

# Analysis of Whole Genome Biomarker Expression in Blood and Brain

Brandi Rollins, Maureen V. Martin, Ling Morgan, and Marquis P. Vawter\*

Department of Psychiatry and Human Behavior, School of Medicine, University of California, Irvine, California

Received 21 May 2009; Accepted 14 December 2009

The consistency of peripheral gene expression data and the overlap with brain expression has not been evaluated in biomarker discovery, nor has it been reported in multiple tissues from the same subjects on a genome wide transcript level. The effects of processing whole blood, transformation, and passaged cell lines on gene expression profiling was studied in healthy subjects using Affymetrix arrays. Ficoll extracted peripheral blood mononuclear cells (PBMCs), Epstein–Barr virus (EBV) transformed lymphocytes, passaged lymphoblastic cell lines (LCLs), and whole blood from Tempus tubes were compared. There were 6,813 transcripts differentially expressed between different methods of blood preparation. Principal component analysis resolved two partitions involving pre- and post-transformation EBV effects. Combining results from Affymetrix arrays, postmortem subjects' brain and PBMC profiles showed co-expression levels of summarized transcripts for 4,103 of 17,859 (22.9%) RefSeq transcripts. In a control experiment, rat hemi-brain and blood showed similar expression levels for 19% of RefSeq transcripts. After filtering transcripts that were not significantly different in abundance between human cerebellum and PBMCs from the Affymetrix exon array the correlation in mean transcript abundance was high as expected ( $r = 0.98$ ). Differences in the alternative splicing index in brain and blood were found for about 90% of all transcripts examined. This study demonstrates over 4,100 brain transcripts co-expressed in blood samples can be further examined by *in vitro* and *in vivo* experimental studies of blood and cell lines from patients with psychiatric disorders. © 2010 Wiley-Liss, Inc.

**Key words:** biomarker; gene expression; whole genome

## INTRODUCTION

The peripheral biomarker discovery process in psychiatric disorders has been rejuvenated by pharmacogenomics and gene expression high throughput data. Advances in the field of biomarkers have gained momentum partially from commercially available microarray high throughput measurements of mRNA. However, the field of biomarker discovery for neuropsychiatric disorders has not yet performed basic experiments to address methodological issues as to what may constitute a biomarker. In the field of cancer pathology there are biopsied samples from tumor progression; however, in psychiatry brain biopsies from psychiatric patients are not feasible for diagnostic purposes, and access to

### How to Cite this Article:

Rollins B, Martin MV, Morgan L, Vawter MP. 2010. Analysis of Whole Genome Biomarker Expression in Blood and Brain. *Am J Med Genet Part B* 153B:919–936.

healthy brain tissue is not possible for living control subjects. The interpretation of postmortem brain gene expression profiles are often complicated by clinical, tissue, gender, aging, and RNA quality factors [Galfalvy et al., 2003; Tomita et al., 2004; Erraji-Benchekroun et al., 2005; Lipska et al., 2006; Atz et al., 2007].

As a potential window into neural function, researchers in neuropsychiatry have focused on sampling other tissues from living subjects such as nasal biopsy [Wolozin et al., 1993; Vawter et al., 1996; Sawa et al., 1999; McCurdy et al., 2006], skin biopsy [Hagenfeldt et al., 1987; Mukherjee et al., 1994; Catts et al., 2006; Tomic et al., 2006], CSF, urine, serum, and blood. Peripheral blood mononuclear cells (PBMCs) are readily available and a commonly studied source of mRNA from living participants. Many studies of complex psychiatric disorders have collected PBMCs and transformed the cells with Epstein–Barr virus (EBV) [Middleton et al., 2004, 2005; Baron et al., 2006; Hu et al., 2006; Philibert et al.,

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: NIMH Conte Center; Grant Number: P50 MH60398; Grant sponsor: Pritzker Family Philanthropic Fund; Grant sponsor: Public Health Service Research Grant; Grant Number: M01 RR00827; Grant sponsor: National Center for Research Resource, Biomarker Research Grant; Grant Number: MH 024370; Grant sponsor: National Institute of Mental Health; Grant sponsor: William Lion Penzner Foundation (Department of Psychiatry).

No authors on this paper have a conflict of interest to declare.

Brandi Rollins and Maureen V. Martin contributed equally to this work.

\*Correspondence to:

Marquis P. Vawter, Functional Genomics Laboratory, Department of Psychiatry and Human Behavior, School of Medicine, University of California, 837 Health Science Drive, Gillespie Neuroscience Facility Room 2119, Irvine, CA 92697-4260. E-mail: mvawter@uci.edu  
Published online 2 February 2010 in Wiley InterScience  
(www.interscience.wiley.com)

DOI 10.1002/ajmg.b.31062

2007a,b] or have reported transcription profile results on freshly isolated PBMCs [Glatt et al., 2005; Segman et al., 2005; Tsuang et al., 2005; Bowden et al., 2006]. The differences in biomarker expression due to these manipulations have not been addressed, nor have temporal, circadian, or longitudinal assessments been made.

The first aim of this study was to estimate the common profile in the global transcriptome within the same subjects before and after EBV transformation. Characterization of gene expression profiles of different cell preparations in healthy controls may be useful for future experimental design and interpretation of peripheral biomarker experiments. There was a report on the effects of EBV transformation on the transcription profile of PBMCs which surveyed about 4,000 genes in lymphoma [Carter et al., 2002], but not the entire transcriptome. We addressed prominent factors that are different between biomarker studies of whole blood, Ficoll extracted PBMCs, EBV transformed lymphocytes, and EBV transformed and passaged lymphoblastic cell lines (LCLs). A within subjects experimental design was followed and subjects' blood draws were performed at the same time of day (to control for circadian effects) following an overnight stay at the University of California, Irvine Medical Center and used in each of these five different preparation methods. The second aim was to estimate the overlap between blood and brain gene expression and alternative splicing of transcripts in matched human and rodent samples. The data from both aims one and two were also contrasted to determine transcripts that are co-expressed in brain and different blood preparations.

## MATERIALS AND METHODS

### Human Subjects

Informed consent was obtained from each control subject using an approved University of California IRB protocol. All work was approved by the University of California, Irvine Human Subjects Committee. Blood samples were drawn at 7 AM, before breakfast after an overnight stay in a UCIMC sleep research center (four males and four females). The outline of the study and the demographics for each of the eight healthy subjects are shown (Supplementary Tables I and II).

We collected, extracted, and ran female subjects in one batch and male subjects in another batch on a separate day. Thus, the Tempus and downstream PBMCs and arrays were performed in separate batches for male and female.

In addition, we assessed pairs of tissues from seven subjects (cerebellum and PBMCs). The RNA samples from each pair were compared using Affymetrix microarray platforms. The RNA samples from six of these subjects' postmortem PMBCs were also compared to the eight healthy control subjects' PBMCs. The demographics of the seven postmortem subjects are shown (Supplementary Table III).

### Animal Subjects

Exon array data were collected for three adult Sprague–Dawley rats using matched blood and brain samples. All work was approved by the UCI Institutional Animal Care and Use Committee. The animals were anesthetized with ketamine/xylazine and 3 ml blood

samples were drawn by needle from the heart and stored overnight in Tempus tubes at 4°C. Rats were transcardially perfused with ~200 ml 1 × phosphate-buffered saline (PBS) to remove remaining blood from the brain. Brains were then removed, frozen on dry ice and stored at –80°C. RNA was extracted from whole blood using a Versagene kit (Minneapolis, MN) and from the entire left hemisphere of rat brain using Trizol. Two micrograms of RNA was labeled and run on Affymetrix Rat Exon Arrays following manufacturer's standard protocol identical to the human exon array procedure. (Please refer to the supplemental material to obtain more detailed discussions of the materials and methods used in this investigation.)

### Tempus Stability Tube Protocol

Blood samples were drawn into Tempus tubes (Applied Biosystems, Foster City, CA) from each healthy control subject (three Tempus tubes containing 9.0 ml total blood). Total RNA was isolated from whole blood with the Versagene RNA Purification System protocol (Gentra Systems, Minneapolis, MN). For more details regarding this protocol, see the Supplementary Methods.

### ACD Tube Protocol

At the time of blood collection in Tempus tubes blood was also collected into four ACD (acid citrate dextrose) tubes (Becton Dickinson, Franklin Lakes, NJ) from each healthy control subject (total of 25–35 ml of whole blood), mixed gently, and stored for 1 hr at room temperature ( $n = 8$ ). Whole blood samples were layered onto Ficoll (Amersham Biosciences, Piscataway, NJ), and PBMCs were separated by density gradient centrifugation at 2,500 rpm (Eppendorf rotor A-4-62) at room temperature for 20 min. The resulting “buffy” coat at the interface was added to 5 ml PBS, pH 7.4, (Invitrogen, Carlsbad, CA) and cell counts were taken with a standard hemocytometer. Cells were centrifuged at 1,000 rpm (Eppendorf rotor A-4-62) for 10 min at room temperature. The resulting pellets were resuspended in 1 ml Trizol and stored at –80°C to be used for RNA extractions.

### EBV Transformations

Freshly isolated PBMCs from each subject were transformed with the standard Coriell Institute (Camden, NJ) EBV protocol using phytohemagglutinin (Invitrogen) ( $n = 14$ ). PBMCs were expanded in RPMI-1640 media (Invitrogen) supplemented with 15% fetal bovine serum (heat-inactivated) (Invitrogen), 2 mM L-glutamine (MP Biomedicals, Solon, OH), and 25 mg of gentamicin (Invitrogen). Cultures were incubated at 37°C, 5% CO<sub>2</sub> and grown to an approximate cell density of  $1 \times 10^6$  cells/ml in 25-cm<sup>2</sup> flasks in a volume of 10 ml. Cultures were then transferred to 75-cm<sup>2</sup> flasks and fed media to a final volume of 50 ml. When cells reached  $\sim 1 \times 10^6$  cells/ml density the cultures were split, half of the lymphocytes were frozen down for RNA extractions (original PBMC transformation RNA) and half were fed equal volume of RPMI and left to grow to  $\sim 1 \times 10^6$  cells/ml density. At that time the cultures were again split, half frozen down for RNA extractions (passage 1 RNA) and half were fed an equal volume of RPMI media.

Lymphocytes were grown to  $\sim 1 \times 10^6$  cells/ml density then frozen for RNA extractions (passage 2 RNA).

### Postmortem PBMC Isolation

Postmortem blood was collected using cardiac puncture into ACD tubes from postmortem subjects no later than 24 hr post-time of death ( $n = 7$ ). The whole blood samples were processed the same as the samples above from living subjects. The resulting pellets were resuspended in 1 ml Trizol and stored at  $-80^\circ\text{C}$  to be used for RNA extractions.

### Trizol Extraction From PBMC and LCL Samples

RNA was extracted from frozen PBMCs ( $n = 7$ ) and EBV-transformed LCLs ( $n = 21$ ) using the standard Trizol isolation protocol (Invitrogen) as previously described [Vawter et al., 2004a]. (See the Supplementary Methods for further detail.) The total RNA was cleaned by passing over silica-based mini-spin columns (Qiagen RNeasy PlusMini Kit, Valencia, CA) and analyzed on a 2100 Bioanalyzer (Agilent, Palo Alto, CA) for quantification of 28S and 18S ribosomal RNA peaks and RIN numbers.

### Trizol Extraction From Brain Samples

All cerebellar samples were taken from the lateral aspect of one cerebellar hemisphere and included lateral portions of the middle lobe, containing lateral parts of the superior and inferior semilunar lobules. The cerebellar samples were microdissected to remove all but a thin ribbon of underlying white matter and the predominant gray matter sample was used for RNA extractions. RNA was extracted in TRIzol reagent using the above manufacturer's protocol. Total RNA was purified using a Qiagen RNeasy Mini Kit (Qiagen) and 2  $\mu\text{g}$  of RNA was used for each rat exon array.

### Oligonucleotide Microarrays

The Affymetrix oligonucleotide microarray chip (Human GeneChip Exon 1.0 ST) expression profiling experiments were carried out following the manufacturer's technical protocol (Affymetrix, Santa Clara, CA). (See the Supplementary Methods for more detail.) Arrays were washed, stained, and scanned on the Affymetrix Fluidics Station and G7 Affymetrix high-resolution scanner using GCOS 1.3. The cel files derived from the 1.0 ST chips were analyzed with a robust multiarray condensation algorithm (RMA) [Irizarry et al., 2003]. The postmortem PBMC samples were identically processed as above ( $n = 7$ ).

RMA expression values were subjected to a within array (subject)  $z$ -score transformation [Cheadle et al., 2003a,b]. The  $z$ -score transformation method [Cheadle et al., 2003a,b] was previously shown to be highly robust in detecting differential gene expression in a transcript spike-in experiment. The  $z$ -score normalized data were used to evaluate whether cell preparations if normalized on a chip basis would change the results. After application of the  $z$ -score transformation, the genes that passed Bonferroni cut-off were nearly identical,  $>97\%$  overlap between  $z$ -score transformed data and RMA normalized data. The data shown in this article focuses on the results using RMA normalized data. The quality control of each

exon array was evaluated by using ExACT (Affymetrix software) for all cel files. The Affymetrix Human 1.0 ST Exon Arrays used for the eight control subjects for five preparations yielded a total of 38 arrays because a single subject's passaged cell lines did not yield sufficient total RNA for two preparations. The same array type (Affymetrix Human 1.0 ST Exon Array) was used to compare four matched brain and PBMC pairs of samples from the same subjects. For a separate comparison of brain and lymphocytes, three post-mortem donor subjects were additionally matched for brain and PBMC pairs and analyzed on the Affymetrix U133 Plus chips using previously published methods [Vawter et al., 2006a]. The overlap between both platforms for the two independent brain and PBMC experiments was evaluated and reported. The same array type (Affymetrix Human 1.0 ST Exon Array) was also used to compare six PBMC samples from postmortem subjects to the eight PBMC samples from healthy controls. The gene expression results for PBMCs from living subjects and postmortem subjects were compared using the exon array.

Differential gene expression between tissues was analyzed by ANOVA with subject as a random effect and tissue type (brain vs. blood) as a fixed effect. The criteria to establish that gene expression differences were not significant for the purpose of this analysis was defined as a gene that displayed a non-significant ANOVA ( $P < 0.05$ ) for tissue type, a fold change within twofold, and minimum expression above background.

### Real-Time Quantitative PCR

Real-time quantitative PCR (qPCR) with SybrGreen dye was used to replicate the microarray results using methods as previously described [Vawter et al., 2004a,b, 2006a]. DNA was removed from each total RNA sample extracted from Tempus tubes using TURBO DNase-Free Kit (Ambion, Austin, TX) following the manufacturer's protocol for rigorous DNase treatment. For additional detail, see the Supplementary Methods. Although each RNA sample was either DNase treated or Qiagen column purified, to further increase gene specificity one primer in each pair was ideally designed (Primer Express, ABI, Foster City, CA) to span two exons and primers were not 3' UTR biased. The primers were tested for amplification of any residual genomic DNA contamination and primer sequences for each gene are listed in the Supplementary Methods. The dissociation curves of real-time PCR were monitored for primer-dimer pairings, which interfere with SybrGreen fluorescence measurements. A melting curve analysis was used for all primers on pooled cDNA, and any amplicons that showed evidence of double peaks in the qPCR dissociation curves or amplification of genomic DNA  $<35 C_t$  were discarded and redesigned.

Real-time qPCR was conducted with an Applied Biosystems 7900 sequence detection system (ABI) according to the manufacturer's protocol for SybrGreen PCR using a 25  $\mu\text{l}$  reaction volume and 5  $\mu\text{l}$  of diluted cDNA template. Samples were run in duplicate and the average  $C_t$  (crossing threshold) was calculated for each sample. The  $\Delta C_t$  calculation was used for relative fold change and a significance level of  $P < 0.05$  (two-tailed with unequal variance) was adopted as statistical evidence of qPCR technical validation of the microarray results. For some genes, there were no differences between the five blood preparations using microarray

results and these samples were similarly tested by qPCR for no difference by *t*-test. For normalization of the  $C_b$ , SLC9A1, previously shown in lymphocytes to be a stable reference gene [Vawter et al., 2006a], was used in qPCR calculations and was stable between preparations. The results of all attempted qPCR validations are shown.

## Alternative Splicing Analysis

Splicing changes were detecting by computing splice indexes, which are comparisons in the signal between different regions in the gene model [Clark et al., 2002]. Splice indexes were calculated using AltAnalyze, an application designed to detect tissue differences in alternative splicing events, using Affymetrix exon 1.0 ST array data. AltAnalyze calculates, organizes, filters, and summarizes transcript tiling data and then calculates a splice index for Ensembl transcripts by subtracting the log<sub>2</sub> expression of each probeset examined from the mean expression of all constitutive aligning probe sets for that array [Salomonis et al., 2009]. Probesets were not filtered based on detection above background (DABG) criteria to avoid potential loss of probesets not expressed between tissues. Alternative splicing differences between tissues were defined as those with *P*-values <0.05 when comparing differences in splice index between tissues by *t*-test.

## Quality Control

All of the quality control parameters (Supplementary Table IV) of the exon arrays for each batch (male and female) had relatively small differences in downstream PBMC RNA quality and quality control analysis receiver operator curve (exon/intron). By using the batch (gender) effect, it did not appreciably change our results as 93% of the same genes passed Bonferroni correction compared with not using the batch (gender) effect. Randomly assigning subjects to each batch irrespective of gender may be a more desirable experimental design; however, since the subjects were collected and extracted in real-time within 1 hr of collection, we carried this batch effect through the experiment. The quality control values for the exon arrays all exceeded the manufacturer's recommendation of 0.8 for all exon arrays (Supplementary Table IV). The minor differences in microarray quality across batches were further assessed by reanalysis of the entire dataset with a *z*-transformation, which minimizes absolute signal intensity differences among chips due to hybridization signal differences while maintaining within array ordinal values. The results and conclusions of downstream statistical analysis were very similar between the *z*-transformed RMA data and RMA dataset. Therefore, only the RMA downstream data analysis is reported.

## RESULTS

### Demographics

The quality of RNA and total white blood cell count are shown for each subject by five different preparation methods (Supplementary Table II). In all preparations, the minimum RNA quality (RNA Integrity Number, RIN) was greater than 8.5 for all samples and the mean RIN for each preparation was >9.0 which showed high

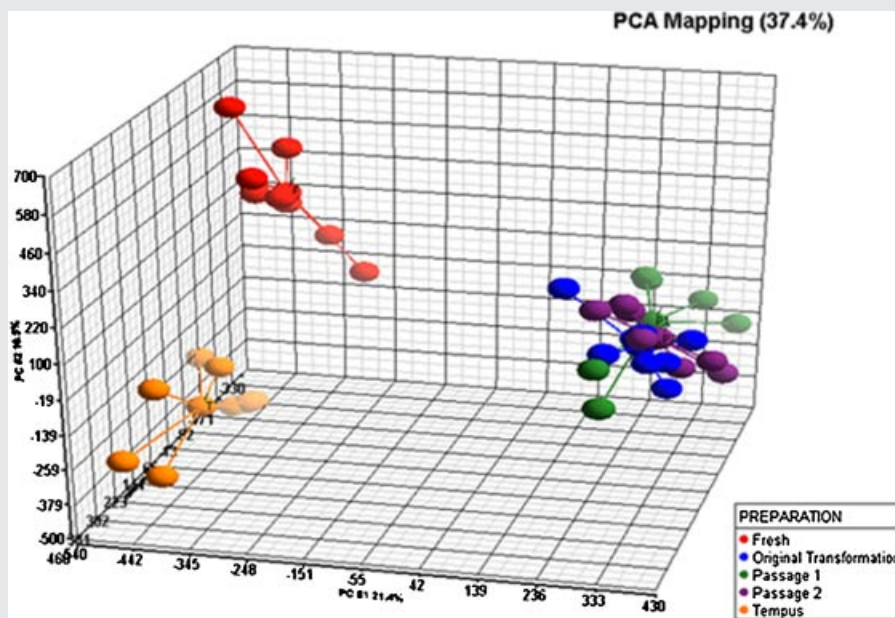
quality RNA was obtained. For reference, the maximum RIN is 10.0 for the highest quality RNA which was shown for many samples, while lower numbers represent slightly degraded but acceptable RNA. One subject showed a low number of cells in passages one and two, thus this subject's data was not available for two preparations. The statistical analysis compared the five different cell preparations from the same blood draw using eight unrelated healthy controls and 38 of 40 samples.

For the postmortem comparison of cerebellum and lymphocytes, a lower overall quality of total RNA was extracted due to the postmortem interval (PMI) and agonal factor effects, which negatively influence RNA stability (Supplementary Table III). We learned that routine transformation of postmortem PBMCs was problematic and transformation of PBMCs collected after ~20 hr PMI is wholly unreliable as the cells are not viable (data not shown). Transformation attempts of postmortem PBMCs used the same methods in our laboratory that have resulted in successful transformations of over 800 fresh samples for genetic studies from living subjects. Thus, our transformation methods work reasonably well but PBMCs were not viable when drawn after long PMIs and with current methods we could not perform the five methods of preparations reliably on postmortem samples.

### Experiment #1: Transformation Induced Transcript Differences Between Preparations

A principal component analysis (PCA) revealed no outliers within a preparation's cluster (Fig. 1). The PCA further showed that three different transformed PBMC preparation clustered together with no overlap between whole blood and Ficoll isolated PBMCs preparations. There were two chief partitions based upon EBV transformation status: partition one (pre-transformation) consisted of whole blood and Ficoll isolated PBMC preparations shown on the left side of the PCA plot (Fig. 1) and partition two (post-transformation) which consisted of the initially transformed PBMC group and the two passaged LCLs (shown on the right of Fig. 1). The post-transformed arrays showed high similarity on the PCA plot and some samples are superimposed on the right side of Figure 1.

The ANOVA for main effect of preparations showed that 6,813 transcripts were differentially expressed between one or more preparations following Bonferroni *P*-value correction for testing of all transcripts. Although it could be argued that Bonferroni correction is too stringent, Bonferroni correction was used in order to identify robust, replicable gene expression differences between preparations from the same subjects due to different cell types. This corresponds to a raw uncorrected *P*-value <  $2.79 \times 10^{-6}$ . Of these 6,813 transcripts there were 6,510 that showed large fold changes of a  $\pm 1.75$ -fold between one or more preparations. The largest number of transcript differences in expression (4,919 RefSeq transcripts) was induced by the transformation of PBMCs compared to fresh PBMC preparation on Ficoll. Significant differences in expression of 2,060 RefSeq transcripts were found between whole blood transcriptome (Tempus tube) and Ficoll isolated PBMC preparations. This indicated that the cellular composition of PBMCs and whole blood (Tempus tube) likely caused the large difference in transcription profiling. Somewhat surprising was the observation that once cells are transformed, the passage of



**FIG. 1.** The principal component analysis result of five preparations of fresh blood draws for each subject shows a virtual separation between pre-transformed (whole blood and PBMC Ficoll prepared cells) mRNA expression and post-transformation expression (P0, P1, and P2). This main effect was further analyzed in ANOVA to find transcript differences that passed Bonferroni corrections. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

transformed cell lines at confluency did not induce significant differential gene expression as only two transcripts passed Bonferroni *P*-value correction when comparing passaged LCLs. This low number of differentially expressed transcripts suggested that the immediate transformation effects were persistent through early passage.

### Genes With Stable Expression Among Preparations

The next analysis identified genes with stable expression among all of the five preparations. A criterion for stable expression was that a transcript must not show a significant difference by ANOVA (after Bonferroni correction *P*-value <0.05) for the main effect of preparation and any 10 post hoc mean comparisons of individual preparations. Further, all transcripts must show at least a minimum expression (raw expression intensity of 16) in each of the five preparations. When a single significant post hoc comparison occurred among preparations, the transcript expression was considered not stable, and was eliminated. Among the five preparations, there were 10,438 transcripts identified (that could be annotated to 8,448 unique genes) which did not show significant differences in post hoc comparisons. To determine whether genes that were not differentially expressed between preparations could be affected by an artifact due to data normalization, a different method of normalization (*z*-score transformation) was utilized. The results showed high agreement of 97.3% between different normalization methods. To summarize the effects of cell purification and transformation, three unsupervised hierarchical clusters were run (Fig. 2A–C). Figure 2A shows all RefSeq transcripts (17,859), Figure 2B shows the 10,438 transcripts that were not

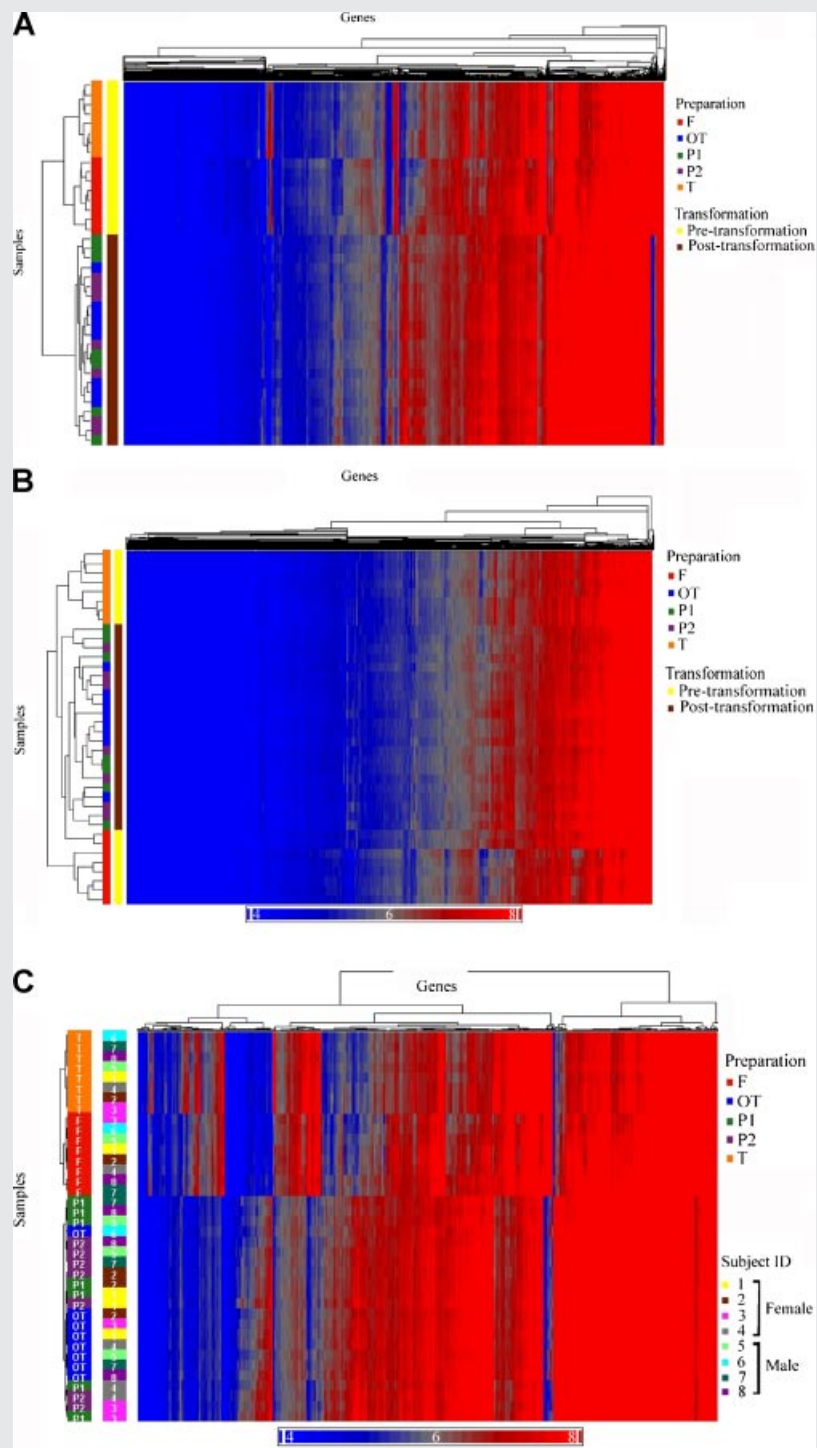
different among preparations, and Figure 2C shows the 6,813 transcripts that were significantly different between preparation methods.

As an external validity test for the effect of EBV transformation on cells, a literature search was conducted for genes that were reported to be associated with EBV. Gene Cards (GeneCards URL for keyword search <http://genecards.org/index.shtml>) was searched for relevant genes by using the keyword EBV which returned 189 gene symbols. Of these 189 gene symbols, 150 symbols matched RefSeq transcripts on the exon array annotation (Supplementary Table V). There were 77 genes (of the 150 reported genes in the literature) that were significant following Bonferroni correction using transformation as a factor (whole blood, PBMCs) compared to post-transformation (EBV transformed, passages 1 and 2). Individual reports [Carter et al., 2002; Nakayama et al., 2002] for the effects of EBV on gene expression represented additional sources of biological support of the validity of the current findings (Supplementary Table VI).

Eleven representative genes (Fig. 3) displaying individual subject's mean values pre- and post-transformation, indicate the large effect that blood purification and transformation had on expression profiles. Further, the transformation and blood purification effects were highly reproducible between subjects.

### Real-Time Quantitative PCR Technical Validation

Ten genes that were significantly different between preparation methods and passed Bonferroni correction were selected for qPCR technical validation. The RNA from cell extracts of the five preparations were DNase treated and primers were used only if genomic



**FIG. 2.** A: The unsupervised hierarchical clustering algorithm used squared Euclidean distance for both rows and columns (individual sample preparations and genes, respectively) and a random seed for the initial cluster. Unsupervised hierarchical clustering of all transcripts shows distinct clusters of five preparations of whole blood from the same subjects based upon all RefSeq transcripts. The distinct clusters are separated by pre- and post-transformation (yellow and brown, respectively) in agreement with the PCA (Fig. 1). All RefSeq transcripts are shown. B: Unsupervised hierarchical clustering is shown for selected transcripts with expression that was not significantly different between preparations. The post-transformed cell preparations from the same subjects show clustering (OT, initial transformation of PBMCs; P1, first passage of OT; and P2, second passage of OT). C: Unsupervised hierarchical clustering is shown for five preparations based upon 6,813 transcripts that were significantly different between preparations (passed Bonferroni correction, uncorrected  $P$ -value  $< 2.79 \times 10^{-6}$ ). All female subjects were collected in one batch (subject ID 1, 2, 3, and 4) while a second batch was processed on a different day for male subjects (subject ID 5, 6, 7, and 8). Clustering by gender (batch) was seen for the whole blood samples, but the samples tended to cluster less by gender (batch) in the subsequent four preparations. A–C: show similar clustering patterns, that is, pre-transformed cells (T, Tempus, whole blood; F, Ficoll prepared PBMCs) cluster together shown in the color bar to the left and legend to right. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

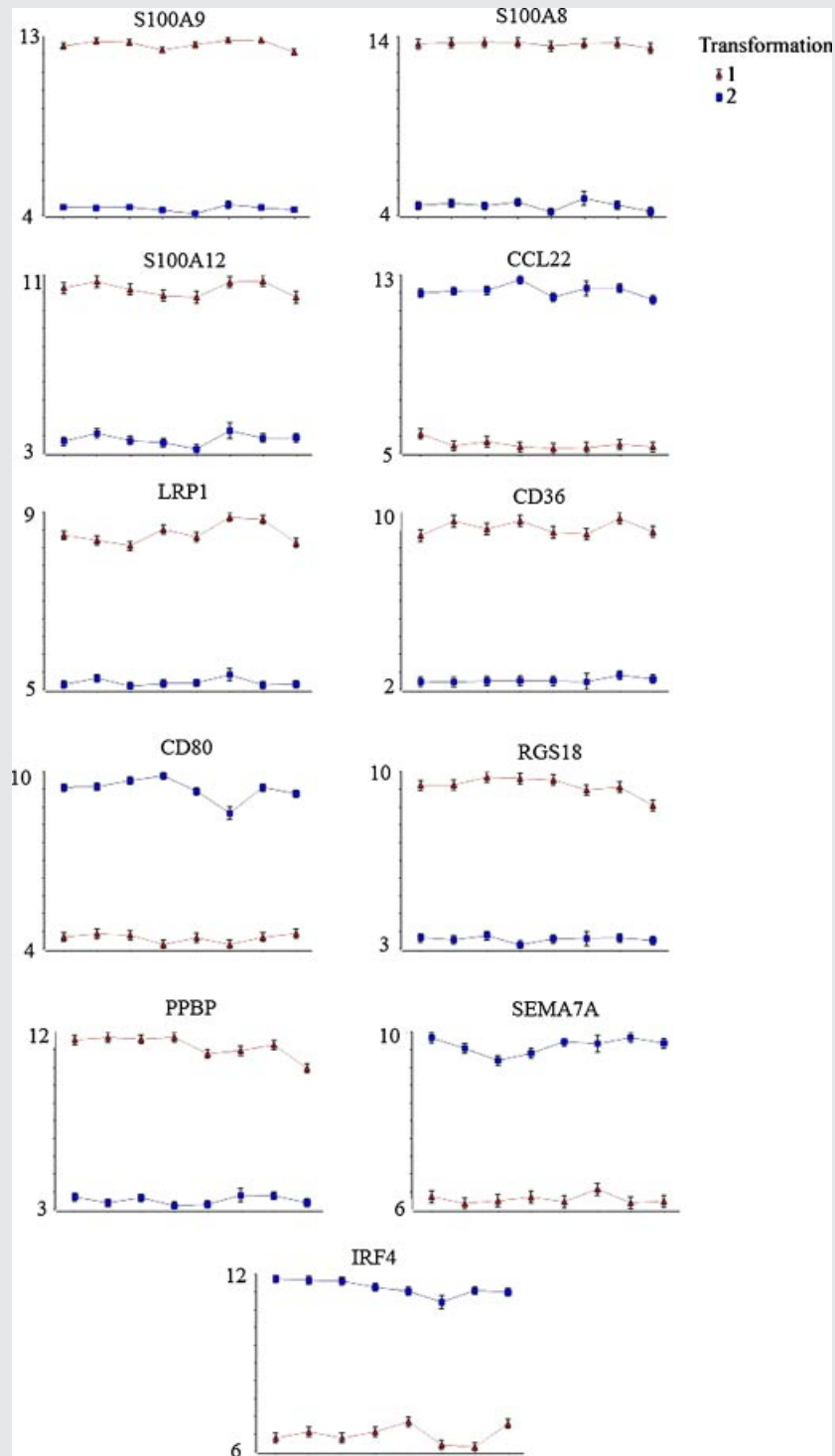


FIG. 3. Eleven genes showing differences in expression values for all eight subjects pre- and post-transformation using the Affymetrix exon array. The individual subjects' (x-axis) average  $\pm$  standard error of mean values before transformation (orange triangle) and after transformation (blue square) are shown. The y-axis represents signal intensity units on a log<sub>2</sub> scale. It can be seen that the gene expression differences are large and consistent for all eight subjects. Seven digit Affymetrix Transcript IDs and associated gene symbols/gene names are as follows: 2359664 [S100A9/Homo sapiens S100 calcium binding protein A9], 2435989 [S100A8/S100A8 Homo sapiens S100 calcium binding protein A8], 2435981 [S100A12/Homo sapiens S100 calcium binding protein A12], 3662687 [CCL22/Chemokine [C-C motif] ligand 22], 3417842 [LRP1/Low density lipoprotein-related protein 1], 3010503 [CD36/CD36 molecule, thrombospondin receptor], 2690900 [CD80/CD80 molecule], 2372719 [RGS18/Regulator of G-protein signaling 18], 2773358 [PPBP/pro-platelet basic protein], 3632907 [SEMA7A/Semaphorin 7A] 2891314 [IRF4/Interferon regulatory factor 4].

DNA did not amplify and cDNA showed positive amplification curves with one dissociation peak for melting curve analysis of the amplicon. Using these previously published stringent methods [Vawter et al., 2004a,b, 2006a,b] we technically validated all 10 transcripts selected (Table I). We have found 100% concordance between our qPCR and microarray fold change results for all attempted confirmations. All expected comparisons between preparations were significant except gene SH2D1A showed a non-significant *t*-test for PBMCs compared to passage 1. These 10 qPCR technical validations showed the microarray results were not due to a technical artifact on the chips or data analysis and that differences persisted after normalization with a reference housekeeping gene (SLC9A1) shown to be stable [Vawter et al., 2006a].

We observed large gene expression differences using the exon array across multiple probesets. In many genes 20–40 independent probesets are tiled across a single transcript and these multiple probesets instilled a confidence comparable to qPCR technical validation since the independent probesets were targeted to different exons of the transcript. Visual inspection of transcripts that passed Bonferroni correction for preparation differences showed that every probeset tiled on the array for the gene was consistently different (Supplementary Fig. 1 depicts exon level information in detail across subjects by preparations to show that for individual exons, they were always different across preparations). Our qPCR validation also showed non-significant differences in gene expression between whole blood and PBMC preparations for several genes which were predicted by microarray results to be similar (HCK, SH2D1A, TRAF1, and CD80, Table I).

### Over-Representation Pathway Analysis

The effects of transformation on gene expression profiles were also examined with an over-representation analysis [Dennis et al., 2003; Hosack et al., 2003]. Transcripts were selected based upon signifi-

cant differences in ANOVA between transformed and non-transformed preparations (see above) and the significant genes were annotated at the DAVID database for pathway selection. The results (Supplementary Table VII and Supplementary Fig. 2) showed over-represented genes involved expected pathways such as hematopoietic cell lineage, cytokine–cytokine receptor interaction, cell adhesion molecules, and T-cell receptor signaling pathways. In addition, a larger number of genes were found to be over-represented in cytobands at multiple genome locations, such as 1p, 1q, 2p, 2q, 7q, and 19q more than expected by chance (Table II). For example, there were 11 transcripts dysregulated in cytoband 19q13.4 following EBV transformation (FPR1, LILRA5, LILRA1, LILRB2, LILRA2, LILRA3, LILRA4, LILRB3, LILRB1, LAIR1, LILRA6). The genome wide distribution graphs (Supplementary Figs. 3 and 4) showed these regional hotspots of gene dysregulation following EBV infection also corresponded to sites of reported EBV integration [Gao et al., 2006]. However, without extensive sequencing of each hotspot we cannot say that the over-representation of gene dysregulation occurs due to cis—effects of viral integration.

### Experiment #2: Transcript Expression Between Brain and Lymphocytes Matched From Same Subjects Using Affymetrix U133 Plus 2.0 and Affymetrix Human Exon 1.0 ST Arrays

We compared PBMCs and brain tissues obtained from the same postmortem subjects in two experiments. In the first experiment, cerebellum tissue (n = 3) and Ficoll isolated PBMCs (n = 3) matched from the same subjects were extracted for RNA, labeled, and hybridized to individual Affymetrix U133 Plus 2.0 microarrays. The gene expression of PBMCs was compared to the same postmortem subjects' cerebellar cortex. The criteria for similar expression in PBMC and cerebellum were that the *t*-test

TABLE I. Ten Transcripts Were Selected for qPCR Technical Validation

Gene	Whole blood (Tempus)		Transformed PBMC		Passage 1		Passage 2	
	Fold change	<i>P</i> -value	Fold change	<i>P</i> -value	Fold change	<i>P</i> -value	Fold change	<i>P</i> -value
PPBP	<b>4.0</b>	<b>6.4E – 03</b>	<b>19,683</b>	<b>6.2E – 06</b>	<b>10,240</b>	<b>1.4E – 05</b>	<b>18,371</b>	<b>1.9E – 04</b>
S100A12	<b>4.3</b>	<b>6.7E – 03</b>	<b>5,869</b>	<b>3.3E – 05</b>	<b>3,130</b>	<b>3.4E – 05</b>	<b>6,023</b>	<b>2.9E – 04</b>
S100A9	<b>1.8</b>	<b>6.8E – 03</b>	<b>5,852</b>	<b>3.6E – 08</b>	<b>5,658</b>	<b>4.6E – 06</b>	<b>5,468</b>	<b>2.5E – 05</b>
TGFBI	<b>–3.3</b>	<b>4.7E – 02</b>	<b>73.3</b>	<b>2.8E – 04</b>	<b>107.5</b>	<b>4.7E – 04</b>	<b>104.3</b>	<b>5.7E – 04</b>
HCK	<b>–1.5</b>	<b>2.2E – 01</b>	<b>15.1</b>	<b>1.9E – 03</b>	<b>17.8</b>	<b>1.9E – 03</b>	<b>19.4</b>	<b>2.6E – 03</b>
SH2D1A	1.1	1.1E – 01	3.2	1.8E – 02	2.7	9.1E – 02	6.7	1.6E – 02
TRAF1	<b>–1.8</b>	<b>8.0E – 02</b>	<b>–45.2</b>	<b>1.0E – 04</b>	<b>–44.6</b>	<b>5.4E – 04</b>	<b>–47.5</b>	<b>7.4E – 04</b>
CD80	<b>–2.1</b>	<b>9.9E – 02</b>	<b>–743</b>	<b>1.16E – 07</b>	<b>–585</b>	<b>2.0E – 05</b>	<b>–966</b>	<b>2.0E – 05</b>
IRF4	<b>–18.5</b>	<b>4.2E – 02</b>	<b>–1,670</b>	<b>1.2E – 03</b>	<b>–492</b>	<b>5.7E – 03</b>	<b>–1,070</b>	<b>6.6E – 06</b>
CCL22	<b>–9.9</b>	<b>6.4E – 04</b>	<b>–3,774,508</b>	<b>3.5E – 08</b>	<b>–2,404,824</b>	<b>7.0E – 08</b>	<b>–1,504,356</b>	<b>5.1E – 06</b>

PPBP, pro-platelet basic protein chemokine [C–X–C motif] ligand 7; S100A12, S100 calcium binding protein A12; S100A9, S100 calcium binding protein A9; TGFBI, transforming growth factor, beta-induced; HCK, hemopoietic cell kinase; SH2D1A, SH2 domain protein 1A, Duncan's disease (lymphoproliferative syndrome); TRAF1, TNF receptor-associated factor 1; CD80, CD80 molecule; IRF4, interferon regulatory factor 4; CCL22, chemokine [C–C motif] ligand 22.

The fold changes for each preparation were compared to the Ficoll prepared PBMCs, that is, a positive fold change indicates higher expression in Ficoll prepared PBMCs, and a negative fold change indicates a decreased expression in the Ficoll prepared PBMCs. All genes tested were technically validated by qPCR, that is, the same RNA samples were used in the microarray and qPCR experiments. There were only modest fold changes between PBMCs and whole blood [Tempus], while the major changes are seen after transformation [transformed PBMC, passage 1, and passage 2]. The *P*-value is from a paired *t*-test comparing the Ficoll isolated PBMCs to each group using the  $\Delta C_t$  method. The target  $C_t$  was normalized to the housekeeping gene SLC9A1 previously shown in lymphocytes to be highly expressed in all five preparations. All bolded text indicates significant differences between PBMCs and preparation method.

TABLE II. Transcripts Differentially Expressed as a Result of EBV Infection

Cytoband	Count	%	Corrected <i>P</i> -value	Differentially expressed genes in cytoband
19q13.4	11	1.34	1.31E-04	FPR1, LILRA5, LILRA1, LILRB2, LILRA2, LILRA3, LILRA4, LILRB3, LILRB1, LAIR1, LILRA6
2p12	6	0.73	3.82E-04	IGKC, CD8A, CD8B, IGKV1-5, EIF2AK3, IGKV1D-13
7q34	6	0.73	2.20E-02	TRBV5-4, TRBV21-1, TRBC1, TRBV19, TRBV4-1, TRBV3-1
1p31	4	0.49	3.32E-02	MSH4, ACADM, RABGGTB, AK5
1q23	4	0.49	3.90E-02	FCGR2A, NDUFS2, FCER1G, FCGR3A
1p31.1	4	0.49	3.90E-02	IFI44, DNAJB4, CTH, IFI44L
2q35	5	0.61	5.46E-02	CTDSP1, DSU, SLC11A1, IL8RB, IL8RA

Two preparations were compared; Ficoll-extracted lymphocytes and lymphoblastic cell lines, and the genes that were over-represented for four different cytogenetic bands are shown.<sup>a</sup>The *P*-values represent Fisher exact tests that were not corrected for multiple testing but are shown for reference to over-represented functional classifications [Dennis et al., 2003]. The differentially expressed transcripts in EBV transformed LCLs compared to PBMCs were analyzed at the NIH DAVID website (<http://david.abcc.ncifcrf.gov/home.jsp>).

for mean difference was not significant ( $P < 0.05$ ), mean expression in each tissue was  $>16$ , and the mean difference between tissues was  $<0.5$ . Bonferroni correction may be too stringent for this experiment due to the small sample size and multiple confounds that have been shown to influence postmortem brain gene expression (i.e., PMI, agonal factors, pH). The results showed that 5,306 unique genes were expressed in both cerebellum and PBMCs at moderate to high signal levels (excluded any probe sets with signal intensity  $<16$ ) for matched subjects' tissues on the Affymetrix U133 Plus 2.0 microarray.

To assess the reliability of our findings, we profiled mRNA from an additional four pairs of matched brain (cerebellum) and Ficoll-extracted PBMC samples on Affymetrix Human Exon 1.0 ST arrays. On the exon array, the data were condensed into the 17,859 well-described RefSeq transcripts, and the same criteria used for the U133 Plus 2.0 chip data comparisons (above) were used. There were

no differences between matched cerebellum and PBMCs from the same subjects for 4,966 unique genes (see Fig. 4 and Supplementary Table VI). This result is in agreement with the U133 Plus 2.0 results. After combining the genes from both Affymetrix experiments there were 4,103 unique genes expressed in both cerebellum and PBMCs at similar expression levels. This overlap of gene expression between tissues showed a high concordance in two independent experiments, run on different platforms, labeling protocols, and independent subjects of brain and blood gene expression.

### Stable Expression of Putative Biomarkers—Intersection of Two Datasets

The intersection of the two main experiments reported above (Experiments 1 and 2, Fig. 5) showed 2,124 genes that were stably expressed among all five preparations and were also present in similar concentrations in both postmortem cerebellum and PBMC pairs derived by independent experiments. This list of putative biomarkers (Supplementary Table VIII) was formed by the intersection of the results from both experiments (8,448 genes that were not differentially expressed in five preparations of samples) and (cerebellum and PBMCs paired samples from the same subjects

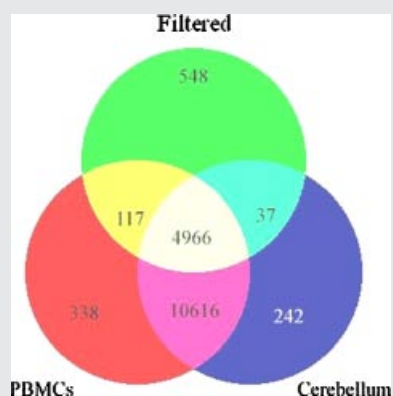


FIG. 4. This Venn diagram shows the overlap between the total number of transcripts with brain exon array expression greater than 4 (right), PBMC exon array expression greater than 4 (left) or filtered with  $P$ -value  $<0.05$  and mean difference  $\pm 0.5$  (after removing duplicated genes from overlapping list) (top). There were no differences between matched cerebellum and PBMCs from the same subjects for 4,966 unique genes. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

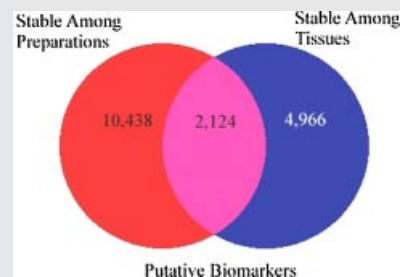


FIG. 5. This Venn diagram shows the overlap of stably expressed genes among five preparations (right) and two tissue types (left). There were 2,124 genes that were stably expressed among all five preparations and present in similar concentrations in postmortem cerebellum and PBMC pairs. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

where 4,103 genes overlapping from the Affymetrix exon array and U133 Plus 2.0 array are expressed at similar levels).

### Expression Correlation in PBMC and Brain Pairs

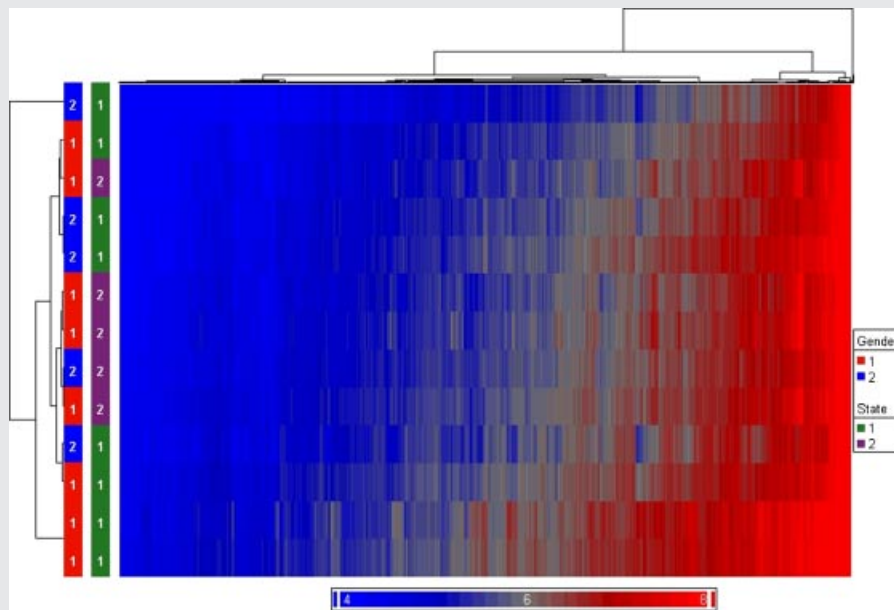
The similarity of expression of proposed biomarkers was next examined by correlations of PBMC and brain expressed transcripts from the same pairs. The simple Pearson correlations of brain-lymphocyte ( $r = 0.64$ ;  $n = 17,859$  transcripts) were performed on the entire transcript dataset from the Affymetrix exon array unfiltered for expression differences or signal intensity (Supplementary Fig. 5A). Transcripts that were not significantly different in abundance between cerebellum and PBMCs from the Affymetrix exon array (after Bonferroni correction) showed correlation in mean transcript abundance much higher as expected ( $r = 0.981$ ,  $n = 6,570$  genes, Supplementary Fig. 5B).

We queried the 4,103 co-expressed transcripts in the cerebellum and PBMC pairs from both Affymetrix platforms with a candidate gene list developed from the literature (Genetic Association Database) [Becker et al., 2004]. After consulting this database of psychiatric disorders for a list of candidate genes and ranking the relevant candidate genes based upon the number of literature references, we selected the top 200 cited candidate genes to determine whether these transcripts were highly expressed in cerebellum and PBMCs from the same pairs. Of these 200 candidate genes there were 252 relevant transcripts annotated on the exon array chip. We queried whether these 252 candidate psychiatric transcripts were expressed in both cerebellum and PBMC pairs from the same subjects (4,103 genes) and found 61 (24%). Examples of these genes are NPY, PSEN1, DTNBP1, CHRNA7, and MECP2, which are

implicated in various neuropsychiatric disorders (Supplementary Fig. 6) and represent potential biomarkers.

### Gene Expression in PBMCs Obtained From Postmortem Versus Living Subjects

We also ran a comparison between postmortem-obtained PBMCs and living subject PBMCs on the Affymetrix exon array platform (Supplementary Tables II and III). Ficoll isolated PBMCs ( $n = 8$ ) and postmortem PBMCs ( $n = 6$ ) were extracted for RNA, labeled, and hybridized to individual Affymetrix Exon 1.0 Arrays (Supplementary Table IV). There was no attempt to match postmortem and living subjects by gender, RIN, or age as the purpose of this preliminary analysis was to determine the upper estimate of genes that were similarly expressed between cell collection conditions. Thus, better matching might ensure a smaller number of genes actually pass the similarity threshold. The criteria for similar expression between postmortem and living gene expression profiles was that the uncorrected  $t$ -test for group mean differences were not significant ( $P < 0.05$ ), the mean expression in each PBMC group was  $> 16$ , and the mean difference between PBMC collection groups was less than  $\pm 1.4$ -fold. The results showed the upper estimate that 14,251 unique transcripts were expressed in approximately equal levels in both groups of PBMCs, postmortem and living subjects, at moderate to high signal levels. An unsupervised clustering showed distinct clusters of these two PBMC preparations reflecting their collection origins (Fig. 6). The similarity of expression of these two groups of PBMCs performed on the filtered transcript dataset from the Affymetrix exon array was  $r = 0.977$  ( $n = 14,251$  transcripts). However, postmortem PBMCs were an unreliable source for EBV



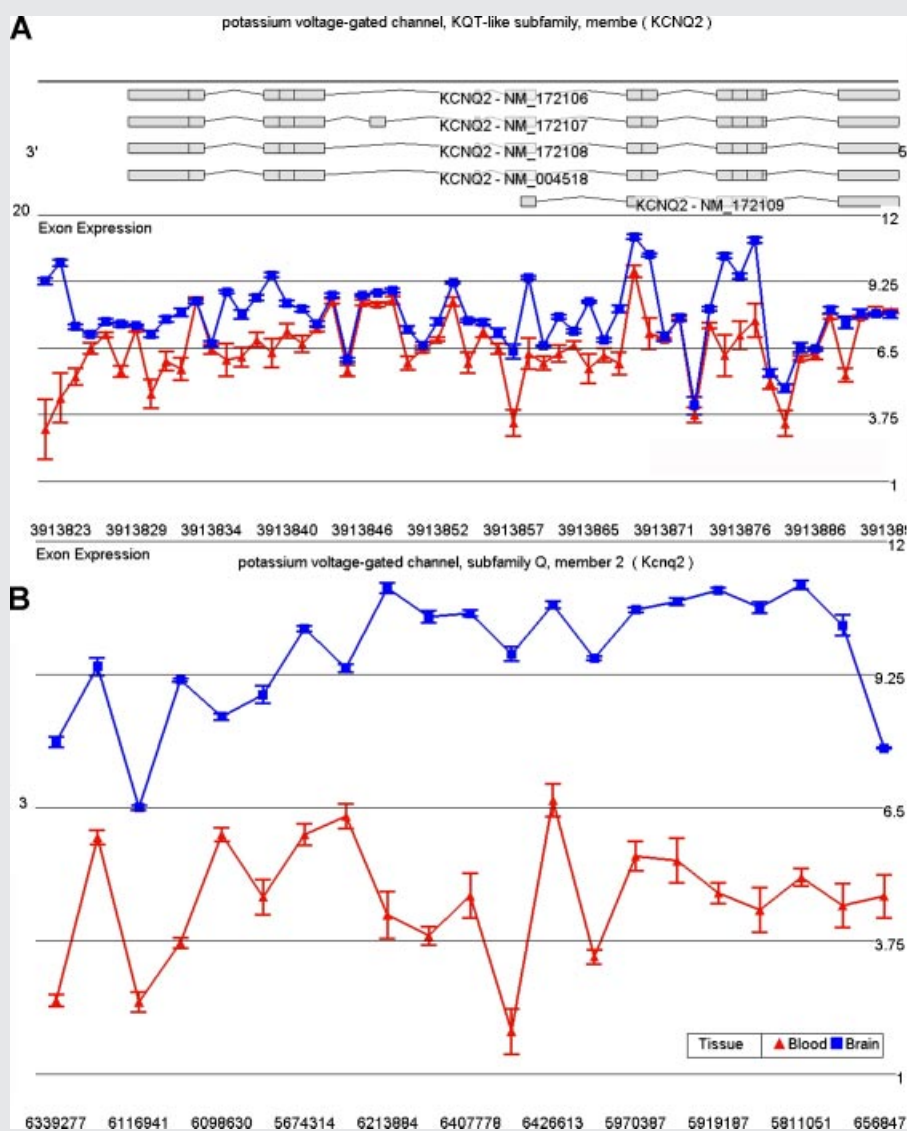
**FIG. 6.** Two preparations of PBMCs from living subjects and postmortem subjects [State 1 and 2, green and purple color bars, respectively] were assayed using Affymetrix exon expression arrays. The transcripts that were not significantly different were clustered ( $n = 14,251$ ). PBMCs nearest neighbors generally appeared by state (right column). Male subjects and female subjects are shown as Gender 1 and 2, red and blue color bars, respectively (left column). [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

transformation due to low cell viability; despite the cell viability differences, the preliminary comparisons show that gene expression profiles were consistent for a large set of transcripts. With better matching and a larger N, there might actually be fewer genes that are similar between postmortem and living PBMCs.

### Exon Array Comparison of Rat Brain and Blood

We examined the effect of tissue type (blood and brain) on gene expression in rat and as expected found a significant effect of tissue on gene expression. There are 8,766 RefSeq annotated transcripts on the rat exon array and 6,106 transcripts were differentially expressed ( $P < 0.05$ ) between blood and brain. The percentage of differentially expressed transcripts in brain and blood (69.7% rat)

was comparable to the percentage of differentially expressed transcripts in humans (72.4%). After removing transcripts with more than a twofold difference between tissue type, and below background mean expression in either brain or blood, there were 1,680 transcripts moderately to highly expressed at comparable levels in both rat tissues. This percentage (19%) of co-expressed transcripts is similar to the 22.9% (4,103/17,859) co-expression in the human array experimental data. The human data may have a larger noise associated with postmortem artifact, and thus not detect differences between tissue types as accurately as in the rat model results. Thus, the human experimental data represents an upper limit of 22.9% to the estimate of the number of transcripts that are comparably expressed between brain and blood. There was a significant difference in brain versus blood or lymphocyte expression of 1875 genes in both species (see Fig. 7 for a representative gene).



**FIG. 7.** There was a significant difference in brain versus blood or lymphocyte expression of 1,875 genes in both species. A representative gene, KCNQ2 is shown in A, which depicts human cerebellum exon array expression (red) and lymphocyte expression (blue). B depicts rat brain exon array expression (red) and blood expression (blue). [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

As an example of biomarker expression that might be useful for monitoring the effects of antipsychotic drug administration, the correlation between blood and saline-perfused brain gene expression was calculated for 44 of the 69 genes found to be influenced by olanzapine treatment [Fatemi et al., 2006] that were represented on the Rat Exon Array. For these 44 genes, the average absolute value of the correlation between blood and brain expression was 0.64. For example, a representative potential biomarker of an antipsychotic regulated gene is RGS19 (regulator of G-Protein signaling 19, Fig. 8). RGS19 expression was altered by antipsychotic treatment [Fatemi et al., 2006] and has highly correlated blood and brain expression ( $r = 0.92$ ) in the rat exon array. A total of 14 of the 44 genes had positive blood–brain correlations above 0.85 (Table III). There were 21 genes overlapping from the list of 1,353 genes with comparable blood and brain gene expression and 8 of these 21 genes were from the list of 14 highly correlated genes, demonstrating that high correlations between blood and brain gene expression are common even when gene expression levels are significantly different. It is possible that genes with higher blood/brain expression correlations will be more amenable to use in biomarker research.

Using an expression value of 7.0 as a cutoff for moderate expression, we found 13% of probesets to be moderately expressed in blood but not brain in the human, and 18% of probesets to be moderately expressed in blood but not brain in the rat, where confounding blood in brain tissue was controlled for. Thus, we do not think blood contamination of human brain tissue influenced our results significantly.

### Alternative Splicing in Blood Versus Brain

There are two or more types of measurements possible when using the exon array. The transcript can be summarized by using the

central tendency of all probesets expression that is often in the range of 20–40 probesets for each transcript. This data has been presented in the cerebellum–PBMC comparison. A second method is to assess alternative splicing. The results of transcript similarity calculated by each method are shown in Table IV. A total of 324,000 probesets representing 14,294 genes were analyzed in terms of expression and alternative splicing in the human. Alternative splicing analyses were performed with AltAnalyze described in the Materials and Methods Section. There were 152,004 alternatively spliced probesets (118,091 passing FDR step-up correction) between tissues based on comparison of splice indices by *t*-test, while there were 197,848 probeset expression differences (175,862 passing FDR step-up correction) between tissues. Approximately half of the probesets examined displayed no significant differences in alternative splicing between tissues. The distribution of alternative splicing indices in blood and brain are shown in Figure 9.

Rat brain was obtained with a short postmortem interval, and following sacrifice was immediately harvested and blood obtained at the time for exon array analysis. A total of 108,979 probesets representing 6,640 genes were analyzed in terms of expression and alternative splicing in the rat. There were 42,634 alternatively spliced probesets (24,210 probesets with *P*-values passing FDR step-up correction) between tissues based on comparison of splice indices by *t*-test, while there were 59,790 expression differences (49,816 probesets with *P*-values passing FDR step-up correction) between tissues. More than half of the probesets examined displayed no significant differences in alternative splicing between tissues.

Examination of alternative splicing differences on the gene level revealed a total of 12,608 differentially alternatively spliced genes in human blood versus brain. There was a total of 4,451 differentially alternatively spliced genes in rat blood versus brain. Of these

**TABLE III. Candidate Biomarkers of Antipsychotic Response Showing Expression Change Following Antipsychotic Treatment in Rat Frontal Cortex [Fatemi et al., 2006]**

Transcript ID	Gene name	Gene symbol	Pearson correlation
7058966	Insulin 2	Ins2	0.84
7185567	Chymotrypsinogen B1	Ctrb1	0.84
7038787	Fibroblast growth factor 3	Fgf3	0.86
7122882	Dentin sialophosphoprotein	Dspp	0.86
7226331	Angiopoietin-like 2	Angptl2	0.87
7300596	Thyroid stimulating hormone receptor	Tshr	0.87
7311776	Timeless homolog ( <i>Drosophila</i> )	Timeless	0.88
7222772	Carbohydrate sulfotransferase 3	Chst3	0.9
7354825	Aldehyde oxidase 1	Aox1	0.9
7027131	Solute carrier family 22 (extraneuronal monoamine transporter), member 3	Slc22a3	0.92
7031596	D site of albumin promoter [albumin D-box] binding protein	Dbp	0.92
7249028	Regulator of G-protein signaling 19	Rgs19	0.94
7382913	Cysteinyl leukotriene receptor 1	Cysltr1	0.94
7174110	Synaptotagmin IV	Syt4	0.997
7115455	Activating transcription factor 3	Atf3	0.998
7149313	Protocadherin 21	Pcdh21	0.999

Genes listed were up- or down-regulated by olanzapine and showed positive blood and brain gene expression correlation  $>0.85$  in rats on Affymetrix Rat Exon Arrays.

**TABLE IV. Transcript Expression and Alternative Splicing in Human and Rat Blood and Brain Were Compared Using Species-Specific Exon Array Results**

	Summarized transcripts (human/rat)	Probeset expression (human/rat)	Alternative splicing of probesets (human/rat)	Alternative splicing of genes (human/rat)	Overlap of genes with differential expression of summarized transcripts and alternative splicing in blood and brain (human/rat)
Total	17,859/8,766 (100%/100%)	324,000/108,879 (100%/100%)	324,000/108,879 (100%/100%)	14,053/6,640 (100%/100%)	12,826/4,074 (100%/100%)
Differential expression or splicing ( $P < 0.05$ ) in blood and brain comparison	13,756/2,255 (77%/26%)	197,848/59,790 (61%/55%)	152,004/42,634 (47%/39%)	12,608/4,451 (90%/67%)	7,719/1,604 (56%/36%)
No differential expression or splicing ( $P > 0.05$ ) in blood and brain comparison	4,103/6,511 (23%/74%)	126,152/49,089 (39%/45%)	171,996/66,243 (53%/61%)	1,445/2,189 (10%/33%)	399/1,014 (10%/16%)

The number of differentially expressed or spliced probesets or genes are listed. There was an increased number of differentially expressed and alternatively spliced probesets and transcripts between blood versus brain in the human compared to the rat. Due to differences in how probesets were realigned to form transcripts in Partek versus AltAnalyze, the number of total human or rat transcripts was not identical across analyses. Alternative splicing analyses detected a higher percentage of transcripts with differences in blood and brain than direct comparison of probesets or transcripts.

4,451 differentially alternatively spliced rat genes, 3,063 were also alternatively spliced in human, showing a high degree of overlap in alternative splicing between species. Approximately 70% of the genes alternatively spliced between rat blood and brain were also alternatively spliced between human blood and brain. There was overlap of transcript expression and alternative splicing between tissues for a large number of transcripts, as well as a moderate degree of overlap between genes displaying differential expression and differentially alternatively spliced genes in blood and brain (56% in humans, 36% in rats) (see Table IV), suggesting the utility of blood-based biomarkers for selected transcripts. There was an increased number of differentially expressed and alternatively spliced probesets and transcripts between blood versus brain in the human compared to the rat. Alternative splicing analyses detected a higher percentage of transcripts with differences in blood and brain than direct comparison of probesets or transcripts. This suggests detection of alternative spliced transcripts between tissues may be a more sensitive method for assessing between tissue differences than comparison of averaged transcript or individual probeset expression differences.

## DISCUSSION

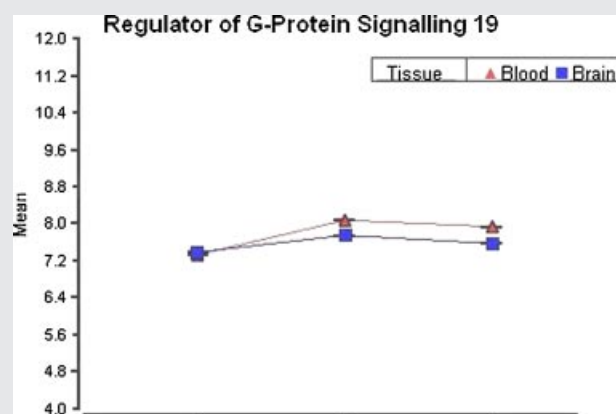
### Differences in Biomarkers Between Preparations

The first aim of this study was to examine the effects of PBMC purification of whole blood and EBV transformation on gene expression in peripheral blood collected from the same subjects. Large differences in gene expression occur following transformation of PBMCs that persist in the early passage of LCLs and appear stable. This strong effect following transformation of PBMCs on gene expression profiles was expected from prior studies. The inverse of this finding is that a large number of transcripts (62%) appear to be expressed in equivalent levels across all preparations (whole blood, PBMC, transformed PBMC, LCL passage one, LCL passage two). For external validity, over one-half of the genes

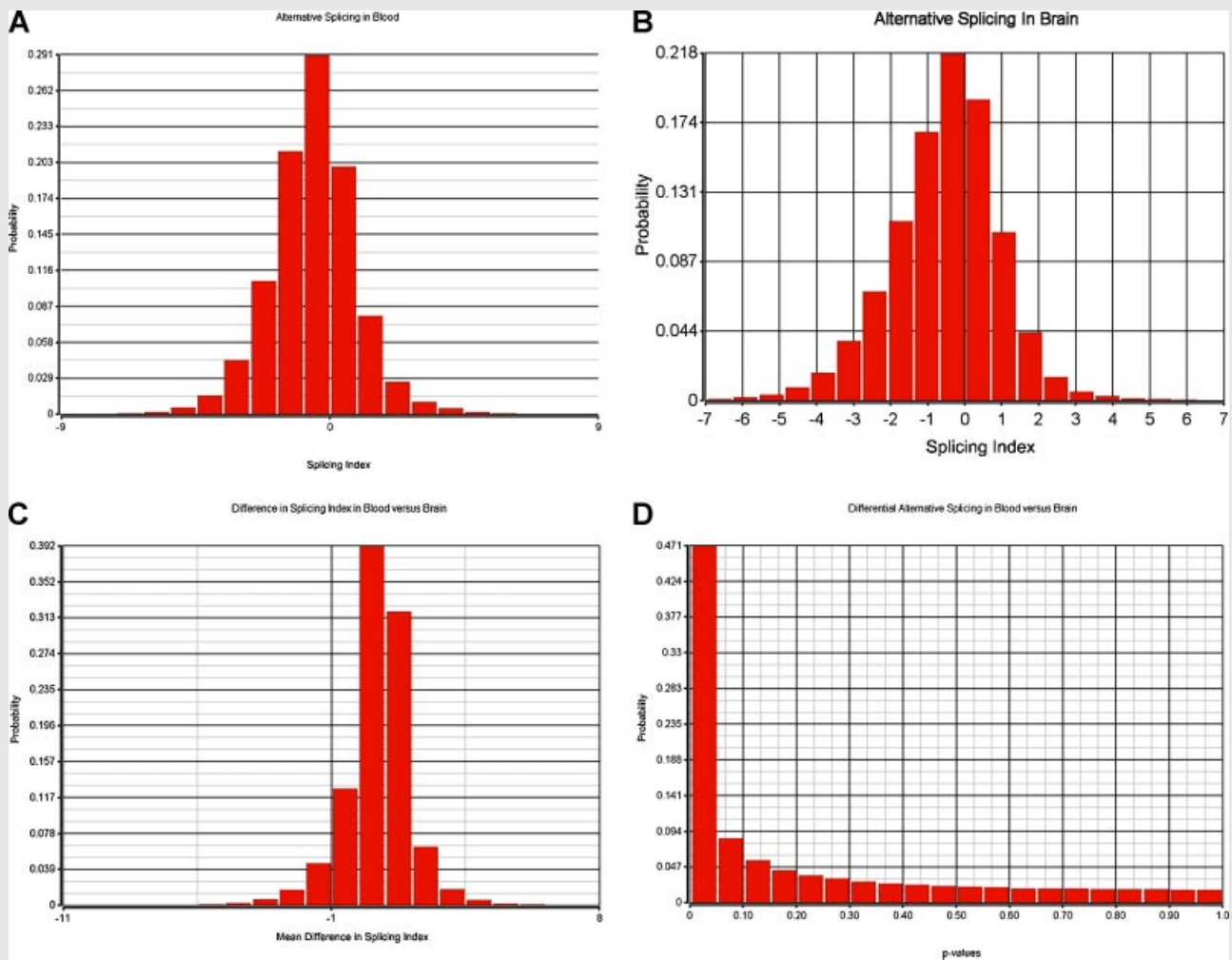
reported to be associated with EBV were significantly changed after transformation. The qPCR results provided technical validation of the microarray results for 100% of the transcripts tested.

### EBV Effects

Genes that were dysregulated among healthy individuals following EBV transformation exhibited low intersubject variability and did not appear to be randomly distributed across the genome. This may be due to multiple factors, including transformation insertion site uniformity, downstream responses of cell populations to EBV



**FIG. 8. RGS19 (regulator of G-protein signaling 19) was highly correlated in blood and brain gene expression (n = 3 independent animals). X-axis represents each animal, and the y-axis is the signal intensity (log<sub>2</sub>) average across probes. The mean ± standard error of the mean for blood and brain are shown. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]**



**FIG. 9.** a: The genome-wide splicing indices in blood was normally distributed. b: The genome-wide splicing indices in brain were normally distributed. c: The mean differences in splicing index were normally distributed. d: A total of 47% of probesets were significantly differentially alternatively spliced in brain compared to blood. The distribution of *P*-values is shown. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

infection, enrichment of a particular cell phenotype, or averaging across multiple clonal cell lines within an individual. These reasons cannot be differentiated without studying individual clonal cell lines from each subject and resequencing multiple genes to locate insertion sites. Prior reports demonstrate the impact of EBV transformation in malignant cells compared to non-malignant cells; however, we did not locate a report comparing in the same healthy individuals the effects of EBV transformed PBMCs to non-transformed PBMCs on a genome wide basis. The effects of a viral infection such as EBV on potential “biomarkers” need to be evaluated since a majority of adults (90%) present with latent EBV infections, for a review [Rickinson and Moss, 1997]. The variability of EBV infection in prior studies of psychiatric disorders is unknown [Schena et al., 1996; Nemanov et al., 1999; Whitney et al., 2003; Gladkevich et al., 2004; Iwamoto et al., 2004; Segman et al., 2005; Takemoto et al., 2005; Wang et al., 2005] except in cases where LCLs were used. Thus, to compare studies of PBMC, whole blood,

and LCLs the present data show that at least one-half of the transcriptome will be significantly altered and could account for lack of replication in biomarker study comparisons.

The effects of EBV transformation on gene expression implicate specific areas in the genome that induce or repress gene expression more often than chance, such as chromosomes 19q13.4, 2p12, 7q34, 1p31, 1q23, and 2q35. Previous studies [Wuu et al., 1996; Takakuwa et al., 2005] implicated several chromosomal regions as sites for integration for the EBV genome. The Raji cell line showed several EBV integration sites (1p, 1q, 2q, 3p, 3q, 4q, 5q, 6q, 7p, 7q, 9q, 11p, 14q, and 15q) through a FISH analysis with 64% of the total signals found at 4q, 2q, 1q, and 7q [Gao et al., 2006]. Chromosomes 1, 2, 4, and 5 frequently carried the EBV genome while chromosomes X and Y had no viral integration in 12 immortalized LCLs [Lestou et al., 1993]. These studies support multiple over represented chromosomal regions also found in the present study on chromosomes 2p12, 7q34, 1p31-31.1, and 1q23 (Supplementary

Figs. 3 and 4). It has been suggested that changes in copy number of genomic DNA might be expected to follow multiple passages of LCLs following EBV integration, as reported on chromosome 1p36.33 [Jeon et al., 2007]. We confined our experiment to early passages although later passages might show changes in copy number.

## Cell Purification

Beyond the direct effects of EBV on PBMCs, there are effects of cell type enrichment following purification and isolation of PBMCs and LCLs compared to whole blood. The differences in cell populations in whole blood, PBMCs, and LCLs affect the gene expression profile comparisons. Specific changes occurring between whole blood (Tempus tubes) and Ficoll isolated PBMCs are not due to transformation, and likely represent changes in cell populations during the isolation and purification of PBMCs. The PBMCs are isolated and frozen within 2 hr of blood draw and the majority is viable during the entire process as monitored with Trypan blue exclusion of cell counts. This result agrees with prior reports that RNA populations fluctuate within hours following blood draw [Rainen et al., 2002; Thach et al., 2003]. The transcriptome of whole blood drawn into Tempus tubes is immediately stabilized due to the proprietary reagents in Tempus tubes [Shou et al., 2005]. PBMCs and whole blood profiles while similar in the PCA have distinct profiles, and the question is raised as to which preparation might show the most relevance for biomarkers [Feezor et al., 2004]. The logical answer to this question is determined by the requirements of each research project, whether it is feasible to quickly isolate and sustain PBMCs for *in vitro* purposes such as drug stimulation or long-term culturing. On the other hand, Tempus whole blood collection preserves the transcriptome profile at the moment of blood draw, and does not require PBMC isolation. However, whole blood in Tempus tubes is not viable for drug stimulation or culturing studies. Ideally, both Tempus tube collection and standard PBMC isolation would offer the greatest range of experimental options.

## Postmortem Brain and PBMCs

The second part of this study compared the complete profile of gene expression in the circulating blood to brain expression using the same subjects. Three separate experiments were conducted to estimate the limit of co-expression of blood and brain transcripts. This estimate was based upon summarized transcript measures in human and rat studies. The results suggest whole blood and brain show co-expression of about 19–22% of the transcriptome. This estimate will be subject to revision as specific brain regions and blood cell components are assayed and compared. The hypothesis that gene expression in the blood is a probe of brain transcriptome was tested by simple overlap between expression profiles. If this hypothesis is proven to be true after more extensive testing, then it is tempting to speculate that gene expression peripheral biomarkers especially relevant to immune system transcripts could monitor disruption of brain function in some psychiatric disorders. The psychiatric conditions might be detected through analysis of transcripts differentially expressed in the blood alone. Transcripts in blood will provide information about the state of the circulating

cells and reflect about 20% of transcripts expressed in brain tissues. Blood cells can be used *in vitro* for experimental studies of brain expressed transcripts in both health and disease.

A major finding of this study was the high correlation (0.98) of gene expression between the matched pairs of PBMCs and cerebellum in a subset (22%) of the total transcriptome (4,103 genes of 17,859 genes). A similar proportion (24%) of frequently cited candidate genes in psychiatric disorders [Becker et al., 2004] shows expression in both PBMCs and cerebellum. Expression of many psychiatric candidate genes in PBMCs agrees with published single gene reports such as DISC1 [Sachs et al., 2005; Maeda et al., 2006], GRM3 splice variants [Sartorius et al., 2006], NPY, PPP3CC [Liu et al., 2007], ATF4 and ATF5 [Kakiuchi et al., 2007], PDLIM5 [Kato et al., 2005; Numata et al., 2007], HTR2A [Fukuda et al., 2006; Padin et al., 2006], D3 receptor [Padin et al., 2006], D4 receptor [Boneberg et al., 2006], MECP2 [Balmer et al., 2002; Jarrar et al., 2003], and CHRNA7. The relevance of our findings of blood and brain transcript similarity for 19–22% of transcripts establishes that future studies with brain tissue and PBMCs or LCLs gene expression will contain gene expression patterns that are in common with brain expression. The present results support the feasibility to probe genes related to neural function in the peripheral blood transcriptome as other researchers have reported gene expression differences in neuropsychiatric disorders in PBMC studies [Gladkevich et al., 2004; Segman et al., 2005; Takemoto et al., 2005; Tsuang et al., 2005; Bowden et al., 2006; Hu et al., 2006; Numata et al., 2007; Philibert et al., 2007a]. Further experimentation will be required to determine if co-regulation of altered blood and brain expression is found in specific disorders.

In the biomarker discovery process, it is anticipated that brain relevant genes would be expressed only in brain; therefore, a gene that is restricted to brain expression would be excluded as a biomarker because on theoretical grounds, it could not be expressed peripherally. Further, some false positives in our biomarker data exist, for example, moderately expressed levels of a biomarker in both blood and brain could reflect a gene that is a contaminant in brain due to residual contamination with blood. We tested for these false positives by comparing matched blood and brain samples that have been perfused via the aorta with saline prior to conducting gene expression analysis. The findings from the animal study suggest that contamination of human brain tissue by blood did not influence the results significantly.

## Postmortem and Living PBMCs

The preliminary results show that postmortem PBMCs are not useful for transformation studies (data not shown); however, postmortem PBMCs were similar to the gene expression profile of PBMCs collected from living subjects. We compared a list of frequently cited candidate genes in psychiatric disorders [Becker et al., 2004] and found that in the top 400 cited candidate genes, a reasonable proportion (53%) show gene expression in both living and postmortem collected PBMCs.

## Alternative Splicing

In the human and rat exon arrays there were approximately equal percentages of transcripts that showed differential alternative

splicing between brain and blood. This differential alternative splicing was found for one-half of the probesets represented on each species' array indicating that the alternative splicing index between blood and brain remained non-significant for one-half of the probesets. Of the 4,033 genes alternatively spliced in rat brain versus blood, ~70% of these were also shown to be alternatively spliced in human brain versus blood. Based upon the blood alternative splicing index, it may be a useful predictor of brain alternatively spliced transcripts for approximately half of the genes in the human transcriptome.

## CONCLUSIONS

The impact of various preparation methods and viral infections will need to be detailed in future discovery experiments aimed at using peripheral gene expression as reliable biomarkers of psychiatric disorders. It will be important to further determine if the peripheral transcriptome shown in this study also is found in other brain regions. Peripheral biomarkers might reflect viral and bacterial infections, adaptations due to molecular signaling, and cellular interchange (extravasation) between the blood-brain compartments. The peripheral gene expression profile might be indirectly associated with a neural disease signature, or participate in a common pathway alteration, or be part of the complex psychoneuroimmune feedback circuitries. The small number of summarized transcripts estimated at 19–22% that are expressed at similar levels in both brain and peripheral blood could represent a pool of genes useful as biomarkers, while the larger pool of transcripts that showed differential alternative splicing indices are also of interest. The alternative splicing index appeared to be significant in 90% of the transcriptome which is consistent with recent reports. Peripheral blood cells can be used for in vitro and in vivo experiments of brain relevant transcripts, and the differences and similarities in the two tissues' transcriptome will be of interest for experiments involving epigenetic, genetic, and environmental measures.

## ACKNOWLEDGMENTS

We acknowledge Jacque Berndt and the investigators and medical examiners at the Orange County Coroner's Office for procurement of brain tissue. We also appreciate the contributions of Dr. Preston Cartagena, Kathleen Burke, and Claudia Cervantes in processing blood and brain samples. F. Warren Lovell, MD, performed a neuropathological evaluation of the postmortem brains. We thank Nathan Salomonis for technical assistance with AltAnalyze. The procurement of brain samples was supported by the NIMH Conte Center Grant P50 MH60398 and the Pritzker Family Philanthropic Fund. This investigation was supported by Public Health Service research grant M01 RR00827 from the National Center for Research Resource, Biomarker research grant MH 024370 from the National Institute of Mental Health (M.P.V.), and the William Lion Penzner Foundation (Department of Psychiatry).

## REFERENCES

- Atz M, Walsh D, Cartagena P, Li J, Evans S, Choudary P, Overman K, Stein R, Tomita H, Potkin S, et al. 2007. Methodological considerations for gene expression profiling of human brain. *J Neurosci Methods* 163(2): 295–309.
- Balmer D, Arredondo J, Samaco RC, LaSalle JM. 2002. MECP2 mutations in Rett syndrome adversely affect lymphocyte growth, but do not affect imprinted gene expression in blood or brain. *Hum Genet* 110(2): 545–552.
- Baron CA, Liu SY, Hicks C, Gregg JP. 2006. Utilization of lymphoblastoid cell lines as a system for the molecular modeling of autism. *J Autism Dev Disord* 36(8):973–982.
- Becker KG, Barnes KC, Bright TJ, Wang SA. 2004. The genetic association database. *Nat Genet* 36(5):431–432.
- Boneberg EM, von Seydlitz E, Propster K, Watzl H, Rockstroh B, Illges H. 2006. D3 dopamine receptor mRNA is elevated in T cells of schizophrenic patients whereas D4 dopamine receptor mRNA is reduced in CD4+ -T cells. *J Neuroimmunol* 173(1–2):180–187.
- Bowden NA, Weidenhofer J, Scott RJ, Schall U, Todd J, Michie PT, Tooney PA. 2006. Preliminary investigation of gene expression profiles in peripheral blood lymphocytes in schizophrenia. *Schizophr Res* 82(2–3): 175–183.
- Carter KL, Cahir-McFarland E, Kieff E. 2002. Epstein-Barr virus-induced changes in B-lymphocyte gene expression. *J Virol* 76(20):10427–10436.
- Catts VS, Catts SV, McGrath JJ, Feron F, McLean D, Coulson EJ, Lutze-Mann LH. 2006. Apoptosis and schizophrenia: A pilot study based on dermal fibroblast cell lines. *Schizophr Res* 84(1):20–28.
- Cheadle C, Cho-Chung YS, Becker KG, Vawter MP. 2003a. Application of z-score transformation to Affymetrix data. *Appl Bioinformatics* 2(4):209–217.
- Cheadle C, Vawter MP, Freed WJ, Becker KG. 2003b. Analysis of microarray data using Z score transformation. *J Mol Diagn* 5(2):73–81.
- Clark TA, Sugnet CW, Ares M Jr. 2002. Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. *Science* 296(5569):907–910.
- Dennis G, Sherman B, Hosack D, Yang J, Gao W, Lane H, Lempicki R. 2003. DAVID: Database for annotation, visualization, and integrated discovery. *Genome Biol* 4:3.
- Erraji-Benchekroun L, Underwood MD, Arango V, Galfalvy H, Pavlidis P, Smyrniotopoulos P, Mann JJ, Sibille E. 2005. Molecular aging in human prefrontal cortex is selective and continuous throughout adult life. *Biol Psychiatry* 57(5):549–558.
- Fatemi SH, Reutiman TJ, Folsom TD, Bell C, Nos L, Fried P, Pearce DA, Singh S, Siderovski DP, Willard FS, et al. 2006. Chronic olanzapine treatment causes differential expression of genes in frontal cortex of rats as revealed by DNA microarray technique. *Neuropsychopharmacology* 31(9):1888–1899.
- Feezor RJ, Baker HV, Mindrinos M, Hayden D, Tannahill CL, Brownstein BH, Fay A, MacMillan S, Laramie J, Xiao W, et al. 2004. Whole blood and leukocyte RNA isolation for gene expression analyses. *Physiol Genomics* 19(3):247–254.
- Fukuda Y, Koga M, Arai M, Noguchi E, Ohtsuki T, Horiuchi Y, Ishiguro H, Niizato K, Iritani S, Itokawa M, et al. 2006. Monoallelic and unequal allelic expression of the HTR2A gene in human brain and peripheral lymphocytes. *Biol Psychiatry* 60(12):1331–1335.
- Galfalvy HC, Erraji-Benchekroun L, Smyrniotopoulos P, Pavlidis P, Ellis SP, Mann JJ, Sibille E, Arango V. 2003. Sex genes for genomic analysis in human brain: Internal controls for comparison of probe level data extraction. *BMC Bioinformatics* 4:37.
- Gao J, Luo X, Tang K, Li X, Li G. 2006. Epstein-Barr virus integrates frequently into chromosome 4q, 2q, 1q and 7q of Burkitt's lymphoma cell line (Raji). *J Virol Methods* 136(1–2):193–199.

- Gladkevich A, Kauffman H, Korf J. 2004. Lymphocytes as a neural probe: Potential for studying psychiatric disorders. *Prog Neuropsychopharmacol Biol Psychiatry* 28:559–576.
- Glatt SJ, Everall IP, Kremen WS, Corbeil J, Sasik R, Khanlou N, Han M, Liew CC, Tsuang MT. 2005. Comparative gene expression analysis of blood and brain provides concurrent validation of SELENBP1 up-regulation in schizophrenia. *Proc Natl Acad Sci USA* 102(43):15533–15538.
- Hagenfeldt L, Venizelos N, Bjerkenstedt L, Wiesel FA. 1987. Decreased tyrosine transport in fibroblasts from schizophrenic patients. *Life Sci* 41(25):2749–2757.
- Hosack DA, Dennis G Jr, Sherman BT, Lane HC, Lempicki RA. 2003. Identifying biological themes within lists of genes with EASE. *Genome Biol* 4(10):R70.
- Hu VW, Frank BC, Heine S, Lee NH, Quackenbush J. 2006. Gene expression profiling of lymphoblastoid cell lines from monozygotic twins discordant in severity of autism reveals differential regulation of neurologically relevant genes. *BMC Genomics* 7:118.
- Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4(2):249–264.
- Iwamoto K, Kakiuchi C, Bundo M, Ikeda K, Kato T. 2004. Molecular characterization of bipolar disorder by comparing gene expression profiles of postmortem brains of major mental disorders. *Mol Psychiatry* 9(4):406–416.
- Jarrar MH, Danko CG, Reddy S, Lee YJ, Bibat G, Kaufmann WE. 2003. MeCP2 expression in human cerebral cortex and lymphoid cells: immunohistochemical characterization of a novel higher-molecular-weight form. *J Child Neurol* 18(10):675–682.
- Jeon JP, Shim SM, Nam HY, Baik SY, Kim JW, Han BG. 2007. Copy number increase of 1p36.33 and mitochondrial genome amplification in Epstein-Barr virus-transformed lymphoblastoid cell lines. *Cancer Genet Cytogenet* 173(2):122–130.
- Kakiuchi C, Ishiwata M, Nanko S, Kunugi H, Minabe Y, Nakamura K, Mori N, Fujii K, Yamada K, Yoshikawa T, et al. 2007. Association analysis of ATF4 and ATF5, genes for interacting-proteins of DISC1, in bipolar disorder. *Neurosci Lett* 417(3):316–321.
- Kato T, Iwayama Y, Kakiuchi C, Iwamoto K, Yamada K, Minabe Y, Nakamura K, Mori N, Fujii K, Nanko S, et al. 2005. Gene expression and association analyses of LIM (PDLIM5) in bipolar disorder and schizophrenia. *Mol Psychiatry* 10(11):1045–1055.
- Lestou VS, De Braekeleer M, Strehl S, Ott G, Gadner H, Ambros PF. 1993. Non-random integration of Epstein-Barr virus in lymphoblastoid cell lines. *Genes Chromosomes Cancer* 8(1):38–48.
- Lipska BK, Deep-Soboslay A, Weickert CS, Hyde TM, Martin CE, Herman MM, Kleinman JE. 2006. Critical factors in gene expression in postmortem human brain: Focus on studies in schizophrenia. *Biol Psychiatry* 60(6):650–658.
- Liu YL, Fann CS, Liu CM, Chang CC, Yang WC, Hung SI, Yu SL, Hwang TJ, Hsieh MH, Liu CC, et al. 2007. More evidence supports the association of PPP3CC with schizophrenia. *Mol Psychiatry* 12(10):966–974.
- Maeda K, Nwulia E, Chang J, Balkissoon R, Ishizuka K, Chen H, Zandi P, McClinnis MG, Sawa A. 2006. Differential expression of disrupted-in-schizophrenia (DISC1) in bipolar disorder. *Biol Psychiatry* 60(9):929–935.
- McCurdy RD, Feron F, Perry C, Chant DC, McLean D, Matigian N, Hayward NK, McGrath JJ, Mackay-Sim A. 2006. Cell cycle alterations in biopsied olfactory neuroepithelium in schizophrenia and bipolar I disorder using cell culture and gene expression analyses. *Schizophr Res* 82(2–3):163–173.
- Middleton FA, Pato MT, Gentile KL, Morley CP, Zhao X, Eisener AF, Brown A, Petryshen TL, Kirby AN, Medeiros H, et al. 2004. Genomewide linkage analysis of bipolar disorder by use of a high-density single-nucleotide-polymorphism (SNP) genotyping assay: A comparison with microsatellite marker assays and finding of significant linkage to chromosome 6q22. *Am J Hum Genet* 74(5):886–897.
- Middleton FA, Pato CN, Gentile KL, McGann L, Brown AM, Trauzzi M, Diab H, Morley CP, Medeiros H, Macedo A, et al. 2005. Gene expression analysis of peripheral blood leukocytes from discordant sib-pairs with schizophrenia and bipolar disorder reveals points of convergence between genetic and functional genomic approaches. *Am J Med Genet Part B* 136B(1):12–25.
- Mukherjee S, Mahadik SP, Schnur DB, Laev H, Reddy R. 1994. Abnormal growth of cultured skin fibroblasts associated with poor premorbid history in schizophrenic patients. *Schizophr Res* 13(3):233–237.
- Nakayama T, Fujisawa R, Izawa D, Hieshima K, Takada K, Yoshie O. 2002. Human B cells immortalized with Epstein-Barr virus upregulate CCR6 and CCR10 and downregulate CXCR4 and CXCR5. *J Virol* 76(6):3072–3077.
- Nemanov L, Ebstein RP, Belmaker RH, Osher Y, Agam G. 1999. Effect of bipolar disorder on lymphocyte inositol monophosphatase mRNA levels. *Int J Neuropsychopharmacol* 2(1):25–29.
- Numata S, Ueno S, Iga J, Yamauchi K, Hongwei S, Hashimoto R, Takeda M, Kunugi H, Itakura M, Ohmori T. 2007. Gene expression in the peripheral leukocytes and association analysis of PDLIM5 gene in schizophrenia. *Neurosci Lett* 415(1):28–33.
- Padin JF, Rodriguez MA, Dominguez E, Dopeso-Reyes IG, Buceta M, Cano E, Sotelo E, Brea J, Caruncho HJ, Isabel Cadavid M, et al. 2006. Parallel regulation by olanzapine of the patterns of expression of 5-HT2A and D3 receptors in rat central nervous system and blood cells. *Neuropharmacology* 51(4):923–932.
- Philibert RA, Crowe R, Ryu GY, Yoon JG, Secrest D, Sandhu H, Madan A. 2007a. Transcriptional profiling of lymphoblast lines from subjects with panic disorder. *Am J Med Genet Part B* 144B(5):674–682.
- Philibert RA, Ryu GY, Yoon JG, Sandhu H, Hollenbeck N, Gunter T, Barkhurst A, Adams W, Madan A. 2007b. Transcriptional profiling of subjects from the Iowa adoption studies. *Am J Med Genet Part B* 144B(5):683–690.
- Rainen L, Oelmueller U, Jurgensen S, Wyrich R, Ballas C, Schram J, Herdman C, Bankaitis-Davis D, Nicholls N, Trollinger D, et al. 2002. Stabilization of mRNA expression in whole blood samples. *Clin Chem* 48(11):1883–1890.
- Rickinson AB, Moss DJ. 1997. Human cytotoxic T lymphocyte responses to Epstein-Barr virus infection. *Annu Rev Immunol* 15:405–431.
- Sachs NA, Sawa A, Holmes SE, Ross CA, DeLisi LE, Margolis RL. 2005. A frameshift mutation in disrupted in Schizophrenia 1 in an American family with schizophrenia and schizoaffective disorder. *Mol Psychiatry* 10(8):758–764.
- Salomonis N, Nelson B, Vranizan K, Pico AR, Hanspers K, Kuchinsky A, Ta L, Mercola M, Conklin BR. 2009. Alternative splicing in the differentiation of human embryonic stem cells into cardiac precursors. *PLoS Comput Biol* 5(11):e1000553.
- Sartorius LJ, Nagappan G, Lipska BK, Lu B, Sei Y, Ren-Patterson R, Li Z, Weinberger DR, Harrison PJ. 2006. Alternative splicing of human metabotropic glutamate receptor 3. *J Neurochem* 96(4):1139–1148.
- Sawa A, Wiegand GW, Cooper J, Margolis RL, Sharp AH, Lawler JF Jr, Greenamyre JT, Snyder SH, Ross CA. 1999. Increased apoptosis of Huntington disease lymphoblasts associated with repeat length-dependent mitochondrial depolarization. *Nat Med* 5(10):1194–1198.

- Schena M, Shalon D, Heller R, Chai A, Brown PO, Davis RW. 1996. Parallel human genome analysis: Microarray-based expression monitoring of 1000 genes. *Proc Natl Acad Sci USA* 93(20):10614–10619.
- Segman RH, Shefi N, Goltser-Dubner T, Friedman N, Kaminski N, Shalev AY. 2005. Peripheral blood mononuclear cell gene expression profiles identify emergent post-traumatic stress disorder among trauma survivors. *Mol Psychiatry* 10(5):500–513, 425.
- Shou J, Dotson C, Qian HR, Tao W, Lin C, Lawrence F, N'Cho M, Kulkarni NH, Bull CM, Gelbert LM, et al. 2005. Optimized blood cell profiling method for genomic biomarker discovery using high-density microarray. *Biomarkers* 10(4):310–320.
- Takakuwa T, Luo WJ, Ham MF, Wada N, Aozasa K. 2005. Identification of Epstein-Barr virus integrated sites in lymphoblastoid cell line (IB4). *Virus Res* 108(1–2):133–138.
- Takemoto M, Kira R, Kusuhara K, Torisu H, Sakai Y, Hara T. 2005. Gene expression profiles in peripheral blood mononuclear cells from patients with subacute sclerosing panencephalitis using oligonucleotide microarrays. *J Neurovirol* 11(3):299–305.
- Thach DC, Lin B, Walter E, Kruzlock R, Rowley RK, Tibbetts C, Stenger DA. 2003. Assessment of two methods for handling blood in collection tubes with RNA stabilizing agent for surveillance of gene expression profiles with high density microarrays. *J Immunol Methods* 283(1–2):269–279.
- Tomita H, Vawter MP, Walsh DM, Evans SJ, Choudary PV, Li J, Overman KM, Atz ME, Myers RM, Jones EG, et al. 2004. Effect of agonal and postmortem factors on gene expression profile: Quality control in microarray analyses of postmortem human brain. *Biol Psychiatry* 55(4):346–352.
- Tosic M, Ott J, Barral S, Bovet P, Deppen P, Gheorghita F, Matthey ML, Parnas J, Preisig M, Saraga M, et al. 2006. Schizophrenia and oxidative stress: Glutamate cysteine ligase modifier as a susceptibility gene. *Am J Hum Genet* 79(3):586–592.
- Tsuang MT, Nossova N, Yager T, Tsuang MM, Guo SC, Shyu KG, Glatt SJ, Liew CC. 2005. Assessing the validity of blood-based gene expression profiles for the classification of schizophrenia and bipolar disorder: A preliminary report. *Am J Med Genet Part B* 133B(1):1–5.
- Vawter MP, Basaric-Keys J, Li Y, Lester DS, Lebovics RS, Lesch KP, Kulaga H, Freed WJ, Sunderland T, Wolozin B. 1996. Human olfactory neuroepithelial cells: Tyrosine phosphorylation and process extension are increased by the combination of IL-1beta, IL-6, NGF, and bFGF. *Exp Neurol* 142(1):179–194.
- Vawter MP, Evans S, Choudary P, Tomita H, Meador-Woodruff J, Molnar M, Li J, Lopez JF, Myers R, Cox D, et al. 2004a. Gender-specific gene expression in postmortem human brain: Localization to sex chromosomes. *Neuropsychopharmacology* 29(2):373–384.
- Vawter MP, Shannon Weickert C, Ferran E, Matsumoto M, Overman K, Hyde TM, Weinberger DR, Bunney WE, Kleinman JE. 2004b. Gene expression of metabolic enzymes and a protease inhibitor in the prefrontal cortex are decreased in schizophrenia. *Neurochem Res* 29(6):1245–1255.
- Vawter MP, Atz ME, Rollins BL, Cooper-Casey KM, Shao L, Byerley WF. 2006a. Genome scans and gene expression microarrays converge to identify gene regulatory loci relevant in schizophrenia. *Hum Genet* 119(5):558–570.
- Vawter MP, Tomita H, Meng F, Bolstad B, Li J, Evans S, Choudary P, Atz M, Shao L, Neal C, et al. 2006b. Mitochondrial-related gene expression changes are sensitive to agonal-pH state: Implications for brain disorders. *Mol Psychiatry* 11(7):615, 663–679.
- Wang Z, Neuburg D, Li C, Su L, Kim JY, Chen JC, Christiani DC. 2005. Global gene expression profiling in whole-blood samples from individuals exposed to metal fumes. *Environ Health Perspect* 113(2):233–241.
- Whitney AR, Diehn M, Popper SJ, Alizadeh AA, Boldrick JC, Relman DA, Brown PO. 2003. Individuality and variation in gene expression patterns in human blood. *Proc Natl Acad Sci USA* 100(4):1896–1901.
- Wolozin B, Lesch P, Lebovics R, Sunderland T. 1993. A. E. Bennett Research Award 1993. Olfactory neuroblasts from Alzheimer donors: Studies on APP processing and cell regulation. *Biol Psychiatry* 34(12):824–838.
- Wuu KD, Chen YJ, Wu SW. 1996. Frequency and distribution of chromosomal integration sites of the Epstein-Barr virus genome. *J Formos Med Assoc* 95(12):911–916.