Disruption in peripheral expression of genes related to cognition among patients with schizophrenia and bipolar disorder and their co-twins

Viehman, Rachael W.

Acceptance Date: 2012

Series: UCLA Electronic Theses and Dissertations

Degree: Ph.D., Psychology 0780 UCLA

Advisor(s): Cannon, Tyrone D

Committee: Ray, Lara, Jentsch, J. David, Smith, Desmond

Permalink: http://escholarship.org/uc/item/8944z06b

Abstract:

Copyright Information:
All rights reserved unless otherwise indicated. Contact the author or original publisher for any necessary permissions. eScholarship is not the copyright owner for deposited works. Learn more at http://www.escholarship.org/help_copyright.html#reuse
Disruption in peripheral expression of genes related to cognition among patients with schizophrenia and bipolar disorder and their co-twins

A Dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Psychology

by

Rachael Willhite Viehman

2012
Disruption in peripheral expression of genes related to cognition among patients with schizophrenia and bipolar disorder and their co-twins

by

Rachael Willhite Viehman

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2012

Professor Tyrone D. Cannon, Chair

Schizophrenia and bipolar disorder are substantially heritable psychiatric illnesses. This study examined peripheral gene expression in an initial sample recruited in Sweden consisting of 35 individuals with schizophrenia and their unaffected co-twins, 24 individuals with bipolar disorder and their unaffected co-twins, and 60 control twins without a history of significant psychiatric illness. Results were confirmed with a secondary sample of 18 individuals with schizophrenia and their unaffected co-twins and 37 non-psychiatric control twins recruited in Finland. Expression in peripheral blood lymphocytes was measured using the Illumina Human WG6 v3.0 gene chip. A cognitive endophenotype was used to screen for genes that varied in relation to cognition out of the genome-wide panel. Results of the mixed-model ANOVA revealed 28 genes showed a significant association with cognitive performance at a corrected $\alpha = 0.05$ and 47 at a corrected $\alpha = 0.10$. Twelve of these genes showed significantly decreased
expression in patients with schizophrenia and five showed decreased expression in patients with bipolar disorder compared to controls. Nine genes also showed decreased expression in the unaffected co-twins of patients with schizophrenia compared to controls. All twelve genes also showed decreased expression in the secondary sample of patients with schizophrenia. Further research is necessary to determine the relationship between decreased peripheral expression of these genes and expression levels in the central nervous system and relationships with other factors such as duration of illness and psychotropic medication use.
The dissertation of Rachael Willhite Viehman is approved.

Lara Ray

J. David Jentsch

Desmond Smith

Tyrone D. Cannon, Committee Chair

University of California, Los Angeles

2012
# Table of Contents

## Introduction Pages

- Abstract ............................................................................................................................ ii
- Committee Page .............................................................................................................. iv
- List of Tables and Figures ............................................................................................. vi
- Acknowledgments ......................................................................................................... vii
- Vita ................................................................................................................................... viii

## Main Text

- Introduction ..................................................................................................................... 1
- Research Design and Methods ..................................................................................... 24
- Results ............................................................................................................................. 33
- Discussion ....................................................................................................................... 37
- Tables and Figures ......................................................................................................... 45
- References ...................................................................................................................... 57
List of Tables and Figures

Table 1. Break-down of recruited Swedish twin-pairs by proband diagnosis…………………..45
Table 2. Demographic characteristics of Swedish sample……………………………………….46
Table 3. Demographic characteristics of Finnish sample………………………………………..47
Table 4. Results of factor analysis of neuropsychological data………………………………….48
Table 5. Loadings of specific neurocognitive domains on factor 1……………………………..49
Table 6. Genes whose expression varies as a result of factor 1 scores (corrected p < 0.05)……50
Table 7. Genes whose expression varies as a result of factor 1 scores (0.05 < corrected p <0.10)…………………………………………………………………………………………………………51
Table 8. Select genes whose expression levels vary as a result of proband membership……52
Figure 1. Peripheral expression of select genes in patients with schizophrenia and unaffected co-twins. Gene expression plotted as percent-change relative to the expression level of members in the control group……………………………………………………………………………53
Figure 2. Peripheral expression of select genes in patients with bipolar disorder and unaffected co-twins. Gene expression plotted as percent-change relative to the expression level of members in the control group…………………………………………………………………54
Table 9. Confirmation of group-level differences in gene expression in a Finnish sample……55
Figure 3. Peripheral expression of select genes in a secondary sample of patients with schizophrenia and unaffected co-twins. Gene expression plotted as percent-change relative to the expression level of members in the control group……………………………………………………………………………56
I would like to thank all of my collaborators in Sweden and the United States, especially my mentor Tyrone Cannon and Christina Hultman from the Karolinska Institute. This work would not have been possible without the help of Desmond Smith and his graduate students including Arshad Khan and Chris Park. I would also like to thank all the members of the Biological Samples Processing Core at the University of California Los Angeles who ran the RNA assays for the project. This work was funded by an NIMH grant RO1 MH052857, Neural Phenotypes for Schizophrenia and Bipolar Disorder and individual fellowships including the University of California Graduate Division Research Mentorship Award (2007-2008), University of California Los Angeles Center for Society and Genetics Fellowship Award (2008-2009), and the University of California Los Angeles Graduate Division Dissertation Year Award (2010-2011).
VITA

EDUCATION

Amherst College (2000-2004). Amherst, Massachusetts

   B.A. Psychology, magna cum laude with distinction, May 2004

University of California Los Angeles (2005- 2012). Los Angeles, California

   M. A. Psychology, December 2006

VA Greater Los Angeles Healthcare System (2011-2012). Los Angeles, California

   Completion of Internship in Clinical Psychology, West Los Angeles VAMC, August 10, 2012

PUBLISHED ARTICLES


SELECTED ABSTRACTS AND PRESENTATIONS


**Genetic Contributions Toward Schizophrenia**

Schizophrenia is substantially heritable. A meta-analysis conducted in 1983 of existing European family studies estimated that risk for psychotic illness in family members increases in a linear fashion with degree of genetic relatedness, with second degree relatives having a lifetime risk for developing a psychotic illness of about 4% and first degree relatives having a lifetime risk of about 10% (McGue, Gottesman, & Rao, 1983). In both cases, there is a substantial increase in risk compared to the estimated lifetime risk for the general population of about 1% (McGue et al., 1983). The authors also tested a model of liability that incorporated the influence of genetic factors, common environment, unique environment and assortative mating and found that liability for schizophrenia is largely due to the influence of genetic factors, whereas the liability towards a broader psychotic disorder diagnosis in family members may also be influenced by common environmental factors and assortative mating (McGue et al., 1983). In general, these findings have been substantiated in later family studies (Beckman, Franzek, & Stober, 1996; Cardno et al., 1999; Gureje, Bamidele, & Aderibigbe, 1994; Kendler et al., 1993; McGuffin, Owen, & Farmer, 1995).

Twin studies have further documented the heritability of psychosis. Concordance rates for MZ and DZ twins have been estimated at about 45% and 15% respectively (Cannon, Kaprio, Lonnqvist, Huttunen, & Koskenvuo, 1998; Cardno & Gottesman, 2000; Davis, Phelps, & Bracha, 1995; McGuffin et al., 1995; Portin & Alanen, 1997). An analysis of adoption studies has provided further support for the notion that schizophrenia is a heritable syndrome (McGuffin et al., 1995; Wynne et al., 2006). More recent studies utilizing structural equation modeling to estimate heritability, rather than twin concordance rates, have found that the bulk of variation in risk for schizophrenia is explained by additive genetics, with unique environment effects.
explaining the rest of variance (Cannon et al., 1998; Cardno et al., 1999). While estimates of high heritability (80-90%) have been supported by a current meta-analysis of all available published twin studies (Cardno et al., 2000), it remains to be seen whether common environmental factors are in play in schizophrenia (Sullivan, Kendler, & Neale, 2003).

High estimates of heritability from family and twin studies have led to the search for individual genes that might confer risk for schizophrenia. Initial results from genetic linkage studies were promising, and many genes in the linked regions have become candidates due to their proposed biological roles, including COMT, neuregulin 1, G72, RGS4, dysbindin, and DISC1 (Chowdari et al., 2002; Chumakov et al., 2002; Harrison & Weinberger, 2005; Hennah et al., 2003; Morris et al., 2004; Schwab et al., 2003; Stefansson et al., 2002; Stefansson et al., 2003; Straub et al., 2002; Thomson et al., 2005; Tunbridge, Harrision, & Weinberger, 2006; Wang et al., 2004; Williams et al., 2004). The exact mechanisms by which these genes may contribute towards risk for schizophrenia is still under investigation; some of them, such as COMT and D-amino acid oxidase activator (DAOA, or G72), are believed to interact with specific neural pathways involved in symptom presentation and functional deficits, and others are known to exert a myriad of effects on the processes of neuronal migration, myelination, neuronal integrity and intracellular transport, all of which are highly implicated in the developmental path of the disorder (Arnold, Talbot, & Hahn, 2005; Lang, Puls, Muller, Strutz-Seebohm, & Gallinat, 2007; Sawamura & Sawa, 2006; Stefansson, Steinthorsdottir, Thorgirsson, Gulcher, & Stefansson, 2004).

However, despite the original promise of linkage studies, association studies have largely failed to find consistent and replicable associations between schizophrenia and specific single-nucleotide polymorphisms (SNPs) in any of the putative risk genes (Sanders et al., 2008). The
failure of even large-scale genome-wide association studies (GWAS) to find significant association signal may imply that the true risk variants for psychosis have yet to be identified, that variation captured on standard SNP chips is not a sufficient method of probing the genome, other types of structural variants not investigated by GWAS such as copy number variants comprise a significant portion of the meaningful genetic risk for psychotic illness, or that effects at any particular variant are so small they have not been able to be detected using current sample sizes.

**Genetic Overlap Between Schizophrenia and Bipolar Disorder**

As the field of psychiatric genetics has evolved, increasing support has been building for there to be a common genetic risk among different psychiatric illnesses, as well as disease-specific genetic variation. In looking for the genetic “causes” of schizophrenia, researchers have shown that there is a large degree of overlap in genetic risk with bipolar disorder (Berretini et al., 2000; Bramon & Sham, 2001; Craddock & Owen, 2005). The diseases are known to co-aggregate in families (Berretini et al., 2000) and twin studies using structural equation modeling have shown that about 50% of the genetic variation in schizophrenia is likely to be shared in genetic risk for bipolar disorder (Cardno et al., 2002; Lichtenstein et al. 2009). Many of the risk genes that have been identified in linkage studies of families with schizophrenia have also been implicated in bipolar disorder (Knight et al., 2009; Karege et al., 2010; Nyegaard et al., 2010; Kucukali et al., 2010; Walker et al., 2010; Li et al., 2011; Liu et al., 2011; Williams et al., 2011a, Williams et al., 2011b). Conversely, genes that were identified primarily as being involved in risk for bipolar disorder have also been shown to potentially play a role in the genetic risk for schizophrenia (Green et al., 2010; Scholz et al., 2010; Christofou et al., 2011; Curtis et al, 2011; Williams et al., 2011). Additionally, two large GWAS studies and a recent meta-analysis of
available GWAS data have supported the shared genetic overlap between schizophrenia and bipolar disorder (Purcell et al. 2009; Moskvina et al., 2009; Wang, Liu, & Aragam, 2010).

**Justification of Gene Expression Studies**

Based on the foregoing, it is generally accepted that the effects of variations in any one gene are not likely to be highly influential in the disease process of psychiatric illnesses. That is, genetic influences in psychotic disorders do not directly alter the phenotype in a simple Mendelian manner, but rather, are complex (multi-genetic) in nature, such that a large network of small genetic variations may add together and interact to produce the disruption in neural architecture and signaling that is behaviorally expressed as the schizophrenia syndrome. A similar analysis applies to all other psychiatric illness in which genetic factors are implicated, and there should be significant overlap between the genetic influences in schizophrenia and bipolar disorder. Theories of multigenetic inheritance in psychiatric illness are not a new. In fact, as far back as the 1980s, single-locus theories of schizophrenia were shown to be inconsistent with the findings of genetic and epidemiological studies of schizophrenia (O’Rourke, Gottesman, Suarez, Rice, & Reich, 1982). Theories of multigenic inheritance for schizophrenia have been further supported by molecular biology research, which has produced some specific evidence that schizophrenia risk genes may be functionally organized in interactive networks (Camargo et al., 2007).

In order for an inherited polymorphism to have a functional effect, there must be some change in gene expression that results in differences in proteins (either structurally or in expressed levels or both). Direct examinations of differences in gene expression may, in effect, cut out the “middle-man” (i.e., structural genomics) and provide a signal of disease promoting differences in gene products, regardless of the exact structural genomic change, thus helping to
resolve some aspects of the current debates in psychiatric genetics about the importance of all these “risk” genes.

The current study will examine the heritability of gene expression in patients with schizophrenia and bipolar disorder and their unaffected co-twins. Using a discordant twin design to examine gene expression allows us to examine whether disruptions in genetic expression are likely to be related to heritable risk for psychiatric illness or may be more tied to illness-related factors and long-term effects of treatment. In fact, Matigian and colleagues (2008) stated in a recent review of gene expression studies that one of the major breakthroughs in this field may come from the utility of twin studies.

**Current literature on alterations in gene expression in schizophrenia and the overlap with bipolar disorder**

Examinations of patterns of gene expression in schizophrenia have already produced some interesting findings. Many of the early studies focused on specific target genes or systems, making use of the prior literature on the neuropathology of schizophrenia. The results have supported the role of genes that relate to dopamine, GABA, glutamate, and serotonin transmission, genes involved in cellular processes, those that code for various steroids, as well as those involved in oligodendrocytes and myelination (Bullock, Cardon, Bustillo, Roberts, & Perrone-Bizzozero, 2008; Byne et al., 2008; Collinge & Curtis, 1991; Eggnan, Hashimoto, & Lewis, 2008; Feldcamp, Souza, Romano-Silva, Kennedy, & Wong, 2008; Ghose et al., 2008; Harrison, McLaughlin, & Kerwin, 1991; Mimmack et al., 2002; Polleskaya & Sokolov, 2002; Tkachev et al., 2003; Vawter et al., 2001; Vawter et al., 2002; Weickert et al., 2008; Williams et al., 2008; O’Connor & Hemby, 2007; Beneyto & Meador-Woodruff, 2008; Chagnon et al., 2008; Eastwood & Harrison, 2008; Kanazawa et al., 2008; Sharma, Grayson, & Gavin, 2008; Ding &
Hedge, 2009; Wasizuka et al., 2009; Benes, 2010; Bousman et al, 2010; Morikawa et al, 2010; Slonimsky et al., 2010; Zhan et al., 2010; Ovadia & Shifman, 2011). Similar research has also been conducted in bipolar disorder, lending support for the role of many of the same genes implicated in schizophrenia including those involved in myelination, cell growth and other cellular processes, phospholipids concentrations, glutamate, dopamine and GABA transmission (Novak & Tallerico, 2006; McCullumsmith et al., 2007; Anitha et al., 2008; Beneyto & Meador-Woodruff, 2008; Eastwood & Harrison, 2008; Barley et al., 2009; Rao, Rapoport, & Kim, 2009; Thompson et al., 2009; Uezato, Meador-Woodruff & McCullumsmith, 2009; Washizuka et al., 2009; Benes, 2010; Bitanihirwe, Lim & Woo, 2010; Bithell et al., 2010; Bousman et al., 2010; Eastwood & Harrison, 2008; Fatemi, Fulsom, & Vasquez, 2010; Kim, Rapoport & Rao, 2010; Karege et al., 2010; Rao et al., 2010; Slonimsky et al., 2010; Zhan et al., 2010; Liu & McNamara, 2011; Ovadio & Shifman, 2011). While such studies are useful, with current debates still very contentious over the validity of identified risk genes, it seems that such narrow testing is unlikely to move the field towards an agreed-upon understanding of the global picture of genetic risk for schizophrenia and bipolar disorder. Advances in genetic technology have allowed genome-wide analysis of gene expression, and studies using this approach seem better poised to answer questions about genomic processes in the disorder. These studies are divided into those using neuronal tissue (i.e., postmortem) and those using peripheral tissue, such as lymphocytes, from which RNA is extracted. As elucidated below, the two types of studies have contrasting strengths and weaknesses, and both have serious methodological challenges including reproducibility, sample size, and criteria used to establish statistical and clinical significance, and procedures to control for type I error.
Postmortem Studies. Mirnics, Middleton, Marquez, Lewis, and Levitt (2000) published the first reported analysis of genome-wide expression in patients with schizophrenia. After checking the reliability of the chip by running a single pair three times and showing that expression differences were reproducible, expression levels from the prefrontal cortex (BA9) of eleven matched pairs of patients and control subjects were examined using a UniGEM-V cDNA microarray. Pairs were matched for gender, and there were no group differences in age, PMI, brain pH or tissue storage. Among the patients, three had a diagnosis of schizoaffective disorder, five had co-morbid alcohol abuse, and two were taking antipsychotics at the time of death. Clinical diagnoses were confirmed with Diagnostic and Statistical Manual for Mental Disorders, Third Edition, Revised (DSM-III-R) diagnostic criteria using patient records. Expression levels were determined to be different if values met criteria for clinical significance (fold change > 1.9) between pairs. Statistical significance was determined by comparing the expression profiles of 250 groups of genes against the total to ascertain which functional groups were differentially expressed in schizophrenia. The authors determined that genes related to the presynaptic secretory machinery, GABA transmission, and glutamate transmission had a different expression profile in schizophrenia (Mirnics et al., 2000). Further analysis of the group of genes that had the largest fold-change (PSYN) showed that eliminating the two most changed genes in the group or the two most different patients did not affect the significant difference in expression profile between that group and the average profile of the entire group of genes that were interrogated (Mirnics et al., 2000). After identifying genes that showed a different expression profile in patients, several procedures were conducted to replicate those findings. The researchers took the two genes that had the biggest decrease in expression in patients (NSF and SYN2) and examined them in the original subjects using in situ hybridization and then again
in a new group of five matched patients and controls. In both cases, the original findings from the microarray were upheld (Mirnics et al., 2000). Then for all the PSYN genes that showed differential expression, a permutation analysis was conducted to determine if different expression profiles would be detectable in the patients regardless of the control subject used for comparison, and an ANCOVA analysis was conducted controlling for age, sex, psychosis diagnosis, co-morbid substance abuse, post-mortem interval, and brain pH. In every analysis it was confirmed that these genes were differentially expressed in patients as compared to controls and that the expression difference could not be attributed to confounding factors (Mirnics et al., 2000).

A second study of gene expression in the prefrontal cortex (BA 9) was conducted several years later (Middleton, Mirnics, Pierri, Lewis, & Levitt, 2002). Ten patients and eleven controls were examined for post-mortem gene expression. Pairs were matched for age, gender, PMI, brain pH and tissue storage. Two of the control subjects had a prior diagnosis of a psychiatric disorder (major depression and alcohol abuse). Diagnoses for all patients were performed using clinical records, toxicology studies and clinical interviews with relatives; all were reported to have met DSM-IIIR criteria for schizophrenia. Two separate arrays were used in the experiment (UniGem and UniGem-V), and therefore expression changes at each gene were normalized for the array to correct for possible differences between arrays. A z-load score was then created for each gene by calculating the product of the mean z score for each subject pair and the number of pair-wise comparisons for the gene that were significantly different between the patient and the control. Normalized expression values for genes involved in metabolism were then compared using ANOVA, with genes organized into functional groups and each gene group tested separately. Post-hoc tests were used to compare the expression profiles of a select gene group with all the expression data available for these subjects. The p values for these post-hoc tests
were then entered into a table and coded for the degree of their effect (i.e. the fold-change) in order to examine group-level differences in gene-expression. Then these p-values were transformed by taking their log and adding a sign based on whether expression was increased or decreased in patients, and interactions between gene-groups were examined using a principle components analysis. Using these procedures, expression levels in several gene groups, including those involved in energy shuttles, amino acid functioning and protein transport, were altered in patients with schizophrenia; however, none of these were significantly different in more than six pairs (Middleton et al., 2002).

A third study of gene expression in BA9 was conducted by another group of researchers using 54 samples from patients with schizophrenia and 50 matched controls; samples were provided by the Stanley Foundation (Prabakaran et al., 2004). RNA expression analysis was conducted using the HG-U133A array from Affymetrix. Genes were determined to be differentially expressed if they met criteria for statistical significance (p < .05) above and beyond a false-discovery rate computed using the Benjamini and Hochberg FDR test (q = 0.5). Using these methods 152 genes were identified as differentially expressed between patients and controls. The differentially expressed genes were then checked against demographic variables (age, gender, race, age of onset, duration of illness, suicide, brain pH, medication, PMI, family history). Demographic variables that correlated with gene expression and disease status were included in an ANCOVA model in order to determine the source of variation in expression levels. Pathway analysis using EASE and GO showed that genes that were significantly different between patients and controls were mainly those involved in mitochondria, glucose and fatty acid metabolism, and oxidative stress (Prabakaran et al., 2004).
Hakak and colleagues (2001) investigated post-mortem expression levels in the dorsolateral prefrontal cortex (BA46) of 16 patients with chronic, intractable schizophrenia as compared to 12 controls with no known neurological or psychiatric disease. Both groups had died of natural causes, and patients were long-time (35 years or more) residents of a local psychiatric hospital who had been evaluated by the researchers antemortem and determined to meet diagnostic criteria for schizophrenia according to the guidelines of the DSM-IV. There was no statistically significant difference between the groups in post-mortem interval (PMI) or age. Microarray was performed in duplicate for every sample using Affymetrix HuGeneFL Chips. Relative quantitative reverse-transcription (RT-PCR) was also performed on a portion of the samples to validate observed differences in expression between the groups all of the genes chosen confirmed the results of the microarray analysis. Genes were identified as differentially expressed between the groups if comparison of mean expression values met for statistical (p<.05 uncorrected) and clinical significance (fold-change> 1.4). This study identified 84 genes that were differentially expressed in the brains of patients with schizophrenia, including genes involved in myelination, neural development/plasticity, GABA signaling, signaling pathways linked to dopamine transmission, neuropeptide functioning and cytoskeleton arrangement (Hakak et al., 2001). Comparison of patients who were on neuroleptics prior to death and those who were not did not reveal any differences in expression of these 89 genes (Hakak et al., 2001).

Shao and Vawter (2008) also examined post-mortem expression in the dorsolateral prefrontal cortex of patients with schizophrenia and bipolar disorder, compared to controls. They used a sample from the Stanley Foundation (n = 35 for each diagnostic group). RNA was analyzed from samples in BA 46 using the Codelink platform. Outliers were detected using a principle components analysis and removed from further investigation. Expression levels were
investigated using an ANCOVA model with gender, age, and brain pH as covariates. No of the genes showed significant expression based on diagnostic group when they used a corrected p-value of 0.05 using a Benjamini-Hochberg FDR method to control for multiple comparisons. However, without any correction for multiple comparisons they found that 327 genes were dysregulated in both patient samples and 1793 were dysregulated only in one patient group but not the other. A secondary analysis looked at only those subjects that had a brain pH above the median. In this group 280 genes were identified as differentially expressed in both patient populations and 3069 were differentially expressed in only one patient population. Combining the results from both analyses, 78 genes showed shared dysregulation in bipolar disorder and schizophrenia. These were analyzed using IPA and were found to be most significantly related to nervous system development, neuroglia processes, cell death, and immune development and response.

Kim and colleagues (2010b) provided the most recent examination of post-mortem gene expression in the dorsolateral prefrontal cortex. They looked at the expression of micro-RNA using tools and assays developed by Taqman. The expression data was log-transformed and analyzed using and ANCOVA model with pH, age, RIN, sex and disease status as covariates. Genes were identified as being differentially expressed if they met for statistical significance with a Benjamini-Hochberg false discovery rate of 10%. Seven genes were identified as differentially expressed between patients with schizophrenia and controls. Fifteen genes were differentially expressed between patients with bipolar disorder and controls. All but one of these genes were confirmed using RT-PCR. Then these genes were analyzed using a network-based approach developed by IP and it was shown that the genes were related to nervous system functioning and might regulate glutamate and dopamine neurotransmission.
Choi et al. (2011) performed a combined analysis of seven expression studies using samples from the Stanley Foundation and an Affymetrix array. They examined expression differences in patients with schizophrenia and bipolar disorder with psychotic features compared to those without psychotic illness (bipolar disorder without psychotic features, depression and controls). They used a p-value of 0.05 uncorrected and a fold change > 1.3 to identify genes that were significantly dysregulated based on diagnosis. After controlling for gender, age, PMI, brain pH, side of the brain, smoking at time of death, and sudden death, a combined sample from all the studies analyzed showed that metallothioneins were consistently up-regulated and neuropeptides were consistently down-regulated in patients with schizophrenia and bipolar disorder with psychotic features.

Kuromitsu and colleagues (2001) looked at post-mortem gene expression in patients with schizophrenia, bipolar disorder, major depression and non-psychiatric controls. Using brains supplied from the Stanley Foundation, all groups (N = 15 per group) were matched for age, gender, PMI, and quality of mRNA stability. Whole-genome expression in the frontal lobes (BA 10) and temporal lobes (BA20) was compared between groups using Affymetrix GENCHIP arrays. For both patients and controls, expression analysis was conducted twice using four randomly selected points in the frontal lobes that did not overlap between sets. In this way schizophrenia sample 1 was compared to control samples 1 & 2 as was schizophrenia sample 2. Genes were selected as being differentially expressed if they met for clinical significance (fold change >2). Several genes had altered expression in two analyses, but only two genes showed consistently different expression in all four analyses: p72 and NPY (Kuromitsu et al., 2001). In order to further examine these differences expression levels of p72 and NPY in the frontal and temporal lobes, 120 samples were examined from all four clinical groups using a quantitative
PCR analysis. Results showed that NPY expression was decreased in the frontal lobes for patients with schizophrenia and bipolar disorder (Kuromitsu et al., 2001).

Moreau et al. (2011) also examined genome-wide expression in the frontal lobes (BA 9) of patients with schizophrenia and bipolar disorder as compared to controls. They used samples provided by the Stanley Medical Research Institute (n = 35 for all three diagnostic groups). Examining the expression of micro-RNAs using tools developed by Taqman, they applied Bayeseian modeling to the expression values to determine whether expression levels were significantly correlated with diagnostic status. They found that 19% of the miRNAs analyzed showed expression levels that were altered based on psychiatric diagnosis. Most of the miRNAs selected in this manner showed decreased expression and they found that every miRNA with altered expression in the schizophrenia sample also had altered expression in the bipolar disorder sample. Furthermore, 24 miRNAs showed expression levels that were altered based on diagnostic status with 95% certainty.

Hemby et al. (2002) examined genome-wide expression in the entorhinal cortex in 8 patients who had been hospitalized for chronic schizophrenia and 9 matched controls with no known psychiatric or neurological illness. All patients had been prospectively recruited and diagnosis was confirmed by the researchers using DSM-IV criteria. In this study genome-wide analysis was conducted using cDNA with a clinical significance cutoff (fold-change >2) and specific target genes chosen based on prior literature were also examined using a statistical significance cutoff (p < 0.05 uncorrected). There were no differences between the two groups in age, PMI or brain weight. Over 2500 genes were identified as being up-regulated in the patients with schizophrenia and more than 1500 were down-regulated (Hemby et al., 2002). The differentially expressed genes identified in this study included those involved in synaptic vesicle
transport, monoamine transmission, GABA transmission and glutamate transmission (Hemby et al., 2002).

Bowden, Scott, & Tooney (2008) published a recent analysis of genome-wide expression patterns in the superior temporal gyri (BA 22) of seven patients with schizophrenia compared to controls. Each patient was matched by age, gender, PMI and brain pH to a control subject to allow for pair-wise comparisons of gene expression. RNA samples were harvested from the post-mortem tissue and oligonucleotide microarrays were conducted using Compugen Version 219K arrays analyzed using GeneSpring 5.0. Relative real-time PCR was also conducted to confirm expression differences in identified genes. Genes were identified as differentially expressed if they met criteria for clinical significance (fold-change > 1.5) in more than 50% of the pairs. A one-class analysis of log 2 transformed expression data was also conducted using SAM version 2.0 and 128 permutations. This method identified 216 genes that were significantly down-regulated and 85 that were significantly up-regulated in schizophrenia with a false discovery rate of 4.7%. (Bowden et al., 2008) Many of the genes identified as differentially expressed had been previously implicated in risk for schizophrenia and were involved in processes such as glutamate neurotransmission, myelination, and intercellular signaling (Bowden et al., 2008).

Several studies have been published examining post-mortem genome-wide expression in the central nervous system of patients with bipolar disorder. As reviewed above, Kiromitsu et al. (2001) found that NPY expression was decreased in the frontal lobes for patients with bipolar disorder, Choi et al. (2011) found that metallothioneins and neuropeptides were significantly dysregulated in the prefrontal cortex of patients with affective and non-affective psychosis, and Shao and Vawter (2008), Mireau et al. (2011), and Kim et al. (2011) all found that most genes
that were dysregulated in patients with schizophrenia were also dysregulated in patients with bipolar disorder. Other studies looking at expression in the frontal lobes have shown that genes involved in cellular growth and neuronal development (Nakatani et al, 2006), mitochondrial functioning and oxidation (Sun et al., 2006) and ubiquitin pathways (Ryan et al., 2006) may all be dysregulated in bipolar disorder.

**Peripheral vs. central expression.** Most of the expression studies to date have focused on an examination of post-mortem levels of mRNA collected from specific brain regions of interest (i.e. the prefrontal cortex, thalamus, hippocampus or cerebellum). Post-mortem studies provide the benefit of testing specific hypotheses about differences in gene expression in the brains of patients suffering from schizophrenia; however they also come with a myriad of confounding factors. Due to the postmortem nature of the studies, expression information usually cannot be correlated with performance measures in patients, and diagnostic status often cannot be independently confirmed by the researchers themselves. Likewise, these samples may not be a representative sample of patients since for the most part they represent individuals who died from unnatural causes related to their illness (i.e. suicide). Furthermore, there is a lot of evidence to suggest that post-mortem RNA is highly affected by the pH levels in the brain, which are in turn affected by the state of the person prior to death. Other factors such as medication status may also have a large impact on expression and may not be easy to document or control for in these studies. Finally, there is of course a limited supply of post-mortem tissue available which results in studies being conducted on very small sample sizes (typically less than 10 patients), which may dramatically affect the ability to identify important expression variants, and expression levels may vary dramatically between different cell types, so the results may not provide any information about broader patterns in gene expression.
In contrast, the current study proposes to examine gene expression variation from peripheral blood samples. This method is much less invasive and will allow for a larger sample size and direct comparison with data collected from patient interviews, questionnaires, brain imaging, and neuropsychological performance measures. Prior research has suggested that peripheral blood RNA is under genetic regulation and thus may be an appropriate means of examining the heritability of complex diseases (Cheung et al; 2003). Peripheral RNA has also been shown to be reliably measured in several studies.

Certain limitations obviously exist in the choice of peripheral mRNA. Perhaps the most obvious and important issue is that circulating levels of mRNA in lymphocytes may not accurately reflect what is expressed in the brain. Indeed, a recent study of mRNA levels in patients with schizophrenia attempted to correlate expression in the brain and lymphocytes and found very little overlap in the genes that were differentially expressed in the separate methods (Matigian et al., 2008). However, as reviewed below, other studies have shown that alterations in the expression of important CNS genes can be detected in both peripheral and central mRNA.

Ilani and colleagues (2001) showed that while the dopamine receptor subtypes D1 and D2 do not show peripheral expression, peripheral expression of D3 receptors was both detectable and elevated in patients with schizophrenia, in accordance with prior post-mortem studies. Later work showed that in a similar fashion, genetically mediated differences in the expression of serotonin receptors in patients with schizophrenia were detectable in peripheral lymphocytes (Hranilovic et al., 2004). Recent work has shown that expression of genes in DISC1-associated networks is detectable using peripheral blood samples and also under genetic control (Ikeda et al., 2008). Many signaling molecules in the immune system that are involved in inflammation, such as cytokines, are capable of crossing the blood-brain barrier and are thus expressed both
centrally and peripherally (Banks, 2005; Kronfol & Remick 2000; Watkins, Maier & Goehler, 1995). Finally a recent review of studies examining peripheral and central expression of neural growth factors found that when comparing expression levels in the CSF and peripheral blood, positive correlations exist for S100B, BDNF and glutamate (van Beveren et al., 2006). Thus, while it is true that there are certain genes that regulate neural functioning that are not expressed peripherally, for all of the major groups of genes that are implicated in schizophrenia (monoamine function, glutamate function, neural growth, myelination, inflammation), there are genes which are also expressed in peripheral lymphocytes, and this is expression is likely to be heritable and may correlate with activity in the CNS.

Peripheral expression in schizophrenia. Vawter et al. (2004) published the first available large-scale study of peripheral expression in patients with schizophrenia. Their sample included expression levels measured from lymphocytes in five patients diagnosed with schizophrenia and nine of their unaffected family members. All subjects were members of a large Western/Northern European pedigree with many members affected with schizophrenia. Patients and family members were matched in terms of gender and age. Expression in lymphocytes was measured following transformation with the Epstein-Barr virus and RNA expression and cDNA synthesis were further confirmed using RT-PCR. A custom array was chosen in order to screen samples for expression differences in genes that were known to have effects on neural development or functioning and which had been previously associated with psychiatric illnesses. Arrays were run twice for each sample in order to correct for instrument error. Gene expression was analyzed using a Student’s T-test, examining differences in mean expression levels between patients and family members. Genes were identified as being differentially expressed if they met the criteria for statistical (p <.05 uncorrected) and clinical significance (fold change > 1.4).
Those genes that met criteria were then analyzed in post-hoc tests, excluding four family members who were below the cut-off age for conversion to psychosis. Nine genes were identified as differentially expressed in patients compared to controls using these criteria (Vawter et al., 2004). The genes that showed decreased expression in patients were APBA3, NPY1R, GNAO1, and PCDH2; those that showed increased expression were PTP4A2, HOXA13, MPP3 and MDH1 (Vawter et al., 2004). Several of these genes have been implicated in prior linkage and association studies. Differences in the expression of NPY1R and GNAO1 were confirmed using RT-PCR, however expression differences in MDH1, APBA3 and PTP4A2 were not consistent when examined using RT-PCR (Vawter et al., 2004). Twenty-one additional genes met criteria for statistical significance but not clinical significance (Vawter et al., 2004). While the authors did not further investigate any of these genes, it should be noted that some of them included genes that have been previously implicated in schizophrenia. Furthermore, there is no real support for the idea that a larger-fold difference in expression is a greater biological difference, as subtle changes in expression may have dramatic effects in certain systems, and so the results of this study are not inconsistent with the role of these genes in schizophrenia.

Middleton et al. (2005) published the next study examining peripheral expression levels in patients with schizophrenia. Their sample included 33 sibling pairs from families with multiple members affected with psychosis and that had shown linkage to 5q in prior studies. RNA was analyzed using a Bioanalyzer RNA chip (Agilent). Genome-wide expression levels were examined using a pair-wise analysis of fold-change and a statistical significance of $p < .05$. Next a pathway analysis was conducted wherein expression levels were scaled and normalized using a robust multi-chip analysis and mean patient/sibling expression ratios were calculated for those genes that showed expression in patients that differed from the median of the dataset.
Functional groups were identified using available information from several online databases and gene ratios were calculated for each group and then a t test was applied to see if the average expression of genes in that group differed significantly from the panel as a whole. The authors identified 2,000 genes that showed differential expression in patients with schizophrenia compared to their siblings (Middleton et al., 2005). A multiple testing correction (Benjamini-Hochberg) was applied and provided a list of 300 genes that still met criteria for statistical significance. The genes identified were involved in several biological systems and pathways including immune and inflammatory functioning, brain development, cell signaling pathways and growth factors (Middleton et al., 2005). Perhaps most interestingly, one previously mentioned schizophrenia risk gene, NRG1, was found to have increased expression in patients compared to siblings in this sample (Middleton et al., 2005).

Bowden and colleagues (2006) conducted genome-wide analysis of peripheral RNA expression in 14 patients with schizophrenia and 14 non-psychiatric controls using a CyScribe cDNA array. Expression differences were calculated by using paired t-tests examining differences between patients and controls that were matched for age and gender. This approach yielded 2500 genes that were down regulated and an additional 1000 that were up regulated in patients with schizophrenia (Bowden et al., 2006). Of these genes, 18 are known to be expressed in the brain and affect functions such as neural development, myelination, metabolism and cellular signaling (Bowden et al., 2006).

Bousman et al. (2009) conducted a genome-wide analysis of peripheral RNA in 24 patients with schizophrenia, 23 patients with bipolar disorder and 24 normal controls recruited in San Diego and Taiwan. Expression was measured in lymphocytes using the Human Exon 1.0 ST arrays from Affymetrix. Genes were labeled as being differentially expressed if they met for a
Benjamini-Hochberg FDR of 10%. While 557 genes were found to be dysregulated in patients with schizophrenia and 917 in patients with bipolar disorder at an uncorrected p value of 0.05, none of the findings remained significant after applying the correction for multiple comparisons. After running the results through an IPA analysis for relevant pathways it was found that the protein ubiquitization pathway showed significant dysregulation in patients with schizophrenia and bipolar disorder in both the San Diego and Taiwan populations.

Most other studies examining peripheral gene expression in patients with bipolar disorder have looked at the role of medications or other long-term illness factors on gene expression. However a study conducted by Padmos et al. (2008) looked at genome-wide expression in monocytes of 42 patients with bipolar disorder, 54 offspring with a parent with bipolar disorder and 25 healthy controls. Expression levels were measured using the U95Av2 microarray platform from Affymetrix. Genes were considered differentially expressed if they met for statistical (p < 0.01 uncorrected) and clinical (fold change> 2.0) significance. This identified 71 genes as being differentially expressed in bipolar disorder. A pathway analysis by IPA identified inflammation, cell movement, and cellular growth and differentiation as the major biological pathways disrupted.

In summary, results from studies examining peripheral expression in patients with schizophrenia have implicated differential expression in many genes that affect neural functioning. Additionally, many of these same genes and pathways have also been shown to be disrupted in patients with bipolar disorder. This evidence confirms that it is possible to identify alterations in gene expression in patients with psychiatric illness using PBLs and suggests that many genes expressed in the brain are also expressed in PBLs. Additional studies have shown that PBLs may allow for both the detection of heritable differences in gene expression between
patients, controls and family members and response to treatment (Na & Kim, 2007; Zhang et al., 2008). These data suggest that this method may in the future allow a means of identifying patients, evaluating treatment response and potentially offering more targeted interventions to correct disruptions in gene expression.

**Gene Expression and Endophenotypes**

While examining patterns of gene peripheral gene expression in patients with schizophrenia may eliminate some of the problems plaguing structural genetics research, psychiatric diagnoses may not be the best phenotype for investigating genetic risk in psychiatric illness. Since the each disruption in gene transcription is not expected to directly cause the complete illness phenotype, intermediate phenotypes may be better poised to screen for the relevant genetic risk. These intermediate phenotypes that link genetic variation to a disease have been termed “endophenotypes.” In order for a phenotype to qualify as a relevant endophenotype for a psychiatric disorder, it must be heritable, associated with the causal (primarily heritable) aspects of the disorder and not the treatment or long-term disease effects, and should vary continuously in the general population (Cannon & Keller, 2006).

Endophenotypes have been studied extensively in schizophrenia. One of the most widely studied and supported endophenotypes in psychosis is neurocognitive dysfunction, which is particularly severe in the domains of short- and long-term memory. Neurocognitive deficits have been found to meet all the criteria necessary for an appropriate endophenotype as they are highly heritable, are present when a patient is asymptomatic, vary continuously in the general population and vary in a direct linear manner with degree of relatedness to a person with psychosis (Cannon et al., 2001; Tuulio-Henrikson et al., 2002; Bertisch et al., 2010; Toulopoulou et al., 2007). Modeling of heritability of neurocognitive functioning in patients with psychosis
has shown that general intelligence and working memory are likely to be the most useful neurocognitive endophenotypes in the study of schizophrenia (Cannon et al., 2001; Aukes et al., 2009; Touloupoulo et al. 2007).

Endophenotypes have also been explored in bipolar disorder. While the neurocognitive deficits present in schizophrenia are not all seen in patients with bipolar disorder, disruptions in cognition are still implicated as potential endophenotypes. A meta-analysis of studies of cognition in bipolar disorder found that the most promising endophenotypes are disruptions in executive functioning, verbal memory and sustained attention (Bora, Yucel, & Pantelis, 2009). However, a more recent large pedigree study of families with bipolar disorder found that psychomotor speed, working memory and declarative memory were the best cognitive endophenotypes for bipolar disorder (Glahn et al., 2010). Therefore, while it is clear that there is some overlap between cognitive endophenotypes for schizophrenia and bipolar disorder, the degree of overlap has not been fully established and disruptions in cognition may be more prominent and wide-reaching in schizophrenia than in bipolar disorder, or more tightly linked to genetic risk for psychosis than for mood disruption.

As cognitive endophenotypes for schizophrenia have become more widely studied and accepted, so has their usefulness in genetic research. A recent linkage study found that performance on cognitive tasks provided a stronger linkage signal than using psychiatric diagnosis (Lien et al., 2010). This adds further support to the idea that cognitive performance is highly heritable; however the genomic region identified was not one that had been previously implicated in schizophrenia risk. Therefore, the findings may just pertain to genetic factors in cognition and not in psychosis even though the sample contained patients with schizophrenia and their unaffected family members. However, a recent gene expression study did clearly show
that using cognitive endophenotypes may enhance the ability to detect meaningful genetic variation that is related to psychotic illness. In that study, a post-mortem sample confirmed that kynurenic acid expression was disrupted in patients with schizophrenia and a second clinical sample showed that a “risk” SNP was associated with cognitive performance in patients with schizophrenia although there was not a significant detectable association with psychiatric diagnosis (Wonodi et al., 2011). This fully supports the notion that cognitive endophenotypes may be better poised to screen for genetic risk, since they are more tightly linked to the genetic disruptions inherent in psychosis than the syndromal diagnosis. Adding even further support to this idea is a recent study from the Consortium on the Genetics of Schizophrenia. Researchers found that a SNP panel enriched with genes of interest to schizophrenia showed multiple associations with cognitive performance on relevant endophenotypes (Greenwood et al., 2011). Therefore, although studies have been underpowered to confirm the initial findings of linkage studies by showing associations between relevant SNPs and psychiatric diagnosis, these SNPs are related to cognitive performance in patients with schizophrenia. Based on the prior literature, the present study uses a cognitive endophenotype to screen for patterns in peripheral gene expression relevant to genetic risk in schizophrenia and bipolar disorder.

**Goals and Hypotheses of the Present Study**

The aims of the present study were to identify variations in RNA transcription that are related to the endophenotype of cognitive performance and to determine whether genes that show cognition-related variations in RNA transcription are also dysregulated in patients with schizophrenia and bipolar disorder and their co-twins.

**Hypothesis 1:** As cognitive performance is highly heritable, there will be many genes whose expression levels vary in relation to cognitive performance, controlling for demographic
factors such as age and gender.

**Hypothesis 2a:** Among the probes whose expression levels vary in relation to cognitive performance, several will show expression patterns that are also be related to genetic risk for psychosis (i.e., differential expression in patients and their non-affected co-twins compared with controls).

**Hypothesis 2b:** Of those genes whose expression patterns vary in relation to genetic risk for psychosis, some will show variation specific to schizophrenia and others will show a general pattern of disruption across schizophrenia and bipolar disorder.

**Hypothesis 2c:** Genes whose expression patterns are found to vary in relation to psychotic illness in the primary study sample will also vary in relation to psychotic illness in a second, independent sample studied using the same methods.

**Research Design and Methods**

**Subjects**

Data collection for the primary study was performed in Sweden: all participants were Swedish citizens. The participants were identified through a linkage of the Swedish Twin Register and the Hospital Discharge Register. We recruited twins who were between the ages of 35 and 65 years at the time of evaluation (i.e., for a five-year study beginning in 2006, those born between 1940-1975, inclusive). A minimum age of 35 was used to avoid sampling pairs in which an unaffected co-twin will develop SZ after the assessment is complete; given the high heritabilities of age-at-onset in SZ, an onset after age 35 in an individual with an already affected co-twin is expected to be rare. Although subjects with clear evidence of dementing disorders were excluded, we used an upper age of 65 years to reduce the likelihood that sub-clinical
disorders of aging would affect the neuroimaging and cognitive data.

In total 170 individuals provided blood samples and were included in this study. Within this sample, 35 individuals had a diagnosis of schizophrenia and 34 were unaffected co-twins of patients with schizophrenia (21 discordant twin pairs, 2 concordant pairs), 24 individuals had a diagnosis of bipolar disorder and 17 were unaffected co-twins of patients with bipolar disorder (15 discordant twin pairs, 2 concordant twin pairs), 4 additional discordant twin pairs came from families where one twin had bipolar disorder and the other twin had schizophrenia, and 60 individuals were normal controls (29 intact twin pairs). A breakdown of concordant and discordant MZ and DZ pairs can be found in Table 1.

Measures

Diagnosis and Clinical Symptom Ratings

Clinical diagnosis was assessed via clinical interview using the Structured Clinical Interview for DSM-IV (SCID-IV) (Spitzer et al., 1994) according to Diagnostic and Statistical Manual for Mental Disorders, Fourth Edition (DSM-IV; American Psychological Association, 1994) criteria. Symptom severity was assessed using the Scale for Assessment of Negative Symptoms (SANS, Andreasen, 1984) and Scale for Assessment of Positive Symptoms (SAPS, Andreasen, 1984).

Neuropsychological Assessment

A standard test battery was administered to all participants. These measures had all been translated into Swedish and back-translated into English and have been validated and used in prior studies (citations). The test battery included the following established tests:

Weschler Intelligence Test (WAIS-R, Weschler D. 1981): The Vocabulary, Similarities, Block Design, Matrix Reasoning and Digit Symbol subtests were used to evaluate participants’ verbal and nonverbal knowledge and reasoning and psychomotor processing speed.
Weschler Memory Scale (WMS-R, Weschler D. 1997): The Digit Span, Spatial Span, Logical Memory, and Visual Reproduction subtest were used to evaluate participants’ complex and working memory in the verbal and spatial domains.

Verbal Fluency Test (Benton A & Hamsher K, 1976): This test was used to evaluate participants’ semantic and phonemic verbal fluency.

Trail Making Tests A and B (Reitan R. & Wolfson D., 1985: These tests were used to evaluate participants’ visuomotor sequencing and set-shifting abilities.

California Verbal Learning Test (CVLT-I, Delis D. et al., 1983): This test was used to evaluate participants’ verbal learning, memory and attention abilities.

Divided Attention Task (Vilkki J. et al. 1996): This test was used to evaluate participants’ visual attention ability.

Swedish Verbal Working Memory Test (Malm et al. 1988): A Swedish computerised version of the Daneman-Carpenter reading span task (Daneman et al., 1980) used to evaluate participants’ verbal working memory capacity.

Procedures

Subject Ascertainment and Study Enrollment

Potential subjects from index pairs were contacted initially through their local health care districts. If they were interested in participating based on this initial contact, they and their co-twin were scheduled for a visit to the Karolinska Institute. Potential subjects from control pairs were sent a letter through the Swedish Twin Registry, and those interested in participating are scheduled for a screening evaluation. Subjects then underwent the informed consent process for the diagnostic, neurocognitive, and neuroimaging procedures of the study, as well as the blood
Subjects could decline to participate in any specific research measure or procedure and still participate in the remaining research measures and procedures. Subjects were invited to participate in the remaining study procedures after their eligibility was determined.

Subjects were included if they meet the following criteria:

1. Have a same-sex co-twin who is potentially available for study.
2. Probands must meet DSM-IV criteria for schizophrenia, schizoaffective disorder, or schizophreniform disorder, or bipolar disorder.
3. Co-twins in discordant pairs must not meet DSM-IV criteria for any psychotic disorder or bipolar disorder as determined by searches of the Swedish Hospital Discharge Register and direct clinical interview.
4. No serious medical illnesses such as cardiovascular, respiratory, endocrine, or neurological disease, including epilepsy.
5. Fluency in Swedish language.
6. Between the ages of 35 and 65 years at the time of evaluation.
7. In healthy control pairs, neither twin can meet DSM-IV criteria for any psychotic disorder or bipolar disorder as determined by searches of the Swedish Hospital Discharge Register and direct clinical interview.

Subjects were excluded for the following reasons:

1. Presence of a neurological disorder
2. History of significant head injury with loss of consciousness
3. Other physical disorder that could affect brain functioning
4. Mental retardation
5. Unable to read or comprehend spoken and written Swedish
6. Lactating or pregnant female

7. Not between the ages of 35 and 65 years at the time of evaluation

The control twins were matched to the index twins in terms of zygosity, gender, handedness, year of birth, parental social class, and duration of cohabitation. All patients were expected to be on maintenance antipsychotic medications under the supervision of their own treating psychiatrists. The study did not involve any changes to patients’ medication regimens.

The samples were representative of the ethnic variation in Sweden (i.e., Swedes, Lapps, and Finns). No subject was excluded on the basis of gender, race, or ethnicity. Analysis of the demographic data revealed that the sample of patients with bipolar disorder and their unaffected co-twins contained significantly more females than the sample of patients with schizophrenia and their unaffected co-twins ($\chi^2 (1) = 9.63\times10^{-13}, p = 7.57\times10^{-8}$). Patients and co-twins with schizophrenia did not differ significantly from control twin pairs in terms of gender distribution ($\chi^2 (1) = 0.212$). There was no significant difference between any of the twin groups with regard to age. Demographic information on the subjects can be found in the Table 2.

The research staff inviting subject participation were trained Psychiatric Nurses, Clinical Psychologists, and licensed psychiatric social workers. Data collection was performed at the Karolinska Institute. Neuropsychology assessments were conducted by Psychology Ph. D. students who had been trained by neuropsychology faculty members at the Karolinska Institute. Subjects were reimbursed fully for their travel expenses and given meal vouchers on days of participation. In addition, subjects were given a fee of $15 US Dollars per hour per procedure (corresponding roughly to $100 for their participation) to help compensate for any lost wages due to their participation. Blood samples and imaging, neurocognitive, and clinical data were de-
identified and sent to UCLA for processing and analysis.

Collection of genetic samples

To facilitate the genetic analysis, a 30 ml (3 x 10 ml each) blood sample was taken from both members of each pair. One of the samples (10 ml) was immediately frozen and held in reserve at the Biobank at the Karolinska institute. DNA and RNA were extracted from the remaining two samples (20 ml), and aliquots of DNA and RNA were sent to UCLA for analysis. Genetic analysis occurred at UCLA under the supervision of Desmond Smith and other members of the Human Genetics Department. RNA analysis was performed using a chip designed by the Illumina Company, Human WG6 v3.0 in the Biological Samples Processing Core under the supervision of Dr. Joseph DeYoung. The following quality control measures were employed in order to ensure the validity of the RNA signal.

1. The first few control samples were run multiple times in order to examine the stability of the RNA signal across different chips.

2. All of the samples were run in duplicate and correlations between samples A and B for each person were evaluated to ensure a high degree of correlation between samples (Intra-class correlation, r > 0.85). All subjects had good correlations between their sample A and sample B data, (ICCs ranged from 0.902 – 0.998). Therefore, no subjects were eliminated at this stage.

3. Industry standard quality-control measures were implemented under the guidance of Dr. Smith and other members of the Human Genetics Department. Namely:
   a. Samples were checked prior to analysis to ensure a useable quantity of RNA was present. All samples met this threshold.
   b. All samples were preprocessed using the Illumina Bead Studio package.
c. Raw sample values were then rank normalized and a background subtraction was conducted to remove background variance.

d. Subjects whose samples did not provide a high enough quality signal, i.e. where either sample A or sample B had an overall average signal strength that was not significantly greater than 0 after rank normalization and background subtraction (analyzed using a Student’s t-test, Bonferroni correction, \( \alpha = 0.05 \)), were excluded from further analysis. This dropped the number of usable probes from 24,526 to 18,559 probes.

4. After quality control, duplicate samples were averaged to produce one vector of expression values for each participant.

5. Average samples were then log-transformed to ensure normalization of the sample distributions.

Secondary study methods

Data collection for the secondary study was conducted in Finland at the National Public Health Institute using the same procedures outlined for the primary study. More information on the procedures and sample characteristics of the secondary study can be found in (Cannon et al., 2000a). Demographic information on the Finnish sample is available in Table 3.

Statistical Design and Analyses

All analyses were performed using the statistical package R (R Development Core Team, 2009).

Aim 1: Identify variations in RNA transcription that are related to the endophenotype of cognitive performance.
**Hypothesis 1:** As cognitive performance is highly heritable, there will be many genes whose expression levels vary in relation to cognitive performance, irrespective of other demographic factors such as age and gender.

In this analysis, the relationship between expression and cognitive performance was examined in the full sample. First, a principal components analysis was conducted in order to find a primary component that would account for the largest share of variance in cognitive performance. Total raw scores from all the neurocognitive tests in the standard battery were entered in the model. Next, the first principal component was selected and a new vector of scores was created for each participant based on the factor loadings for this component. This vector of cognitive performance was then entered into iterative mixed model multiple regressions with family entered as a random variable, RNA expression levels for an individual probe entered as a predictor, and covariates of age and gender used to model the outcome of cognitive performance. A Bonferroni-corrected significance level of $\alpha = 2.69 \times 10^{-6}$ was used in order to evaluate whether specific genetic products (RNA levels) were significantly related to cognitive performance. The Bonferroni correction was computed by dividing the nominal alpha of 0.05 by the number of probes that were retained after quality assurance tests (i.e., 18,559). A secondary analysis was done with a more relaxed threshold $\alpha = 5.39 \times 10^{-6}$, which corresponds to a corrected $p$ value of 0.10, in order to explore a greater number of probes that may be related to cognition and genetic risk for psychosis. Finally, a network analysis was conducted using DAVID (Huang, Sherman, & Lempicki 2009a; Huang, Sherman, & Lempicki 2009b) in order to identify the biological processes related to these genes, so that the global patterns of cognitively relevant expression variation could be more easily viewed.
Aim 2: Determine whether genes that show cognition-related variations in RNA transcription are dysregulated in patients with schizophrenia and their co-twins.

Hypothesis 2a: Among the probes whose expression levels vary in relation to cognitive performance, several will show expression patterns that are also related to genetic risk for psychosis.

In this analysis, the genetic products nominated in Aim 1 as related to cognition were further examined to see if their expression patterns vary as a result of genetic risk for psychotic illness. Iterative t-tests were done to evaluate for each probe whether expression levels were significantly different between controls and the combined sample of all patients and patients’ co-twins. A significance level of $\alpha = 0.05$ was used with a Benjamini-Hochberg correction for multiple comparisons. The Benjamini-Hochberg approach (Benjamini & Hochberg, 1995) is the standard method for controlling the false discovery rate in multiple hypothesis testing, and it provides a measurement of the proportion of false positives found in all the specified tests, based on the distribution of the p values. This test is a less conservative approach than using the family-wise error rate provided by a Bonferroni correction. Given the availability of an independent sample on which to test replication, the decision was made to use a false discovery rate correction, in order to maximize our power to identify expression patterns that were significantly related to genetic risk for psychosis, while still accounting for the number of statistical tests being conducted at this stage.

Hypothesis 2b: Of those genes whose expression patterns vary in relation to genetic risk for psychosis, some will show disease-specific variation and others will show a general pattern of disruption across different diagnostic groups.
In this analysis the genes nominated by the results of hypothesis 2a were further examined to see if variations in gene expression were related to genetic risk for schizophrenia, bipolar disorder, or both. Again, iterative t-tests were used to see if variations in gene expression differed between patients, cotwins, and controls. The significance threshold was kept at $\alpha = 0.05$ with no correction applied, as this step was meant to be a descriptive evaluation of those genes nominated in hypothesis 2a and was not used to nominate additional genes that might be related to specific illnesses.

**Hypothesis 2c:** Genes whose expression patterns are found to vary in relation to psychotic illness in the primary study sample will also vary in relation to illness in a second, independent sample studied using the same methods.

In order to further control for false positives, those genes that were found to be significantly related to genetic risk for psychosis in Hypothesis 2a were examined in an independent sample of Finnish twins who had been tested with the same battery of cognitive tests. Again, t-tests were conducted to examine whether expression of these genetic products varied in relation to diagnostic status. The significance threshold was kept at $\alpha = 0.05$ with no correction for multiple comparisons, since this analysis constituted a replication test and only markers that were significant (after FDR correction) in the first sample were tested.

**Results**

**Hypothesis 1: Variations in peripheral RNA expression are related to cognitive performance.**

The principle components analysis of neurocognitive performance variables revealed five distinct factors (Table 4). The first principal component accounted for 39.6% of the variance in
neurocognitive performance (Eigen value = 9.113); the loadings of the items in the rotated matrix are presented in Table 5. The neurocognitive domains that were most strongly associated with this first principal component were delayed complex visual memory, verbal learning and recall, and verbal fluency. A summary variable reflecting the weighted linear combination of these variables based on the factor loadings was moved forward for subsequent analyses.

Twenty-eight genes showed variation in their peripheral expression values that was significantly associated with the summary neurocognitive variable in the full sample after controlling for gender, age, twinship, and multiple comparisons (i.e., Bonferroni corrected alpha level of \( p < 2.69 \times 10^{-6} \)). These genes are shown in Table 6.

When the significance threshold was relaxed slightly (i.e., Bonferroni corrected alpha level of \( 2.70 \times 10^{-6} < p < 5.39 \times 10^{-6} \)), an additional 19 genes showed variation in their peripheral expression values that was significantly associated with the summary neurocognitive variable in the full sample after controlling for gender, age, and twin-ship. Full information on these results can be found in Table 7. Further examination of all 47 genes showed that they were mostly related to the biological processes of the assembly and regulation of intracellular organelles, RNA processing and transcription regulation, mitochondrial functioning, DNA metabolism and repair, cellular protein metabolism and assembly, purine binding, protein kinase binding, immune system functioning and apoptosis, cell cycle regulation, zinc and ion binding, and cell membrane assembly.

**Hypothesis 2a: Variations in peripheral RNA expression are related to psychotic illness.**

Those genes that were significantly associated with neurocognitive performance were then investigated further to see if their peripheral expression values differed significantly
between patients with schizophrenia or bipolar disorder, their unaffected co-twins, and controls. Of the initial 28 genes that met the most conservative Bonferroni-corrected significance threshold, seven showed expression patterns that varied significantly as a result of being from a proband twinship (i.e., schizophrenia or bipolar disorder): TMEM203, ATIC, HSPA9B, CDC16, MCM3, MAGED1, FTSJ1. For all of these genes, patients and unaffected co-twins showed under-expression of the gene in peripheral blood samples compared to controls. When the examination was expanded to include an additional 19 genes that met the less stringent significance threshold, an additional five genes showed expression patterns that varied significantly as a result of being from a proband twinship: MAP4K1, DNMT1, DSTN, AP4B1, EIF2B1. Again for all of these genes, patients and unaffected co-twins showed under-expression of the gene in peripheral blood samples compared to controls. These results are represented in Table 8.

Hypothesis 2b: Variations in peripheral RNA expression are specifically related to schizophrenia or bipolar disorder.

Further investigations were made to look at the relationship between those 12 genes and specific populations. Eleven of these genes were significantly under-expressed in patients with schizophrenia compared to controls, and the remaining gene (AP4B1) showed marginal significance (0.05 < p < 0.10) for having lower peripheral expression in patients with schizophrenia as compared to controls. Nine of the twelve genes (TMEM203, CDC16, MCM3, MAGED1, DNMT1, DSTN, AP4B1, and EIF2B1) were also significantly under-expressed in unaffected co-twins of patients with schizophrenia compared to controls. Of the remaining three genes, two (HSPA9B and FTSJ1) showed marginal significance for having lower peripheral
expression in unaffected co-twins as compared to controls. Peripheral expression of these twelve genes did not differ significantly between patients with schizophrenia and their unaffected co-twins. These results are represented in Table 9 and Figure 1.

Five of the 12 genes were significantly under-expressed in patients with bipolar disorder compared to controls (ATIC, DNMT1, HSPA9B, MCM3, and FTSJ1). Of the remaining seven genes, five showed marginal significance (0.10 < p < 0.05) for having lower peripheral expression in patients with bipolar disorder as compared to controls (TMEM203, MAP4K1, MAGED1, AP4B1, and EIF2B1). Three of the genes (DNMT1, MCM3, and FTSJ1) also showed significantly lower expression in patients with bipolar disorder as compared to their unaffected co-twins. An additional three genes (HSPA9B, DSTN, EIF2B1). showed marginal significance (0.10 < p < 0.05) for having lower peripheral expression in patients with bipolar disorder compared to their unaffected co-twins. None of the twelve genes showed significantly different expression in unaffected co-twins of patients with bipolar disorder as compared to controls. Results can be found in Table 9 and Figure 2.

**Hypothesis 2c: Confirmation in a second sample of twins with schizophrenia.**

The twelve genes that showed lower peripheral expression values as a result of being from a proband twinship were examined in a second sample of patients and co-twins with schizophrenia collected in Finland. Nine of the twelve genes also showed lower peripheral expression values in patients with schizophrenia compared to controls in this sample: MAP4K1, ATIC, DNMT1, CDC16, DSTN, MCM3, MAGED1, AP4B1, FTSJ1. The remaining three genes (TMEM203, HSPA9B, and EIF2B1) all showed marginal significance for having lower peripheral expression in patients with schizophrenia compared to controls. None of the genes
showed significantly lower peripheral expression values in unaffected co-twins as compared to controls or in patients compared to their unaffected co-twins. These results can be found in Table 10 and Figure 3.

Discussion

This study investigated differences in peripheral gene expression in patients with schizophrenia and bipolar disorder, as well as their unaffected co-twins, compared with control twins who had no prior history of significant psychiatric illness. To our knowledge, this is the first study of gene expression in schizophrenia using the discordant co-twin design. It is also one of the first studies to examine the relationship between peripheral expression patterns and cognition in a sample of people with and without psychiatric illness.

We found that even when applying a strict Bonferroni correction for multiple comparisons, 28 genes showed detectable expression patterns that varied in relation to cognitive performance, after controlling for gender, age, and twinship, with 19 more genes meeting a slightly less conservative threshold. While there are likely to be hundreds or even thousands of genes that are involved in cognition, this study showed that a fraction of those can be reliably detected in expression patterns measured in peripheral lymphocytes, in a fairly small sample of individuals. This lends further support to the idea that peripheral expression patterns contain information that is relevant to the functioning of the central nervous system. As 12 of these genes also showed disrupted expression patterns in patients with schizophrenia and their co-twins, these findings also support the idea that the use of endophenotypes, such as cognitive performance, may greatly enhance our ability to reliably detect meaningful genetic variation that is related to heritable risk for psychiatric illness.
While we cannot rule out the possibility that a number of our genetic findings may be signals of secondary disease processes, rather than true markers of heritable risk of psychotic illness, there are several reasons to believe that these genes are worthy of future study. First, as we screened our genome-wide results for only those genes whose expression was related to cognitive performance, our final panel is likely to include genes that are expressed in the central nervous system and who expression in the CNS is highly correlated with peripheral expression patterns. Indeed, examination of the relevant literature shows that all of the genes that were associated with schizophrenia and bipolar disorder in our sample are expressed in the CNS. Second, since we were able to show confirmation of disruption in the expression of many of these genes in both patients with schizophrenia and bipolar disorder, there is also enhanced support for their involvement in the disease process as it has already been shown that there is a large degree of overlap in genetic risk for those two psychiatric illnesses. Third, as we were also able to show disruption in expression of these genes in unaffected co-twins, at least in the case of schizophrenia, there is even greater reason to believe that these genes might be related to heritable risk for psychotic illness, rather than exclusively to secondary factors associated with these disorders or their treatment. Finally, as the majority of the genes also showed disrupted expression in a second sample of patients with schizophrenia, it is highly unlikely that these genes represent false-positives. In fact, given all of this information and our modest sample size, there is reason to believe that our findings represent only the “tip of the iceberg” of the true differences in peripheral expression that are related to psychotic illness, rather than their being random findings related to our specific sample characteristics that are not meaningful for psychotic illness as a whole.
Several of the genes implicated in our study are involved in broad biological processes that are known to be disrupted in schizophrenia and other psychiatric disorders. MAP4K1 codes for a protein that is involved in the mitogen-activated protein kinase (MAPK) signaling pathway. As this pathway regulates gene transcription and cell cycle progression, it is critical for cellular development and organization. Many genes that have also been implicated in psychotic illness, such as BDNF, TNF, and IL1 are also involved in this pathway. MAP4K1 is also involved in the signaling pathways associated with the hepatocyte growth factor receptor, and so between these two pathways it has a large effect on cellular development and organization. Mitogen-activated protein kinsases have been found to be elevated in the cerebellar vermis in patients with schizophrenia (Kyosseva et al., 1999). A recent study also showed that maternal inflammation during fetal development may disrupt MAPK signaling pathways (Deng et al., 2011). This is potentially relevant to our findings of disrupted MAP4K1 transcription in patients with psychotic illness, since exposure to maternal inflammation during fetal development is also one of the most prominently implicated environmental risk factors for schizophrenia (Buka, Tsuang, & Lipsitt, 1993; Zorrilla et al., 1997; Dalman et al., 2001; Murray et al., 2004; Van Erp et al., 2002; Zornberg, Buka, & Tsuang, 2000; Clarke, Harley, & Cannon, 2006).

CDC16 codes for a cell cycle-regulated ubiquitin ligase that is part of the anaphase promoting complex (APC). This protein controls the progression through mitosis and the G1 phase of the cell cycle. This protein likely interacts with MAP4K1 as the MAPK pathway is also a primary component of this regulatory process. There have not been any studies directly linking expression of this protein to any kind of neurological or psychiatric disorder.
DNMT1 codes for a DNA-methyltransferase. This protein is responsible for maintaining methylation patterns during development and is also involved in DNA transcription, through its interactions with the methylation of histones. It is known to be expressed throughout the body, including in the central nervous system. Recent research has also shown that DNMT1 is likely to play an important role in regulating neuronal functioning, and may be highly involved in the cellular processes of synaptic plasticity, learning and memory (Feng & Fan, 2009). Thus, there is a reason to believe that disruptions in the transcription of this gene could cause broad changes in neurological functioning, similar to the ones observed in psychiatric conditions such as schizophrenia and bipolar disorder. Indeed, prior studies have shown disruptions in DNMT1 expression in the brain and in peripheral lymphocytes of patients with schizophrenia, as compared to healthy controls (Veldic et al., 2004, Zhubi et al., 2009). However, it should be noted that in this instance DNMT1 was found to be over-expressed in patients, whereas we observed under-expression in our population.

DSTN codes for an actin-depolymerizing protein. This protein breaks actin filaments and binds to actin monomers. The depolymerizing of actin filaments is a crucial part of cytoskeleton and organelle organization. Recent research has shown that levels of actin-depolymerizing proteins may play a crucial role in the process of synaptic plasticity (Gu et al., 2010). This provides support for a potential role for this gene in cognition and psychotic illness.

HSPA9 codes for a heat shock protein, mortalin, that is involved in RNA-degradation. As such, this protein regulates cellular aging and cell proliferation. Prior research has shown that mortalin expression may be disrupted in the brains of patients with schizophrenia and may also be effected by dopamine activity (Gabriele et al., 2010). Mortalin expression has also been implicated in the genetics of Parkinson’s Disease (Shi et al., 2008), furthering pointing towards
an important interplay between expression of this heat shock protein and dopamine activity in the brain. Additional research has shown that levels of mortalin may alter the body’s response to hypoxia (Luo et al., 2010). This may also be a mechanism by which expression of this gene is important in the pathophysiology of schizophrenia and other psychotic illness, as fetal hypoxia is a predominant environmental risk factor for schizophrenia, but only in families with a known genetic risk for the disorder (Cannon et al., 2000b).

MAGED1 codes for a ring finger protein that is involved in apoptosis. MAGED1 is specifically involved in neural apoptotic response after the binding of nerve growth factor (NGF) (Salehi et al., 2000). MAGED1 may also regulate DLX family member proteins and p53 transcriptional activity. While this protein has not been implicated in any neurological or psychiatric conditions, numerous studies have shown that expression of MAGED1 influences neuronal development and growth (Reddy et al., 2010, Lim et al., 2006, Kendall et al., 2005, Bragason et al., 2005). Additionally, recent research has shown that this gene may be involved in regulating circadian rhythms (Wang et al., 2010), providing some support for its role in bipolar disorder.

Three of the genes identified in our study are involved in known neurological conditions. ATIC codes for an enzyme AICAR transformylase/IMP cyclohydrolase that is involved in purine metabolism. Disruptions in this enzyme have been known to cause a neurological condition called AICA-ribosuria due to ATIC deficiency (Marie et al., 2004). This condition is caused by a build-up of AICA-riboside in the erythrocytes and fibroblasts that cannot be metabolized (Marie et al., 2004). Clinical features of this disorder include profound mental retardation, epilepsy, and congenital blindness (Marie et al., 2004). FTSJ codes for a protein believed to be involved in transcription regulation. Disruptions in the expression of this gene have been known
to cause x-linked mental retardation (Gong et al., 2008, Takano et al. 2008, Freude et al., 2004). X—linked mental retardation type 44 is characterized by a decline in general intelligence and specific impairments in adaptive behavior (Freude et al., 2004). This type of cognitive disruption is also seen in individuals with fragile X disorder. EIF2B1 codes for translation initiation factor EIF2B. Disruptions in this protein are known to cause a neurological condition that results in leukodystrophy with vanishing white matter (Maletkovic et al., 2008). This results in progressive diffuse cerebral hemispheric leukencephalopathy, cerebellar ataxia, spasticity, and inconstant optic atrophy. The syndrome can also present with seizures and memory difficulties. Therefore, while ATIC, FTSJ, and EIF2B1 have not been implicated in genetic risk for schizophrenia or other psychiatric disorders, it is clear that they are expressed in the human brain and that disruptions in these pathways can cause profound neurological and possibly psychological deficits.

Three other genes identified in our study do not seem to have as clear a relationship to psychiatric or neurological functioning. TMEM203 is a gene on chromosome 9 that codes for a transmembrane protein integral to membrane functioning. This protein is expressed in the brain, as well as in peripheral tissues. The exact functioning of this protein has not been elucidated and it has not been linked to any known diseases or disorders. However, this gene is located in the q34.3 region of chromosome 9, which is the same part of the genome containing GRIN1, PTGDS, and NOTCH1, all of which have been implicated in the genetic risk for schizophrenia. Therefore, it is possible that disruption in the transcription of this protein is not playing a primary role in the biology of psychotic disorders but is a marker of change in the structure of chromosome 9 for patients with psychotic illness, possibly in the histones, that also affects the transcription of nearby proteins, which are more relevant to the development of a psychiatric
condition. MCM3 codes for an acetylase that inhibits the initiation of DNA replication. This allows DNA to undergo only one replication per cell cycle and thus is involved in cellular division. MCM3 is expression is found in a variety of tissues throughout the body, including in the central nervous system. AP4B1 codes for a subunit of either a clatrin or non-clathrin associated protein coat that is involved in targeting proteins from the trans-Golgi network to the endosomal-lysosomal system. This protein is found in the soma and dendritic shafts of neuronal cells. Neither gene has not been implicated in any known neurological or psychiatric conditions.

All of these findings will need to be verified at a later date using RT-PCR in order to confirm that transcription was disrupted in patients and co-twins. Further, it should be noted that while the analyses of gene expression in relation to cognitive functioning adjusted for the correlated nature of data from twin pairs, the analyses of gene expression in relation to group status did not. This correction was not applied in the latter analyses to preserve power.

Additionally, as this study used peripheral expression and the relationship between decreased expression in white blood cells and expression in the central nervous system is not completely transparent, future research must be conducted to examine whether expression of these proteins is altered in the brain and if so, in what regions. However, the fact that the genes were identified initially due to their relationship with cognition implies that disruptions in transcription likely are apparent in the central nervous system and effect neuronal functioning.

Overall, 9 of the 12 genes identified in our study appear to have a clear and plausible relationship to the pathophysiology of psychotic illness and/or neurological functioning. Given the fact that transcription of these genes was found to be disrupted in patients and co-twins with schizophrenia in our first sample and the findings were replicated in a second sample, there appears to be support for the continued investigation of these genes in heritable risk for
psychosis. Future studies should investigate the mechanism by which these genes may be involved in psychotic illness and provide a more rigorous examination of potential confounding factors such as co-morbid medical illness and medications, which may be effecting transcription levels in PBLs. However, based on this study there is some reason to believe that expression of various genes in PBLs may be able to be used as a plausible, and relatively inexpensive, biomarker for psychosis in the future.
Table 1. Break-down of recruited Swedish twin-pairs by proband diagnosis.

<table>
<thead>
<tr>
<th>Sample # of pairs</th>
<th>Concordant Schizophrenia</th>
<th>Discordant Schizophrenia</th>
<th>Concordant Bipolar Disorder</th>
<th>Discordant Bipolar Disorder</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MZ</td>
<td>DZ</td>
<td>MZ</td>
<td>DZ</td>
<td>MZ</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>14</td>
<td>6</td>
<td>9</td>
<td>12</td>
</tr>
</tbody>
</table>
Table 2. Demographic characteristics of Swedish sample.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Men</th>
<th>Women</th>
<th>% Men</th>
<th>Average Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schizophrenia</td>
<td>19</td>
<td>16</td>
<td>54.3%</td>
<td>49.78</td>
</tr>
<tr>
<td>SZ cotwins</td>
<td>17</td>
<td>17</td>
<td>50%</td>
<td>50.81</td>
</tr>
<tr>
<td>Bipolar Disorder</td>
<td>7</td>
<td>17</td>
<td>29.2%</td>
<td>49.06</td>
</tr>
<tr>
<td>BP cotwins</td>
<td>5</td>
<td>12</td>
<td>29.4%</td>
<td>52.02</td>
</tr>
<tr>
<td>Controls</td>
<td>29</td>
<td>31</td>
<td>48.3%</td>
<td>49.25</td>
</tr>
</tbody>
</table>
Table 3. Demographic characteristics of Finnish sample.

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
<th>% Men</th>
<th>Average Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schizophrenia</td>
<td>5</td>
<td>13</td>
<td>27.8%</td>
<td>47.33</td>
</tr>
<tr>
<td>SZ co-twins</td>
<td>5</td>
<td>13</td>
<td>27.8%</td>
<td>47.32</td>
</tr>
<tr>
<td>Controls</td>
<td>12</td>
<td>25</td>
<td>32.4%</td>
<td>49.37</td>
</tr>
</tbody>
</table>
Table 4. Results of factor analysis of neuropsychological data.

<table>
<thead>
<tr>
<th>Component</th>
<th>Initial Sums of Squares Loading</th>
<th>Rotated Sums of Squares Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eigen Value</td>
<td>% Variance</td>
</tr>
<tr>
<td>1</td>
<td>9.113</td>
<td>39.62</td>
</tr>
<tr>
<td>2</td>
<td>2.649</td>
<td>11.52</td>
</tr>
<tr>
<td>3</td>
<td>1.975</td>
<td>8.588</td>
</tr>
<tr>
<td>4</td>
<td>1.507</td>
<td>6.550</td>
</tr>
<tr>
<td>5</td>
<td>1.268</td>
<td>5.514</td>
</tr>
</tbody>
</table>
Table 5. Loadings of specific neurocognitive domains on factor 1.

<table>
<thead>
<tr>
<th>Neurocognitive Domain</th>
<th>Test</th>
<th>Average Score/Standard Deviation</th>
<th>Factor 1 initial factor loading</th>
<th>Factor 1 rotated factor loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal Knowledge</td>
<td>WAIS Vocabulary</td>
<td>25.07 (4.93)</td>
<td>0.606</td>
<td>0.206</td>
</tr>
<tr>
<td>Perceptual Organization</td>
<td>WAIS Block Design</td>
<td>39/13 (11.7)</td>
<td>0.606</td>
<td>0.077</td>
</tr>
<tr>
<td>Immediate Complex Visuospatial Memory</td>
<td>WMS Visual Reproduction I</td>
<td>81.55 (14.4)</td>
<td>0.661</td>
<td>0.232</td>
</tr>
<tr>
<td>Delayed Complex Visuospatial Memory</td>
<td>WMS Visual Reproduction II</td>
<td>65.58 (22.1)</td>
<td>0.679</td>
<td>0.303</td>
</tr>
<tr>
<td>Fine Motor Dexterity</td>
<td>Purdue Pegboard Dominant</td>
<td>78.01 (21.5)</td>
<td>-0.480</td>
<td>-0.209</td>
</tr>
<tr>
<td></td>
<td>Purdue Pegboard Nondominant</td>
<td>87.75 (27.4)</td>
<td>-0.502</td>
<td>-0.174</td>
</tr>
<tr>
<td>Verbal learning</td>
<td>CVLT Trials 1-5 Total words</td>
<td>50.73 (12.08)</td>
<td>0.770</td>
<td>0.868</td>
</tr>
<tr>
<td>Verbal recall</td>
<td>CVLT Short delay free recall</td>
<td>10.52 (3.21)</td>
<td>0.779</td>
<td>0.885</td>
</tr>
<tr>
<td></td>
<td>CVLT Short delay cued recall</td>
<td>11.11 (2.87)</td>
<td>0.718</td>
<td>0.921</td>
</tr>
<tr>
<td></td>
<td>CVLT Long delay free recall</td>
<td>11.05 (3.36)</td>
<td>0.757</td>
<td>0.925</td>
</tr>
<tr>
<td></td>
<td>CVLT Long delay cued recall</td>
<td>11.15 (3.08)</td>
<td>0.754</td>
<td>0.934</td>
</tr>
<tr>
<td>Auditory attention and working memory</td>
<td>WAIS Digit Span Forward</td>
<td>8.71 (1.79)</td>
<td>0.512</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>WAIS Digit Span Backwards</td>
<td>5.83 (2.09)</td>
<td>0.587</td>
<td>0.173</td>
</tr>
<tr>
<td>Visuomotor Sequencing</td>
<td>Trail Making Test A</td>
<td>34.65 (12.2)</td>
<td>-0.549</td>
<td>-0.194</td>
</tr>
<tr>
<td>Visuomotor sequencing and set-shifting</td>
<td>Trail Making Test B</td>
<td>84.19 (31.5)</td>
<td>-0.655</td>
<td>-0.273</td>
</tr>
<tr>
<td>Visuospatial attention and working memory</td>
<td>WMS Spatial Span Forward</td>
<td>8.13 (1.93)</td>
<td>0.535</td>
<td>0.164</td>
</tr>
<tr>
<td></td>
<td>WMS Spatial Span Backwards</td>
<td>7.56 (1.93)</td>
<td>0.582</td>
<td>0.076</td>
</tr>
<tr>
<td>Phonemic Fluency</td>
<td>FAS</td>
<td>13.04 (4.16)</td>
<td>0.627</td>
<td>0.343</td>
</tr>
<tr>
<td>Semantic Fluency</td>
<td>Animal Naming</td>
<td>14.71 (3.92)</td>
<td>0.638</td>
<td>0.486</td>
</tr>
<tr>
<td>Verbal working memory</td>
<td>Working memory R</td>
<td>37.11 (23.4)</td>
<td>0.662</td>
<td>0.189</td>
</tr>
<tr>
<td>Verbal working memory</td>
<td>Working memory # words</td>
<td>30.59 (20.3)</td>
<td>0.680</td>
<td>0.174</td>
</tr>
<tr>
<td>Verbal working memory</td>
<td>Working memory correct order</td>
<td>28.97 (19.1)</td>
<td>0.685</td>
<td>0.177</td>
</tr>
<tr>
<td>Attention</td>
<td>Dual task</td>
<td>0.3714 (0.121)</td>
<td>-0.085</td>
<td>0.105</td>
</tr>
</tbody>
</table>
Table 6. Genes whose expression varies as a result of factor 1 scores (corrected p < 0.05).

<table>
<thead>
<tr>
<th>Illumina Probe</th>
<th>corr. p-value</th>
<th>Gene</th>
<th>Chromosome</th>
<th>Description/Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILMN_1814122</td>
<td>8.86E-04</td>
<td>MDC1</td>
<td>6</td>
<td>mediator of DNA damage</td>
</tr>
<tr>
<td>ILMN_1671257</td>
<td>2.90E-03</td>
<td>DKC1</td>
<td>X</td>
<td>dyskerin</td>
</tr>
<tr>
<td>ILMN_1666364</td>
<td>4.20E-03</td>
<td>COQ10A</td>
<td>12</td>
<td>coenzyme Q10 homolog A</td>
</tr>
<tr>
<td>ILMN_1727389</td>
<td>4.64E-03</td>
<td>CDC16</td>
<td>13</td>
<td>cell division cycle 16 homolog</td>
</tr>
<tr>
<td>ILMN_1800787</td>
<td>5.19E-03</td>
<td>RAFTLIN</td>
<td>3</td>
<td>lipid raft linker 1</td>
</tr>
<tr>
<td>ILMN_1673991</td>
<td>7.64E-03</td>
<td>ATIC</td>
<td>2</td>
<td>5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase</td>
</tr>
<tr>
<td>ILMN_1738656</td>
<td>7.90E-03</td>
<td>C17orf25</td>
<td>17</td>
<td>glyoxalase domain containing 4</td>
</tr>
<tr>
<td>ILMN_2224143</td>
<td>8.22E-03</td>
<td>MCM3</td>
<td>6</td>
<td>minichromosome maintenance deficient 3</td>
</tr>
<tr>
<td>ILMN_2339796</td>
<td>0.0145</td>
<td>CDC16</td>
<td>13</td>
<td>cell division cycle 16 homolog</td>
</tr>
<tr>
<td>ILMN_1775522</td>
<td>0.0156</td>
<td>MAGED1</td>
<td>X</td>
<td>melanoma antigen family D</td>
</tr>
<tr>
<td>ILMN_1662334</td>
<td>0.0173</td>
<td>DNAJA3</td>
<td>16</td>
<td>Dna J homolog, subfamily A, member 3</td>
</tr>
<tr>
<td>ILMN_1684628</td>
<td>0.0179</td>
<td>ZFP90</td>
<td>16</td>
<td>zinc finger protein 90 homolog</td>
</tr>
<tr>
<td>ILMN_1686871</td>
<td>0.0183</td>
<td>PARP1</td>
<td>1</td>
<td>poly (ADRP-ribose) polymerase family</td>
</tr>
<tr>
<td>ILMN_1812559</td>
<td>0.0190</td>
<td>SLC7A6</td>
<td>16</td>
<td>solute carrier family 7</td>
</tr>
<tr>
<td>ILMN_1751072</td>
<td>0.0250</td>
<td>SRPRB</td>
<td>3</td>
<td>signal recognition particle receptor</td>
</tr>
<tr>
<td>ILMN_1814859</td>
<td>0.0275</td>
<td>DDX47</td>
<td>12</td>
<td>DEAD box polypeptide 47</td>
</tr>
<tr>
<td>ILMN_2073012</td>
<td>0.0302</td>
<td>TMEM203</td>
<td>9</td>
<td>transmembrane protein 203</td>
</tr>
<tr>
<td>ILMN_2401822</td>
<td>0.0302</td>
<td>FTSJ1</td>
<td>X</td>
<td>FtsJ homolog 1</td>
</tr>
<tr>
<td>ILMN_1757317</td>
<td>0.0345</td>
<td>LARS</td>
<td>5</td>
<td>leucyl-tRNA synthetase</td>
</tr>
<tr>
<td>ILMN_1733696</td>
<td>0.0385</td>
<td>IMP3</td>
<td>15</td>
<td>u3 small nucleolar ribonucleoprotein</td>
</tr>
<tr>
<td>ILMN_1681590</td>
<td>0.0389</td>
<td>LARP1</td>
<td>5</td>
<td>La ribonucleoprotein domain family</td>
</tr>
<tr>
<td>ILMN_1705247</td>
<td>0.0395</td>
<td>ACSL5</td>
<td>10</td>
<td>acyl-CoA synthetase long-chain family member 5</td>
</tr>
<tr>
<td>ILMN_2188722</td>
<td>0.0415</td>
<td>GLS</td>
<td>2</td>
<td>glutaminase</td>
</tr>
<tr>
<td>ILMN_1693664</td>
<td>0.0420</td>
<td>POMGNT1</td>
<td>1</td>
<td>protein O-linked mannosyltransferase 2-N-acetylglucosaminyltransferase</td>
</tr>
<tr>
<td>ILMN_1679209</td>
<td>0.0431</td>
<td>HSPA9B</td>
<td>5</td>
<td>heat shock protein 70kDa 9B</td>
</tr>
<tr>
<td>ILMN_1751400</td>
<td>0.0465</td>
<td>SCAP1</td>
<td>17</td>
<td>src family associated phosphoprotein 1</td>
</tr>
<tr>
<td>ILMN_1721978</td>
<td>0.0472</td>
<td>CARD11</td>
<td>7</td>
<td>caspase recruitment domain family member 11</td>
</tr>
<tr>
<td>ILMN_1748018</td>
<td>0.0481</td>
<td>GORASP2</td>
<td>2</td>
<td>golgi reassembly stacking protein 2</td>
</tr>
</tbody>
</table>
Table 7. Genes whose expression varies as a result of factor 1 scores (0.05 < corrected p < 0.1)

<table>
<thead>
<tr>
<th>Illumina Probe</th>
<th>corrected p-value</th>
<th>Gene</th>
<th>Chromosome</th>
<th>Description/Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILMN_1720270</td>
<td>0.0509</td>
<td>CDR2</td>
<td>16</td>
<td>cerebellar degeneration-related protein 2</td>
</tr>
<tr>
<td>ILMN_1799367</td>
<td>0.0510</td>
<td>TXNDC14</td>
<td>11</td>
<td>thioredoxin domain containing 14</td>
</tr>
<tr>
<td>ILMN_1709132</td>
<td>0.0515</td>
<td>SSTATIP1</td>
<td>18</td>
<td>elongation protein 2 homolog</td>
</tr>
<tr>
<td>ILMN_1679476</td>
<td>0.0622</td>
<td>GART</td>
<td>21</td>
<td>phosphoribosylglycinamide formyltransferase,</td>
</tr>
<tr>
<td>ILMN_1665943</td>
<td>0.0629</td>
<td>MAP4K1</td>
<td>19</td>
<td>mitogen-activated protein kinase 1 (MAP4K1)</td>
</tr>
<tr>
<td>ILMN_1663954</td>
<td>0.0652</td>
<td>TH1L</td>
<td>20</td>
<td>TH1-like transcript variant</td>
</tr>
<tr>
<td>ILMN_1725707</td>
<td>0.0703</td>
<td>ATG16L1</td>
<td>2</td>
<td>autophagy related 16-like 1</td>
</tr>
<tr>
<td>ILMN_1778032</td>
<td>0.0746</td>
<td>SURF6</td>
<td>9</td>
<td>surfeit 6</td>
</tr>
<tr>
<td>ILMN_1760201</td>
<td>0.0753</td>
<td>DNMT1</td>
<td>19</td>
<td>DNA (cytosine-5')-methyltransferase 1</td>
</tr>
<tr>
<td>ILMN_2087528</td>
<td>0.0758</td>
<td>CPSF3</td>
<td>2</td>
<td>cleavage and polyadenylation specific factor 3</td>
</tr>
<tr>
<td>ILMN_1788363</td>
<td>0.0804</td>
<td>MLH1</td>
<td>3</td>
<td>mutL homolog 1</td>
</tr>
<tr>
<td>ILMN_1679797</td>
<td>0.0819</td>
<td>ADARB1</td>
<td>21</td>
<td>adenosine deaminase, RNA-specific, B1</td>
</tr>
<tr>
<td>ILMN_2375418</td>
<td>0.0829</td>
<td>DPH2</td>
<td>1</td>
<td>DPH2 homolog, transcript variant 1</td>
</tr>
<tr>
<td>ILMN_1682738</td>
<td>0.0881</td>
<td>SMAD3</td>
<td>15</td>
<td>SMAD family member 3</td>
</tr>
<tr>
<td>ILMN_1753716</td>
<td>0.0882</td>
<td>EIF2B1</td>
<td>12</td>
<td>eukaryotic translation initiation factor 2B, subunit 1 alpha</td>
</tr>
<tr>
<td>ILMN_1706426</td>
<td>0.0910</td>
<td>DSTN</td>
<td>20</td>
<td>destrin transcript variant 1</td>
</tr>
<tr>
<td>ILMN_1669377</td>
<td>0.0944</td>
<td>AP4B1</td>
<td>1</td>
<td>adaptor-related protein complex 4, beta 1 subunit</td>
</tr>
<tr>
<td>ILMN_2258774</td>
<td>0.0946</td>
<td>MRPL43</td>
<td>10</td>
<td>mitochondrial ribosomal protein L43</td>
</tr>
<tr>
<td>ILMN_2138435</td>
<td>0.0967</td>
<td>MRPS27</td>
<td>5</td>
<td>mitochondrial ribosomal protein S27</td>
</tr>
</tbody>
</table>
Table 8. Select genes whose expression levels vary as a result of proband membership.

<table>
<thead>
<tr>
<th>Gene</th>
<th>family history p-value corrected</th>
<th>SZ cases v. controls p-value</th>
<th>SZ cases v. co-twins p-value</th>
<th>BP cases v. controls p-value</th>
<th>BP cases v. co-twins p-value</th>
<th>BP co-twins v. controls p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMEM203</td>
<td>0.0388</td>
<td>2.75 E-03</td>
<td>0.361</td>
<td>0.0132</td>
<td>0.0563</td>
<td>0.169</td>
</tr>
<tr>
<td>MAP4K1</td>
<td>0.0398</td>
<td>3.49 E-03</td>
<td>0.215</td>
<td>0.129</td>
<td>0.0641</td>
<td>0.282</td>
</tr>
<tr>
<td>ATIC</td>
<td>0.0398</td>
<td>0.0119</td>
<td>0.279</td>
<td>0.113</td>
<td>0.0289</td>
<td>0.182</td>
</tr>
<tr>
<td>DNMT1</td>
<td>0.0398</td>
<td>1.71 E-03</td>
<td>0.672</td>
<td>0.0305</td>
<td>0.0157</td>
<td>0.0448</td>
</tr>
<tr>
<td>HSPA9B</td>
<td>0.0398</td>
<td>0.0182</td>
<td>0.586</td>
<td>0.0549</td>
<td>0.0206</td>
<td>0.0599</td>
</tr>
<tr>
<td>CDC16</td>
<td>0.0398</td>
<td>0.0202</td>
<td>0.574</td>
<td>0.0195</td>
<td>0.144</td>
<td>0.304</td>
</tr>
<tr>
<td>DSTN</td>
<td>0.0398</td>
<td>3.77 E-04</td>
<td>0.377</td>
<td>0.0285</td>
<td>0.133</td>
<td>0.0513</td>
</tr>
<tr>
<td>MCM3</td>
<td>0.0398</td>
<td>3.15 E-03</td>
<td>0.391</td>
<td>0.0215</td>
<td>0.0435</td>
<td>0.0336</td>
</tr>
<tr>
<td>MAGED1</td>
<td>0.0398</td>
<td>5.61 E-03</td>
<td>0.581</td>
<td>0.0195</td>
<td>0.0889</td>
<td>0.182</td>
</tr>
<tr>
<td>AP4B1</td>
<td>0.0398</td>
<td>0.070</td>
<td>0.836</td>
<td>0.0409</td>
<td>0.0904</td>
<td>0.264</td>
</tr>
<tr>
<td>FTSJ1</td>
<td>0.0398</td>
<td>2.87 E-03</td>
<td>0.334</td>
<td>0.0549</td>
<td>0.0317</td>
<td>0.0375</td>
</tr>
<tr>
<td>EIF2B1</td>
<td>0.0398</td>
<td>0.0146</td>
<td>0.834</td>
<td>0.0172</td>
<td>0.0586</td>
<td>0.0601</td>
</tr>
</tbody>
</table>
Figure 1. Peripheral expression of select genes in patients with schizophrenia and unaffected co-twins. Gene expression plotted as percent-change relative to the expression level of members in the control group.
Figure 2. Peripheral expression of select genes in patients with bipolar disorder and unaffected co-twins. Gene expression plotted as percent-change relative to the expression level of members in the control group.
Table 9. Confirmation of group-level differences in gene expression in a Finnish sample.

<table>
<thead>
<tr>
<th>Gene</th>
<th>SZ cases v. controls p -value</th>
<th>SZ cases v. co-twins p-value</th>
<th>SZ co-twins v. controls p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGC14327</td>
<td>0.0513</td>
<td>0.168</td>
<td>0.888</td>
</tr>
<tr>
<td>MAP4K1</td>
<td>0.0384</td>
<td>0.176</td>
<td>0.874</td>
</tr>
<tr>
<td>ATIC</td>
<td>0.0122</td>
<td>0.0629</td>
<td>0.944</td>
</tr>
<tr>
<td>DNMT1</td>
<td>0.0397</td>
<td>0.101</td>
<td>0.957</td>
</tr>
<tr>
<td>HSPA9B</td>
<td>0.0564</td>
<td>0.149</td>
<td>0.993</td>
</tr>
<tr>
<td>CDC16</td>
<td>0.0406</td>
<td>0.172</td>
<td>0.768</td>
</tr>
<tr>
<td>DSTN</td>
<td>0.0313</td>
<td>0.0893</td>
<td>0.988</td>
</tr>
<tr>
<td>MCM3</td>
<td>0.0278</td>
<td>0.112</td>
<td>0.865</td>
</tr>
<tr>
<td>MAGED1</td>
<td>0.0119</td>
<td>0.0852</td>
<td>0.886</td>
</tr>
<tr>
<td>AP4B1</td>
<td>0.0488</td>
<td>0.190</td>
<td>0.834</td>
</tr>
<tr>
<td>FTSJ1</td>
<td>0.0294</td>
<td>0.133</td>
<td>0.905</td>
</tr>
<tr>
<td>EIF2B1</td>
<td>0.0735</td>
<td>0.160</td>
<td>0.994</td>
</tr>
</tbody>
</table>
Figure 3. Peripheral expression of select genes in a secondary sample of patients with schizophrenia and unaffected co-twins. Gene expression plotted as percent-change relative to the expression level of members in the control group.
References


also confers risk of recurrent depression and of schizophrenia. *Molecular Psychiatry, 15*: 1016-1022.


Hennah, W., Varilo, T., Kestila, M. Paunio, T., Arajarvi, R., Haukka, J., et al. (2003). Haplotype transmission analysis provides evidence of association for DISC1 to schizophrenia and suggests sex-dependent effects. Human Molecular Genetics, 12, 3151-3159.


lymphoblast gene expression profiles in schizophrenia: Are non-neural cells informative?

*PLoS ONE, 3*(6), e2412. Doi: 10.1371/journal.pone.0002412.


Straub, R. E., Jiang, Y., MacLean, C. J., Ma, Y., Webb, B. T., Myakishev, M. V., et al. (2002). Genetic variation in the 6p22.3 gene DTNBP1, the human ortholog of the mouse
The dysbindin gene, is associated with schizophrenia. *American Journal of Human Genetics, 71*, 337-348.


of neurocognitive functions in families with schizophrenia. *American Journal of Medical Genetics Neuropsychiatric Genetics, 114:* 483-490.


Veldic, M, Caruncho, H J, Liu, W S, et al. (2004). DNA-methyltransferase 1 mRNA is selectively overexpressed in telencephalic GABAergic interneurons of schizophrenia


