

## SHORT REPORT

# Expression QTL analysis of top loci from GWAS meta-analysis highlights additional schizophrenia candidate genes

Simone de Jong<sup>1,2,6</sup>, Kristel R van Eijk<sup>1</sup>, Dave WLH Zeegers<sup>1</sup>, Eric Strengman<sup>1,3</sup>, Esther Janson<sup>1</sup>, Jan H Veldink<sup>4</sup>, Leonard H van den Berg<sup>4</sup>, Wiepke Cahn<sup>2</sup>, René S Kahn<sup>2,5</sup>, Marco PM Boks<sup>2,5</sup>, Roel A Ophoff<sup>\*,2,3</sup> and The PGC Schizophrenia (GWAS) Consortium<sup>7</sup>

There is genetic evidence that schizophrenia is a polygenic disorder with a large number of loci of small effect on disease susceptibility. Genome-wide association studies (GWASs) of schizophrenia have had limited success, with the best finding at the MHC locus at chromosome 6p. A recent effort of the Psychiatric GWAS consortium (PGC) yielded five novel loci for schizophrenia. In this study, we aim to highlight additional schizophrenia susceptibility loci from the PGC study by combining the top association findings from the discovery stage (9394 schizophrenia cases and 12 462 controls) with expression QTLs (eQTLs) and differential gene expression in whole blood of schizophrenia patients and controls. We examined the 6192 single-nucleotide polymorphisms (SNPs) with significance threshold at  $P < 0.001$ . eQTLs were calculated for these SNPs in a sample of healthy controls ( $n = 437$ ). The transcripts significantly regulated by the top SNPs from the GWAS meta-analysis were subsequently tested for differential expression in an independent set of schizophrenia cases and controls ( $n = 202$ ). After correction for multiple testing, the eQTL analysis yielded 40 significant *cis*-acting effects of the SNPs. Seven of these transcripts show differential expression between cases and controls. Of these, the effect of three genes (*RNF5*, *TRIM26* and *HLA-DRB3*) coincided with the direction expected from meta-analysis findings and were all located within the MHC region. Our results identify new genes of interest and highlight again the involvement of the MHC region in schizophrenia susceptibility. *European Journal of Human Genetics* (2012) 20, 1004–1008; doi:10.1038/ejhg.2012.38; published online 21 March 2012

**Keywords:** expression; schizophrenia; eQTL; MHC; SNP

## INTRODUCTION

Schizophrenia is a severe mental disorder, affecting about 1% of the population worldwide. Heritability is estimated to be around 80%, but the underlying genes are largely unknown.<sup>1</sup> Large-scale, genome-wide studies have identified rare genomic microdeletions as well as common variants associated with the disease.<sup>2–7</sup>

In a recent study, Purcell *et al.*<sup>2</sup> demonstrated that data from genome-wide association studies (GWASs) for schizophrenia are compatible with a very large number of loci with common alleles ( $N > 3000$ ), each with a very small contribution to disease susceptibility (odds ratios  $< 1.05$ ). Alternative approaches may be necessary to decipher the genetic basis of schizophrenia and related disorders.<sup>8</sup> We therefore aim to combine different layers of genomic information to uncover genetic signal from common variants that would not be identified by current GWAS approaches.

A recent meta-analysis comparing 9394 cases to 12 462 controls resulted in identification of numerous common variants<sup>9</sup> with sub threshold association with schizophrenia (6192 single-nucleotide polymorphisms (SNPs) with  $P < 0.001$ ). However, as Purcell *et al.*<sup>2</sup>

showed, variants associated with the disease may not reach genome-wide significance. Therefore, it is likely that there are more true positives in the top 6192 SNPs than were identified by performing a standard case–control association analysis. Subtle effects of these SNPs on gene expression could be a functional mechanism by which they confer risk for development of schizophrenia.<sup>10,11</sup> Recently, it has been shown that true GWAS hits are enriched for expression QTLs (eQTLs).<sup>11–14</sup> Therefore, variations influencing gene expression are more likely to be contributing to the phenotype. To this end, we generated eQTLs for the top 6192 SNPs ( $P < 0.001$  in meta-analysis). Next, we tested whether the identified transcripts are differentially expressed between patients and healthy controls. These analyses have the potential to provide further support in the involvement of these SNPs in schizophrenia and may highlight additional schizophrenia candidate genes that have not been identified using genome-wide significance thresholds.

Although gene expression in whole blood is only moderately correlated with gene expression in brain tissue,<sup>15–17</sup> several studies suggest that gene expression in blood could serve as a marker of

<sup>1</sup>Department of Medical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands; <sup>2</sup>Department of Psychiatry, Rudolf Magnus Institute of Neuroscience, University Medical Center, Utrecht, The Netherlands; <sup>3</sup>Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, CA, USA; <sup>4</sup>Department of Neurology, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, Utrecht, The Netherlands; <sup>5</sup>Julius Centre for Health Sciences and Primary Care, University Medical Center, Utrecht, The Netherlands

\*Correspondence: Professor RA Ophoff, Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California, Gonda #4335B, Box 951761, 695 Charles E. Young Drive South, Los Angeles, CA 90095-1761, USA. Tel: +1 310 794 9602; Fax: +1 310 794 9613; E-mail: ROphoff@mednet.ucla.edu

<sup>6</sup>Current address: UCLA Center for Neurobehavioral Genetics (affiliation 3).

<sup>7</sup>The Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium authors are listed in Supplementary Table S1.

Received 4 October 2011; revised 30 January 2012; accepted 2 February 2012; published online 21 March 2012

brain-related disease states, including schizophrenia.<sup>16–22</sup> Therefore, gene expression profiling in blood may provide additional insight into the etiology of the disease. We performed the gene expression analyses using whole-blood samples of a relatively large sample of schizophrenia patients and controls.

## MATERIALS AND METHODS

### eQTL analysis in controls

We calculated the eQTLs for the top SNPs<sup>9</sup> in a sample of 437 healthy controls for which genotypes (Illumina 370k array) and whole-blood gene expression data (Illumina H-12 beadchip) was available as described before.<sup>11,23</sup> In short, this dataset consists of 244 males and 193 females with a mean age of 62 years, who were recruited as controls in a study of gene expression in amyotrophic lateral sclerosis. These control subjects were selected for being in good general health and unaffected for neurological and neurodegenerative diseases; no separate screen for psychiatric disorders was performed for these subjects.

Of the 6192 SNPs, 1336 were already available on the array. Imputation was performed by BEAGLE version 3.0.4<sup>24</sup> using the Hapmap phased founder set, release 2, phase 3 (The International HapMap Project, <http://www.hapmap.org>). A  $R^2$  cutoff of 0.90 resulted in 4073 successfully imputed SNPs yielding a total of 5409 SNPs for analysis. The gene expression data of these controls were quantile-normalized and log<sub>2</sub>-transformed using the PreprocessCore package in R.<sup>25</sup> Expression probes were then filtered for mean detection value <0.90 as by manufacturer protocol, leaving 12 990 high-quality probes for analysis.

The 12 990 expression probes were taken as quantitative traits and tested for association with the 5409 available SNPs using a linear association of allele dosage with age and gender as covariates in PLINK.<sup>26</sup> To adjust for significant differences in mean age of this control sample and the schizophrenia sample described below, we included age (and gender) as covariates. As *trans*-effects are difficult to identify in a study of this size due to limited power, we focused on *cis*-effects only, that is, 1 MB around the probe center position on either side. We used Bonferroni correction for multiple testing, setting the significance thresholds for *cis*-effects  $0.05/5409=9.24E-6$ .

### Differential expression schizophrenia versus controls

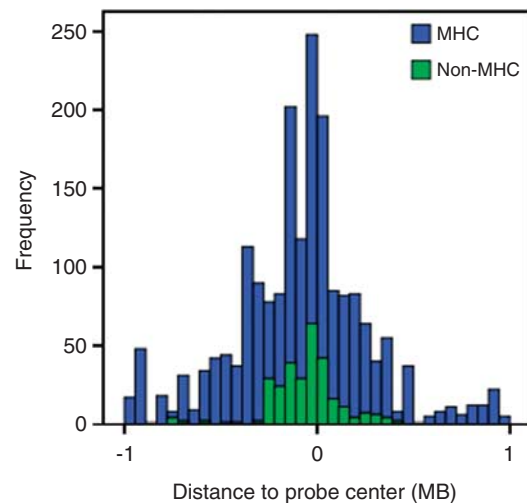
We examined whether the probes associated with the top SNPs are also related to schizophrenia disease status. This set consists of 106 schizophrenia cases and 96 healthy controls including 118 male and 84 female subjects, with an average age of 39 years. Diagnoses were determined by Standardized Psychiatric interviews, either The Comprehensive Assessment of Symptoms and History (CASH) or the Composite international diagnostic interview (CIDI) by trained clinicians. Schizophrenia was defined by a DSM-IV-TR diagnosis of #295.0–295.89, and #298.9. This study was approved by Medical Research Ethics Committee (METC) of the University Medical Center Utrecht, The Netherlands. The data was normalized (robust spline normalization), transformed (variance-stabilizing transformation) and filtered according to the Lumi procedure as described previously.<sup>27</sup>

We used the Limma package<sup>28</sup> in R to generate a regression model with selected expression values as dependent and status as independent values. We included age and gender as covariates. We took FDR-corrected  $P$ -value of 0.05 as significance threshold.

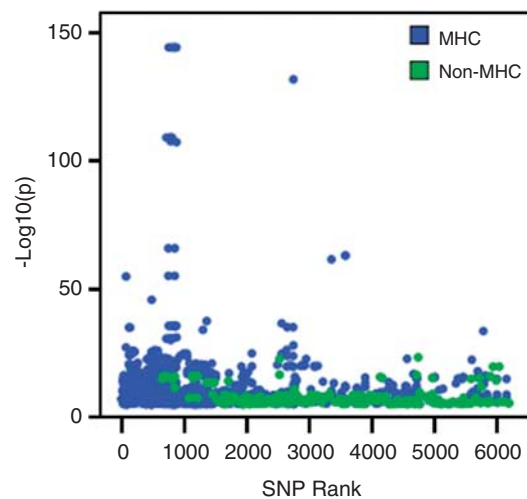
## RESULTS

### eQTL analysis in controls

We identified 40 unique transcripts from MHC and non-MHC regions regulated by the top 5409 SNPs. The distribution of eQTL results within the *cis*-region is displayed in Figure 1. As we found the signal from the MHC region in the eQTL results to be substantial, these results will be discussed separately. In total, 1664 significant *cis*-effects on 23 unique expression probes from 578 unique SNPs in the MHC region were found. Outside the MHC, 166 *cis*-effects, were identified, representing 249 unique SNPs and 17 unique expression probes. In this instance, we define unique to indicate different SNPs, without considering linkage disequilibrium (LD) between them. The eQTL effects are plotted against the rank of the SNPs in the top list



**Figure 1** Distribution of eQTL effects. A region of 1 MB around the center of the expression probe was taken as a threshold for *cis*-effects. The plot shows a histogram of frequency of eQTL effects (SNP-probe combinations) in this region. Effects in the MHC region are displayed in blue, effects elsewhere on the genome in green.



**Figure 2** eQTL effects versus SNP rank. The  $-\log_{10}(p)$  value of eQTL effects (SNP-probe combinations) are plotted against the original rank of the SNP in the schizophrenia meta-analysis top 6192. Effects in the MHC region are displayed in blue, effects elsewhere on the genome in green.

in Figure 2. This shows the effects of the MHC region to be stronger ( $P$ -values ranging from  $4.6E-145$  to  $9.1E-6$ ) than the rest of the genome ( $P$ -values ranging from  $4.3E-24$  to  $9.1E-6$ ). In addition, the highest ranking genes are dominated by effects from this region. All significant eQTL effects are given in Supplementary Table 2.

To formally test whether the MHC is overrepresented in the eQTL results, we performed a Fisher's exact test for count data. Of the 5409 SNPs tested, 1071 are located within or are associated to a probe in the extended MHC region (chromosome 6, 26–34 MB). Even though that already represents a large part of the input, the overrepresentation in the eQTL results (unique SNPs influencing one or more expression probes) is significant when compared with SNPs located elsewhere on the genome (Fisher's  $P < 2.2E-16$ ). The fact that this region is known to have complicated and extended LD structures<sup>29</sup> may in part explain the fact that more SNPs are related to the same gene expression probes.

### Differential expression schizophrenia *versus* controls

The levels of 40 unique transcripts from MHC and non-MHC regions that we identified were tested for association with schizophrenia in an independent gene expression dataset of schizophrenia cases and controls. Seven of the transcripts were differentially expressed in cases. Table 1 presents the results together with the corresponding eQTL results. Five of the seven transcripts are from genes located in the MHC (*TRIM26*, *RNF5*, *TUBB*, *HLA-DRB3*, *HIST1H2BD*). Two differentially expressed probes are located elsewhere on the genome, *C16ORF61* on chromosome 16 and *CRELD2* on chromosome 22. All but two probes are downregulated in schizophrenia patients compared with healthy controls (*C16ORF61* and *HIST1H2BD*). Although 18 of the 106 schizophrenia patients in this dataset are antipsychotic-free (eg, have not been treated with antipsychotic drugs during the 6-month-period before blood sampling), the effects of medication on gene expression in blood in the overall sample might have contributed to our results.

Next, we examined whether the direction of differential expression coincided with the original meta-analysis results. When the risk allele is associated with upregulation or downregulation of expression, this should be mirrored in the differential expression results. We therefore assume that the eQTL itself is not different between cases and controls, but the frequency of the allele associated with changes in expression levels is. This was the case for three out of the seven genes; *TRIM26*, *RNF5* and *HLA-DRB3*, all of which are located within the MHC region. *TRIM26* is one of the genes identified as susceptibility locus in the original meta-analysis after the replication stage.<sup>9</sup>

For our top genes (*TRIM26*, *RNF5* and *HLA-DRB3*) we examined whether the same eQTLs are also present in brain, using a publically available human brain expression dataset of 144 individuals (Gibbs *et al*<sup>30</sup> GEO; GSE15745). Data are available for four different brain regions: cerebellum, frontal cortex, temporal cortex and pons. eQTL analysis was performed as described in blood, with the addition of covariates post-mortem interval, batch, and source of collection. Ring finger protein 5 (*RNF5*) is not expressed at detectable levels in any of these four regions, *HLA-DRB3* and *TRIM26* are expressed in all four brain regions. The top SNP regulating *HLA-DRB3* (rs9268858) was not available for the brain dataset. However, rs2395185 is in strong LD ( $R^2=0.98$ , based on HapMap3) and shows significant eQTL results for *HLA-DRB3* in frontal cortex ( $P=5.7E-7$ ,  $\beta=-0.27$ ,  $t=-5.27$ ), pons ( $P=5.6E-3$ ,  $\beta=-0.22$ ,  $t=-2.82$ ), and cerebellum ( $P=5.7E-7$ ,  $\beta=-0.17$ ,  $t=-3.08$ ). Similarly, the top SNP-regulating *TRIM26* (rs2844766) in blood was not available for the brain samples but rs1264616, with moderate LD ( $R^2=0.45$ , based on HapMap3) shows evidence of an eQTL in cerebellum ( $P=9.0E-3$ ,  $\beta=0.09$ ,  $t=2.64$ ). These eQTLs show the same directionality in both blood and brain.

### DISCUSSION

Our results are noteworthy in two respects: First, we show that by combining different layers of genetic information we can successfully identify potential candidate genes for schizophrenia. By using eQTL analysis we have identified SNPs that actually have an effect on expression of transcripts that differ between patients and controls. Second, we found that the top list that we took as a starting point harbored significantly more eQTLs in the MHC region compared with elsewhere in the genome. Strong eQTLs in this region have been observed in a previous study in lymphoblastoid cell lines, derived from peripheral blood lymphocytes.<sup>31</sup> Specifically, *cis* eQTLs have been found for *TUBB*<sup>32</sup> and *HLA-DRB3*,<sup>33,34</sup> whereas *RNF5* has been found to be one of the most variable genes between individuals.<sup>35</sup> The fact that eQTLs in the MHC region are overrepresented in our study could

**Table 1** eQTL and differential expression results for potential schizophrenia candidate genes

Gene symbol	Array address ID	Location	SNP name	Rank top 6.192	SNP-probe center (bp)	Expression QTL		Rank range	Risk allele	Direction of xp	AVG xp	Differential expression		
						P-value	# SNPs					Log FC	t	Adjusted P-value
<b>TRIM26</b>	<b>2370138</b>	<b>6p21.3</b>	<b>rs2844766</b>	<b>445</b>	<b>78138</b>	<b>4.5E-06</b>	<b>4</b>	<b>445-617</b>	<b>T</b>	↓	<b>8.86</b>	<b>-0.24</b>	<b>-6.84</b>	<b>4.0E-09</b>
C16ORF61	6400025	16q23.2	rs4888116	2642	51339	5.5E-06	1	2642	C	↓	9.10	0.24	5.08	1.8E-05
<b>RNF5</b>	<b>2690603</b>	<b>6p21.3</b>	<b>rs3132947</b>	<b>3215</b>	<b>28549</b>	<b>3.2E-08</b>	<b>22</b>	<b>671-5578</b>	<b>G</b>	↑	<b>8.40</b>	<b>-0.11</b>	<b>-4.78</b>	<b>4.8E-05</b>
CRELD2	6560390	22q13.33	rs138897	2183	89937	2.4E-08	26	1945-6118	C	↑	9.21	-0.13	-3.75	2.4E-03
TUBB	6580474	6p21.33	rs3094127	895	4464	3.4E-11	60	22-1120	A	↑	10.84	-0.12	-3.11	1.8E-02
HIST1H2BD	6200669	6p21.3	rs9379827	1949	5056	2.9E-18	5	1713-2018	A	↓	8.55	0.16	2.82	3.2E-02
<b>HLA-DRB3</b>	<b>1770504</b>	<b>6p21.3</b>	<b>rs9268858</b>	<b>748</b>	<b>25356</b>	<b>2.4E-14</b>	<b>33</b>	<b>66-5596</b>	<b>C</b>	↓	<b>9.77</b>	<b>-0.25</b>	<b>-2.82</b>	<b>3.2E-02</b>

Abbreviations: SNP, single-nucleotide polymorphism; QTL, quantitative trait locus.

Of the 40 probes identified by eQTL analysis to be regulated by the top 5409 SNPs, 7 were also found to be differentially expressed between schizophrenia cases and controls. Of these, the direction of *TRIM26*, *RNF5* and *HLA-DRB3* coincided with meta-analysis results (printed in bold-face). The first three columns contain gene symbol, Illumina array address ID of the expression probes and location of the genes. The next part of the table contains details of the eQTL analysis, with the top SNP, its rank in the original meta-analysis, distance to associated expression probe center and eQTL P-value. The number of SNPs associated to the expression level and rank range in the meta-analysis is given, respectively. Finally, the direction of expression associated with the risk allele of the top SNP is given. The last part of the table contains the differential expression results, with average expression, the log-fold change (LFC) between cases and controls, the *t*-statistic and the corresponding *P*-value (adjusted for testing of 40 probes).

be due to the type of tissue studied. Blood has an important function in the (primary) immune response and eQTLs can be tissue-specific.<sup>36</sup> However, it is also possible that genetic control of this region is stronger than elsewhere on the genome, perhaps related to the extended LD patterns.

The MHC harbors relatively many genes and is highly polymorphic. It regulates the immune response and has been associated to a large number of immune phenotypes and diseases.<sup>37</sup> The association of SNPs in these MHC genes can result from mere physical closeness to the schizophrenia variant within the MHC region. However, that would not explain the difference in gene expression of MHC genes presented here, as expression transcripts are not in LD. We find that the eQTL SNPs are often related to multiple probes of different genes within the *cis*-region (up to 6 per SNP). This could point to important regulatory regions in the MHC. We therefore propose that eQTL analysis could aid in the refinement of the MHC region.

Although association of schizophrenia with this area has been consistent, the functional explanation is still ambiguous. Inflammation is suggested to be an inherent part of schizophrenia. Reports show that schizophrenia patients have either higher or lower prevalence of some autoimmune disorders than expected<sup>38</sup> and that the efficacy of anti-inflammatory drugs differs in patients.<sup>39</sup> The latter observations could support the viral hypothesis of schizophrenia.<sup>40</sup> However, of the three identified candidate genes in the MHC region, only one has a known immune function. This is *HLA-DRB3* (major histocompatibility complex, class II, DR beta 3). Class II molecules are expressed in antigen-presenting cells. They have a central role in the immune system by presenting peptides derived from extracellular proteins. The function of *TRIM26* (tripartite motif-containing 26) in the classical class I subregion is unknown; however, it is thought to have DNA-binding activity.<sup>29</sup> *RNF5* in the classical class III subregion is involved in cell motility. It has been shown to be a regulator of breast cancer progression through its effect on actin cytoskeletal alterations.<sup>41</sup> This suggests that the MHC signals associated with schizophrenia are not necessarily immune-related, but could also stem from genes in the region that are involved in non-immune-related pathways.

As schizophrenia is a brain-related disease, it was useful to replicate our findings in brain tissue. Publically available brain expression data indicated that eQTLs for *TRIM26* and *HLA-DRB3* are also present in specific brain regions. Although these results fit well with a possible involvement of these loci in schizophrenia, the lack of sufficient number of available brain tissue from patients has prevented us to examine brain-specific differential gene expression related to disease.

In summary, we have identified three genes, *TRIM26*, *RNF5* and *HLA-DRB3* that are regulated by the most significant SNPs in a recent meta-analysis and of which expression is associated with schizophrenia disease status. These effects are small but significant, indicating that the current approach allows to detect the small functional effects that may have a role in schizophrenia susceptibility.<sup>2</sup> The results of this study indicate that signal not only stems from the SNPs with the most significant *P*-values in the top 6192 list, but can be linked to less significant SNPs as well. In addition to confirming one of the top findings in the meta-analysis, *TRIM26*, we also identify *RNF5* and *HLA-DRB3* as potential candidate genes for schizophrenia. The results of this study gives further insight into the relationship of SNPs with gene expression, highlight the importance of the MHC region for schizophrenia susceptibility and indicate that genetic causal variants for schizophrenia might act through regulation of expression. This approach can be fruitful in identifying phenotypic effects of SNPs highlighted by GWASs.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

We thank the participants and their relatives and staff at the various recruitment centers. This work was supported by TI-pharma project T5-203 and funding from the US National Institutes of Health MH078075 and NS058980 (RAO) and the Amyotrophic Lateral Sclerosis Association (RAO and LH5dB). Statistical analyses were carried out on the Genetic Cluster Computer (<http://www.geneticcluster.org>), which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003) along with a supplement from the Dutch Brain Foundation.

- 1 Sullivan PF, Kendler KS, Neale MC: Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry* 2003; **60**: 1187–1192.
- 2 Purcell SM, Wray NR, Stone JL *et al*: Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
- 3 Shi J, Levinson DF, Duan J *et al*: Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 2009; **460**: 753–757.
- 4 Stefansson H, Ophoff RA, Steinberg S *et al*: Common variants conferring risk of schizophrenia. *Nature* 2009; **460**: 744–747.
- 5 Stefansson H, Rujescu D, Cichon S *et al*: Large recurrent microdeletions associated with schizophrenia. *Nature* 2008; **455**: 232–236.
- 6 Buizer-Voskamp JE, Muntjewerff JW, Strengman E *et al*: Genome-wide analysis shows increased frequency of copy number variation deletions in Dutch schizophrenia patients. *Biol Psychiatry* 2011; **70**: 655–662.
- 7 Vrijenhoek T, Buizer-Voskamp JE, van der Stelt I *et al*: Recurrent CNVs disrupt three candidate genes in schizophrenia patients. *Am J Hum Genet* 2008; **83**: 504–510.
- 8 Manolio TA, Collins FS, Cox NJ *et al*: Finding the missing heritability of complex diseases. *Nature* 2009; **461**: 747–753.
- 9 Ripke S, Sanders AR, Kendler KS *et al*: Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 2011; **43**: 969–976.
- 10 Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M: Mapping complex disease traits with global gene expression. *Nat Rev Genet* 2009; **10**: 184–194.
- 11 Dubois PC, Trynka G, Franke L *et al*: Multiple common variants for celiac disease influencing immune gene expression. *Nat Genet* 2010; **42**: 295–302.
- 12 Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ: Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet* 2010; **6**: e1000888.
- 13 Richards AL, Jones L, Moskvina V *et al*: Schizophrenia susceptibility alleles are enriched for alleles that affect gene expression in adult human brain. *Mol Psychiatry* 2011; **17**: 193–201.
- 14 Zhong H, Beaulaurier J, Lum PY *et al*: Liver and adipose expression associated SNPs are enriched for association to type 2 diabetes. *PLoS Genet* 2010; **6**: e1000932.
- 15 Cai C, Langfelder P, Fuller TF *et al*: Is human blood a good surrogate for brain tissue in transcriptional studies? *BMC Genomics* 2010; **11**: 589.
- 16 Jasinska AJ, Service S, Choi OW *et al*: Identification of brain transcriptional variation reproduced in peripheral blood: an approach for mapping brain expression traits. *Hum Mol Genet* 2009; **18**: 4415–4427.
- 17 Rollins B, Martin MV, Morgan L, Vawter MP: Analysis of whole genome biomarker expression in blood and brain. *Am J Med Genet B Neuropsychiatr Genet* 2010; **153B**: 919–936.
- 18 Iwamoto K, Kakiuchi C, Bundo M, Ikeda K, Kato T: Molecular characterization of bipolar disorder by comparing gene expression profiles of postmortem brains of major mental disorders. *Mol Psychiatry* 2004; **9**: 406–416.
- 19 Tsuang MT, Nossova N, Yager T *et al*: Assessing the validity of blood-based gene expression profiles for the classification of schizophrenia and bipolar disorder: a preliminary report. *Am J Med Genet B Neuropsychiatr Genet* 2005; **133B**: 1–5.
- 20 Gladkevich A, Kauffman HF, Korf J: Lymphocytes as a neural probe: potential for studying psychiatric disorders. *Prog Neuropsychopharmacol Biol Psychiatry* 2004; **28**: 559–576.
- 21 Glatt SJ, Everall IP, Kremen WS *et al*: Comparative gene expression analysis of blood and brain provides concurrent validation of SELENBP1 up-regulation in schizophrenia. *Proc Natl Acad Sci USA* 2005; **102**: 15533–15538.
- 22 Kuzman MR, Medved V, Terzic J, Krainc D: Genome-wide expression analysis of peripheral blood identifies candidate biomarkers for schizophrenia. *J Psychiatr Res* 2009; **43**: 1073–1077.
- 23 Saris CG, Horvath S, van Vught PW *et al*: Weighted gene co-expression network analysis of the peripheral blood from Amyotrophic Lateral Sclerosis patients. *BMC Genomics* 2009; **10**: 405.
- 24 Browning BL, Browning SR: A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet* 2009; **84**: 210–223.
- 25 Bolstad BM, Irizarry RA, Astrand M, Speed TP: A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 2003; **19**: 185–193.

- 26 Purcell S, Neale B, Todd-Brown K *et al*: PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
- 27 Du P, Kibbe WA, Lin SM: Lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 2008; **24**: 1547–1548.
- 28 Smyth GK: Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004; **3**: Article3.
- 29 Horton R, Wilming L, Rand V *et al*: Gene map of the extended human MHC. *Nat Rev Genet* 2004; **5**: 889–899.
- 30 Gibbs JR, van der Brug MP, Hernandez DG *et al*: Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS genetics* 2010; **6**: e1000952.
- 31 Dixon AL, Liang L, Moffatt MF *et al*: A genome-wide association study of global gene expression. *Nat Genet* 2007; **39**: 1202–1207.
- 32 Stranger BE, Nica AC, Forrest MS *et al*: Population genomics of human gene expression. *Nat Genet* 2007; **39**: 1217–1224.
- 33 Monks SA, Leonardson A, Zhu H *et al*: Genetic inheritance of gene expression in human cell lines. *Am J Hum Genet* 2004; **75**: 1094–1105.
- 34 Goring HH, Curran JE, Johnson MP *et al*: Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nat Genet* 2007; **39**: 1208–1216.
- 35 Cheung VG, Jen KY, Weber T *et al*: Genetics of quantitative variation in human gene expression. *Cold Spring Harb Symp Quant Biol* 2003; **68**: 403–407.
- 36 Heinzen EL, Ge D, Cronin KD *et al*: Tissue-specific genetic control of splicing: implications for the study of complex traits. *PLoS Biol* 2008; **6**: e1.
- 37 Vandiedonck C, Knight JC: The human major histocompatibility complex as a paradigm in genomics research. *Brief Funct Genomic Proteomic* 2009; **8**: 379–394.
- 38 Eaton WW, Byrne M, Ewald H *et al*: Association of schizophrenia and autoimmune diseases: linkage of Danish national registers. *Am J Psychiatry* 2006; **163**: 521–528.
- 39 Laan W, Grobbee DE, Selten JP, Heijnen CJ, Kahn RS, Burger H: Adjuvant aspirin therapy reduces symptoms of schizophrenia spectrum disorders: results from a randomized, double-blind, placebo-controlled trial. *J Clin Psychiatry* 2010; **71**: 520–527.
- 40 Doorduyn J, de Vries EF, Willemsen AT, de Groot JC, Dierckx RA, Klein HC: Neuroinflammation in schizophrenia-related psychosis: a PET study. *J Nucl Med* 2009; **50**: 1801–1807.
- 41 Bromberg KD, Kluger HM, Delaunay A *et al*: Increased expression of the E3 ubiquitin ligase RNF5 is associated with decreased survival in breast cancer. *Cancer Res* 2007; **67**: 8172–8179.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)