Neurobehavioral HIV Medicine

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Exploration of a genomic expression and pathway analysis approach to neurocognitive performance: preliminary findings

This article was published in the following Dove Press journal:
Neurobehavioral HIV Medicine
10 July 2010
Number of times this article has been viewed

Chad A Bousman[1]
Gursharan Chana[2]
Stephen J Glatt[3]
Sharon D Chandler[1]
Todd May[1]
James Lohr[1]
Ian P Everall[2]
William S Kremen[1]
Ming T Tsuang[1]

[1]Psychiatry, University of California, San Diego, CA, USA; [2]Psychiatry, University of Melbourne, Melbourne, Victoria, Australia; [3]Psychiatry and Behavioral Sciences, SUNY Upstate Medical University, Syracuse, NY, USA

Abstract: Identification of genomic biomarkers for neurocognitive performance could revolutionize screening, diagnosis, staging, and/or prognosis practices for HIV-associated neurocognitive disorders (HAND). This study sought to explore the relationship between blood-based gene expression and neurocognitive performance. 8 healthy adults were recruited. Subjects were non-smokers, reported taking no medications, and were free of any psychiatric disorders. Correlations adjusting for education and ancestry were conducted to generate lists of genes significantly correlated with scores on neurocognitive test from the National Institute of Mental Health's Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) consensus cognitive battery (MCCB). Ingenuity pathway analysis was used to identify canonical pathways. For each of the 10 MCCB tests, large effect sizes ($r > 0.49$) were observed. 65% of the genes and 80% of canonical pathways were unique to 1 of the 10 MCCB tests, albeit none survived a 10% FDR correction. Minimal gene overlap was observed across tests; however, several overlapping canonical pathways were observed. Analysis of the relationship between gene expression alterations and neurocognitive performance may provide the potential for improving our ability to identify genomic markers of HAND as well as provide guidance to physicians and researchers in HIV medicine and other disciplines in their pursuit of viable genomic biomarkers for neurocognitive impairment.

Keywords: gene expression, cognitive, methodology, MATRICS

Introduction

The examination of genetic attributes of neurocognition (ie, cognitive neurogenetics) to understand and uncover the underlying biological mechanisms of the pathogenesis as well as clinically viable biomarkers for early identification of HIV-associated neurocognitive disorders (HAND) has become an exciting aspect of NeuroAIDS research.[1] Genomic biomarkers for HAND could assist physicians in their determination of whether further neurological investigation would be valuable and/or assist in tailoring pharmacological interventions. In addition, genomic biomarkers have the potential to improve screening, diagnostic, staging and/or prognostic procedures.[2,3] To date, the majority of genetic work related to neurocognitive performance, HAND, and other neurobehavioral disorders has been candidate-gene driven and concerned primarily with testing genotypic associations for neurocognitive performance among a small pool of candidate genes.[1,4] To our knowledge, no study has attempted to examine gene expression and/or pathway associations with neurocognitive performance. This type of analysis may be advantageous in that it could identify novel candidate genes and/or biological pathways from which additional candidate genes could be extracted

Correspondence: Chad A Bousman
Center for Behavioral Genomics, Department of Psychiatry, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92039, USA
Email cbousman@ucsd.edu.

and further investigated for their significance. In this study, we sought to explore a blood-based genomic expression and pathway analysis approach to neurocognitive performance among a group of healthy adults rather than those with HAND to avoid confounds related to medication or disease status and to provide preliminary data for a novel method that could be used in future biomarker discovery efforts.

Methods

Subjects

Ten healthy volunteers were recruited from the University of California, San Diego Psychopharmacology Research Initiatives Center for Excellence participant network and clinically assessed using the Diagnostic Interview for Genetic Studies (DIGS)[5] as described elsewhere.[6] Participants were free of: (1) substance abuse or dependence in the past year; (2) neurologic problems (eg, stroke); (3) systemic medical illnesses (eg, diabetes, HIV); (4) history of head injury with documented loss of consciousness lasting longer than 10 minutes; (5) pregnancy; (6) physical disabilities; (7) use of prescribed medications; (8) current tobacco use; or (9) a personal or family history of a schizophrenia, bipolar disorder, major depressive disorder, or a cluster-A (schizotypal, schizoid, or paranoid) personality disorder, all of which may affect gene expression. All study procedures were approved by the Institutional Review Board at University of California, San Diego.

Neurocognitive assessment

Participants were administered the National Institute of Mental Health's Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) consensus cognitive battery (MCCB).[7,8] The MCCB contains 10 tests representing 7 domains, takes approximately 60 minutes to administer (for more details on the battery see references 7 and 8), and contains neurocognitive tests within the domains required for diagnosing HAND[9] (Table 1). The MCCB was administered by a trained Master's-level research assistant, on a single occasion, and in the order listed in Table 1.

Blood collection and processing

Whole blood (10 mL) was collected in the morning after subjects fasted overnight but directly prior (< 30 minutes) to administration of the neurocognitive battery. Blood samples were immediately transferred to an RNase-free laboratory, where all subsequent procedures (ie, stabilization, isolation, storage) took place.[6] Prior to hybridization, two subjects were excluded for low RNA integrity (RIN < 6.0).[10] The remaining eight samples were then transcribed to cDNA and hybridized to GeneChip® Human Exon 1.0 ST Arrays (Affymetrix, Inc., Santa Clara, CA, USA) per the "Whole Transcript (WT) Sense Target Labeling Assay" protocol (Affymetrix, 2006) using 1 μg of total RNA from each sample.

Microarray data analyses

The principal analyses of these data were designed to explore gene expression and pathway correlates among the 10 neurocognitive tests in the battery described above, as well as to potentially demonstrate the utility of a more robust biomarker discovery method for neurocognitive performance. To accomplish this, microarray data were imported into Partek Genomics Suite software (Partek Inc., St Louis, MO, USA)

Table 1 National Institute of Mental Health's Measurement and Treatment Research to Improve Cognition In Schizophrenia (MATRICS) consensus cognitive battery

Test	Domain	Test score used
Trail making test, Part A	Speed of information processing	Time to completion
Brief assessment of cognition in SCZ (BACS), symbol coding	Speed of information processing	Total number correct
Hopkins verbal learning test (HVLT)-revised, immediate recall subtest	Verbal learning	Total number words (all three trials)
Wechsler memory scale (WMS), 3rd ed., spatial span task	Working memory	Sum of raw scores on forward and backward conditions
Letter-number (LN) span test	Working memory	Number of correct trials
Neuropsychological assessment battery, mazes subtest	Reasoning and problem solving	Total raw score
Brief visuospatial memory test (BVMT)-revised	Visual learning	Total recall over three trials
Category fluency test, animal naming	Speed of information processing	Total number of animals named
Mayer-Salovey-Caruso emotional intelligence test (MSCEIT), managing emotions subtest	Social cognition	Branch score using general consensus scoring
Continuous performance test (CPT)	Attention/vigilance	Mean d' value across 2-, 3-, and 4-digit conditions

Note: Table adapted with permission from Nuechterlein KH, Barch DM, Gold JM, Goldberg TE, Green MF, Heaton RK. Identification of separable cognitive factors in schizophrenia. *Schizophr Res*. 2004;72(1):29–39.[13]
Abbreviation: SCZ, schizophrenia.

using robust multichip average to analyze probe intensities and to determine gene expression correlates of each of the 10 neurocognitive tests (for more detail on data processing, see Glatt et al[6]). Pearson's partial correlations (adjusting for education and ancestry) between raw scores on each of the neurocognitive tests and the mean expression level on a gene-by-gene basis were conducted.

Pathway analysis

Following correlational analysis of individual gene transcripts, gene lists containing all nominally ($P < 0.05$) significantly correlated genes for each of the 10 neurocognitive tests were generated. These lists were then imported into Ingenuity Pathway Analysis (IPA) software (Ingenuity Systems®, Redwood City, CA, USA) to associate correlated genes with their representative canonical pathways. Canonical pathways that were most significant to the selected gene lists were identified by querying the IPA library of canonical pathways. The significance of the association between the datasets and the canonical pathways was measured in two ways: (1) using the Fischer's exact test, we calculated the probability that the association between the gene list and the canonical pathway was explained by chance alone; and (2) we calculated a ratio of the number of genes from the gene list that mapped to the pathway divided by the total number of molecules that exist in the canonical pathway. Pathways with a high ratio and a low *P*-value were interpreted as indicative of potentially good candidates for further exploration.

Results

Participant characteristics

Participant characteristics and raw neurocognitive test scores for each of the eight participants are shown in Table 2. Participants ranged in age from 37–53 years ($M = 45$, $sd = 7$), were predominantly Caucasian (63%) and male (63%), and had completed between 12 and 17 years ($M = 15$, $sd = 2$) of education. Mean raw scores for each neurocognitive test among participants were in the normative range of those reported among healthy subjects recruited to determine norms for the MCCB.[8]

Gene expression correlates of neurocognitive performance

Correlation analyses identified 6496 of a possible 21,866 genes whose expression was significantly correlated (nominal $P < 0.05$) with one or more of the MCCB tests. The number of genes significantly correlated with each test ranged from 65 (Mayer-Salovey-Caruso Emotional

Table 2 Participant characteristics and raw neurocognitive test scores

Subject	Demographics				Reason and prob. solv.	Verbal learning	Working memory		Visual learning	Speed of processing			Attention/ vigilance	Social cognition
	Age	Sex	Ethnic	Edu	Mazes	Hopkins	WMS	LN Span	BVMT	Cat	BACS	Trails A	CPT	MSCEIT
S1	49	M	CEU	16	19	20	14	16	32	20	37	30	3.05	100
S2	44	M	CEU	13	20	30	18	14	25	27	63	17	3.13	103
S3	49	M	CEU	17	26	31	18	13	31	31	44	36	3.68	80
S4	53	F	HISP	14	21	27	12	11	19	24	54	24	2.14	92
S5	52	M	CEU	12	23	25	15	14	24	30	60	34	2.04	83
S6	38	M	AA	15	9	23	14	16	23	26	53	17	2.78	109
S7	37	F	Asian	14	22	30	19	17	35	26	75	30	4.24	109
S8	38	F	CEU	16	17	29	17	17	26	23	76	15	4.15	96
Mean (s.d)	**45 (7)**	**–**	**–**	**15 (2)**	**20 (5)**	**27 (4)**	**16 (3)**	**15 (2)**	**27 (5)**	**26 (4)**	**58 (14)**	**25 (8)**	**3.151 (0.83)**	**96 (11)**

Abbreviations: CEU, Caucasian; HISP, Hispanic; AA, African-American; NAB, Neuropsychological Assessment Battery; HVLT, Hopkins Verbal Learning Test; WMS, Wechsler Memory Scale-spatial span; LN span, letter-number span; BVMT, Brief Visuospatial Memory Test; Cat, category test-animals; BACS, Brief Assessment of Cognition in Schizophrenia; CPT, Continuous Performance Test; MSCEIT, Mayer-Salovey-Caruso Emotional Intelligence Test.

Intelligence Test: MSCEIT) to 1425 (Hopkins Verbal Learning Test: HVLT) (Table 3). 63% (4116/6496) of these genes were correlated with a single neurocognitive test only. None of the genes was significantly correlated with more than three tests. Interestingly, the letter-number span, animal-naming, and MSCEIT tests showed the most unique gene expression correlates in that the unique proportion (ie, number of genes only significant for the test/total number of significant genes for the test) of correlated genes for these tests was 68%, 85%, and 100%, respectively (Table 3). In contrast, the other tests showed less than 50% (range: 22%–47%) uniqueness and thus shared a greater majority of their gene correlates with other tests in the MCCB. It should be noted that due to the small sample size, the large number of comparisons made, and the consequent severity of the penalty for multiple testing, no individual gene correlate remained significant after correction (ie, 10% FDR). However, correlation coefficients (r) ranged from 0.98 to –0.98 and large effect sizes ($r > \pm 0.49$)[11] were observed for a majority of the nominally significant genes for most of the individual tests and domains (Table 3).

Canonical pathway correlates of neurocognitive performance

Examination of the top 10 canonical pathways populated by all the correlated genes for each of the 10 MCCB tests (Supplemental Table 1) showed that 80% of pathways were unique to one of the 10 tests. No pathways overlapped with all MCCB tests. Although, NF-kß signaling, N-Glycan degradation, one carbon pooling by folate, and phenylalanine metabolism canonical pathways were the most represented pathways among the MCCB, with each of these pathways mapping onto 3 or more of the 10 tests. Table 4 provides a comprehensive summary of gene and canonical pathway overlap among the tests of the MCCB.

Discussion

This study sought to explore the relationship between blood-based gene expression and neurocognitive performance among healthy adults to avoid confounds related to medication or disease status and provide preliminary data for a novel method that could be used in future biomarker discovery efforts in HAND as well as other neurocognitive-related disorders. We found that nominally significant ($P < 0.05$) neurocognitive performance-associated genes mapped onto several canonical pathways, particularly those involved in cell signaling or metabolism. Among the top 10 pathways across the 10 tests (k = 100), 49% were signaling and 17% were metabolism pathways. It should be noted that these pathways have extensive interactions with a variety of biological processes such as cell growth, proliferation, and apoptosis,[12] so it is not surprising that these pathways are represented in associations with neurocognitive performance. In fact, several studies of HAND have investigated gene polymorphisms in signaling (eg, chemokine and dopamine receptors) and to lesser extent metabolism pathways (eg, dopamine metabolism).[1]

Table 3 Gene expression correlates, uniqueness, and effects by MCCB test and domain

Test/domain	# of sig. correlated genes	# Unique to test/domain	Unique proportion	% of genes $r > 0.49$
NAB-mazes	720	349	48.47%	52.22%
Hopkins verbal learning	1,425	667	46.81%	62.53%
WMS-spatial span	1,234	476	38.57%	56.89%
LN Span	929	629	67.71%	61.46%
Working memory domain	633	36	5.69%	51.18%
BVMT	1,225	487	39.76%	66.53%
Animals	1,137	963	84.70%	66.31%
BACS	962	97	10.08%	63.83%
Trails A	533	183	34.33%	48.41%
Speed of information processing domain	866	113	13.05%	59.12%
CPT	558	51	9.14%	50.72%
Social cognition	65	65	100.00%	61.54%

Note: r = Pearson's partial correlation coefficient.

Abbreviations: MCCB, National Institute of Mental Health's Measurement and Treatment Research to Improve Cognition in Schizophrenia consensus cognitive battery; NAB, Neuropsychological Assessment Battery; HVLT, Hopkins Verbal Learning Test; WMS, Wechsler Memory Scale; LN span, letter-number span; BVMT, Brief Visuospatial Memory Test; Cat, category test-animals; BACS, Brief Assessment of Cognition in Schizophrenia; CPT, Continuous Performance Test; MSCEIT, Mayer-Salovey-Caruso Emotional Intelligence Test.

Table 4 Summary of overlapping genes (above diagonal) and canonical pathways (below diagonal) among the MCCB tests and domains

	Reason	Vrb. lrn	Working memory			Vis. lrn	Processing speed				Att/vig	Soc.Cog.
	Mazes	Hopkins	WMS	LN span	Wrk mem	BVMT	Category	BACS	Trails A	SIP	CPT	MSCEIT
Mazes		0	0	267	0	0	142	0	0	0	0	0
Hopkins	0		758	0	0	0	0	0	0	0	0	0
WMS	0	3		0	0	0	0	0	0	0	0	0
LN span	2	1	0		0	0	61	0	0	0	0	0
Wrk mem	2	0	0	0		37	0	175	223	329	380	0
BVMT	1	1	1	0	0		0	704	0	438	0	0
Category	2	1	0	1	0	2		0	0	0	0	0
BACS	0	1	1	0	0	6	1		0	560	44	0
Trails A	0	0	0	0	4	0	0	0		17	344	0
SIP	0	0	0	0	1	1	0	1	1		122	0
CPT	2	0	0	1	5	0	0	0	3	1		0
MSCEIT	1	0	0	0	1	0	0	0	0	1	1	

Abbreviations: MCCB, National Institute of Mental Health's Measurement and Treatment Research to Improve Cognition in Schizophrenia consensus cognitive battery; HVLT, Hopkins Verbal Learning Test; SSpan, Wechsler Memory Scale-spatial span; LN span, letter-number span; BVMT, Brief Visuospatial Memory Test; Cat, category test-animals; BACS, Brief Assessment of Cognition in Schizophrenia; CPT, Continuous Performance Test; MSCEIT, Mayer-Salovey-Caruso Emotional Intelligence Test.

From a gene expression and pathway vantage point, our results also suggest that tests within established neurocognitive domains[13] may be more unique than similar. This was observed for two of the domains (ie, working memory and processing speed) which were represented by more than one test. Whereas, examination of between-domain overlap revealed that several domains had multiple gene expression and pathway correlates in common (eg, visual learning [BVMT] and processing speed [BACS] tests). These findings suggest neurocognitive tests, particularly those in the MCCB, may not hold to their established neurocognitive domain structure[13] when examined from a blood-based genomic perspective. This is particularly relevant to future biomarker discovery efforts in HAND in that the current diagnosis of HAND requires assessment of at least five areas of neurocognitive functioning known to be affected by HIV (eg, executive functioning, episodic memory, speed of information processing, motor skills, attention/working memory, language, and sensoriperception) but does not stipulate the use of any particular neurocognitive tests.[9] From a genomic perspective, HAND diagnosed using the current criteria may be very heterogeneous and result in difficulties with biomarker discovery and replication. Thus, if our results are replicated in a larger sample, this ultimately could advocate for restructuring of domains, examination of neurocognitive tests independent of domain membership, or adoption of a specific core group of neurocognitive tests for future genetic research related to biomarker discovery in HAND.

It must be acknowledged that these explanations are both speculative and post hoc. Moreover, there are several areas of overlap in cognitive component processes across tests that do not coincide with gene or pathway overlap. Further exploration is clearly needed, but the present results provide preliminary evidence that this approach might be useful in elucidating genetic biomarkers for cognitive processing and ultimately HAND. In so doing, this approach might also lead to novel ways to organize or select cognitive tests for HAND diagnosis. Indeed, it is worth noting that the cognitive tests used in the present study as well as many similar tests commonly employed were not designed for genetic studies,[14] and current ways of organizing cognitive measures may not map neatly onto gene expression or canonical pathways.

Several caveats should be considered when interpreting these results and addressed in future investigations. First, it remains unclear whether molecular signatures in blood accurately reflect those found in the brain. However, several investigators have reported that the circulating blood may act as a "sentinel tissue",[15] "neural probe",[16] or "surrogate"[17] for underlying pathophysiology in brain disorders. In fact, Liew and colleagues[15] reported an 80% overlap between blood and brain gene expression. Furthermore, blood-based approaches allow for better standardization of technical procedures and the ability to profile human subjects in a relatively non-invasive manner.[18,19] Second, we attempted to limit the influence of different subtypes of cells in blood by focusing on leukocytes. Within this cell category, several cell types (ie, neutrophils, eosinophils, basophils, lymphocytes, monocytes, and macrophages) with varying roles in the blood exist and may influence gene expression

profiles.[20] Future blood-based studies may find it advantageous to isolate specific lymphocyte subtypes in an effort to achieve greater sensitivity in detecting gene expression changes. Third, the sample size was small, which may have prohibited detection of effects that would have retained statistical significance in a larger sample, even after correcting for multiple testing. In fact, large expression effects ($r > \pm 0.49$)[11] were observed for a majority of genes for each of the MCCB tests suggesting many of these genes may have been significant in a larger sample and justify further investigation. Fourth, pathway-level analysis was done using one of the largest knowledge bases of biological networks (IPA, Redwood City, CA, USA), a considerable strength. However, IPA is manually curated and relies on previously published findings on mammalian biology. Thus, in some cases cellular component annotation can be missing or incomplete due to the lack of information in protein databases to which IPA is linked (eg, UniProt) and ultimately may underestimate extracellular entities (eg, metabolites, hormones).[21] Fifth, expression profiles were based on samples collected immediately (< 30 minutes) prior to cognitive testing. Thus, expression and pathway profiles observed in this study are not a result of cognitive performance tests but rather potential markers for performance ability. Future longitudinal research that examines expression profiles before and after neurocognitive testing could further elucidate the biological mechanisms underlying neurocognitive performance. Sixth, this study utilized the MATRICS consensus cognitive battery which was designed to evaluate cognitive-enhancing agents and other interventions to treat core cognitive deficits of schizophrenia and not necessarily for the identification of genomic markers for neurocognitive performance or HAND. Thus, it is possible that use of other neurocognitive tests and/or batteries will yield different results. Finally, in an effort to conserve power, potentially influential covariates (eg, age, gender, diet, and exercise) were not adjusted for in our analysis. Future studies with larger samples that have the power to include more covariates are warranted.

Despite these limitations, our work provides preliminary data for the application of genomic expression and pathway analysis to tests of neurocognitive performance. This approach may provide the potential for improving our ability to identify genomic markers of HAND as well as provide guidance to physicians and researchers in HIV medicine and other disciplines in their pursuit of viable genomic biomarkers for neurocognitive impairment.

Acknowledgments

This work was primarily supported by a National Institutes of Health grant R21MH075027 (MTT) but also supported from grants R01DA012846, R01DA018662, R01MH065562, and R01MH071912 (MTT), R01MH079881, R25MH074508, R25MH081482, and R41MH079728 (IPE), R01AG018386, R01AG022381, and R01AG022982 (WSK), P50MH081755 (Eric Courchesne/SJG), the UCSD Center for AIDS Research Genomics Core, and a NARSAD Young Investigator Award (SJG).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Levine AJ, Singer EJ, Shapshak P. The role of host genetics in the susceptibility for HIV-associated neurocognitive disorders. *AIDS Behav*. 2009;13(1):118–132.
2. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. 2001;69(3):89–95.
3. Quinones MP, Kaddurah-Daouk R. Metabolomics tools for identifying biomarkers for neuropsychiatric diseases. *Neurobiol Dis*. 2009;35(2):165–176.
4. Savitz J, Solms M, Ramesar R. The molecular genetics of cognition: dopamine, COMT and BDNF. *Genes Brain Behav*. 2006;5(4):311–328.
5. Nurnberger JI Jr, Blehar MC, Kaufmann CA, et al. Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. *Arch Gen Psychiatry*. 1994;51(11):849–859; discussion 863–864.
6. Glatt SJ, Chandler SD, Bousman CA, et al. Alternatively spliced genes as biomarkers for schizophrenia, bipolar disorder and psychosis: a blood-based spliceome-profiling exploratory study. *Curr Pharmacogenomics Person Med*. 2009;7(3):164.
7. Nuechterlein KH, Green MF, Kern RS, et al. The MATRICS consensus cognitive battery, part 1: test selection, reliability, and validity. *Am J Psychiatry*. 2008;165(2):203–213.
8. Kern RS, Nuechterlein KH, Green MF, et al. The MATRICS consensus cognitive battery, part 2: co-norming and standardization. *Am J Psychiatry*. 2008;165(2):214–220.
9. Woods SP, Moore DJ, Weber E, Grant I. Cognitive neuropsychology of HIV-associated neurocognitive disorders. *Neuropsychol Rev*. 2009;19(2):152–168.
10. Schroeder A, Mueller O, Stocker S, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol*. 2006 Jan 31;7:3.
11. Cohen J. A power primer. *Psychol Bull*. 1992;112(1):155–159.
12. Whistler T, Taylor R, Craddock RC, Broderick G, Klimas N, Unger ER. Gene expression correlates of unexplained fatigue. *Pharmacogenomics*. 2006;7(3):395–405.
13. Nuechterlein KH, Barch DM, Gold JM, Goldberg TE, Green MF, Heaton RK. Identification of separable cognitive factors in schizophrenia. *Schizophr Res*. 2004;72(1):29–39.
14. Goldberg TE, Weinberger DR. Genes and the parsing of cognitive processes. *Trends Cogn Sci*. 2004;8(7):325–335.
15. Liew CC, Ma J, Tang HC, Zheng R, Dempsey AA. The peripheral blood transcriptome dynamically reflects system wide biology: a potential diagnostic tool. *J Lab Clin Med*. 2006;147(3):126–132.

16. Gladkevich A, Kauffman HF, Korf J. Lymphocytes as a neural probe: potential for studying psychiatric disorders. *Prog Neuropsychopharmacol Biol Psychiatry*. 2004;28(3):559–576.
17. Sullivan PF, Fan C, Perou CM. Evaluating the comparability of gene expression in blood and brain. *Am J Med Genet B Neuropsychiatr Genet*. 2006;141B(3):261–268.
18. Tsuang MT, Nossova N, Yager T, et al. Assessing the validity of blood-based gene expression profiles for the classification of schizophrenia and bipolar disorder: a preliminary report. *Am J Med Genet B Neuropsychiatr Genet*. 2005;133B(1):1–5.
19. Mohr S, Liew CC. The peripheral-blood transcriptome: new insights into disease and risk assessment. *Trends Mol Med*. 2007;13(10):422–432.
20. Chana G, Glatt SJ, Everall IP, Tsuang MT. *Blood and Brain Gene Expression in Major Psychiatric Disorders: A Search for Biomarkers. Biomarkers for Psychiatric Disorders*. New York, NY: Springer; 2008.
21. Pospisil P, Iyer LK, Adelstein SJ, Kassis AI. A combined approach to data mining of textual and structured data to identify cancer-related targets. *BMC Bioinformatics*. 2006;7:354.

Supplementary material

Table S1 Top 10 canonical pathways by neurocognitive domain and test

Domain	Test	Canonical pathway	*Log p*	Ratio
Reasoning and problem solving	Mazes	Protein ubiquitination pathway	3.47E+00	6.44E-02
		Coagulation system	2.09E+00	1.08E-01
		Chemokine signaling	1.53E+00	6.67E-02
		Cell cycle: G2/M DNA damage checkpoint regulation	1.22E+00	6.98E-02
		Purine metabolism	1.19E+00	3.12E-02
		cAMP-mediated signaling	1.16E+00	4.40E-02
		Inositol metabolism	1.12E+00	8.00E-02
		Phenylalanine metabolism	9.12E-01	2.80E-02
		Tight junction signaling	8.48E-01	3.66E-02
		One carbon pool by folate	8.21E-01	5.26E-02
Verbal learning	Hopkins	Clatrin-mediated endocytosis	3.04E+00	1.15E-01
		NF-κB signaling	2.37E+00	1.10E-01
		PPAR signaling	2.10E+00	1.12E-01
		Ceramide signaling	1.96E+00	1.22E-01
		G-protein coupled receptor signaling	1.92E+00	9.41E-02
		Estrogen receptor signaling	1.85E+00	1.02E-01
		Circadian rhythm signaling	1.65E+00	1.56E-01
		Role of PKR in interferon induction and antiviral response	1.63E+00	1.28E-01
		TGF-β signaling	1.61E+00	1.08E-01
		LXR/RXR activation	1.34E+00	9.41E-02
Working memory	WMS	NF-κB signaling	3.79E+00	1.24E-01
		Role of pattern recognition receptors in recognition of bacteria and viruses	2.30E+00	1.18E-01
		Toll-like receptor signaling	2.08E+00	1.30E-01
		Role of PKR in interferon induction and antiviral response	1.85E+00	1.28E-01
		Death receptor signaling	1.61E+00	1.08E-01
		Nucleotide excision repair pathway	1.57E+00	1.43E-01
		Inositol phosphate metabolism	1.53E+00	7.51E-02
		G-protein coupled receptor signaling	1.49E+00	7.92E-02
		Activation of IRF by cytosolic pattern recognition receptors	1.47E+00	9.46E-02
		GABA receptor signaling	1.43E+00	1.09E-01
	LN span	JAK/stat signaling	2.55E+00	1.19E-01
		TGF-β signaling	2.43E+00	9.64E-02
		α-adrenergic signaling	2.22E+00	8.41E-02
		One carbon pool by folate	1.79E+00	1.05E-01
		EGF signaling	1.73E+00	1.06E-01
		Erythropoietin signaling	1.59E+00	8.00E-02
		IL-2 signaling	1.56E+00	9.43E-02
		Neuregulin signaling	1.54E+00	7.53E-02
		SAPK/JNK signaling	1.54E+00	7.53E-02
		Chemokine signaling	1.44E+00	8.00E-02
	Domain	Neurotrophin/TRK signaling	2.16E+00	6.58E-02
		Complement system	1.73E+00	8.33E-02
		Hypoxia signaling in the cardiovascular system	1.65E+00	5.63E-02
		Nicotinate and nicotinamide metabolism	1.53E+00	3.88E-02
		N-glycan degradation	1.39E+00	6.67E-02
		Phenylalanine metabolism	1.23E+00	2.80E-02
		NRF2-mediated oxidative stress response	1.12E+00	3.28E-02
		Airway inflammation in asthma	1.08E+00	1.67E-01
		Histidine metabolism	1.05E+00	2.70E-02
		Purine metabolism	1.03E+00	2.40E-02

(*Continued*)

Table S1 (*Continued*)

Domain	Test	Canonical pathway	*Log p*	Ratio
Visual learning	BVMT	Axonal guidance signaling	3.45E+00	7.42E-02
		Integrin signaling	2.33E+00	8.29E-02
		Sonic hedgehog signaling	2.19E+00	1.61E-01
		Calcium signaling	1.92E+00	6.93E-02
		Regulation of actin-based motility by rho	1.67E+00	8.70E-02
		Synthesis and degradation of ketone bodies	1.53E+00	1.58E-01
		Lysine biosynthesis	1.47E+00	4.69E-02
		IL-10 signaling	1.33E+00	8.45E-02
		Coagulation system	1.29E+00	1.08E-01
		NF-κB signaling	1.24E+00	6.90E-02
Speed of processing	Category	cAMP-mediated signaling	1.84E+00	8.18E-02
		Valine, leucine and isoleucine degradation	1.74E+00	8.41E-02
		Synthesis and degradation of ketone bodies	1.49E+00	1.58E-01
		Phototransduction pathway	1.40E+00	9.52E-02
		TGF-β signaling	1.37E+00	8.43E-02
		Calcium signaling	1.17E+00	5.94E-02
		Glycerophospholipid metabolism	1.10E+00	5.62E-02
		Cell cycle: G2/M DNA damage checkpoint regulation	1.10E+00	9.30E-02
		Caveolar-mediated endocytosis	9.41E-01	7.41E-02
		Pentose phosphate pathway	9.04E-01	5.62E-02
	BACS	Sonic hedgehog signaling	1.89E+00	1.29E-01
		Lysine biosynthesis	1.78E+00	4.69E-02
		Calcium signaling	1.69E+00	5.45E-02
		Aminosugars metabolism	1.69E+00	5.77E-02
		Cell cycle: G1/S checkpoint regulation	1.57E+00	8.62E-02
		NF-κB signaling	1.54E+00	6.21E-02
		Glycolysis/gluconeogenesis	1.51E+00	5.67E-02
		Axonal guidance signaling	1.25E+00	4.35E-02
		Wnt/β-catenin signaling	1.24E+00	5.45E-02
		Integrin signaling	1.15E+00	5.18E-02
	Trails A	N-glycan degradation	2.01E+00	1.00E-01
		Pantothenate and CoA biosynthesis	1.53E+00	4.76E-02
		Neurotrophin/TRK signaling	1.51E+00	5.41E-02
		Glycine, serine and threonine metabolism	1.39E+00	3.47E-02
		Dopamine receptor signaling	1.18E+00	4.40E-02
		Valine, leucine and isoleucine biosynthesis	1.13E+00	4.44E-02
		NRF2-mediated oxidative stress response	1.12E+00	3.31E-02
		ERK/MAPK signaling	1.06E+00	3.28E-02
		Hypoxia signaling in the cardiovascular system	9.63E-01	4.23E-02
		Folate biosynthesis	9.35E-01	2.63E-02
	Domain	N-glycan degradation	2.34E+00	1.33E-01
		Glycosphingolipid biosynthesis – ganglioseries	1.37E+00	6.90E-02
		Aminosugars metabolism	1.06E+00	6.73E-02
		Glycosphingolipid biosynthesis – globoseries	9.76E-01	7.32E-02
		Glycerolipid metabolism	9.01E-01	4.14E-02
		Eicosanoid signaling	8.75E-01	5.95E-02
		Regulation of actin-based motility by rho	8.06E-01	5.43E-02
		C21-steroid hormone metabolism	7.97E-01	5.71E-02
		T helper cell differentiation	7.82E-01	7.89E-02
		Phototransduction pathway	7.59E-01	7.94E-02
Attention/vigilance	CPT	Neurotrophin/TRK signaling	1.61E00	5.13E-02
		CTLA4 signaling in cytotoxic T lymphocytes	1.36E00	4.49E-02
		Phenylalanine metabolism	1.31E00	2.8E-02
		NRF2-mediated oxidative stress response	1.27E00	3.24E-02
		N-glycan degradation	1.21E00	6.67E-02

(*Continued*)

Table S1 (*Continued*)

Domain	Test	Canonical pathway	*Log p*	Ratio
		Phenylalanine, tyrosine and tryptophan biosynthesis	1.18E00	3.12E-02
		Airway inflammation in asthma	1.12E00	1.67E-01
		One carbon pool by folate	1.1E00	5.26E-02
		Angiopoietin signaling	1.09E00	4.17E-02
		Pyrimidine metabolism	1.09E00	2.63E-02
Social cognition	MSCEIT	Tyrosine metabolism	1.71E00	1.08E-02
		Dendritic cell maturation	1.46E00	1.21E-02
		O-glycan biosynthesis	1.17E00	2.33E-02
		Serotonin receptor signaling	1.08E00	2.17E-02
		Retinol metabolism	1.02E00	1.61E-02
		Phenylalanine metabolism	9.94E-01	9.35E-03
		Role of cytokines in mediating communication between immune cells	9.79E-01	1.79E-02
		Lymphotoxin β receptor signaling	9.71E-01	1.69E-02
		C21-steroid hormone metabolism	9.43E-01	1.45E-02
		FGF signaling	8.02E-01	1.14E-02

Abbreviations: NAB, Neuropsychological Assessment Battery; HVLT, Hopkins Verbal Learning Test; SSpan, Wechsler Memory Scale-Spatial span; LN span, letter-number span; BVMT, Brief Visuospatial Memory Test; Cat, category test-animals; BACS, Brief Assessment of Cognition in Schizophrenia; CPT, Continuous Performance Test; MSCEIT, Mayer-Salovey-Caruso Emotional Intelligence Test.