Genes and Immunity (2004), 1–7

www.nature.com/gene

FULL PAPER

Expression levels for many genes in human peripheral blood cells are highly sensitive to ex vivo incubation

EC Baechler[1], FM Batliwalla[3], G Karypis[2], PM Gaffney[1], K Moser[1], WA Ortmann[1], KJ Espe[1], S Balasubramanian[1], KM Hughes[1], JP Chan[4], A Begovich[4], S-YP Chang[4], PK Gregersen[3] and TW Behrens[1]

[1]*Department of Medicine, University of Minnesota, Minneapolis, MN, USA;* [2]*Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN, USA;* [3]*Robert S Boas Center for Genomics and Human Genetics, North Shore Long Island Jewish Research Institute, Manhasset, NY, USA;* [4]*Celera Diagnostics, Alameda, CA, USA*

Monitoring of gene and protein expression in peripheral blood cells has significant potential for improving the diagnosis and therapy of many human diseases. As genomic-scale microarray and proteomic technologies are applied to peripheral blood, it is important to consider the variables that may affect interpretation of data. Here we report experiments performed to identify genes that are particularly sensitive to ex vivo handling prior to RNA extraction for gene expression microarrays or quantitative real-time RT-PCR assays. We examined Affymetrix gene expression in samples from eight normal individuals where blood was processed for RNA either immediately after blood draw or the next day following overnight incubation. These studies identified hundreds of genes that are sensitive to ex vivo handling of blood, and suggest that this is an important variable to consider when designing and interpreting human PBMC experiments.
Genes and Immunity advance online publication, 3 June 2004; doi:10.1038/sj.gene.6364098

Keywords: *microarray; gene expression; peripheral blood mononuclear cells; RNA; stress response*

Introduction

The ability to profile gene expression on a genomic scale has facilitated rapid progress in identifying molecular pathways associated with human malignancies and other diseases.[1,2] Many of these studies have relied on tumor and other biopsy material from affected and unaffected individuals. However, such tissues are not always readily available for harvesting, and procurement can be expensive and invasive.

We, and others, have recently explored the use of peripheral blood cells as a readily 'biopsied' source of material for gene expression analyses in human disease.[3–7] Our work to date has focused on the examination of blood cells in patients with rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE).[5] RA is a chronic, inflammatory arthritis associated with progressive damage to synovial joints together with extra-articular manifestations.[8] SLE is a systemic autoimmune disorder characterized by autoantibodies that damage organs such as the skin, kidneys, joints, lungs, various blood elements, and the central nervous system.[9] The severity of disease, the spectrum of clinical involvement, and the response to therapies are highly variable among patients with both diseases, leading to significant challenges in diagnosis and management. Our gene expression studies were initiated with the hope of identifying new biomarkers to help guide the treatment of these disorders, as well as novel therapeutic targets.

During the course of these experiments, we identified several variables inherent to gene profiling in peripheral blood mononuclear cells (PBMCs) that have significant potential to influence data interpretation. These include differences in peripheral blood cell populations, circadian and/or cyclic hormonal influences on gene expression, as well as issues relating to the handling of samples *ex vivo* prior to the extraction of RNA. Here, we focus on the latter, and describe experiments performed to address the issue of the sensitivity of blood cell gene expression to *ex vivo* handling prior to RNA isolation and subsequent microarray or real-time PCR analyses. These experiments identify many genes that are sensitive to *ex vivo* incubation, suggesting that caution should be exercised before assigning roles for these genes in disease pathophysiology.

Results

In a recent study, we isolated RNA from the PBMCs of 48 SLE patients and 42 controls and then performed Affymetrix microarrays to identify differentially expressed genes.[5] During data analysis, we noticed clusters of genes that were strongly up- and downregulated specifically in samples where there was a significant

Correspondence: Dr TW Behrens, Center for Immunology, University of Minnesota Medical School, 6-126 BSBE Building, 312 Church Street SE, Minneapolis, MN 55455, USA. E-mail: behre001@umn.edu
Received 19 March 2004; accepted 05 April 2004

delay in blood cell processing. Some of the genes in the upregulated cluster included early response genes such as junB, v-fos, and immediate early response 3, as well as heat shock genes and others. Figure 1 shows unsupervised hierarchical clustering of genes that best discriminated normal control samples that were processed immediately after blood draw (0 h) from samples where blood was collected at an outlying clinic and then sent by overnight carrier to the laboratory (Figure 1a—O/N). Included in this cluster is one blood sample that was incubated overnight in the laboratory at room temperature and processed the next day (O/N*).

An intermediate gene expression signature was also observed in samples that were either drawn locally and delivered to the laboratory by courier, or drawn in the lab but not processed immediately (Figure 1a—2 h +). Other controls, from whom blood was drawn fresh in the laboratory and processed immediately (within 30 min), clustered separately from the samples that experienced a delay in processing (Figure 1a—0 h).

These data suggested the possibility that the expression of some genes in blood cells might be sensitive to the amount of time elapsed between blood draws and processing of the samples. When the same set of controls was clustered using genes that were absent from the *ex vivo* incubation signature, the samples no longer clustered according to their time delay in processing (Figure 1b).

We next designed a formal experiment to identify the genes that undergo significant changes in gene expression when blood cells are incubated for extended periods of time *ex vivo* (Figure 2a). Blood was drawn from eight healthy control individuals. Four of these controls were recruited and drawn at the University of Minnesota, while the other four were recruited and drawn at North Shore Long Island Jewish Research Institute. At each site, PBMCs were isolated immediately from two CPT tubes and resuspended in RNA*later*® (Ambion, Austin, TX, USA). RNA*later*® is a stabilizing solution that permeabilizes cells and protects the mRNA from degradation. The RNA preserved in RNA*later*® and two additional CPT tubes containing whole blood were then shipped overnight to the collaborating laboratory. PBMCs were isolated from the overnight-shipped CPT tubes, and RNA was then extracted from both samples and prepared for hybridization according to a standard Affymetrix protocol (see Materials and methods). Thus, from each of eight controls we obtained global gene expression profiles both from a fresh PBMC sample and from PBMCs isolated from blood shipped overnight, with both samples representing the same blood draw.

The array data were processed using Affymetrix Microarray Suite 4.0 software. Unsupervised hierarchical clustering of the 6414 genes found to be expressed in PBMCs demonstrated that the gene expression profiles of fresh samples were strikingly different from the

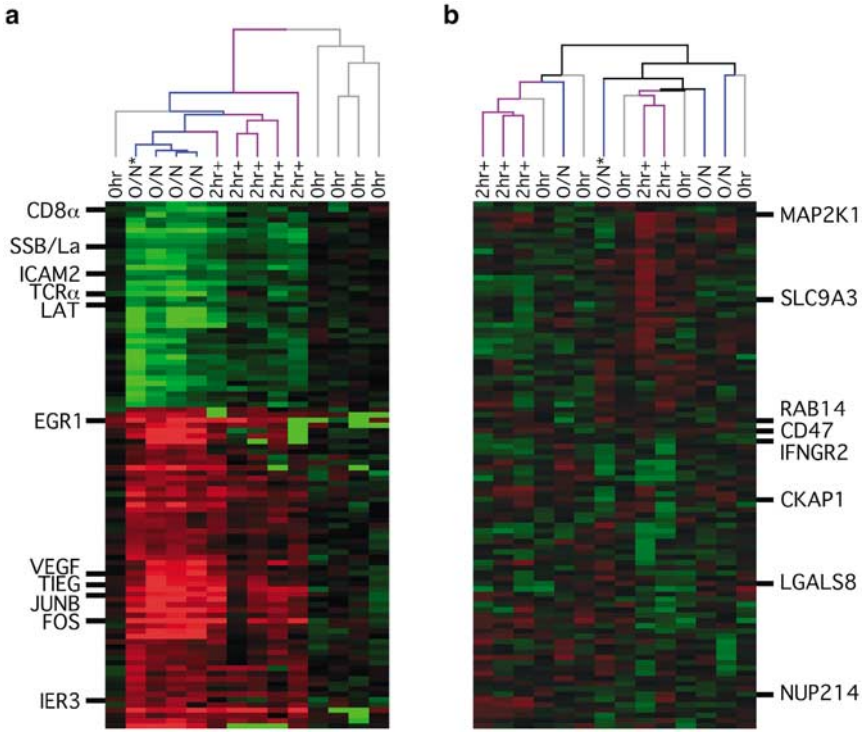

Figure 1 Delays in blood processing result in significant changes in gene expression. (a) PBMC samples from normal subjects were processed for microarrays immediately after isolation (0 hr), with a delay of 2 h or longer (2 hr +), or following overnight incubation of the sample (O/N). Hierarchical clustering was then used to visualize expression levels for a subset of 100 genes that were found to be dysregulated in overnight samples ($P < 0.01$ by paired *t*-test). This analysis resulted in the segregation of samples according to the time delay in processing of blood. SSB/LA, Sjogrens syndrome antigen B; ICAM2, intercellular adhesion molecule 2; TCRα, T-cell receptor alpha; LAT, linker for activation of T cells; EGR1, early growth response 1; VEGF, vascular endothelial growth factor; TIEG, TGF-beta inducible early growth response; IER3, immediate early response 3. (b) Clustering of the same samples as shown in panel a, using a subset of 100 genes that showed no significant changes in expression level after *ex vivo* incubation of blood. MAP2K1, mitogen-activated protein kinase kinase 1; SLC9A3, solute carrier family 9, isoform 3; IFNGR2, interferon γ receptor 2; CKAP1, cytoskeleton-associated protein 1; LGALS8, galectin 8; NUP214, nucleoporin 214 kDa.

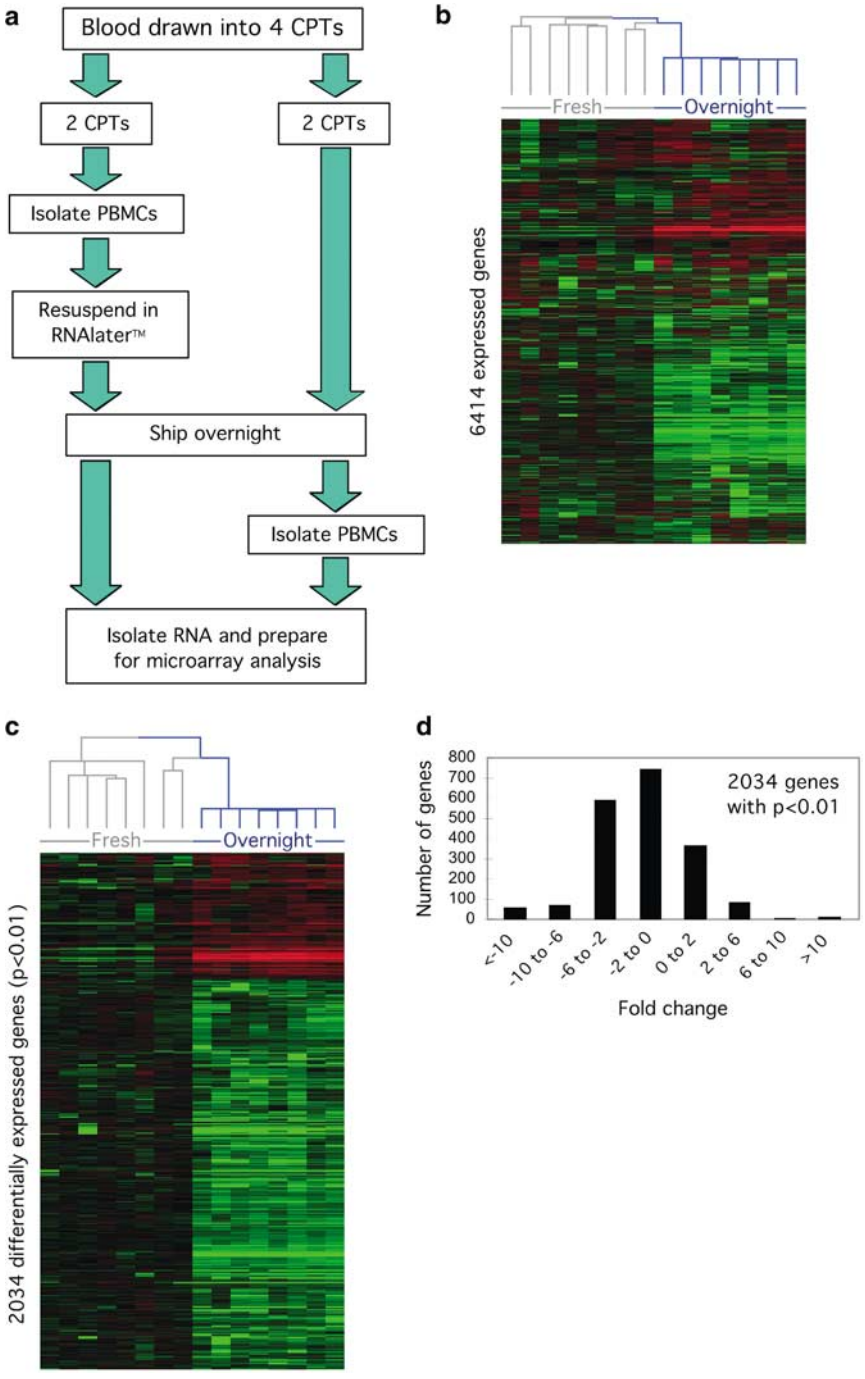

Figure 2 Identifying the genes that are sensitive to overnight *ex vivo* incubation of blood. (a) Overview of the study design. (b) Unsupervised hierarchical clustering of 6414 genes expressed in PBMCs separates overnight samples from freshly isolated samples. Gray bars, fresh samples; blue bars, samples shipped overnight. (c) Clustering of 2034 genes (of the 6414 expressed genes; 32% of total) that show altered expression levels during overnight shipment ($P<0.01$). Samples colored as in (b). (d) From the list of 2034 genes that exhibited significant changes in gene expression after *ex vivo* incubation ($P<0.01$), bars indicate the number of genes falling within the indicated range of fold changes.

profiles of samples shipped overnight (Figure 2b). In fact, the clustering program completely segregated fresh samples from their overnight counterparts. We conclude that many genes expressed in blood cells are very sensitive to *ex vivo* sample processing.

A paired *t*-test comparing the gene expression profiles of fresh blood *vs* blood shipped by overnight carrier identified 2034 genes that undergo significant changes in expression under these environmental stresses ($P<0.01$, Figure 2c). The majority of these genes exhibited an average fold change of less than 6 (increase or decrease), although several genes underwent even more dramatic changes in expression (Figure 2d). The array analysis program SAM (Significance of Microarrays[10]) similarly identified 97% of these 2034 genes as differentially expressed between the two groups.

Genes that were differentially expressed in the overnight samples belong to a variety of biological pathways (Figure 3). Many of these genes are known to participate in stress-induced pathways: immediate early genes, early growth response genes, heat shock proteins, etc. For example, the oncogenes jun and fos are stress-responsive,[11] and were expressed at many-fold higher levels in the overnight-shipped samples as compared to the fresh samples. We also observed elevated expression of the protein phosphatase DUSP1, which is upregulated during conditions of oxidative stress and heat shock.[12] Vascular endothelial growth factor (VEGF) also participates in stress responses[13] and showed higher expression levels after overnight incubation of blood. Other known stress-response genes found to be upregulated during overnight shipment included TGF-beta-inducible early growth response gene (TIEG) and growth arrest and DNA-damage-inducible alpha (GADD45A) and beta (GADD45B).

Many of the genes sensitive to *ex vivo* incubation are involved in basic cellular processes such as transcriptional regulation, cell cycle progression, and apoptosis. The transcriptional regulator nuclear factor kappa-B (NFKB) activates the transcription of many downstream genes that mediate inflammatory responses.[14] Two proteins that inhibit activation of NFKB (NFKBIA and NFKBIE) were increased in expression in the overnight samples, while NFKB2 (a subunit of the NFKB complex) was downregulated (Figure 3—Transcription). Other notable transcription factors in the list of differentially expressed genes include the signal transducer and activator of transcription (STAT) family members STAT1 and STAT4, and the breast and ovarian cancer susceptibility gene BRCA1.

The cell division cycle (CDC) proteins in yeast carry out various functions promoting progression of the cell cycle.[15] Two human homologs of this family, CDC23 and CDC42, were dysregulated in the overnight-shipped blood samples (Figure 3—Cell Cycle). In addition, two cyclin-dependent kinase inhibitors (CDKN1A and CDKN2D), which regulate cell cycle progression through G1, were also differentially expressed. Transcripts for three members of the origin recognition complex (ORC2L, ORC3L, and ORC5L), which is essential for the initiation of DNA replication, were found at lower levels in the overnight samples.

Several genes that promote Fas-dependent apoptosis were decreased in expression after overnight incubation (Figure 3—TNF-related, Apoptosis). Interaction of the death receptor Fas with its ligand (FasL/TNFSF6) initiates apoptosis in many cell types, including lymphocytes. Both the receptor and its ligand are candidate contributory genes in SLE, as mutations in either of these genes result in autoimmunity in mouse models of SLE.[16,17] Expression of FasL was decreased in the overnight samples relative to the fresh samples. Fas-associated via death domain (FADD), an important adaptor protein in the signaling pathway downstream

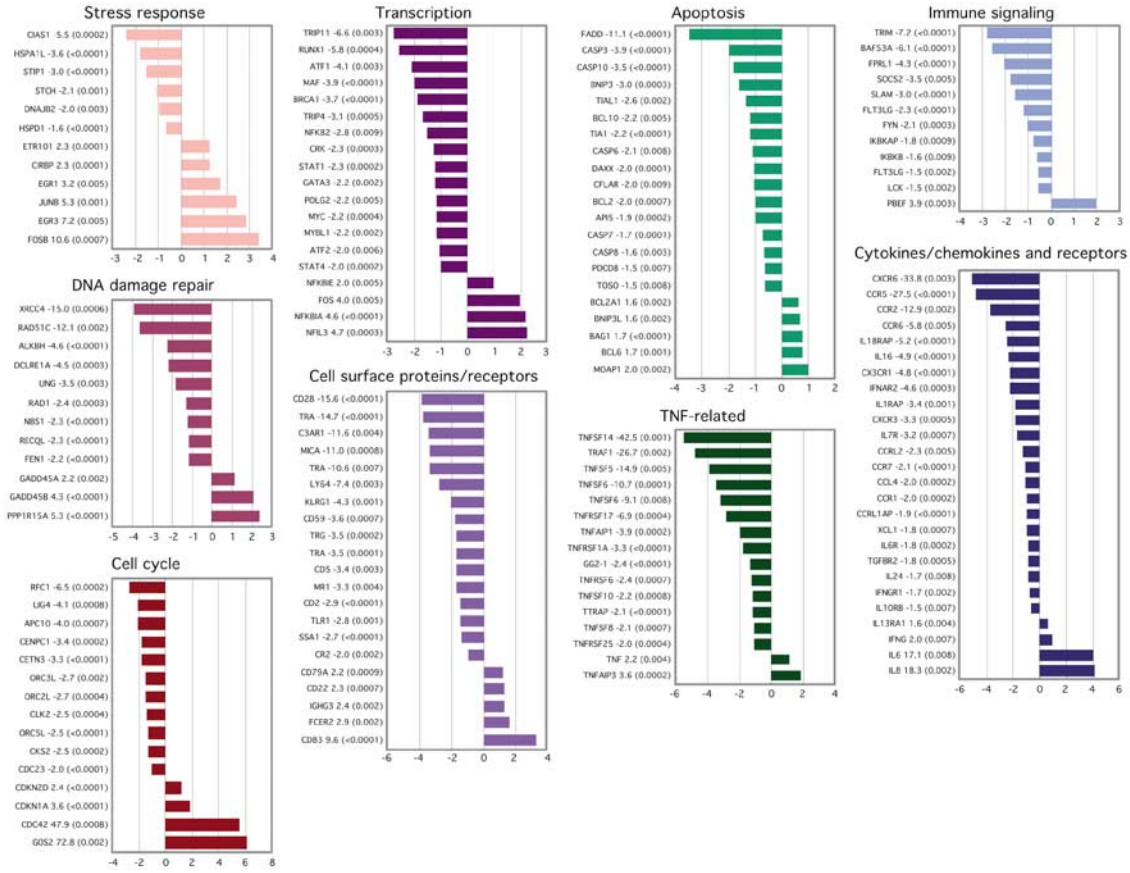

Figure 3 Functional classification of selected genes whose expression is sensitive to overnight *ex vivo* incubation of blood. Following the gene symbol is the average linear fold change (increase or decrease) of eight overnight samples *vs* eight fresh samples. In parentheses is the *P*-value from a paired *t*-test. Bars represent $\log_2$ of the average fold change.

of this receptor, also showed decreased expression. Interaction of FADD with the Fas receptor leads to activation of the caspase proteolytic cascade. Caspase 10 (CASP10), one of the earliest cysteine proteases in the cascade, was downregulated in overnight samples.

Importantly, many of the genes found to be sensitive to *ex vivo* stress encode proteins that perform functions essential to the immune response. In particular, the T-cell signaling molecules LCK (lymphocyte-specific protein tyrosine kinase), FYN (FYN oncogene related to SRC, FGR, YES), and SLAM (signaling lymphocytic activation molecule) were expressed at lower levels in the overnight samples (Figure 3—Immune Signaling). Transcripts for two key regulators of inflammatory signaling (IKBKAP and IKBKB) were also decreased in overnight samples. The majority of the cytokines and chemokines that were differentially expressed exhibited decreased expression in overnight samples; notable exceptions include IL8 and IL6, which were dramatically upregulated (Figure 3—Cytokines/Chemokines and Receptors). Many receptors for cytokines and chemokines were also among the differentially expressed genes. Other important cell surface proteins include the complement receptors CR2 and C3AR1, the autoantigen SSA1, toll-like receptor TLR1, and the T-cell receptor alpha and gamma chains (Figure 3—Cell Surface Proteins/Receptors). In addition, the expression of many TNF-related genes was sensitive to *ex vivo* incubation time. TNF itself was expressed at increased levels in overnight samples, while several other members of the TNF and TNFR superfamilies were downregulated.

These data indicate that a variety of signaling pathways are activated in peripheral blood cells following overnight shipment. We also saw evidence for these changes in samples that were drawn locally but processed two or more hours after blood draw (Figure 1b), suggesting that the problem is not simply one of overnight shipment. To test this, we used quantitative real-time RT-PCR to measure gene expression levels from blood drawn into CPT or EDTA tubes and incubated on the lab bench for 0, 2, 6, 12, 24, and 48 h. This allowed us to observe gene expression changes in PBMCs (isolated from CPT tubes) as well as total white blood cells (isolated from EDTA tubes). We observed significant changes in the expression of many genes over this time course (Figure 4). In some cases, changes were already apparent at the earliest timepoint measured (2 h post-draw), and the trends generally continued throughout the period of measurement. The surface glycoprotein ICAM1 (intercellular adhesion molecule 1; CD54) and the amino-acid transporter SLC7A5 (solute carrier family 7, member 5) were dramatically upregulated throughout the course of the experiment, particularly in total white blood cells. The cytokine interleukin-16 (IL-16) was downregulated during the period of incubation, while the expression of cytokine IFN-Gamma (interferon-gamma) was highly variable over the time course. These results suggest that significant stress responses occur as early as 2 h after blood draw and can influence gene expression data in real time RT-PCR quantitative gene expression experiments.

Discussion

Gene expression profiling in the blood cells of patients has the potential to advance our understanding of many human diseases, both by providing biomarkers to better diagnose and clinically manage these disorders, and by suggesting new therapeutic targets. Diseases characterized by inflammation and by activated immune responses are obvious targets for initial experiments, and indeed a number of groups, including our own, have initiated such studies.[3,5,18,19] In order for the information from peripheral blood cell gene expression to be optimally useful for translational clinical studies, it will be important to attempt to control as many variables as possible.

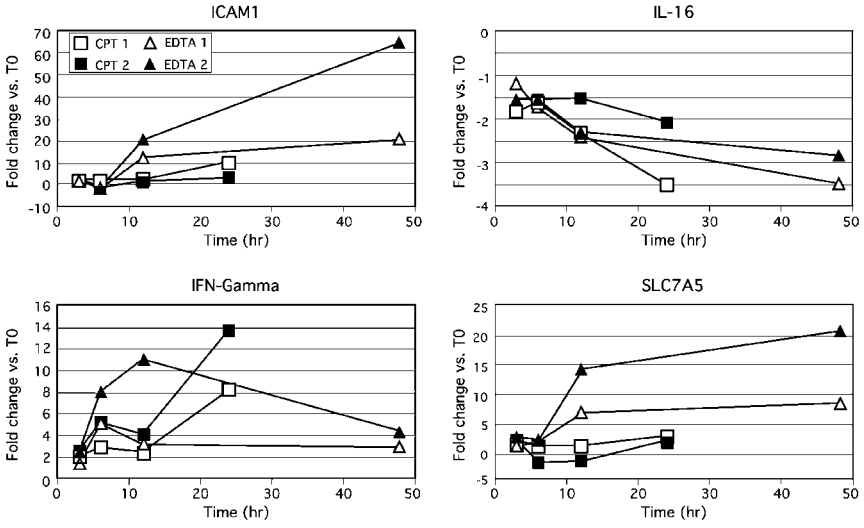

Figure 4 Gene expression changes in PBMCs and total WBCs are evident as early as 2 h post-draw. Blood was drawn into CPT and EDTA tubes and RNA was extracted after 0, 2, 6, 12, 24, and 48 h of incubation at room temperature. Data are presented as the fold change *vs* time zero (T0) for two control individuals. Transcript levels of representative genes were measured by real-time RT-PCR. ICAM1, intracellular adhesion molecule 1. SLC7A5, solute carrier family 7, member 5. IFN-Gamma, interferon-γ. IL16, interleukin 16. Data are representative of 17 genes examined. See supplementary materials for additional data.

Many factors are known to dynamically influence the numbers and function of peripheral blood cells and the relative proportions of cells at a given point in time. For instance, some disease states such as SLE are characterized by lymphopenia,[20] while infectious diseases are often characterized by changes in the percentages of WBC subsets in blood (eg increased neutrophils in bacterial diseases, increased lymphocytes in viral diseases). Since microarrays are sensitive to the percentages of cell types in a sample, rather than the raw cell number, one would ideally like to control for the composition of the cell subsets in blood at the time of the blood draw. We now routinely obtain a complete blood count (CBC) at the time of the blood draw, and additionally perform flow cytometric analysis on samples to quantify all major cell subsets. We are also working to develop analytic approaches to deal with the heterogeneity in cell populations that take cell subset information into account. Some key issues regarding normal variability in blood cell gene expression profiles have been recently addressed.[4]

The experiments described here identify a large number of genes that show significant changes in gene expression after overnight shipping of whole blood. The majority of these transcripts are present at lower levels in the samples that were shipped overnight, compared with fresh samples. This could reflect the onset of mRNA decay, resulting in fewer intact transcripts and thus a reduced level of expression relative to fresh samples. Alternatively, this may represent negative regulation of transcription as an active response to cellular stress. Our data also suggest that many of these same genes are dysregulated by much shorter *ex vivo* incubations of blood, as short as 1–3 h. We believe that this point underscores the extreme sensitivity of blood cells to *ex vivo* handling. As many of the genes upregulated by *ex vivo* handling of samples are also potentially involved in inflammatory responses, it is possible that genes may be falsely implicated in disease states. For this reason, we previously chose to exclude these '*ex vivo* stress' genes from our PBMC datasets when comparing SLE with control samples.[5] Recently, methods have been developed to stabilize peripheral blood cell mRNA immediately at the time of blood draw.[21] Our recent experience with the PAXgene tube system (Qiagen/Becton-Dickinson) (unpublished data) suggests that this method has significant potential to circumvent issues relating to changes in blood cell gene expression after prolonged incubation of blood *ex vivo*.

Materials and methods

Sample processing and chip hybridization

Following informed consent, approximately 30 ml of blood was drawn from normal donors into four Vacutainer CPT™ tubes (mononuclear cell preparation tubes, Becton-Dickinson, Franklin Lakes, NJ, USA). Two of the tubes were processed immediately, and the freshly isolated PBMCs were placed in RNA*later*®. These PBMCs, with their RNA preserved in RNA*later*®, and the two unprocessed tubes were sent overnight to the corresponding institution. PBMCs were then isolated from the unprocessed tubes and RNA was isolated from all samples using Trizol (Gibco-BRL, Invitrogen, Carlsbad, CA, USA), followed by RNeasy® cleanup (Qiagen, Valencia, CA, USA). In all, 5–10 μg of total RNA was used to prepare biotinylated cRNA for hybridization using a standard Affymetrix protocol (Expression Analysis Technical Manual, Affymetrix, Inc., 2000). A measure of 10 μg of each labeled cRNA were used for hybridization.

Data processing

After scanning the array, Affymetrix Microarray Suite (MAS) 4.0 software was used to generate expression values (referred to as an 'average difference', or AD) for each gene. Raw data for the 6414 genes expressed in PBMCs are provided in Web Table A (see Supplementary Information). Each chip was scaled to an overall intensity of 1500 to correct for minor differences in overall chip hybridization intensity, and to allow comparison between chips. A threshold of 20 AD units was assigned to any gene that was called 'Absent' by MAS. Furthermore, any gene with an AD less than 20 was also assigned this threshold. Processed microarray data are provided in Web Table B (see Supplementary Information).

Comparison analyses and hierarchical clustering

The individual gene expression levels of fresh samples and overnight-shipped samples were compared using a paired Student's *t*-test. We identified 2034 genes that exhibited significant differences in gene expression between the two groups ($P < 0.01$). In total, 97% of these genes were also called significant by SAM,[10] which identified a total of 2254 genes as differentially expressed ($\delta = 2.0$, median number of false positives = 29.7). Hierarchical clustering was performed using Cluster and visualized using TreeView.[22] Before clustering, the expression level for each gene was divided by the mean expression level of the fresh samples. The $\log_2$ of this ratio was used for input into the clustering program.

Kinetic RT-PCR

Whole blood from each donor was collected into five CPT™ tubes (Becton-Dickinson), and the tubes were inverted eight times immediately after collection. For 0 h samples, the peripheral mononuclear cells (PBMCs) were prepared immediately (see below). For other samples, whole blood was stored in CPT™ tubes after collection at room temperature for 3, 6, 12 and 24 h with gentle shaking until processed. CPT™ tubes were centrifuged at 1500–1800 RCF for 30 min at room temperature, and the plasma and PBMC layers were mixed and poured into a 15 ml Falcon tube. PBMCs were washed twice with phosphate-buffered saline (PBS) and resuspended in 600 μl of RLT buffer (RNeasy® kit, Qiagen). RNA was then extracted using the RNeasy® mini kit with an on-column DNase treatment (30 min). Total RNA was quantified using the RiboGreen® Quantitation assay (Molecular Probes, Eugene, OR, USA), and the integrity of the RNA was confirmed by agarose gel electrophoresis and subsequent staining with SYBR® Gold (Molecular Probes). In all, 5 ng of total RNA was used in each RT-PCR reaction on the ABI Prism® 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA, USA). PCR conditions were described previously.[23] The data from this experiment are available in Web Table C (see Supplementary Information).

Acknowledgements

This study was financially supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases and the Alliance for Lupus Research.

References

1 Ramaswamy S, Golub TR. DNA microarrays in clinical oncology. *J Clin Oncol* 2002; **20**: 1932–1941.
2 Staudt LM. Gene expression profiling of lymphoid malignancies. *Annu Rev Med* 2002; **53**: 303–318.
3 Maas K, Chan J, Parker J *et al*. Cutting edge: molecular portrait of human autoimmune disease. *J Immunol* 2002; **169**: 5–9.
4 Whitney AR, Diehn M, Popper SJ *et al*. Individuality and variation in gene expression patterns in human blood. *Proc Natl Acad Sci USA* 2003; **100**: 1896–1901.
5 Baechler EC, Batliwalla FM, Karypis G *et al*. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc Natl Acad Sci USA* 2003; **100**: 2610–2615.
6 Bennett L, Palucka AK, Arce E *et al*. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J Exp Med* 2003; **197**: 711–723.
7 Gregersen PK, Behrens TW. Fine mapping the phenotype in autoimmune disease: the promise and pitfalls of DNA microarray technologies. *Genes Immunity* 2003; **4**: 175–176.
8 Harris EDJ. Rheumatoid arthritis. Pathophysiology and implications for therapy. *N Engl J Med* 1990; **322**: 1277–1289.
9 Wallace DJ. *Dubois' Lupus Erythematosus*. Williams & Wilkins: Philadelphia, 2002.
10 Tusher V, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 2001; **98**: 5116–5121.
11 Bukh A, Martinez-Valdez H, Freedman S, Freedman M, Cohen A. The expression of c-fos, c-jun, and c-myc genes is regulated by heat shock in human lymphoid cells. *J Immunol* 1990; **144**: 4835–4840.
12 Keyse SM, Emslie EA. Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine phosphatase. *Nature* 1992; **359**: 644–647.
13 Shweiki D, Neeman M, Itin A, Keshet E. Induction of vascular endothelial growth factor expression by hypoxia and by glucose deficiency in multicell spheroids: implications for tumor angiogenesis. *Proc Natl Acad Sci USA* 1995; **92**: 768–772.
14 Green DR. Death and NF-kappaB in T cell activation: life at the edge. *Mol Cell* 2003; **11**: 551–552.
15 Reed SI. Ratchets and clocks: the cell cycle, ubiquitylation and protein turnover. *Nat Rev Mol Cell Biol* 2003; **4**: 855–864.
16 Watanabe-Fukunaga R, Brannan CI, Copeland NG, Jenkins NA, Nagata S. Lymphoproliferation disorder in mice explained by defects in Fas antigen that mediates apoptosis. *Nature* 1992; **356**: 314–317.
17 Wu J, Wilson J, He J, Xiang L, Schur PH, Mountz JD. Fas ligand mutation in a patient with systemic lupus erythematosus and lymphoproliferative disease. *J Clin Invest* 1996; **98**: 1107–1113.
18 Gu J, Marker-Hermann E, Baeten D *et al*. A 588-gene microarray analysis of the peripheral blood mononuclear cells of spondyloarthropathy patients. *Rheumatology (Oxford)* 2002; **41**: 759–766.
19 Rus V, Atamas SP, Shustova V *et al*. Expression of cytokine- and chemokine-related genes in peripheral blood mononuclear cells from lupus patients by cDNA array. *Clin Immunol* 2002; **102**: 283–290.
20 Scheinberg M, Cathcart E. B cell and T cell lymphopenia in systemic lupus erythematosus. *Cell Immunol* 1974; **12**: 309–314.
21 Rainen L, Oelmueller U, Jurgensen S *et al*. Stabilization of mRNA expression in whole blood samples. *Clin Chem* 2002; **48**: 1883–1890.
22 Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 1998; **95**: 14863–14868.
23 Rogge L, Bianchi F, Biffi M *et al*. Transcript imaging of the development of human T helper cells using oligonucleotide arrays. *Nat Genet* 2000; **25**: 96–101.

Supplementary Information accompanies the paper on Genes and Immunity's website (http://www.nature.com/gene)