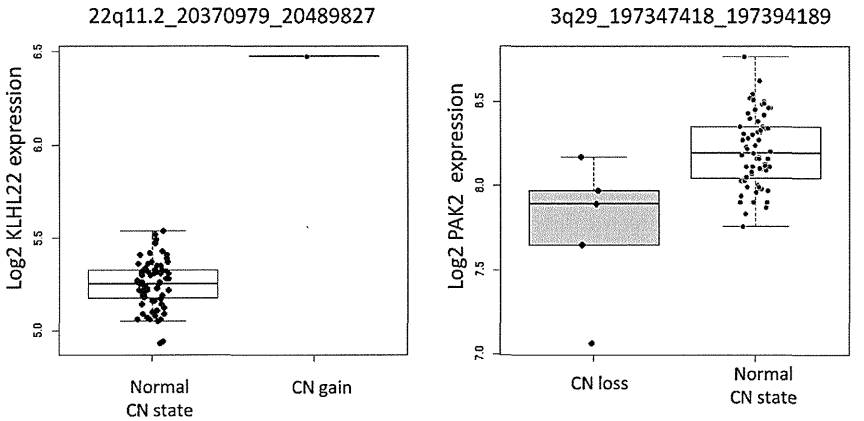

Fig. 5. Significant eCNVRs previously shown to be associated with schizophrenia or autism-spectrum disorders within candidate loci 3q29 and 22q11.2.

Please cite this article in press as: Mehta, D., et al., Comprehensive survey of CNVs influencing gene expression in the human brain and its implications for pathophysiology. Neurosci. Res. (2013), http://dx.doi.org/10.1016/j.neures.2013.10.009

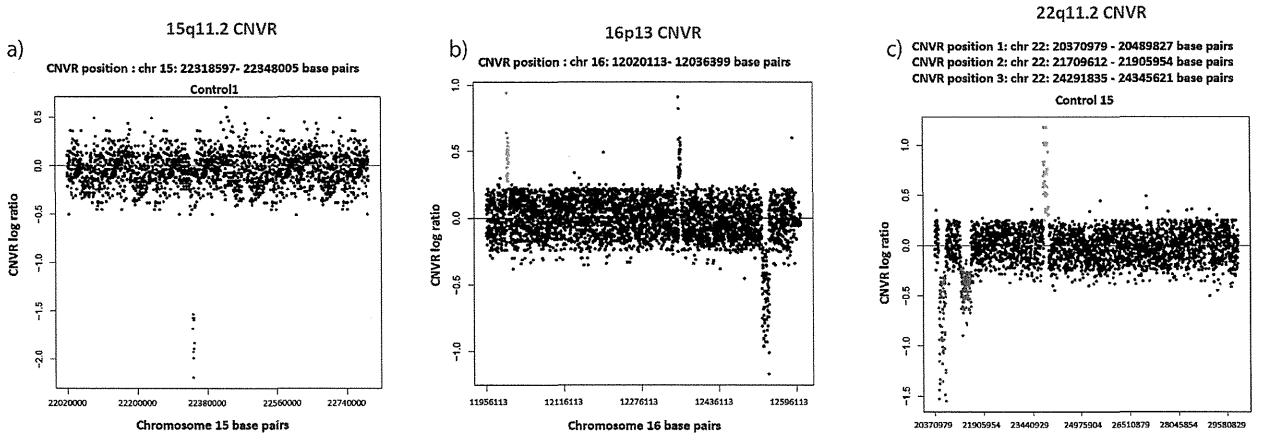

Fig. 6. Log ratio plots of 3 examples of significant eCNVRs within candidate loci 15q11.2, 16p13 and 22q11.2 previously associated with psychiatric disorders. Red dots indicate losses and green dots indicate gains. Blue dots indicate other CNVRs within this individual which were detected but these were not significantly associated with gene expression levels. (a) 15q11.2, (b) 16p13, and (c) 22q11.2.

429 transcripts were significantly associated with CNVR state after corrections for multiple testing and permutation. This corresponded to 583 CNVR-probe pairs (293 unique eCNVRs). Among the eCNVRs, a significant over-representation of rare/low frequency CNVRs ($p = 1.087 \times 10^{-10}$) and gene-harboring/genic CNVRs ($p = 1.4 \times 10^{-6}$) was observed. Overrepresentation of rare/low frequency CNVs among eCNVRs is interesting from an evolutionary point of view. A significant proportion of variance in gene expression could be explained by the eCNVR, with an average of 26% variance across the transcripts. A large proportion of negative correlations observed, demonstrated the complex relationship between CNVs and gene expression. Regulatory mechanisms such as epistasis or auto-regulatory feedback mechanisms at the level of the gene might explain the negative correlations. For instance, deletions that affect silencers or insulator elements can result in increased gene expression of the transcript (Weischenfeldt et al., 2013). Comparisons of the brain eCNVRs identified in the current study to previously reported eCNVRs yielded a 10% overlap, thereby providing a replication for these eCNVRs despite the differences in samples and study design between the studies.

Functional annotation of transcripts associated with CNVRs revealed a significant enrichment of corticotrophin-releasing hormone pathway across all samples, and also upon stratification by cases and controls. However, genes overlapping CNVRs only in cases but not in controls were enriched for glutathione metabolism and oxidative stress. Glutathione is a major antioxidant in the brain and plays a crucial role in protecting against oxidative damage. It is reported that glutathione levels were decreased (Gawryluk et al., 2001) and oxidative stress is enhanced (Ng et al., 2008) in schizophrenia and bipolar disorder, and mood stabilizers increases glutathione S-transferase (Wang et al., 2004). Thus, altered glutathione and oxidative stress pathways due to CNV might be related to pathophysiology of bipolar disorder and schizophrenia.

To test whether the eCNVRs were located within psychiatric phenotype-associated loci, we performed a literature search to identify CNVRs robustly associated with psychiatric diseases and systematically checked these loci ($n = 8$ unique loci). The 293 significant eCNVRs identified in this study included 72 (24.6%) psychiatric-disorder associated eCNVRs within these 8 loci, indicating that copy number variants in these loci might be directly involved in transcriptional regulation in the brain. These eCNVRs encompassed 7 (16p11.2, 1q21.1, 22q11.2, 3q29, 15q11.2, 17q12 and 16p13.1) of the 8 tested loci. Of the 72 eCNVRs, 19 CNVRs were identified only in schizophrenia patients and 21 CNVRs were observed only in bipolar disorder patients. A total of 19 CNVRs were found only in controls of which 7 were found in a single control individual (C_15). For C_15, the CNVR burden was within the range of that detected across all other samples and by technically validating two CNVRs harbored by this individual, we excluded the possibility of sample contamination or hybridization artifacts. This control individual however due to unknown or unexplained reasons harbored several of the known bipolar disorder and schizophrenia-associated CNVRs. For the 15q13.3 region, we did not identify any CNVRs associated with gene expression levels.

Recently, Ye and colleagues identified that CNVs in 1q21.1 and 22q11.2 were significantly associated with expression levels of nearby transcripts in dorsolateral prefrontal cortex (Ye et al., 2012). We found an association between a CNVR in 1q21.1 and CHD1L as reported by Ye and colleagues and in the current study the same CNVR was also associated with gene expression levels of FMO5, PRKAB2, RNF115 and ITGA10. These CNVRs were present in 2 bipolar disorder patients and control C_15. Additionally, we identified 8 further CNVRs in 1q21.1 (6 only in bipolar patients, one in 2 bipolar patients, one control and control C_15 and one only in control C_15) significantly associated with gene expression levels of PDE4DIP, SEC22B, RBM8A, PRKAB2 and ITGA10. In line with Ye et al. (2012), we observed a significant association of a 22q11.2 CNVR in control C_15 with COMT gene expression for two separate gene expression probes. In addition, our data pointed also toward the PI4KA gene within this locus whose expression was significantly associated with 2 CNVRs (both in control C_15) in the 22q11.2 locus. The initial study by Saito and colleagues (Saito et al., 2003) identified a link between PI4KA and 22q11.2-linked psychiatric disorders. The PI4KA gene encodes a phosphatidylinositol (PI) 4-kinase which catalyses the first committed step in the biosynthesis of phosphatidylinositol 4,5-bisphosphate. Incorporating the results of all association of PI4KA with schizophrenia till date has yielded mixed results and the link between PI4KA and psychiatric disorders remains unclear (Kanahara et al., 2009; Saito et al., 2003; Vorstman et al., 2009). This is the first report highlighting a functional link between CNVRs within the 22q11.2 locus and PI4KA gene expression in the human brain, suggesting that PI4KA might indeed be related to 22q11.2-related psychiatric diseases. In summary, results of the current study replicate the findings by Ye and colleagues that 1q21.1 and 22q11.2 may be involved in pathophysiology of psychiatric disorders by affecting gene expression levels in the brain.

For an additional five candidate regions reported to be associated with schizophrenia and/or autism-spectrum disorders (3q29, 15q11.2, 16p11.2, 16p13.1 and 17q12), for the first time we

Please cite this article in press as: Mehta, D., et al., Comprehensive survey of CNVs influencing gene expression in the human brain and its implications for pathophysiology. Neurosci. Res. (2013), http://dx.doi.org/10.1016/j.neures.2013.10.009

identified significant functional influence of CNVRs on prefrontal cortex gene expression, implicating that these loci confer a risk of psychiatric disorders by affecting gene expression in the brain. Of note was the finding of 6 CNVRs within the 16p11.2 locus that significantly influenced gene expression profiles of several transcripts including CORO1A, TAOK2, DOC2A, SEPHS2 and CDIPT transcripts in the human prefrontal cortex. Both deletions and duplications within the 16p11.2 region have been significantly associated with schizophrenia, autism and autism-spectrum disorders in several studies (Levinson et al., 2011; Luo et al., 2012; McCarthy et al., 2009; Weiss et al., 2008).

The current study has several strengths and limitations. On one hand, due to the small sample size, the power of this study is limited and replication of these findings in larger cohorts is warranted. Nonetheless, several of the results reported in this study overlap with previous reports, hence for these findings our study provides a replication of the previous results. Furthermore, 9.3% of autosomal CNVRs detected on the SNP array were successfully detected in the same individuals using the CNV arrays, thereby providing a technical validation of these data. Also, we acknowledge that possible confounding effects of medication or smoking or other illness-related factors are difficult to account for and might influence the gene expression profiles. To the best of our knowledge this is the most comprehensive genome-wide CNV-gene expression association analysis performed so far and the first genome-wide hypothesis-free study assessing the influence of rare/low frequency CNVs on gene expression in the human brain. Other strengths of this study include assessment of brain tissue which is more relevant for psychiatric diseases and utilization of brain samples with high pH levels to increase reliability of the data.

In conclusion, we used a hypothesis-free approach to identify brain CNVRs which significantly influence genome-wide gene expression levels of nearby transcripts. Such an integrative approach is important to prioritize functional CNVs which exhibit downstream consequences at the gene expression level over other CNVs. This study demonstrates that CNVRs influencing gene expression in the human prefrontal cortex are significantly enriched for rare/low frequency CNVs and gene harboring CNVs. Our results replicate previous findings of associations at 1q21.1 and 22q11.2 regions and suggest the possible role of candidates within the 3q29, 15q11.2, 16p11.2, 16p13.1 and 17q12 loci in schizophrenia and bipolar disorder. Future studies surveying different types of genetic variation in diverse tissues are required to fully comprehend human phenotypic diversity and disease.

Conflict of interest

None declared.

Acknowledgements

We thank the Research Resource Center, RIKEN BSI, for technical assistance.

This study is supported by a Grant-in-aid for Scientific Research on Innovative Areas (Comprehensive Brain Science Network) from the Ministry of Education, Science, Sports and Culture of Japan, Grant-in-aid from Ministry of Health, Labour and Welfare, and CREST from the Japan Science and Technology Agency.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.neures.2013.10.009.

References

Altar, C.A., Hunt, R.A., Jurata, L.W., Webster, M.J., Derby, E., Gallagher, P., Lemire, A., Brockman, J., Laeng, P., 2008. Insulin, IGF-1, and muscarinic agonists modulate schizophrenia-associated genes in human neuroblastoma cells. Biol. Psychiatry 64, 1077–1087.

Gawryluk, J.W., Wang, J.F., Andreazza, A.C., Shao, L., Young, L.T., 2001. Decreased levels of glutathione, the major brain antioxidant, in post-mortem prefrontal cortex from patients with psychiatric disorders. Int. J. Neuropsychopharmacol. 14, 123–130.

Gibbs, J.R., van der Brug, M.P., Hernandez, D.G., Traynor, B.J., Nalls, M.A., Lai, S.L., Arepalli, S., Dillman, A., Rafferty, I.P., Troncoso, J., Johnson, R., Zielke, H.R., Ferrucci, L., Longo, D.L., Cookson, M.R., Singleton, A.B., 2010. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. PLoS Genet. 6, e1000952.

Guilmatre, A., Dubourg, C., Mosca, A.L., Legallic, S., Goldenberg, A., Drouin-Garraud, V., Layet, V., Rosier, A., Briault, S., Bonnet-Brilhault, F., Laumonnier, F., Odent, S., Le Vacon, G., Joly-Helas, G., David, V., Bendavid, C., Pinoit, J.M., Henry, C., Impallomeni, C., Germano, E., Tortorella, G., Di Rosa, G., Barthelemy, C., Andres, C., Faivre, L., Frebourg, T., Saugier Veber, P., Campion, D., 2009. Recurrent rearrangements in synaptic and neurodevelopmental genes and shared biologic pathways in schizophrenia, autism, and mental retardation. Arch. Gen. Psychiatry 66, 947–956.

Haraksingh, R.R., Abyzov, A., Gerstein, M., Urban, A.E., Snyder, M., 2011. Genome-wide mapping of copy number variation in humans: comparative analysis of high resolution array platforms. PLoS ONE 6, e27859.

Hedges, D.J., Hamilton-Nelson, K.L., Sacharow, S.J., Nations, L., Beecham, G.W., Kozhekbaeva, Z.M., Butler, B.L., Cukier, H.N., Whitehead, P.L., Ma, D., Jaworski, J.M., Nathanson, L., Lee, J.M., Hauser, S.L., Oksenberg, J.R., Cuccaro, M.L., Haines, J.L., Gilbert, J.R., Pericak-Vance, M.A., 2012. Evidence of novel fine-scale structural variation at autism spectrum disorder candidate loci. Mol. Autism 3, 2.

Ingason, A., Rujescu, D., Cichon, S., Sigurdsson, E., Sigmundsson, T., Pietilainen, O.P., Buizer-Voskamp, J.E., Strengman, E., Francks, C., Muglia, P., Gylfason, A., Gustafsson, O., Olason, P.I., Steinberg, S., Hansen, T., Jakobsen, K.D., Rasmussen, H.B., Giegling, I., Moller, H.J., Hartmann, A., Crombie, C., Fraser, G., Walker, N., Lonnqvist, J., Suvisaari, J., Tuulio-Henriksson, A., Bramon, E., Kiemeney, L.A., Franke, B., Murray, R., Vassos, E., Toulopoulou, T., Muhleisen, T.W., Tosato, S., Ruggeri, M., Djurovic, S., Andreassen, O.A., Zhang, Z., Werge, T., Ophoff, R.A., Investigators, G., Rietschel, M., Nothen, M.M., Petursson, H., Stefansson, H., Peltonen, L., Collier, D., Stefansson, K., St Clair, D.M., 2011. Copy number variations of chromosome 16p13.1 region associated with schizophrenia. Mol. Psychiatry 16, 17–25.

International Schizophrenia Consortium., 2008. Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature 455, 237–241.

Ionita-Laza, I., Rogers, A.J., Lange, C., Raby, B.A., Lee, C., 2009. Genetic association analysis of copy-number variation (CNV) in human disease pathogenesis. Genomics 93, 22–26.

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

Iwamoto, K., Ueda, J., Bundo, M., Kojima, T., Kato, T., 2011. Survey of the effect of genetic variations on gene expression in human prefrontal cortex and its application to genetics of psychiatric disorders. Neurosci. Res. 70, 238–242.

Kanahara, N., Iyo, M., Hashimoto, K., 2009. Failure to confirm the association between the PIK4CA gene and schizophrenia in a Japanese population. Am. J. Med. Genet. B: Neuropsychiatr. Genet. 150B, 450–452.

Lanktree, M.B., Anand, S.S., Yusuf, S., Hegele, R.A., Investigators, S., 2009. Replication of genetic associations with plasma lipoprotein traits in a multiethnic sample. J. Lipid Res. 50, 1487–1496.

Levinson, D.F., Duan, J., Oh, S., Wang, K., Sanders, A.R., Shi, J., Zhang, N., Mowry, B.J., Olincy, A., Amin, F., Cloninger, C.R., Silverman, J.M., Buccola, N.G., Byerley, W.F., Black, D.W., Kendler, K.S., Freedman, R., Dudbridge, F., Pe'er, I., Hakonarson, H., Bergen, S.E., Fanous, A.H., Holmans, P.A., Gejman, P.V., 2011. Copy number variants in schizophrenia: confirmation of five previous findings and new evidence for 3q29 microdeletions and VIPR2 duplications. Am. J. Psychiatry 168, 302–316.

Lipson, D., Aumann, Y., Ben-Dor, A., Linial, N., Yakhini, Z., 2006. Efficient calculation of interval scores for DNA copy number data analysis. J. Comput. Biol. 13, 215–228.

Luo, R., Sanders, S.J., Tian, Y., Voineagu, I., Huang, N., Chu, S.H., Klei, L., Cai, C., Ou, J., Lowe, J.K., Hurles, M.E., Devlin, B., State, M.W., Geschwind, D.H., 2012. Genome-wide transcriptome profiling reveals the functional impact of rare de novo and recurrent CNVs in autism spectrum disorders. Am. J. Hum. Genet. 91, 38–55.

McCarthy, S.E., Makarov, V., Kirov, G., Addington, A.M., McClellan, J., Yoon, S., Perkins, D.O., Dickel, D.E., Kusenda, M., Krastoshevsky, O., Krause, V., Kumar, R.A., Grozeva, D., Malhotra, D., Walsh, T., Zackai, E.H., Kaplan, P., Ganesh, J., Krantz, I.D., Spinner, N.B., Roccanova, P., Bhandari, A., Pavon, K., Lakshmi, B., Leotta, A., Kendall, J., Lee, Y.H., Vacic, V., Gary, S., Iakoucheva, L.M., Crow, T.J., Christian, S.L., Lieberman, J.A., Stroup, T.S., Lehtimaki, T., Puura, K., Haldeman-Englert, C., Pearl, J., Goodell, M., Willour, V.L., Derosse, P., Steele, J., Kassem, L., Wolff, J., Chitkara, N., McMahon, F.J., Malhotra, A.K., Potash, J.B., Schulze, T.G., Nothen, M.M., Cichon, S., Rietschel, M., Leibenluft, E., Kustanovich, V., Lajonchere, C.M., Sutcliffe, J.S., Skuse, D., Gill, M., Gallagher, L., Mendell, N.R., Wellcome Trust Case Control Consortium, Craddock, N., Owen, M.J., O'Donovan, M.C., Shaikh, T.H., Susser, E., Delisi, L.E., Sullivan, P.F., Deutsch, C.K., Rapoport, J., Levy, D.L., King, M.C., Sebat, J., 2009. Microduplications of 16p11.2 are associated with schizophrenia. Nat. Genet. 41, 1223–1227.

Please cite this article in press as: Mehta, D., et al., Comprehensive survey of CNVs influencing gene expression in the human brain and its implications for pathophysiology. Neurosci. Res. (2013), http://dx.doi.org/10.1016/j.neures.2013.10.009

Moreno-De-Luca, D., Consortium, S., Mulle, J.G., Simons Simplex Collection Genetics Consortium, Kaminsky, E.B., Sanders, S.J., GeneStar, Myers, S.M., Adam, M.P., Pakula, A.T., Eisenhauer, N.J., Uhas, K., Weik, L., Guy, L., Care, M.E., Morel, C.F., Boni, C., Salbert, B.A., Chandrareddy, A., Demmer, L.A., Chow, E.W., Surti, U., Aradhya, S., Pickering, D.L., Golden, D.M., Sanger, W.G., Aston, E., Brothman, A.R., Gliem, T.J., Thorland, E.C., Ackley, T., Iyer, R., Huang, S., Barber, J.C., Crolla, J.A., Warren, S.T., Martin, C.L., Ledbetter, D.H., 2010. Deletion 17q12 is a recurrent copy number variant that confers high risk of autism and schizophrenia. Am. J. Hum. Genet. 87, 618–630.

Ng, F., Berk, M., Dean, O., Bush, A.I., 2008. Oxidative stress in psychiatric disorders: evidence base and therapeutic implications. Int. J. Neuropsychopharmacol. 11, 851–876.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575.

Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W., Cho, E.K., Dallaire, S., Freeman, J.L., Gonzalez, J.R., Gratacos, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J.R., Marshall, C.R., Mei, R., Montgomery, L., Nishimura, K., Okamura, K., Shen, F., Somerville, M.J., Tchinda, J., Valsesia, A., Woodwark, C., Yang, F., Zhang, J., Zerjal, T., Zhang, J., Armengol, L., Conrad, D.F., Estivill, X., Tyler-Smith, C., Carter, N.P., Aburatani, H., Lee, C., Jones, K.W., Scherer, S.W., Hurles, M.E., 2006. Global variation in copy number in the human genome. Nature 444, 444–454.

Saito, T., Stopkova, P., Diaz, L., Papolos, D.F., Boussemart, L., Lachman, H.M., 2003. Polymorphism screening of PIK4CA: possible candidate gene for chromosome 22q11-linked psychiatric disorders. Am. J. Med. Genet. B: Neuropsychiatr. Genet. 116B, 77–83.

Schlattl, A., Anders, S., Waszak, S.M., Huber, W., Korbel, J.O., 2011. Relating CNVs to transcriptome data at fine resolution: assessment of the effect of variant size, type, and overlap with functional regions. Genome Res. 21, 2004–2013.

Stefansson, H., Rujescu, D., Cichon, S., Pietilainen, O.P., Ingason, A., Steinberg, S., Fossdal, R., Sigurdsson, E., Sigmundsson, T., Buizer-Voskamp, J.E., Hansen, T., Jakobsen, K.D., Muglia, P., Francks, C., Matthews, P.M., Gylfason, A., Halldorsson, B.V., Gudbjartsson, D., Thorgeirsson, T.E., Sigurdsson, A., Jonasdottir, A., Jonasdottir, A., Bjornsson, A., Mattiasdottir, S., Blondal, T., Haraldsson, M., Magnusdottir, B.B., Giegling, I., Moller, H.J., Hartmann, A., Shianna, K.V., Ge, D., Need, A.C., Crombie, C., Fraser, G., Walker, N., Lonnqvist, J., Suvisaari, J., Tuulio-Henriksson, A., Paunio, T., Toulopoulou, T., Bramon, E., Di Forti, M., Murray, R., Ruggeri, M., Vassos, E., Tosato, S., Walshe, M., Li, T., Vasilescu, C., Muhleisen, T.W., Wang, A.G., Ullum, H., Djurovic, S., Melle, I., Olesen, J., Kiemeney, L.A., Franke, B., Group Sabatti, C., Freimer, N.B., Gulcher, J.R., Thorsteinsdottir, U., Kong, A., Andreassen, O.A., Ophoff, R.A., Georgi, A., Rietschel, M., Werge, T., Petursson, H., Goldstein, D.B., Nothen, M.M., Peltonen, L., Collier, D.A., St Clair, D., Stefansson, K., 2008. Large recurrent microdeletions associated with schizophrenia. Nature 455, 232–236.

Stranger, B.E., Forrest, M.S., Dunning, M., Ingle, C.E., Beazley, C., Thorne, N., Redon, R., Bird, C.P., de Grassi, A., Lee, C., Tyler-Smith, C., Carter, N., Scherer, S.W., Tavare, S., Deloukas, P., Hurles, M.E., Dermitzakis, E.T., 2007. Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 315, 848–853.

Vorstman, J.A., Chow, E.W., Ophoff, R.A., van Engeland, H., Beemer, F.A., Kahn, R.S., Sinke, R.J., Bassett, A.S., 2009. Association of the PIK4CA schizophrenia-susceptibility gene in adults with the 22q11.2 deletion syndrome. Am. J. Med. Genet. B: Neuropsychiatr. Genet. 150B, 430–433.

Wang, J.F., Shao, L., Sun, X., Young, L.T., 2004. Glutathione S-transferase is a novel target for mood stabilizing drugs in primary cultured neurons. J. Neurochem. 88, 1477–1484.

Weischenfeldt, J., Symmons, O., Spitz, F., Korbel, J.O., 2013. Phenotypic impact of genomic structural variation: insights from and for human disease. Nat. Rev. Genet. 142, 125–138.

Weiss, L.A., Shen, Y., Korn, J.M., Arking, D.E., Miller, D.T., Fossdal, R., Saemundsen, E., Stefansson, H., Ferreira, M.A., Green, T., Platt, O.S., Ruderfer, D.M., Walsh, C.A., Altshuler, D., Chakravarti, A., Tanzi, R.E., Stefansson, K., Santangelo, S.L., Gusella, J.F., Sklar, P., Wu, B.L., Daly, M.J., Autism, C., 2008. Association between microdeletion and microduplication at 16p11.2 and autism. N. Engl. J. Med. 358, 667–675.

Ye, T., Lipska, B.K., Tao, R., Hyde, T.M., Wang, L., Li, C., Choi, K.H., Straub, R.E., Kleinman, J.E., Weinberger, D.R., 2012. Analysis of copy number variations in brain DNA from patients with schizophrenia and other psychiatric disorders. Biol. Psychiatry 72, 651–654.

Zhang, D., Cheng, L., Qian, Y., Alliey-Rodriguez, N., Kelsoe, J.R., Greenwood, T., Nievergelt, C., Barrett, T.B., McKinney, R., Schork, N., Smith, E.N., Bloss, C., Nurnberger, J., Edenberg, H.J., Foroud, T., Sheftner, W., Lawson, W.B., Nwulia, E.A., Hipolito, M., Coryell, W., Rice, J., Byerley, W., McMahon, F., Schulze, T.G., Berrettini, W., Potash, J.B., Belmonte, P.L., Zandi, P.P., McInnis, M.G., Zollner, S., Craig, D., Szelinger, S., Koller, D., Christian, S.L., Liu, C., Gershon, E.S., 2009. Singleton deletions throughout the genome increase risk of bipolar disorder. Mol. Psychiatry 14, 376–380.

Please cite this article in press as: Mehta, D., et al., Comprehensive survey of CNVs influencing gene expression in the human brain and its implications for pathophysiology. Neurosci. Res. (2013), http://dx.doi.org/10.1016/j.neures.2013.10.009

Neuropharmacology xxx (2014) 1–7

Contents lists available at ScienceDirect

Neuropharmacology

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

Invited review

Comprehensive DNA methylation and hydroxymethylation analysis in the human brain and its implication in mental disorders

Tadafumi Kato [a,*], Kazuya Iwamoto [b]

[a] Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan
[b] Department of Molecular Psychiatry, Graduate School of Medicine, University of Tokyo, Tokyo 113-8655, Japan

ARTICLE INFO

Article history:
Received 3 October 2013
Received in revised form
14 December 2013
Accepted 18 December 2013

Keywords:
Epigenetics
DNA methylation
DNA hydroxymethylation
Mental disorders

ABSTRACT

Covalent modifications of nucleotides, such as methylation or hydroxymethylation of cytosine, regulate gene expression. Early environmental risk factors play a role in mental disorders in adulthood. This may be in part mediated by epigenetic DNA modifications. Methods for comprehensive analysis of DNA methylation and hydroxymethylation include DNA modification methods such as bisulfite sequencing, or collection of methylated, hydroxymethylated, or unmethylated DNA by specific binding proteins, antibodies, or restriction enzymes, followed by sequencing or microarray analysis. Results from these experiments should be interpreted with caution because each method gives different result. Cytosine hydroxymethylation has different effects on gene expression than cytosine methylation; methylation of CpG islands is associated with lower gene expression, whereas hydroxymethylation in intragenic regions is associated with higher gene expression. The role of hydroxymethylcytosine is of particular interest in mental disorders because the modification is enriched in the brain and synapse related genes, and it exhibits dynamic regulation during development. Many DNA methylation patterns are conserved across species, but there are also human specific signatures. Comprehensive analysis of DNA methylation shows characteristic changes associated with tissues, brain regions, cell types, and developmental states. Thus, differences in DNA methylation status between tissues, brain regions, cell types, and developmental stages should be considered when the role of DNA methylation in mental disorders is studied. Several disease-associated changes in methylation have been reported: hypermethylation of *SOX10* in schizophrenia, hypomethylation of *HCG9* (HLA complex group 9) in bipolar disorder, hypermethylation of *PRIMA1*, hypermethylation of *SLC6A4* (serotonin transporter) in bipolar disorder, and hypomethylation of *ST6GALNAC1* in bipolar disorder. These findings need to be replicated in different patient populations to be generalized. Further studies including animal experiments are necessary to understand the roles of DNA methylation in mental disorders.

This article is part of a Special Issue entitled 'Neuroepigenetic disorders'.

1. Introduction

It is well known that DNA contains the code for the amino acid sequence of proteins. In addition, the DNA molecule has information on the regulation of gene expression, which is mediated by DNA-protein or DNA–RNA interactions. The amino acid sequence is determined by the sequence of four nucleotides. Similarly, DNA-protein interactions that mediate gene expression regulation are regulated by the covalent modifications of nucleotides. The most studied covalent modification of nucleotides in mammals is methylation of the cytosine residue (Suzuki and Bird, 2008). The process of DNA methylation has been well studied; however, mechanisms of DNA demethylation are still not completely understood (Franchini et al., 2012).

DNA damage produces several types of oxidative DNA adducts including 8-oxoguanine and 8-hydroxyguanine (Cadet et al., 2003), and 5-hydroxymethylcytosine (5hmC) had also been regarded as one of such DNA adducts. Identification of ten-eleven translocation (TET) proteins as the enzymes that catalyze hydroxymethylation (Ito et al., 2010; Tahiliani et al., 2009) and enrichment of 5hmC in brain cells (Kriaucionis and Heintz, 2009) suggested its role in brain function and neuropsychiatric diseases. 5hmC is involved in the demethylation of cytosine (Cortellino et al., 2011; Guo et al., 2011; Hackett et al., 2013; He et al., 2011; Ito et al., 2011; Shen et al., 2013), it also plays a functional role by binding to methyl CpG biding protein 3 (Mbd3) (Yildirim et al., 2011), methyl CpG biding protein 3 (MeCP2) (Mellen et al., 2012), and Uhrf2 (Spruijt et al., 2013). TET proteins further oxidize 5-hmC into 5-formylcytosine (5fC) and 5-carboxylcytosine

* Corresponding author. Tel.: +81 48 467 6949; fax: +81 48 467 6947.
E-mail address: kato@brain.riken.jp (T. Kato).

0028-3908/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.neuropharm.2013.12.019

Please cite this article in press as: Kato, T., Iwamoto, K., Comprehensive DNA methylation and hydroxymethylation analysis in the human brain and its implication in mental disorders, Neuropharmacology (2014), http://dx.doi.org/10.1016/j.neuropharm.2013.12.019

(5caC) (Ito et al., 2011). Recently, 5fC was found to be enriched at poised enhancers in mouse embryonic stem (ES) cells, suggesting its functional significance in gene regulation (Song et al., 2013).

The role of DNA methylation in mental disorders has long been suggested (Petronis, 2010). Epidemiological studies show a role for both genetic and environmental factors in mental disorders. Among environmental factors, early adversities such as childhood abuse or maltreatment are suggested in depression or post-traumatic stress disorder, whereas perinatal problems such as virus infection, malnutrition, and perinatal complication, are suggested in psychoses. However, it is unknown how these early environmental factors affect behavioral phenotypes in adulthood. DNA methylation can be affected by environmental factors, and methylation remains relatively stable over time. Thus, the role of DNA methylation as a mechanism of the effect of early environmental factors on adult mental disorders has drawn attention.

However, there have been no well-replicated findings of altered DNA methylation of candidate genes in mental disorders. It is suggested that genetic association studies of candidate genes frequently encounter false positive findings (Hirschhorn et al., 2002). Recently, there has been a greater focus on genome-wide association analysis rather than candidate gene approaches. The genome-wide approach can also be applied to the study of DNA methylation and hydroxymethylation.

In this review, recent studies on the comprehensive analysis of DNA methylation and hydroxymethylation in the human brain are summarized. A particular focus on the roles for methylation in mental disorders is given. Therefore, animal experiments and studies regarding brain tumors are not discussed here.

2. Methods for the comprehensive analysis of DNA methylation and hydroxymethylation

There are several approaches for genome-wide analysis of DNA methylation or hydroxymethylation. A common method utilizes the modification of cytosine to uracil by sodium bisulfite (Hayatsu et al., 1970). Methylcytosine (mC) is not converted into uracil by sodium bisulfite, allowing the identification of methylated cytosine. Bisulfite sequencing has been widely used for DNA methylation analysis since its discovery in 1970. However, this method cannot discriminate 5hmC from mC. A modified method, however, enables this discrimination (Fig. 1). For mC-specific analysis, called oxidative bisulfite sequencing (oxBS-Seq), 5hmC is selectively oxidized to 5fC, which is then converted to uracil after bisulfite treatment (Booth et al., 2012). Tet-assisted bisulfite sequencing (TAB-Seq) specifically analyzes 5hmC (Yu et al., 2012). In this method, all cytosine modifications except for glucose-protected 5hmC are converted to uracil by first treating with the Tet enzyme followed by bisulfite modification. In addition, specific analysis of 5caC has also been developed by protecting 5caC with 1-ethyl-3-[3-dimethylaminopropyl]-carbodiimide hydrochloride before bisulfite modification (chemical modification assisted bisulfite sequencing; CAB-seq) (Lu et al., 2013). Similarly, a method for base-pair level analysis of 5fC has also been reported wherein 5fC is protected with O-ethylhydroxylamine before modification with bisulfite (Song et al., 2013). Bisulfite-modified DNA is subject to analysis by next-generation sequencing or bead arrays (Bibikova et al., 2006). Bead arrays can determine predefined representative CpG sites for each gene. Reduced representation bisulfite sequencing (RRBS), which can selectively analyze CpG-rich regions, is also often used. RRBS is popular because the cost of whole genome bisulfite sequencing analysis is still high (Meissner et al., 2005). For the analysis of specific CpG sites, other methods such as Sanger sequencing, pyrosequencing, or mass-spectrometry are used.

In the other type of comprehensive analysis method, methylated, hydroxymethylated, or unmethylated DNA is collected using specific binding proteins or antibodies. For example, MBD2b conjugated beads are used to collect methylated DNA. Similarly, DNA containing 5hmC can be collected using streptavidin magnetic beads after glucosylation of 5hmC and subsequent biotinylation. To collect unmodified DNA, unmethylated DNA-specific binding

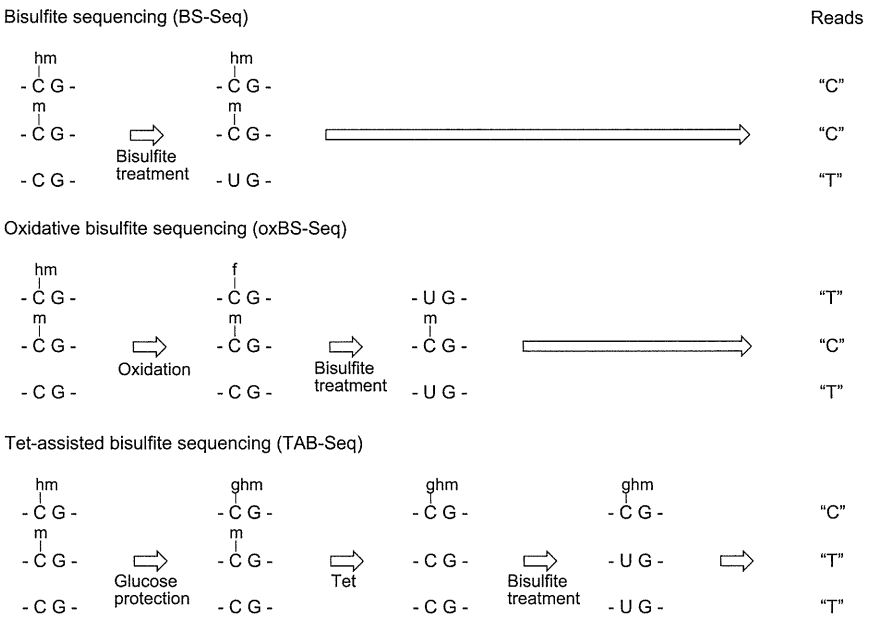

Fig. 1. Specific analysis of methylcytosine and hydroxymethylcytosine. Recently developed modified versions of bisulfite (BS) sequencing enable mC or 5hmC-specific analysis at the base pair resolution. In oxidative BS sequencing (oxBS-Seq) (Booth et al., 2012), 5hmC is selectively oxidized to 5fC. 5fC is converted to uracil after BS treatment. 5hmC reads as T instead of C when sequenced. In TAB-Seq (Yu et al., 2012), genomic DNA is pretreated with glucosyltransferase to modify 5hmC with glucose, which is resistant to BS treatment. Genomic DNA is then treated with the Tet enzyme. Sequence reads of all cytosine modifications except 5hmC would be T, and 5hmC would be C. Note that after BS treatment two other cytosine modifications, 5-carboxylmethycytosine and 5-formylcytosine, would be read as T. mC, methylcytosine; 5hmC, 5-hydroxymethycytosine; C, cytosine; T, thymine; BS, bisulfite; oxBS-Seq, oxidative bisulfite sequencing; TAB-Seq, Tet-assisted bisulfite sequencing.

Please cite this article in press as: Kato, T., Iwamoto, K., Comprehensive DNA methylation and hydroxymethylation analysis in the human brain and its implication in mental disorders, Neuropharmacology (2014), http://dx.doi.org/10.1016/j.neuropharm.2013.12.019

proteins can be used. Similarly, anti-mC antibody or anti-5hmC antibody can be used to collect methylated or hydroxymethylated DNA, respectively. Selective digestion of non-methylated DNA by methylation-sensitive restriction enzymes can also be used to enrich methylated DNA. On the other hand, McrBC, a restriction enzyme that can selectively digest methylated DNA, is used for the enrichment of unmethylated DNA. Glucosylation of 5hmC inhibits the activity of methylation sensitive restriction enzymes, and thus this can also be used to discriminate 5hmC and mC. After the collection of DNA, the samples are subjected to tiling arrays or deep sequencing analysis. The resolution of these methods is not at the base pair level; therefore, tiling arrays and deep sequencing would be equally useful for this analysis.

Both bisulfite sequencing and collection of methylated DNA have strengths and weakness. Bisulfite sequencing can reveal DNA modifications (methylation or hydroxymethylation) at the base pair level, but cannot discriminate between the types of modifications through a standard method. Collection of methylated DNA followed by sequencing can analyze methylation and hydroxymethylation separately, but not at the base pair level. Thus, both methods can be combined to obtain a genome-wide picture of DNA methylation and hydroxymethylation. As discussed above, sodium bisulfite sequencing cannot discriminate DNA methylation and hydroxymethylation, and, therefore, it would be appropriate to refer to these results as "DNA modification" or "DNA (hydroxy)methylation." In the following sections, the results of this method are written as "DNA methylation" for simplicity.

In summary, a number of different experimental techniques are available to analyze DNA methylation and hydroxymethylation status. The results of these analyses are dependent on the methodologies used. Methodological differences should be carefully considered when results are interpreted and compared across different studies.

3. DNA methylation

3.1. Functional significance

DNA methylation has important roles in the regulation of gene expression, imprinting, and X-chromosome inactivation (Bird, 1980). Cytosine methylation predominantly occurs at, but is not restricted to, the CpG site in genomic DNA in mammals. CpG site is less frequent than simple mathematical probability predicts because methylated cytosine can be mutated to thymine during evolution (Bird, 1980). CpG-rich genomic regions, called CpG islands, are frequently found around the transcription start sites. CpG islands of house-keeping genes are generally unmethylated; lower DNA methylation at CpG islands on the promoter is usually associated with higher gene expression (Suzuki and Bird, 2008). Methylation at the region surrounding a CpG island, called the CpG island shore, is involved in tissue differentiation (Doi et al., 2009). DNA methylation of CpG islands in intragenic or intergenic regions is associated with alternative promoter usage (Maunakea et al., 2010). DNA methylation in gene body is related to enhanced transcription (Ball et al., 2009).

In conclusion, DNA methylation can affect a wide-range of cellular functions, and it is hypothesized to play a role in diseases.

3.2. DNA methylation signature of tissues, brain regions, and cell types

An early study using a BAC (bacterial artificial chromosome) microarray identified tissue-specific DNA methylation of *SHANK3*. *SHANK3* was found to be unmethylated and highly expressed in the human brain but not in peripheral blood lymphocyte (Ching et al., 2005).

Using Restriction Landmark Genomic Scanning (RLGS), Ghosh and colleagues searched for brain-specific DNA methylation differences and identified loci showing differential methylation in the human brain. RLGS is a traditional method of comprehensive DNA methylation analysis involving digestion using methylation-sensitive restriction enzymes, followed by two-dimensional electrophoresis. This study demonstrated that *LHX2* is methylated and *CNPY1* is hypomethylated in cerebellum (Ghosh et al., 2010). This study also found clear differential methylation of several loci between gray matter and white matter. The authors suggest that this might be mediated by differential methylation between neurons and glial cells (Ghosh et al., 2010). However, the loci showing differential methylation between gray and white matter were not identified.

Using bead arrays that can examine 1505 CpG sites from 807 genes, Ladd-Acosta and colleagues studied the DNA methylation status of 76 human brain samples including patients with autism and bipolar disorder (Ladd-Acosta et al., 2007). By hierarchical clustering analysis, they clearly showed that DNA methylation status is different between brain regions including cerebral cortex, cerebellum, and pons. DNA methylation differences of five genes, *RASSF1*, *HDAC7A*, *GABRB3*, *EN2*, and *HTR2A* between the cerebral cortex and cerebellum were confirmed in an independent cohort.

To identify the differences in methylation signature between neurons and non-neuronal cells such as glial cells, we separated neurons and non-neurons from human postmortem brains and performed comprehensive DNA methylation analyses using bead arrays of bisulfite modified DNA and tiling array analysis of DNA collected by MBD-conjugated beads (Iwamoto et al., 2011). We found that neurons are hypomethylated, and the DNA methylation status of bulk cortex mostly reflects non-neurons. Genes expressed in astrocytes were methylated in neurons, and genes related to neuronal function were methylated in non-neurons. Interestingly, inter-individual difference of DNA methylation is larger in neurons than in non-neurons. This difference might reflect environment-dependent changes of DNA methylation in neurons.

Considering the differences in DNA methylation status between tissues, brain regions, and cell types, the tissues and cell types to be analyzed are crucial when the role of DNA methylation in mental disorders is studied.

3.3. Developmental aspects

DNA methyltransferases (DNMTs) and MBDs play an important role for de novo and maintenance DNA methylation as well as the recruitment of proteins involved in transcriptional regulation. Expression of these genes undergo complex regulation from early neuronal development to the adult brain, establishing a developmental and cell-type-specific DNA methylation signature in the brain (Yao and Jin, 2013). Importantly, mutations within DNMTs or MBDs are known to cause neurological disorders. For example, mutations in *MECP2* causes Rett syndrome (Chahrour and Zoghbi, 2007) and those in *DNMT3B* lead to immunodeficiency-centromeric instability-facial anomalies syndrome, which is characterized by mental retardation (Hansen et al., 1999). DNA methylation profiles in brain are drastically altered throughout development. Siegmund and colleagues performed a real-time PCR-based quantitative methylation assay of 50 genes in 125 postmortem brains. They identified four typical patterns of changes during development: age-dependent linear increase, biphasic distribution, stochastic accumulation, and a decrease in DNA methylation (Siegmund et al., 2007). This study also found higher DNA methylation of *PAX8* in patients with schizophrenia than in

Please cite this article in press as: Kato, T., Iwamoto, K., Comprehensive DNA methylation and hydroxymethylation analysis in the human brain and its implication in mental disorders, Neuropharmacology (2014), http://dx.doi.org/10.1016/j.neuropharm.2013.12.019

controls (Siegmund et al., 2007). Numata and colleagues examined about 27,000 CpG sites from 14,500 genes using bead arrays in the prefrontal cortex of 108 human subjects of various ages from fetal to elderly. DNA methylation showed drastic changes during the prenatal period, but showed continuous changes during aging. Typical alterations were characterized by prenatal demethylation and increase of methylation with aging (Numata et al., 2012). It was suggested that sex differences in methylation observed in this study are attributable to cross reactions to the sex chromosomes (Chen et al., 2012). Recently, Lister and colleagues performed genome-wide bisulfite sequencing analysis of the mouse and human brain (Lister et al., 2013). This study identified developmentally regulated DNA methylation changes, and found that genome-wide reconfiguration of the DNA methylation pattern occurs during the fetal to young adult stage. They also identified age-dependent accumulations of non-CpG methylation in neurons, but not in non-neurons. Although presence of non-CpG methylation in brain has been previously suggested (Xie et al., 2012; Varley et al., 2013), finding a specific accumulation in neuronal cells implies a unique epigenetic regulation in the brain. This underscores the importance for the consideration of the complexity of brain cell-types in future studies.

These studies show that data should be interpreted in the context of developmental- and aging-associated changes when we study DNA methylation in mental disorders.

3.4. Evolutionary aspects

Xin and colleagues performed comprehensive DNA methylation analysis in human and mouse brains by digestion using methylation-sensitive restriction enzymes, followed by deep sequencing. They identified that DNA methylation is evolutionally conserved in CpG dense regions, regardless of sequence conservation across species (Xin et al., 2011). DNA methylation patterns on the CpG island shore of promoters were different between the prefrontal cortex and auditory cortex. The authors of this study built a database named "MethylomeDB" with their data of DNA methylation in human and mouse brains (Xin et al., 2012). Wang and colleagues identified 150 differentially methylated regions (DMRs) between human and rhesus macaque using the Chip-Seq approach (Wang et al., 2012a). Through extensive validation experiments, they identified four DMRs (*K6IRS2*, *ProSAPiP1*, *ICAM1*, and *RNF32*). Among them, *ICAM1* and *ProSAPiP1* encode neuronal function-related proteins. Another study compared whole-genome bisulfite sequencing data of the prefrontal cortex between humans and chimpanzees (Zeng et al., 2012). They revealed extensive differences in the DNA methylation profile. These changes mostly consisted of hypomethylated genes in the human brain. Importantly, they found enrichment of DMRs in genes related to neurological and psychological disorders.

These studies show partial conservation of DNA methylation patterns across species; however, there are also human specific signatures. Comparative evolution studies of DNA methylation profiles will not only provide insight into the evolution of human-specific traits, but also important candidate genes for neuropsychiatric disorders.

3.5. Disease-associated changes

DNA methylation analysis of the candidate genes have been widely performed using postmortem brains of patients with mental disorders. These included genes coding for BDNF, COMT, serotonin receptors, glutamate receptors, dopamine transporters, and serotonin transporters. Results from these studies have been reviewed elsewhere (Dempster et al., 2013; Nishioka et al., 2012).

Comprehensive gene expression analyses in patients with schizophrenia consistently identified downregulation of oligodendrocyte-related genes. Thus we searched for DNA methylation changes of transcription factors that can explain the global downregulation of oligodendrocyte genes. We found that higher DNA methylation of *SOX10* is related to lower gene expression of many oligodendrocyte-related genes. DNA methylation of *SOX10* was higher in the gray matter than in the white matter (Iwamoto et al., 2005). Consistent with the initial findings, subsequent analysis showed a marked difference in DNA methylation status of *SOX10* between neuronal and non-neuronal cells (Iwamoto et al., 2011).

Mill and colleagues performed comprehensive DNA methylation analysis using DNA microarrays to study human postmortem brains obtained from patients with schizophrenia and bipolar disorder as well as control subjects (Mill et al., 2008). CpG island microarray analysis of DNA after restriction enzyme-based enrichment revealed disease-specific methylation differences in numerous loci, including genes involved in glutamatergic and GABAergic neurotransmission and brain development. Genes involved in mitochondrial function, brain development, and stress response were differentially methylated between groups. The strongest candidate gene obtained from this comprehensive analysis was HLA complex group 9 (*HCG9*). The authors confirmed lower DNA methylation of *HCG9* in bipolar disorder (Kaminsky et al., 2012). Sabunciyan and colleagues performed a comprehensive analysis of DNA methylation in the frontal cortex of patients with major depression ($N = 39$) and controls ($N = 26$) using Comprehensive High-throughput Arrays for Relative Methylation (CHARM), a methylation-sensitive restriction enzyme-based method (Sabunciyan et al., 2012). Among the 224 genes showing robust differential methylation, genes related to neuronal growth and development were enriched. Among the 10 genes that were experimentally validated by pyrosequencing, hypermethylation of *PRIMA1* in patients under depression was most robust. *PRIMA1* encodes a protein that anchors acetylcholinesterase in the neuronal membrane, and thus its decrease might cause enhanced cholinergic neurotransmission. The authors confirmed that acetylcholinesterase-like immunoreactivity was decreased in postmortem brains of patients with major depression. These findings are compatible with the cholinergic hypothesis of depression (Sabunciyan et al., 2012).

Several candidate genes in the brain have been identified from the comprehensive analysis of DNA methylation differences between monozygotic twins discordant for mental disorders. We enriched methylated DNA using MBD-conjugated beads and searched for DNA methylation differences between monozygotic twins discordant for bipolar disorder using tiling arrays (Sugawara et al., 2011). We found that the CpG island shore of *SLC6A4*, which encodes a serotonin transporter, was differentially methylated between twins. Hypomethylation of *SLC6A4* was verified in lymphoblastoid cells and postmortem brain samples of patients with bipolar disorder. Dempster and colleagues found altered DNA methylation of *ST6GALNAC1*, which encodes an enzyme that transfers sialic acid to O-linked N-acetylgalactosamine residues, in monozygotic twins discordant for bipolar disorder or schizophrenia (Dempster et al., 2011). Hypomethylation of this gene was also found in postmortem brains of patients.

As discussed above, comprehensive DNA methylation studies found interesting candidate genes. However, these studies are only a start point to identify the pathophysiological significance of these methylation changes. These findings need to be replicated in different patient populations to be generalized. Further studies

Please cite this article in press as: Kato, T., Iwamoto, K., Comprehensive DNA methylation and hydroxymethylation analysis in the human brain and its implication in mental disorders, Neuropharmacology (2014), http://dx.doi.org/10.1016/j.neuropharm.2013.12.019

including animal experiments should be performed to understand the roles of DNA methylation in mental disorder.

4. Hydroxymethylation

4.1. Changes during development

In contrast to DNA methylation studies, the functional and pathophysiological roles of hydroxymethylation have only recently been proposed. Therefore, there are few studies in human brain on hydroxymethylation. The majority of studies on 5hmC focus on ES cells (Ficz et al., 2011; Pastor et al., 2011; Stroud et al., 2011; Wu et al., 2011; Yu et al., 2012). ES cells were found to contain high levels of 5hmC that decreases after differentiation (Kinney et al., 2011; Szwagierczak et al., 2010). 5hmC increases with age in neuronal cells (Szulwach et al., 2011b).

4.2. Location of 5hmC in genome

Glucosylation-mediated enrichment of hydroxymethylated DNA and subsequent deep sequencing has been used to examine DNA derived from mouse cerebellum. The authors found that 5hmC is enriched in gene bodies and proximal upstream and downstream regions relative to transcription start sites, transcription termination sites, and distal regions (Song et al., 2011). Higher hydroxymethylation in intragenic and proximal regions is associated with higher gene expression. Hydroxymethylation in these genomic regions was higher in the cerebellum of adults compared to postnatal day 7 mice. Increases in hydroxymethylation during aging were enriched in genes related to neurodegenerative disorders, angiogenesis, and hypoxia response. These findings suggest that hydroxymethylation might play a role in age-related neurodegeneration.

Jin and colleagues mapped 5hmC in the frontal lobe by immunoprecipitation with an anti-5hmC antibody (Jin et al., 2011). In human brains, 5hmC was enriched at promoters and gene bodies but absent in non-genic regions. Enrichment of 5hmC in gene bodies was correlated with higher gene expression. This correlation was more prominent than that between mC and gene expression.

4.3. 5hmC in the brain

Consistent with the initial report that 5hmC is enriched in brain cells (Kriaucionis and Heintz, 2009), 5hmC was most abundant in the brain than in other human tissues (Li and Liu, 2011).

Szulwach and colleagues mapped 5hmC using a glucosylation-based enrichment method (Szulwach et al., 2011a) in the human and mouse cerebellum. The level of 5hmC was increased with development, from around 1% (postnatal day 7) to 2.5–5% (one year) in adult mice. 5hmC was enriched in the 5′-UTR (untranslated region) and exons but was depleted in introns. 5hmC was affected by the gene dosage of *MeCP2*. These studies were confirmed and extended in the developing human cerebellum (Wang et al., 2012b). This study found that 5hmC is enriched in exons and 5′-UTRs but depleted in introns. Fetus-specific or adult-specific differentially hydroxymethylated regions overlapped with genes that are enriched with the target sequence of FMRP (fragile X mental retardation protein) and CpG island shores.

Immunohistochemistry analysis detected 5hmC in various cell types in the brain (Orr et al., 2012). However, while 5hmC is robustly detected in neuronal nuclei, some oligodendrocyte nuclei lack in 5hmC immunoreactivity.

Khare and colleagues examined the genomic distribution of 5hmC using enzyme digestion of glucosylated DNA followed by microarray analysis. They found that 5hmC is enriched in genes with synapse related functions in the human and mouse brain. They also found tissue-specific differential distribution of 5hmC at the exon–intron boundary. Constitutive exons contained higher levels of 5hmC than alternatively spliced exons (Khare et al., 2012).

In summary, the role of 5hmC is of particular interest in mental disorders because it is enriched in the brain and in synapse-related genes. Hydroxymethylation of cytosine can occur in positions of the genome different from methylation, and it can regulate gene expression in several ways. 5hmC modifications are also regulated by development and aging. The possible role of 5hmC in mental disorders is a contemporary area of research.

5. Future directions

The role of DNA methylation in neuropsychiatric disorders is currently an active area of investigation. However, the role of hydroxymethylation and other cytosine modifications in neuropsychiatric disorders should be studied as well. As discussed above, patterns and regulation of cytosine modifications in brain cells are more complex than previously expected. Comprehensive studies of cytosine modifications in the human brain have recently begun. The role of epigenetic regulation in the physiology and pathology of the brain should be further studied in detail in the coming decade.

References

Ball, M.P., Li, J.B., Gao, Y., Lee, J.H., LeProust, E.M., Park, I.H., Xie, B., Daley, G.Q., Church, G.M., 2009. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat. Biotechnol. 27, 361–368.

Bibikova, M., Lin, Z., Zhou, L., Chudin, E., Garcia, E.W., Wu, B., Doucet, D., Thomas, N.J., Wang, Y., Vollmer, E., Goldmann, T., Seifart, C., Jiang, W., Barker, D.L., Chee, M.S., Floros, J., Fan, J.B., 2006. High-throughput DNA methylation profiling using universal bead arrays. Genome Res. 16, 383–393.

Bird, A.P., 1980. DNA methylation and the frequency of CpG in animal DNA. Nucleic Acids Res. 8, 1499–1504.

Booth, M.J., Branco, M.R., Ficz, G., Oxley, D., Krueger, F., Reik, W., Balasubramanian, S., 2012. Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science 336, 934–937.

Cadet, J., Douki, T., Gasparutto, D., Ravanat, J.L., 2003. Oxidative damage to DNA: formation, measurement and biochemical features. Mutat. Res. 531, 5–23.

Chahrour, M., Zoghbi, H.Y., 2007. The story of Rett syndrome: from clinic to neurobiology. Neuron 56, 422–437.

Chen, Y.A., Choufani, S., Grafodatskaya, D., Butcher, D.T., Ferreira, J.C., Weksberg, R., 2012. Cross-reactive DNA microarray probes lead to false discovery of autosomal sex-associated DNA methylation. Am. J. Hum. Genet. 91, 762–764.

Ching, T.T., Maunakea, A.K., Jun, P., Hong, C., Zardo, G., Pinkel, D., Albertson, D.G., Fridlyand, J., Mao, J.H., Shchors, K., Weiss, W.A., Costello, J.F., 2005. Epigenome analyses using BAC microarrays identify evolutionary conservation of tissue-specific methylation of SHANK3. Nat. Genet. 37, 645–651.

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., Abramowitz, L.K., Bartolomei, M.S., Rambow, F., Bassi, M.R., Bruno, T., Fanciulli, M., Renner, C., Klein-Szanto, A.J., Matsumoto, Y., Kobi, D., Davidson, I., Alberti, C., Larue, L., Bellacosa, A., 2011. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell 146, 67–79.

Dempster, E., Viana, J., Pidsley, R., Mill, J., 2013. Epigenetic studies of schizophrenia: progress, predicaments, and promises for the future. Schizophr. Bull. 39, 11–16.

Dempster, E.L., Pidsley, R., Schalkwyk, L.C., Owens, S., Georgiades, A., Kane, F., Kalidindi, S., Picchioni, M., Kravariti, E., Toulopoulou, T., Murray, R.M., Mill, J., 2011. Disease-associated epigenetic changes in monozygotic twins discordant for schizophrenia and bipolar disorder. Hum. Mol. Genet. 20, 4786–4796.

Doi, A., Park, I.H., Wen, B., Murakami, P., Aryee, M.J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S., Miller, J., Schlaeger, T., Daley, G.Q., Feinberg, A.P., 2009. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nat. Genet. 41, 1350–1353.

Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T.A., Marques, C.J., Andrews, S., Reik, W., 2011. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature 473, 398–402.

Franchini, D.M., Schmitz, K.M., Petersen-Mahrt, S.K., 2012. 5-Methylcytosine DNA demethylation: more than losing a methyl group. Annu. Rev. Genet. 46, 419–441.

Ghosh, S., Yates, A.J., Fruhwald, M.C., Miecznikowski, J.C., Plass, C., Smiraglia, D., 2010. Tissue specific DNA methylation of CpG islands in normal human adult somatic tissues distinguishes neural from non-neural tissues. Epigenetics 5, 527–538.

Please cite this article in press as: Kato, T., Iwamoto, K., Comprehensive DNA methylation and hydroxymethylation analysis in the human brain and its implication in mental disorders, Neuropharmacology (2014), http://dx.doi.org/10.1016/j.neuropharm.2013.12.019

Guo, J.U., Su, Y., Zhong, C., Ming, G.L., Song, H., 2011. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell 145, 423–434.

Hackett, J.A., Sengupta, R., Zylicz, J.J., Murakami, K., Lee, C., Down, T.A., Surani, M.A., 2013. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. Science 339, 448–452.

Hansen, R.S., Wijmenga, C., Luo, P., Stanek, A.M., Canfield, T.K., Weemaes, C.M., Gartler, S.M., 1999. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. Proc. Natl. Acad. Sci. U. S. A. 96, 14412–14417.

Hayatsu, H., Wataya, Y., Kai, K., Iida, S., 1970. Reaction of sodium bisulfite with uracil, cytosine, and their derivatives. Biochemistry 9, 2858–2865.

He, Y.F., Li, B.Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., Sun, Y., Li, X., Dai, Q., Song, C.X., Zhang, K., He, C., Xu, G.L., 2011. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science 333, 1303–1307.

Hirschhorn, J.N., Lohmueller, K., Byrne, E., Hirschhorn, K., 2002. A comprehensive review of genetic association studies. Genet. Med. 4, 45–61.

Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., Zhang, Y., 2010. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature 466, 1129–1133.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., Zhang, Y., 2011. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science 333, 1300–1303.

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

Iwamoto, K., Bundo, M., Yamada, K., Takao, H., Iwayama-Shigeno, Y., Yoshikawa, T., Kato, T., 2005. DNA methylation status of SOX10 correlates with its down-regulation and oligodendrocyte dysfunction in schizophrenia. J. Neurosci. 25, 5376–5381.

Jin, S.G., Wu, X., Li, A.X., Pfeifer, G.P., 2011. Genomic mapping of 5-hydroxymethylcytosine in the human brain. Nucleic Acids Res. 39, 5015–5024.

Kaminsky, Z., Tochigi, M., Jia, P., Pal, M., Mill, J., Kwan, A., Ioshikhes, I., Vincent, J.B., Kennedy, J.L., Strauss, J., Pai, S., Wang, S.C., Petronis, A., 2012. A multi-tissue analysis identifies HLA complex group 9 gene methylation differences in bipolar disorder. Mol. Psychiatry 17, 728–740.

Khare, T., Pai, S., Koncevicius, K., Pal, M., Kriukiene, E., Liutkeviciute, Z., Irimia, M., Jia, P., Ptak, C., Xia, M., Tice, R., Tochigi, M., Morera, S., Nazarians, A., Belsham, D., Wong, A.H., Blencowe, B.J., Wang, S.C., Kapranov, P., Kustra, R., Labrie, V., Klimasauskas, S., Petronis, A., 2012. 5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary. Nat. Struct. Mol. Biol. 19, 1037–1043.

Kinney, S.M., Chin, H.G., Vaisvila, R., Bitinaite, J., Zheng, Y., Esteve, P.O., Feng, S., Stroud, H., Jacobsen, S.E., Pradhan, S., 2011. Tissue-specific distribution and dynamic changes of 5-hydroxymethylcytosine in mammalian genomes. J. Biol. Chem. 286, 24685–24693.

Kriaucionis, S., Heintz, N., 2009. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 324, 929–930.

Ladd-Acosta, C., Pevsner, J., Sabunciyan, S., Yolken, R.H., Webster, M.J., Dinkins, T., Callinan, P.A., Fan, J.B., Potash, J.B., Feinberg, A.P., 2007. DNA methylation signatures within the human brain. Am. J. Hum. Genet. 81, 1304–1315.

Li, W., Liu, M., 2011. Distribution of 5-hydroxymethylcytosine in different human tissues. J. Nucleic Acids 2011, 870726.

Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D., Lucero, J., Huang, Y., Dwork, A.J., Schultz, M.D., Yu, M., Tonti-Filippini, J., Heyn, H., Hu, S., Wu, J.C., Rao, A., Esteller, M., He, C., Haghighi, F.G., Sejnowski, T.J., Behrens, M.M., Ecker, J.R., 2013. Global epigenomic reconfiguration during mammalian brain development. Science 341, 1237905.

Lu, X., Song, C.X., Szulwach, K., Wang, Z., Weidenbacher, P., Jin, P., He, C., 2013. Chemical modification-assisted bisulfite sequencing (CAB-Seq) for 5-carboxylcytosine detection in DNA. J. Am. Chem. Soc. 135, 9315–9317.

Maunakea, A.K., Nagarajan, R.P., Bilenky, M., Ballinger, T.J., D'Souza, C., Fouse, S.D., Johnson, B.E., Hong, C., Nielsen, C., Zhao, Y., Turecki, G., Delaney, A., Varhol, R., Thiessen, N., Shchors, K., Heine, V.M., Rowitch, D.H., Xing, X., Fiore, C., Schillebeeckx, M., Jones, S.J., Haussler, D., Marra, M.A., Hirst, M., Wang, T., Costello, J.F., 2010. Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 466, 253–257.

Meissner, A., Gnirke, A., Bell, G.W., Ramsahoye, B., Lander, E.S., Jaenisch, R., 2005. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res. 33, 5868–5877.

Mellen, M., Ayata, P., Dewell, S., Kriaucionis, S., Heintz, N., 2012. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell 151, 1417–1430.

Mill, J., Tang, T., Kaminsky, Z., Khare, T., Yazdanpanah, S., Bouchard, L., Jia, P., Assadzadeh, A., Flanagan, J., Schumacher, A., Wang, S.C., Petronis, A., 2008. Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. Am. J. Hum. Genet. 82, 696–711.

Nishioka, M., Bundo, M., Kasai, K., Iwamoto, K., 2012. DNA methylation in schizophrenia: progress and challenges of epigenetic studies. Genome Med. 4, 96.

Numata, S., Ye, T., Hyde, T.M., Guitart-Navarro, X., Tao, R., Wininger, M., Colantuoni, C., Weinberger, D.R., Kleinman, J.E., Lipska, B.K., 2012. DNA methylation signatures in development and aging of the human prefrontal cortex. Am. J. Hum. Genet. 90, 260–272.

Orr, B.A., Haffner, M.C., Nelson, W.G., Yegnasubramanian, S., Eberhart, C.G., 2012. Decreased 5-hydroxymethylcytosine is associated with neural progenitor phenotype in normal brain and shorter survival in malignant glioma. PLoS One 7, e41036.

Pastor, W.A., Pape, U.J., Huang, Y., Henderson, H.R., Lister, R., Ko, M., McLoughlin, E.M., Brudno, Y., Mahapatra, S., Kapranov, P., Tahiliani, M., Daley, G.Q., Liu, X.S., Ecker, J.R., Milos, P.M., Agarwal, S., Rao, A., 2011. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature 473, 394–397.

Petronis, A., 2010. Epigenetics as a unifying principle in the aetiology of complex traits and diseases. Nature 465, 721–727.

Sabunciyan, S., Aryee, M.J., Irizarry, R.A., Rongione, M., Webster, M.J., Kaufman, W.E., Murakami, P., Lessard, A., Yolken, R.H., Feinberg, A.P., Potash, J.B., 2012. Genome-wide DNA methylation scan in major depressive disorder. PLoS One 7, e34451.

Shen, L., Wu, H., Diep, D., Yamaguchi, S., D'Alessio, A.C., Fung, H.L., Zhang, K., Zhang, Y., 2013. Genome-wide analysis reveals TET- and TDG-Dependent 5-Methylcytosine Oxidation dynamics. Cell 153, 692–706.

Siegmund, K.D., Connor, C.M., Campan, M., Long, T.I., Weisenberger, D.J., Biniszkiewicz, D., Jaenisch, R., Laird, P.W., Akbarian, S., 2007. DNA methylation in the human cerebral cortex is dynamically regulated throughout the life span and involves differentiated neurons. PLoS One 2, e895.

Song, C.X., Szulwach, K.E., Dai, Q., Fu, Y., Mao, S.Q., Lin, L., Street, C., Li, Y., Poidevin, M., Wu, H., Gao, J., Liu, P., Li, L., Xu, G.L., Jin, P., He, C., 2013. Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. Cell 153, 678–691.

Song, C.X., Szulwach, K.E., Fu, Y., Dai, Q., Yi, C., Li, X., Li, Y., Chen, C.H., Zhang, W., Jian, X., Wang, J., Zhang, L., Looney, T.J., Zhang, B., Godley, L.A., Hicks, L.M., Lahn, B.T., Jin, P., He, C., 2011. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. Nat. Biotechnol. 29, 68–72.

Spruijt, C.G., Gnerlich, F., Smits, A.H., Pfaffeneder, T., Jansen, P.W., Bauer, C., Munzel, M., Wagner, M., Muller, M., Khan, F., Eberl, H.C., Mensinga, A., Brinkman, A.B., Lephikov, K., Muller, U., Walter, J., Boelens, R., van Ingen, H., Leonhardt, H., Carell, T., Vermeulen, M., 2013. Dynamic readers for 5-(hydroxy) methylcytosine and its oxidized derivatives. Cell 152, 1146–1159.

Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S., Jacobsen, S.E., 2011. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. Genome Biol. 12, R54.

Sugawara, H., Iwamoto, K., Bundo, M., Ueda, J., Miyauchi, T., Komori, A., Kazuno, A., Adati, N., Kusumi, I., Okazaki, Y., Ishigooka, J., Kojima, T., Kato, T., 2011. Hypermethylation of serotonin transporter gene in bipolar disorder detected by epigenome analysis of discordant monozygotic twins. Transl. Psychiatry 1, e24.

Suzuki, M.M., Bird, A., 2008. DNA methylation landscapes: provocative insights from epigenomics. Nat. Rev. Genet. 9, 465–476.

Szulwach, K.E., Li, X., Li, Y., Song, C.X., Han, J.W., Kim, S., Namburi, S., Hermetz, K., Kim, J.J., Rudd, M.K., Yoon, Y.S., Ren, B., He, C., Jin, P., 2011a. Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. PLoS Genet. 7, e1002154.

Szulwach, K.E., Li, X., Li, Y., Song, C.X., Wu, H., Dai, Q., Irier, H., Upadhyay, A.K., Gearing, M., Levey, A.I., Vasanthakumar, A., Godley, L.A., Chang, Q., Cheng, X., He, C., Jin, P., 2011b. 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. Nat. Neurosci. 14, 1607–1616.

Szwagierczak, A., Bultmann, S., Schmidt, C.S., Spada, F., Leonhardt, H., 2010. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. Nucleic Acids Res. 38, e181.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., Rao, A., 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324, 930–935.

Varley, K.E., Gertz, J., Bowling, K.M., Parker, S.L., Reddy, T.E., Pauli-Behn, F., Cross, M.K., Williams, B.A., Stamatoyannopoulos, J.A., Crawford, G.E., Absher, D.M., Wold, B.J., Myers, R.M., 2013. Dynamic DNA methylation across diverse human cell lines and tissues. Genome Res. 23, 555–567.

Wang, J., Cao, X., Zhang, Y., Su, B., 2012a. Genome-wide DNA methylation analyses in the brain reveal four differentially methylated regions between humans and non-human primates. BMC Evol. Biol. 12, 144.

Wang, T., Pan, Q., Lin, L., Szulwach, K.E., Song, C.X., He, C., Wu, H., Warren, S.T., Jin, P., Duan, R., Li, X., 2012b. Genome-wide DNA hydroxymethylation changes are associated with neurodevelopmental genes in the developing human cerebellum. Hum. Mol. Genet. 21, 5500–5510.

Wu, H., D'Alessio, A.C., Ito, S., Wang, Z., Cui, K., Zhao, K., Sun, Y.E., Zhang, Y., 2011. Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. Genes. Dev. 25, 679–684.

Xie, W., Barr, C.L., Kim, A., Yue, F., Lee, A.Y., Eubanks, J., Dempster, E.L., Ren, B., 2012. Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. Cell 148, 816–831.

Xin, Y., Chanrion, B., O'Donnell, A.H., Milekic, M., Costa, R., Ge, Y., Haghighi, F.G., 2012. MethylomeDB: a database of DNA methylation profiles of the brain. Nucleic Acids Res. 40, D1245–D1249.

Xin, Y., O'Donnell, A.H., Ge, Y., Chanrion, B., Milekic, M., Rosoklija, G., Stankov, A., Arango, V., Dwork, A.J., Gingrich, J.A., Haghighi, F.G., 2011. Role of CpG context and content in evolutionary signatures of brain DNA methylation. Epigenetics 6, 1308–1318.

Yao, B., Jin, P., 2013. Cytosine modifications in neurodevelopment and diseases. Cell. Mol. Life Sci. Aug 3 [Epub ahead of print].

Yildirim, O., Li, R., Hung, J.H., Chen, P.B., Dong, X., Ee, L.S., Weng, Z., Rando, O.J., Fazzio, T.G., 2011. Mbd3/NURD complex regulates expression of 5-

Please cite this article in press as: Kato, T., Iwamoto, K., Comprehensive DNA methylation and hydroxymethylation analysis in the human brain and its implication in mental disorders, Neuropharmacology (2014), http://dx.doi.org/10.1016/j.neuropharm.2013.12.019

hydroxymethylcytosine marked genes in embryonic stem cells. Cell 147, 1498–1510.

Yu, M., Hon, G.C., Szulwach, K.E., Song, C.X., Zhang, L., Kim, A., Li, X., Dai, Q., Shen, Y., Park, B., Min, J.H., Jin, P., Ren, B., He, C., 2012. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell 149, 1368–1380.

Zeng, J., Konopka, G., Hunt, B.G., Preuss, T.M., Geschwind, D., Yi, S.V., 2012. Divergent whole-genome methylation maps of human and chimpanzee brains reveal epigenetic basis of human regulatory evolution. Am. J. Hum. Genet. 91, 455–465.

Please cite this article in press as: Kato, T., Iwamoto, K., Comprehensive DNA methylation and hydroxymethylation analysis in the human brain and its implication in mental disorders, Neuropharmacology (2014), http://dx.doi.org/10.1016/j.neuropharm.2013.12.019

Available online at www.sciencedirect.com
SciVerse ScienceDirect
Journal of Genetics and Genomics 40 (2013) 325–329
JGG

HIGHLIGHT

Epigenetic Regulation of Serotonin Transporter in Psychiatric Disorders

Hiroko Sugawara [a,b], Miki Bundo [c], Jun Ishigooka [b], Kazuya Iwamoto [c], Tadafumi Kato [a,*]

[a] *Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*
[b] *Department of Psychiatry, Tokyo Women's Medical University, Tokyo 162-8666, Japan*
[c] *Department of Molecular Psychiatry, Graduate School of Medicine, University of Tokyo, Tokyo 113-8655, Japan*

Received 29 March 2012; revised 12 October 2012; accepted 15 October 2012
Available online 24 November 2012

SLC6A4 (solute carrier family 6, member 4) gene encodes a serotonin transporter (5-hydroxytryptamine transporter, HTT), which transports synaptic serotonin into presynaptic terminal. *SLC6A4* is known to be the target of antidepressants such as selective serotonin reuptake inhibitors (SSRIs). Inhibition of HTT increases synaptic serotonin concentration and thereby exerts antidepressant efficacy. A large number of genetic studies suggest the contribution of genetic variations of *SLC6A4* to various psychiatric disorders. The most studied genetic variation, HTT-linked polymorphic region (HTTLPR), is located at the promoter region of *SLC6A4*. It consists of two major alleles: short (S) and long (L). Each allele contains further variations (Nakamura et al., 2000). The HTTLPR has been reported to affect the gene expression level of *SLC6A4* (Heils et al., 1996; Lesch et al., 1996; Bradley et al., 2005), and individuals carrying the low-expressing S allele of HTTLPR revealed anxiety-related personality trait (Lesch et al., 1996). Furthermore, it was reported that the HTTLPR moderates the influence of stressful life event on depression (Caspi et al., 2003; Kendler et al., 2005). These results suggest the contribution of gene–environment (G × E) interaction involving *SLC6A4* to psychiatric disorders.

Epigenetic factors also contribute to the mechanism of G × E interaction. DNA methylation is affected by environmental factors (Feinberg, 2007; Petronis, 2010). Epigenetic gene regulation by DNA methylation contributes to long-lasting gene expression changes (Bird, 2002). Here, we searched for recent articles relevant to DNA methylation of *SLC6A4* (Table 1), and focused on recent progress in the research on the roles of epigenetic regulation of *SLC6A4* by DNA methylation of *SLC6A4* in psychiatric disorders such as mood and anxiety disorders.

THE INTERACTION OF DNA METHYLATION AND GENOTYPE ON GENE EXPRESSION LEVEL OF *SLC6A4*

The majority of DNA methylation occurs at the fifth position of cytosine residue in the dinucleotides CpG sequences in mammals. While cytosine residues in the dinucleotides are generally methylated, CpG-rich regions, which are called "CpG island" and located within and around the regulatory promoter regions, are less methylated. Usually, the extent of methylation at the promoter region CpG island inversely correlates with the extent of gene expression.

The S allele of HTTLPR has been shown to have the lower promoter activity compared with L allele, which is associated with decreased mRNA expression (Heils et al., 1996; Lesch et al., 1996; Bradley et al., 2005). Philibert et al. (2007) examined the relationship between DNA methylation at the promoter region CpG island and gene expression level of *SLC6A4* using lymphoblastoid cell lines (LCLs). There was no significant association between total DNA methylation and mRNA levels. However, DNA methylation was associated with decreased mRNA levels under the control of HTTLPR genotype (Philibert et al., 2007). They could not replicate this finding in the second study (Philibert et al., 2008). On the other hand, in infant rhesus macaques, carriers of the S allele exhibited higher methylation of CpG island, and this was associated with lower gene expression of *SLC6A4* in peripheral blood mononuclear cells (PBMCs) (Kinnally et al., 2010).

* Corresponding author. Tel: +81 48 467 6949, fax: +81 48 467 6947.
E-mail address: kato@brain.riken.jp (T. Kato).

http://dx.doi.org/10.1016/j.jgg.2012.10.002

Table 1
Articles relevant to *SLC6A4* methylation

Reference	Subject	*N*	Source	Examination		
				Environmental factors	Biological factors	Methods
Wang et al., 2012	Healthy adult male	25	T cells and monocytes	Childhood-limited aggression	Methylation	Pyrosequencing
Vijayendran et al., 2012	Female adoptee	158	LCLs	Child abuse	Methylation	Beadchip
					HTTLPR	Real-time PCR
					Expression	RT-PCR
Sugawara et al., 2011a	MZ twins discordant for BD	2 pairs	LCLs	–	Methylation	Tiling array, bisulfite sequencing, pyrosequencing
	BD and C	BD = 20, C = 20	LCLs	–	Methylation	Pyrosequencing
					HTTLPR	PCR
					Expression	RT-PCR
	BD and C	BD = 35, C = 35	Brains	–	Methylation	Pyrosequencing
Park et al., 2011	AD and C	AD = 27, C = 15	PBMCs	–	Methylation	Pyrosequencing
Koenen et al., 2011	PTSD	100	PBMCs	Traumatic events	Methylation	Beadchip, pyrosequencing
					HTTLPR	RFLP
Kinnally et al., 2011	Female bonnet macaques	20	PBMCs	Early life stress	Methylation	Pyrosequencing
Beach et al., 2011	Adoptee	192	LCLs	Child sex abuse	Methylation	Mass spectroscopy
Devlin et al., 2010	Pregnant woman and infant	82	Maternal peripheral leukocytes, umbilical cord leukocytes	–	Methylation	Pyrosequencing
van Ijzendoorn et al., 2010	Adoptee	143	LCLs	Unresolved loss or trauma	Methylation	Mass spectroscopy
					HTTLPR	PCR
Wong et al., 2010	MZ and DZ twins	MZ twin-pairs = 46, DZ twin pairs = 45	Buccal cells	–	Methylation	Mass spectroscopy
Kinnally et al., 2010	Infant rhesus macaques	87	PBMCs	Early life stress	Methylation	Pyrosequencing
					HTTLPR	RFLP
					Expression	RT-PCR
Olsson et al., 2010	MD and C	MD = 25, C = 125	Buccal cells	–	Methylation	Mass spectroscopy
					HTTLPR	PCR
Beach et al., 2010	Adoptee	155	LCLs	Child abuse	Methylation	Mass spectroscopy
					HTTLPR	PCR
Philibert et al., 2008	Adoptee	192	LCLs	–	Methylation	Mass spectroscopy
					HTTLPR	PCR
					Expression	RT-PCR

(*continued on next page*)