Transcriptional Response of Peripheral Lymphocytes to Early Fibrosarcoma: A Model System for Cancer Detection Based on Hybridization Signatures

Márcia M. C. Marques,* Cristina M. Junta,* Carlos R. Zárate-Blades,† Elza Tiemi Sakamoto-Hojo,‡ Eduardo A. Donadi,§ and Geraldo A. S. Passos*,||,[1]

*Molecular Immunogenetics Group, Department of Genetics, Faculty of Medicine of Ribeirão Preto, University of São Paulo (USP), 14040-900, Ribeirão Preto, SP, Brazil; †Department of Biochemistry and Immunology, Faculty of Medicine of Ribeirão Preto, USP, 14040-900, Ribeirão Preto, SP, Brazil; ‡Department of Biology (FFCLRP), USP, 14040-900, Ribeirão Preto, SP, Brazil; §Department of Medicine, Division of Clinical Immunology, Faculty of Medicine Ribeirão Preto, USP, 14040-900, Ribeirão Preto, SP, Brazil; and ||Department of Morphology (DMEF), Faculty of Dentistry of Ribeirão Preto, USP, 14040-904, Ribeirão Preto, SP, Brazil

Since circulating leukocytes, mainly B and T cells, continuously maintain vigilant and comprehensive immune surveillance, these cells could be used as reporters for signs of infection or other pathologies, including cancer. Activated lymphocyte clones trigger a sensitive transcriptional response, which could be identified by gene expression profiling. To assess this hypothesis, we conducted microarray analysis of the gene expression profile of lymphocytes isolated from immunocompetent BALB/c mice subcutaneously injected with different numbers of tumorigenic B61 fibrosarcoma cells. Flow cytometry demonstrated that the number of circulating T ($CD3^+CD4^+$ or $CD3^+CD8^+$) or B ($CD19^+$) cells did not change. However, the lymphocytes isolated from tumor cell–injected animals expressed a unique transcriptional profile that was identifiable before the development of a palpable tumor mass. This finding demonstrates that the transcriptional response appears before alterations in the main lymphocyte subsets and that the gene expression profile of peripheral lymphocytes can serve as a sensitive and accurate method for the early detection of cancer.
Exp Biol Med 234:802–812, 2009

Key words: immune surveillance; early cancer detection; lymphocyte gene profiling; hybridization signatures; cDNA microarray; fibrosarcoma

Introduction

The early detection of tumor cells *in vivo* could significantly enhance the success of cancer treatments. Peripheral blood lymphocytes (B and T cells), which have a cell surface receptor–based recognition system (surface immunoglobulins [Ig] and the T-cell receptors [TCR], respectively), represent a system naturally engaged in this detection through immunologic surveillance. The recognition of infectious agents or cancer cells by the immune system may trigger different patterns of gene expression that are associated with the nature and site of the insult. This possibility has been referred to as a "pathognomonic gene expression signature" that could, therefore, be used to diagnose hidden or nascent diseases.

B and T cells have sensitive signal transduction systems that are triggered by surface receptors that recognize free antigen (surface Ig on B cells) or major histocompatibility complex (MHC)-antigenic peptide complexes (the TCR on T cells) and allow lymphocytes to respond to injury at any site in the body (1–3). The recognition of cancer cells by lymphocytes may result in a different pattern of gene expression that can be associated with the nature and site of an early tumor (4). Our group has pursued the use of blood lymphocytes as diagnostic reporters of disease, including

MMCM was recipient of a FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil) fellowship. This project was funded by FAPESP and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil).

[1] To whom correspondence should be addressed at Molecular Immunogenetics Group, Department of Genetics, Faculty of Medicine of Ribeirão Preto, University of São Paulo (USP), 14040-900, Ribeirão Preto, SP, Brazil.
E-mail: passos@rge.fmrp.usp.br

Received October 29, 2008.
Accepted March 10, 2009.

DOI: 10.3181/0810-RM-311
1535-3702/09/2347-0802$15.00

autoimmune diseases in man and experimental cancer in mice, through gene expression profiling (5–7).

Microarray technology is the ideal method for the determination of gene expression profiles because of its sensitivity and robustness, and it allows for the measurement of thousands of genes in a single array. Moreover, the bioinformatic analyses of microarrays that are currently available allow for the accurate hierarchical organization of experimental samples and genes.

Our aim in this study was to evaluate the gene expression profiles in peripheral lymphocytes in response to a small number of fibrosarcoma cells in immunocompetent BALB/c mice before the development of a visible tumor mass. We used an *in vivo* model system in which BALB/c mice received subcutaneously injections of transformed B61 cells (3T3 fibroblasts transfected with a mutant p21 *Ha-ras* oncogene) (8, 9), which, over time, develop into fibrosarcoma at the site of injection (5).

For the gene expression profiling, we used a glass slide cDNA microarray containing 4500 sequences that represent the main biological functions of the murine immune system, including interleukins and their receptors; genes involved in cell cycle control, metabolism, signal transduction, DNA repair/recombination and apoptosis; and expressed sequence tags (ESTs) of unknown function.

Using a significance analysis of microarrays (SAM) algorithm (10) and ANOVA (11), we were able to statistically validate the microarray data and to identify genes that were induced or repressed. Hierarchical clustering was useful in comparing the different hybridization signatures observed between the mice subjected to bacterial infection or inflammation and the mice injected with B61 fibrosarcoma cells.

By comparing the flow cytometric analysis of peripheral blood lymphocytes with the microarray data, we demonstrate that specific gene expression signatures could be associated with each treatment group, despite the absence of changes in the number of T ($CD3^+CD4^+$ or $CD3^+CD8^+$) and B ($CD19^+$) cells. These data demonstrate that the blood lymphocyte transcriptional response, as determined by the microarray assay, is a sensitive and specific method for the early detection of tumor cells *in vivo*.

Materials and Methods

Ha-ras-1–Transformed B61 Cell Line. Mouse BALB/3T3 fibroblasts were transfected with a plasmid expressing a mutated human c-*Ha-ras-1* oncogene (pEJ, which expresses the p21 ras protein with a G→V substitution at position 12). This transformed cell line, designated B61, contains 50–100 copies of the active mutant c-*Ha-ras-1* oncogene and was generated by Kovary and colleagues (8, 9). The B61 adherent cells were cultured in DEMF10 medium supplemented with 10% fetal bovine serum plus antibiotics in an incubator with 5% CO_2 at 37°C for 1 week. The cultures were trypsinized and supplemented with fresh medium at 3-day intervals.

Injection of B61 Fibrosarcoma Cells into Mice, Peripheral Lymphocyte Separation, and Extraction of Total RNA. We injected 0.1 ml of sterile nonpyrogenic saline containing 10^3 or 10^5 B61 cells subcutaneously (sc) in the scapular region of 4- to 6-week-old female immunocompetent BALB/c mice.

Three days after injection, before the emergence of any tumor mass at the site of injection, peripheral blood was collected (approximately 0.9 ml per animal) and pooled for a total of 10 ml. The mononuclear cells were separated by centrifugation on a Ficoll-Hypaque (GE Healthcare, Uppsala, Sweden) cushion, resuspended in DEM medium, and incubated at 37°C in a tissue culture plate (Corning, Monterey, Mexico). The nonadherent cells, consisting mainly of lymphocytes, were collected, and total RNA was isolated by using the Trizol reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. All RNA samples were extracted 3 days after tumor inoculation. The integrity of the RNA samples was evaluated by conventional agarose gel electrophoresis, and only those preparations free of contaminating DNA, proteins, or phenol were used (data not shown). The gene expression experiments were independently repeated three times, as required by the SAM program (10), thereby reducing the false-discovery rate. All manipulations of mice were approved by the Animal Research Ethical Committee of the University of São Paulo, Campus of Ribeirão Preto, Brazil.

Bacterial Infection, Inflammation, and Control Groups. Total RNA was also extracted from the peripheral lymphocytes isolated from groups of BALB/c mice subjected to bacterial infection or inflammation. Mice in the bacterial infection group received an injection given sc of a culture of total bacteria from mouse feces (10^7 colony forming units [cfu] in 0.1 ml sterile apyrogenic saline per animal), and mice in the inflammation group received an injection of Zymozan (Sigma, St. Louis, MO) (500 μg in 0.1 ml sterile apyrogenic saline per animal) given sc. Mice that received an. injection of 0.1 ml sterile apyrogenic saline given sc served as controls. As described for animals injected with tumors cells, all RNA samples were extracted 3 days after injection.

Flow Cytometry Analysis. Peripheral blood mononuclear cells (1×10^7 cells per milliliter) from mice in each group were incubated for 40 mins at 4°C with Fc block (1 μg per 10^6 cells; Pharmingen, San Diego, CA). Cells were incubated with the appropriate monoclonal antibody (0.75 μg per 10^6 cells) for 30 mins at 4°C in the dark. The following anti–mouse cell surface marker antibodies were used: (i) PE-conjugated anti-CD8 (clone 53–6.7), anti-CD19 (clone 1D3), and anti-IgG2a (G155–178); and (ii) FITC-labeled anti-CD3 (clone 145–2C11), anti-CD4 (RM4–5), and anti-IgG2a (clone R35–95). All antibodies were purchased from Pharmingen and used according to the

manufacturer's instructions. The cells were analyzed with a FACsVantage flow cytometer (Becton Dickinson, Foster City, CA) using the FACSort software (Becton Dickinson). The cells were defined according to size (forward scatter), granularity (side scatter), and fluorescence intensity. The number of $CD3^+$, $CD4^+$, $CD8^+$, and $CD19^+$ cells was then determined. The flow cytometry experiments were independently repeated 10 times.

cDNA Microarray Method. The gene expression of lymphocytes was assessed by using glass slide cDNA microarrays according to standard protocols (12). The arrays were prepared on silane-coated UltraGAPS slides (#40015, Corning, New York, NY) containing a total of 4500 target cDNA sequences. These sequences were from the Soares thymus 2NbMT normalized library, represented EST cDNA clones prepared from the thymus of a 4-week-old male C57BL/6J mouse, and are available at the IMAGE Consortium (http://image.llnl.gov/image/html/iresources.shtml). The cDNA inserts were homogeneous in size (near 1 kb), cloned into three vectors (pT7T3D, pBluescript, and Lafmid), and amplified in 384- or 96-well plates by using vector-PCR amplification with the following primers that are specific for the three vectors: LBP 1S GTGGAATTGTGAGCGGATACC forward and LBP 1AS GCAAGGCGATTAAGTTGG reverse.

The microarrays were prepared by using PCR products from the cDNA clones and a Generation III Array Spotter (Amersham Molecular Dynamics, Sunnyvale, CA).

Complex cDNA Probe Preparation and Hybridization. The cDNA complex probes derived from the total RNA isolated from lymphocytes of each group were prepared by reverse transcription using 10 μg of total RNA and were labeled with Cy3 fluorochrome using the CyScribe postlabeling kit (GE Healthcare). The hybridization took place for 15 hrs and was followed by washing with an automatic slide processor system (ASP, Amersham Biosciences). The washed microarrays were then scanned with a Generation III laser scanner (Amersham Biosciences). As a reference for the hybridization procedure, we used equimolar quantities of cDNA generated from unrelated total RNA (mouse thymus total RNA). This approach allowed us to estimate the amount of cDNA target in each microarray spot.

A complete file providing all genes and ESTs present in the microarrays used in this study is available online (www.rge.fmrp.usp.br/passos/mmu_array).

cDNA Microarray Data Analysis. Microarray image quantification was performed by using Spotfinder software (http://www.tm4.org/spotfinder.html). The normalization process was carried out by using the R platform (http://www.r-project.org), and the statistical data were analyzed by the Multiexperiment Viewer (MeV) software (version 3.1; available online at http://www.tm4.org/mev.html) (13).

To analyze the gene expression patterns, we used an unsupervised hierarchical clustering method that grouped genes on the vertical axis and samples on the horizontal axis on the basis of similarity in their expression profiles. The similarities and dissimilarities in gene expression are presented as dendrograms (in which the pattern and length of the branches reflect the relatedness of the samples or genes) and as heat maps that were generated by Cluster version 3.0 and Java Tree View (http://rana.lbl.gov/EisenSoftware.htm).

Oligonucleotide Primer Design and Quantitative Real-Time Polymerase Chain Reaction (PCR) for Microarray Data Confirmation. Two genes that were induced and two that were repressed in the mice given injections of 10^5 B61 cells were selected on the basis of the hierarchical clustering–based expression pattern. The cDNA sequences of the selected genes were retrieved from the National Center for Biotechnology Information GenBank database (http://www.ncbi.nlm.nih.gov) by using the following accession numbers: *Ada*, NM_007398.3; *Cdk4*, NM_009870; and *Parp3*, NM_145619.2. The Primer3 web tool (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) was used to select pairs of oligonucleotide primers that spanned an intron and exon junction and had an optimal melting temperature of 60°C. The following pairs of primers were used: *Ada* (3′ CCTATGAGGGCGCAGTAAAG 5′ and 3′ TAACCATGTCCCACCCTTC 5′), *Cdk4* (3′ CAATGTTGTACGGCTGATGG 5′ and 3′ GGAGGTGCTTTGTCCAGGTA 5′), and *Parp3* (3′ CCTGCTGATAATCGGGTCAT 5′ and 3′ TTGTTGTTGCCGATGTT 5′). The cDNA samples were prepared by using the SuperScript II enzyme as recommended by the manufacturer (Invitrogen). The expression of the genes mentioned above was quantified by a 7500 Real-Time PCR System (Applied Biosystems). The expression was normalized to the expression of the housekeeping gene *Gapdh* (NM_008084; 3′ GGGTGTGAACCACGAGAAAT 5′ and 3′ CCTTCCACAATGCCAAAGTT 5′). The real-time PCR experiments were independently repeated 10 times.

Data Mining. Data mining, including the gene ontology, of the differentially expressed genes (induced or repressed) that characterized each group of mice was performed by using the SOURCE (http://source.stanford.edu/cgi-bin/source/sourceSearch) and DAVID (http://apps1.niaid.nih.gov/david) databases.

Results

Peripheral Lymphocyte Subsets. As shown in Figure 1, there were no significant changes in the T ($CD3^+CD4^+$ or $CD3^+CD8^+$) and B ($CD19^+$) cell subsets among mice given sc injections of Zymozan (inflammation), the bacterial suspension (infection), or B61 fibrosarcoma cells (10^3 or 10^5 cells per animal) and those in the control group, which was given injections of only sterile saline. All analyses were carried out 3 days after the respective injections. This time point was long enough to elicit local inflammation in the Zymozan-injected mice and peritoneal

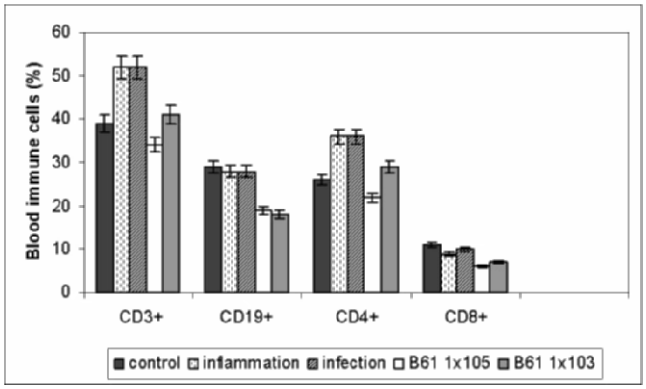

Figure 1. Flow cytometric analysis to determine the number of $CD3^+$, $CD4^+$, $CD8^+$, and $CD19^+$ cells in the blood of mice given sc injections of different numbers of B61 tumor cells and the control groups. No significant changes were detected (ANOVA, $P \leq 0.05$).

infection in the infection mice, but no palpable or biopsy-identifiable tumor was found at the site of injection in the B61 cell–injected mice.

Peripheral Lymphocyte Transcriptional Response. From the set of the 4500 cDNA sequences present in the microarray used to determine the differential gene expression profile, we selected only those sequences that were present in 80% of the microarray hybridizations ($n = 3600$) for statistical analysis. To determine the variation in the gene expression in the full data set and to identify genes that were differentially expressed by the different groups of mice, we used two different statistical methods: a SAM algorithm (10) and a one-way ANOVA (11).

Using the multiclass SAM algorithm with a 5% false-discovery rate, we identified 425 genes that were significantly and differentially expressed (Fig. 2A). The resulting data were further analyzed by using an unsupervised hierarchical clustering method, which discriminated the groups of mice on the basis of their respective expression signatures. The first SAM sample cluster included the control group (saline-injected mice) and the group of mice given injections of 10^3 B61 tumor cells. The second SAM sample cluster included the infection and inflammation groups, which were positioned together because of their similar expression signatures. Finally, the third SAM sample cluster was the only exclusive cluster and contained the group of mice given injections of 10^5 B61 tumor cells.

Using the ANOVA approach (bootstrapping with 5000 permutations), we identified 711 informative genes that exhibited $P < 0.05$ (Fig. 2B). On the basis of this analysis, we again used the unsupervised hierarchical clustering method to analyze the data.

Similar to the results obtained with the SAM data, we obtained three sample clusters. The first ANOVA sample cluster included the control group (saline-injected mice) and the group given injections of 10^3 B61 tumor cells, which clustered together. The second ANOVA sample cluster included the infection and inflammation groups. Although these two groups displayed disparate expression signatures, the dissimilarities were not discriminatory, and the groups still clustered together. Finally, the group given injections of 10^5 B61 tumor cells comprised the third ANOVA sample cluster. The animals from this group featured an expression signature significantly different from those of all of the other groups, forming a distinct cluster.

Overall, both hierarchical clustering analyses positioned the groups of mice into three large clusters (Fig. 2A and B), and both statistical approaches demonstrated that gene expression of peripheral lymphocytes was different according to the treatment. Although the number of genes was distinct for each statistical approach, both analyses yielded similar results, further emphasizing the relevance of the data.

A one-way ANOVA was applied to evaluate the variation in the gene expression in the full data set. Bootstrapping from 1000 permutations and a significance level of $P < 0.05$ were used to select 711 genes. The data from these 711 genes were then analyzed by using unsupervised clustering. Using the same full data set and the SAM algorithm (false-discovery rate < 0.05 and 1000 bootstrapped permutations), we found 425 significantly and differentially expressed genes. Interestingly, 246 significantly and differentially expressed genes as determined by the SAM algorithm were included among the 711 informative genes as determined by the ANOVA analysis. This result statistically validates the microarray data. A file providing the entire quantitative normalized microarray data set obtained in this study is available online (http://rge.fmrp.usp.br/passos/earlyB61/dataset).

Gene Expression Assessed by Real-Time PCR. The *Ada*, *Cdk4*, and *Parp3* genes were selected to confirm the results obtained by the cDNA microarray method. These genes were chosen on the basis of their statistical significance determined by the SAM algorithm. The real-time PCR analysis confirmed that these genes were induced (*Ada* and *Cdk4*) or repressed (*Parp3*) in the group of mice in which 10^5 B61 tumor cells were injected (Fig. 3).

Two major conclusions may be drawn from these results. First, the different treatments produced specific expression signatures, as detected by two distinct hierarchical clustering analyses. Second, at least 246 differentially expressed genes were detected and validated by two distinct statistical approaches. Some of these validated genes, which could be informative for the presence of the B61 cells, were chosen for discussion (Table 1).

Discussion

There is great interest in uncovering new methods for accurate detection of early cancer in asymptomatic patients. To illustrate this interest, Sharma *et al.* (14) considered breast cancer, where mammographic screening is the most consistent method to detect tumors. Although effective, it has important limitations. For example, in absence of

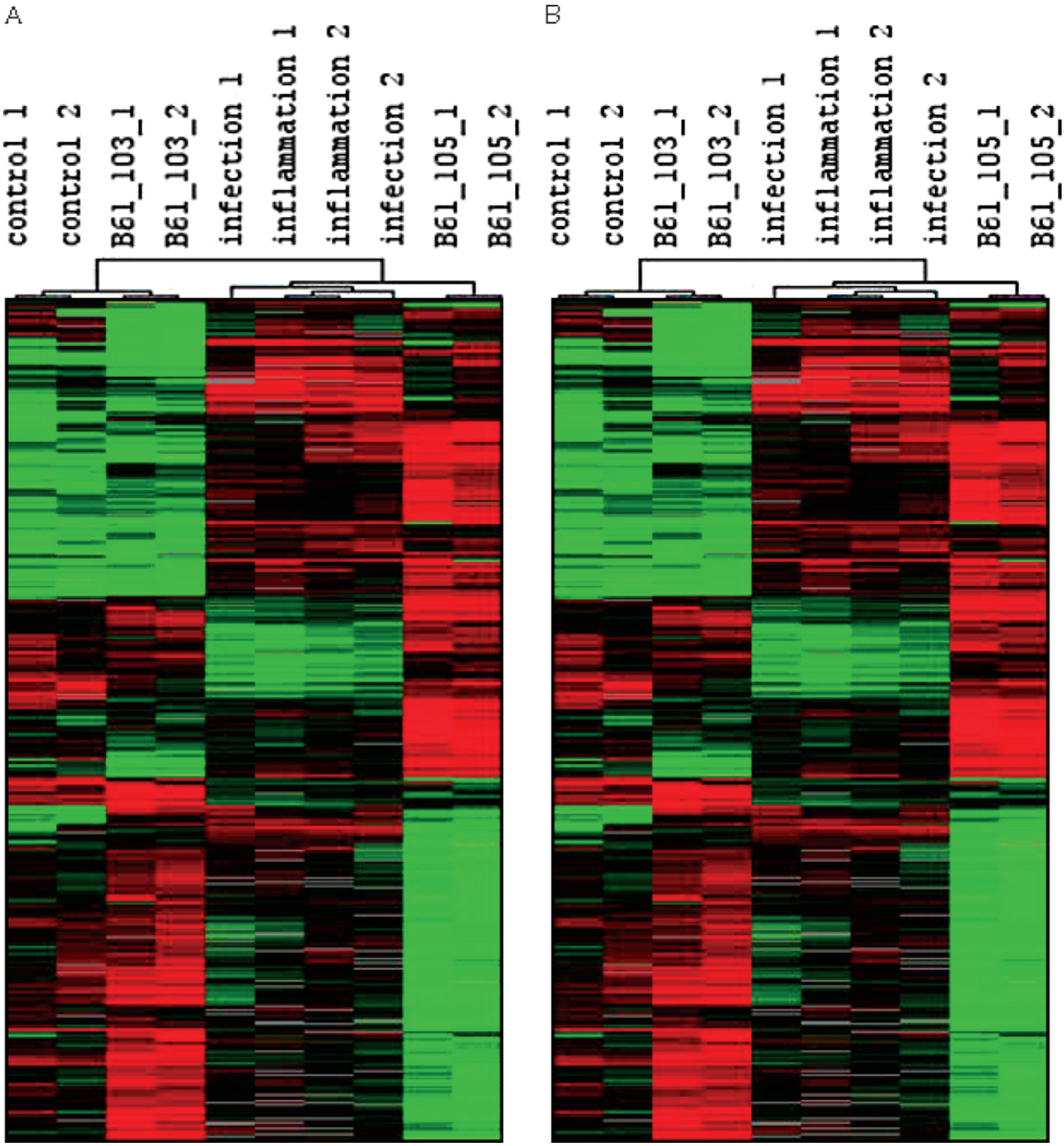

Figure 2. Dendrograms of samples constructed on the basis of the ANOVA (A) or the SAM algorithm (B) confirming the separation of the different groups of mice given injections of different substances (B61 tumor cells, Zymozan [inflammation] or bacterial suspension [infection]) and control mice on the basis of gene expression signatures.

microcalcification, mammography often fails to detect tumors that are less than 5 mm in size. Therefore, these authors developed a method for breast cancer detection based on microarray gene profiling of peripheral blood cells in asymptomatic individuals.

Most published results of cancer diagnosis/prognosis based on gene expression have involved clinical samples (15). We agree with the authors who believe that obtaining samples for clinical purposes requires a prior knowledge of both their presence and their location in the body. Therefore, a gene expression-based assay not requiring samples from the diseased area remains a focus of research.

We previously demonstrated that a primary immune system organ such as the thymus exhibits a differential

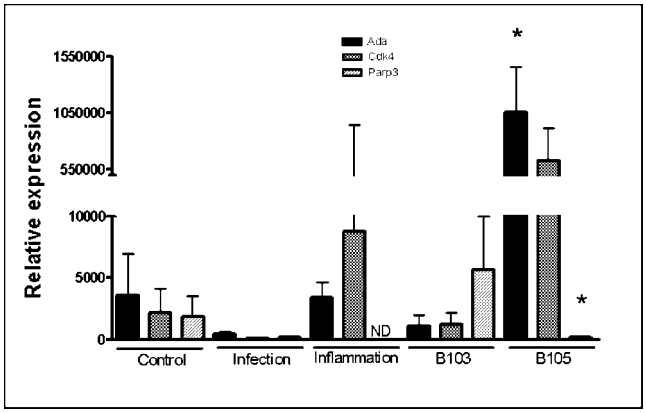

Figure 3. Quantitative real-time PCR was used to confirm the induction of *Ada* and *Cdk4* and the repression of *Parp3* in the mice into which 10^5 B61 cells were injected. The expression was compared with that in the B61 10^3, inflammation, infection, and control groups of mice. The expression levels were normalized to *Gapdh* expression ($n = 10$; mean ± standard error of the mean; one-way ANOVA, * $P < 0.001$).

transcriptional response in the presence of fibrosarcoma developing at a distant site in the body (5). This observation led us to search for an assay based on the gene expression profile of immune system cells for the detection of cancer.

Because B61 tumor cells require at least 6 days to form a palpable tumor mass at the site of injection (5), we chose 3 days after the injection of tumor cells as a single time point to analyze the gene expression profile. This time point assured the presence of tumor cells in the animals in absence of a tumor mass.

Circulating blood cells, mainly lymphocytes, represent a logical choice of cells to investigate because of their vigilant and comprehensive surveillance for signs of infection or other threats, including cancer (16). In addition, blood is easily obtained in a noninvasive manner.

Interestingly, we have observed differential gene expression profiles as early as 3 days after injection of tumor cells. T-cell responses often take 7 days to develop, whereas innate immune responses develop faster but involve other cells. Therefore, the expression profile observed may be implicated in the molecular mechanisms of such cells before the cellular response.

A recent study using a proteomic approach based on plasma protein fractionation demonstrated that an in-depth proteomic analysis may provide a useful strategy for early cancer detection in a genetically engineered mouse model and in patients with pancreatic cancer (17). In the present study, we demonstrate a simpler strategy based on microarray transcriptome profiling, which may be of more practical use. Our aim was to define specific hybridization signatures that characterized mice injected with fibrosarcoma tumor cells. The infection and inflammation groups were also assayed to determine the specificity of the observed hybridization signature. Because different tumors produce particular antigens, the lymphocyte gene expression profile in response to different tumor cell lines, although not included in this study, needs be investigated.

Another interest was to investigate whether at an early time point there were alterations in the frequency of the main sets of circulating lymphocytes, which could possibly explain the alterations in gene expression. Figure 1 demonstrates that the number of T ($CD3^+CD4^+$ or $CD3^+CD8^+$) and B ($CD19^+$) cells did not significantly change between the control, inflammation, infection, and tumor cell-injected groups of mice.

The unsupervised analysis of the differentially expressed genes allowed for the detection of specific hybridization signatures and, consequently, hierarchical clustering of these groups. These findings strongly suggest that the differential hybridization signatures observed is a result of the early modulation of gene expression in peripheral lymphocytes.

The SAM algorithm (10), which is based in the classic *t* test, was specifically adapted for the high-throughput analysis of microarray data and is widely used for the identification of differentially expressed genes. ANOVA (11) was used in parallel as a second method to validate the results obtained by the SAM algorithm (Fig. 2).

Although the microarray used in this study is not widely used by researchers because it was prepared in our laboratory, it identified genes that were expressed at significantly different levels. Furthermore, the data from this array were statistically and experimentally validated by traditional methods such as the SAM algorithm (10) and quantitative real-time PCR.

Figure 2 demonstrates that the group of mice into which 10^3 B61 tumor cells were injected clustered together with the control group, which indicates a similar hybridization signature. This result suggests that 10^3 B61 tumor cells did not elicit a specific gene response in peripheral lymphocytes.

The infection and inflammation groups of mice were positioned in the same cluster and exhibited similar hybridization signatures. Although these two groups were tested to better define the specificity of the hybridization signatures, a general conclusion may be drawn from this result: host responses to bacterial infection or inflammatory stimuli share common mechanisms and/or effector immune cells that modulate a common set of genes. Table 1 lists the genes identified by both statistical approaches for each group of mice.

Finally, we were able to identify a set of genes that were exclusively modulated in the lymphocytes of mice injected with 10^5 B61 tumor cells, which were characterized by their specific hybridization signature (Fig. 2A and B). On the basis of the statistical significance of their expression profile, some genes were selected for discussion.

Among the induced genes, we have highlighted *Rps11* (ribosomal protein S11, accession number NM_013725), which encodes a protein constituent of the ribosome, and *Ppp2r5c* (protein phosphatase 2, regulatory subunit B

Table 1. Genes Differentially and Significantly Expressed in Peripheral Blood Lymphocytes of BALB/c Mice at an Early Stage of Fibrosarcoma, as Detected by SAM or ANOVA Statistical Tests ($P \leq 0.05$)

Gene name	Symbol	Clone ID	Biological function
Reticulon 4	Rtn4	640358	Angiogenesis
B-cell leukemia/lymphoma 6	Bcl6	573624	Apoptosis
BH3 interacting domain death agonist	Bid	583029	Apoptosis
BH3 interacting domain death agonist	Bid	640130	Apoptosis
Caspase 2	Casp2	573760	Apoptosis
Caspase 8, apoptosis-related cysteine peptidase	CASP8	5228776	Apoptosis
Ganglioside-induced differentiation-associated-protein 2	Gdap2	641674	Apoptosis
Nuclear factor of kappa light chain gene enhancer in B-cells 1, p105	Nfkb1	575033	Apoptosis
Purinergic receptor P2X, ligand-gated ion channel, 1	P2rx1	576538	Apoptosis
Transcription factor 7, T-cell specific	Tcf7	576200	Apoptosis
DEAD (Asp-Glu-Ala-Asp) box polypeptide 51	Ddx51	639801	ATP binding
DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 3, X-linked	Ddx3x	639697	ATP binding
DEAH (Asp-Glu-Ala-His) box polypeptide 38	Dhx38	582814	ATP binding
RAB27A, member RAS oncogene family	Rab27a	640750	Blood coagulation
RAB27A, member RAS oncogene family	Rab27a	640294	Blood coagulation
Plakophilin 1	Pkp1	583760	Cell adhesion
Cyclin-dependent kinase 4	Cdk4	640338	Cell cycle control
Dual specificity phosphatase 1	Dusp1	583905	Cell cycle control
E2F transcription factor 2	E2f2	640754	Cell cycle control
ELK4, member of ETS oncogene family	Elk4	583566	Cell cycle control
Microtubule-associated protein, RP/EB family, member 2	Mapre2	572898	Cell cycle control
Multiple endocrine neoplasia 1	Men1	576049	Cell cycle control
Protein phosphatase 2	Ppp2ca	583384	Cell cycle control
Sema domain, immunoglobulin domain (Ig), and GPI membrane anchor, (semaphorin) 7[a]	Sema7a	583574	Cell differentiation
B-cell translocation gene 1, anti-proliferative	Btg1	574654	Cell proliferation
Follistatin-like 1	Fstl1	583623	Cell proliferation
Protein kinase, cAMP dependent regulatory, type II beta	Prkar2b	640704	Cell proliferation
Integrin, alpha 5	ITGA5	135671	Cell-matrix adhesion
Centromere protein H	Cenph	640013	Cellular division
RAS-related C3 botulinum substrate 2	Rac2	640180	Chemotaxis
Histone cluster 1, H2bc	Hist1h2bc	640662	Chromosome organization
Moesin	Msn	583860	Cytoskeletal protein binding
Myosin, heavy polypeptide 11	Myh11	583021	Cytoskeleton organization
Annexin A4	Anxa4	583430	Development
Nucleoporin 50	Nup50	639698	Development
High mobility group nucleosomal binding domain 2	Hmgn2	640742	DNA packaging
Excision repair cross-complementing rodent repair deficiency, complementation group 5	ERCC5	1308118	DNA repair
Fanconi anemia, complementation group G	FANCG	1705194	DNA repair
Meiotic recombination 11 homolog A	Mre11A	639936	DNA repair
Poly (ADP-ribose) polymerase family, member 3	Parp3	640154	DNA repair
PRP19/PSO4 pre-mRNA processing factor 19 homolog	Prpf19	583380	DNA repair
Serine (or cysteine) peptidase inhibitor, clade B, member 2	Serpinb2	573970	Endopeptidase inhibitor activity
EST	EST	583835	Function unknown
CUE domain containing 2	Cuedc2	640434	Function unknown
DNA segment, Chr 14, ERATO Doi 449, expressed	D14Ertd449e	640010	Function unknown
DNA segment, Chr 4, Wayne State University 53, expressed	D4Wsu53e	583358	Function unknown
DNA segment, Chr 4, Wayne State University 53, expressed	D4Wsu53e	583830	Function unknown
EST	EST	583622	Function unknown
Inositol polyphosphate-5-phosphatase F	Inpp5f	640224	Function unknown
Jumonji domain containing 3	Jmjd3	583712	Function unknown
Melanocyte proliferating gene 1	Myg1	583514	Function unknown
Melanocyte proliferating gene 1	Myg1	640034	Function unknown
Metadherin	Mtdh	583311	Function unknown
EST	EST	582828	Function unknown
Ribosomal protein S25	Rps25	583911	Function unknown
EST	EST	639950	Function unknown
EST	EST	639771	Function unknown
EST	EST	583885	Function unknown
EST	EST	640297	Function unknown
EST	EST	583789	Function unknown
EST	EST	640044	Function unknown
EST	EST	640678	Function unknown
EST	EST	576048	Function unknown

Table 1. (Continued)

Gene name	Symbol	Clone ID	Biological function
EST	EST	641854	Function unknown
EST	EST	583861	Function unknown
EST	EST	583020	Function unknown
EST	EST	640356	Function unknown
EST	EST	583600	Function unknown
EST	EST	640875	Function unknown
EST	EST	640249	Function unknown
EST	EST	639701	Function unknown
Sex comb on midleg-like 4	Scml4	583715	Function unknown
Spastic paraplegia 20	Spg20	573449	Function unknown
Transcription factor 25	Tcf25	574026	Function unknown
EST	EST	575372	Function unknown
EST	EST	582865	Function unknown
EST	EST	582888	Function unknown
EST	EST	582892	Function unknown
EST	EST	582894	Function unknown
EST	EST	583044	Function unknown
EST	EST	583144	Function unknown
EST	EST	583192	Function unknown
EST	EST	583193	Function unknown
EST	EST	583200	Function unknown
EST	EST	583216	Function unknown
EST	EST	583240	Function unknown
EST	EST	583241	Function unknown
EST	EST	583265	Function unknown
EST	EST	583265	Function unknown
EST	EST	583270	Function unknown
EST	EST	583289	Function unknown
EST	EST	583312	Function unknown
EST	EST	583313	Function unknown
EST	EST	583330	Function unknown
EST	EST	583337	Function unknown
EST	EST	583337	Function unknown
EST	EST	583361	Function unknown
EST	EST	583361	Function unknown
EST	EST	583385	Function unknown
EST	EST	583428	Function unknown
EST	EST	583438	Function unknown
EST	EST	583481	Function unknown
EST	EST	583528	Function unknown
EST	EST	583581	Function unknown
EST	EST	583598	Function unknown
EST	EST	583598	Function unknown
EST	EST	583688	Function unknown
EST	EST	583691	Function unknown
EST	EST	583810	Function unknown
EST	EST	583855	Function unknown
EST	EST	583899	Function unknown
EST	EST	639721	Function unknown
EST	EST	639730	Function unknown
EST	EST	639737	Function unknown
EST	EST	639775	Function unknown
EST	EST	640333	Function unknown
EST	EST	640396	Function unknown
EST	EST	640416	Function unknown
EST	EST	640417	Function unknown
EST	EST	640428	Function unknown
EST	EST	640445	Function unknown
EST	EST	640539	Function unknown
EST	EST	640562	Function unknown
EST	EST	640586	Function unknown
EST	EST	640612	Function unknown
EST	EST	640657	Function unknown
EST	EST	640657	Function unknown
EST	EST	640660	Function unknown
EST	EST	640791	Function unknown
EST	EST	640794	Function unknown
EST	EST	640815	Function unknown

Table 1. (Continued)

Gene name	Symbol	Clone ID	Biological function
EST	EST	640816	Function unknown
EST	EST	640841	Function unknown
EST	EST	640861	Function unknown
EST	EST	640862	Function unknown
EST	EST	640900	Function unknown
EST	EST	640956	Function unknown
EST	EST	641098	Function unknown
Adenosine deaminase	Ada	577879	Immune response
B-cell receptor-associated protein 31	Bcap31	574060	Immune response
Beta-2 microglobulin	B2m	576472	Immune response
Beta-2 microglobulin	B2m	576493	Immune response
CD28 antigen	Cd28	576501	Immune response
CD4 antigen	Cd4	640273	Immune response
CD4 antigen	Cd4	640273	Immune response
Chemokine (C-C motif) ligand 2	Ccl2	573898	Immune response
Chemokine (C-C motif) ligand 25	Ccl25	575567	Immune response
Chemokine (C-C motif) receptor 9	Ccr9	573572	Immune response
Chemokine (C-X3-C motif) ligand 1	Cxcl1	1329058	Immune response
Chemokine (C-X-C motif) ligand 4	Cxcl4	582960	Immune response
Chemokine (C-X-C motif) ligand 4	Cxcl4	573339	Immune response
Complement component 2	C2	583642	Immune response
Fas	FAS	5222802	Immune response
Fc receptor, IgG, alpha chain transporter	Fcgrt	573972	Immune response
Guanylate nucleotide binding protein 2	Gbp2	583808	Immune response
Guanylate nucleotide binding protein 2	Gbp2	583832	Immune response
Interleukin 16	Il 16	573050	Immune response
Interleukin 4	Il 4	640300	Immune response
Lymphocyte antigen 6 complex, locus A	Ly6a	581749	Immune response
SAM domain and HD domain, 1	Samhd1	573999	Immune response
Tumor necrosis factor receptor superfamily, member 13b	Tnfrsf13b	574040	Immune response
Tnf receptor-associated factor 6	Traf6	1330069	Immune response
RIKEN cDNA A430107D22 gene	A430107D22Rik	574964	Inhibitory regulator of the ras-cyclic amp pathway
Tyrosine 3-monooxygenase	Ywhah	640494	Intracellular protein transport
Neutrophil cytosolic factor 4	Ncf4	576061	Intracellular signaling cascade
Signal transducer and activator of transcription 1	Stat1	583302	Intracellular signaling cascade
Prosaposin	Psap	1264859	Lipid metabolism
Alcohol dehydrogenase, 1	Adhfe1	640899	Metabolism
Aminolevulinic acid synthase 1	Alas1	583710	Metabolism
Aminolevulinic acid synthase 1	Alas1	583638	Metabolism
Ariadne homolog 2	Arih 2	583839	Metabolism
Arylformamidase	Afmid	582985	Metabolism
Glutathione S-transferase, theta 2	Gstt2	639705	Metabolism
GTP cyclohydrolase 1	Gch1	575924	Metabolism
Lanosterol synthase	Lss	641986	Metabolism
Kinesin family member 11	Kif11	583152	Mitotic centrosome separation
Heterogeneous nuclearribonucleoprotein C	Hnrpc	640481	mRNA processing
Splicing factor, arginine/serine-rich 1	Sfrs1	582848	mRNA processing
Splicing factor 1	Sfrs1	583015	mRNA processing
Splicing factor 1	Sfrs1	640898	mRNA processing
Splicing factor 1	Sfrs1	583038	mRNA processing
Splicing factor 1	Sfrs1	583110	mRNA processing
Poly(rC) binding protein 3	Pcbp3	583766	Nucleic acid binding
Deoxythymidylate kinase	Dtymk	583035	Nucleotide biosynthesis
PTK2 protein tyrosine kinase 2 beta	Ptk2b	573298	Protein amino acid phosphorylation
SCY1-like 1	Scyl1	583831	Protein amino acid phosphorylation
STE20-like kinase	Slk	583214	Protein amino acid phosphorylation
Ena-vasodilator stimulated phosphoprotein	Ev1	582896	Protein binding
SH3-domain GRB2-like 1	Sh3gl1	582991	Protein binding
Ribonuclease P/MRP 30 subunit	Rpp30	638902	Protein biosynthesis
Ribosomal protein L12	Rpl12	639953	Protein biosynthesis
Ribosomal protein L18A	Rpl18a	583588	Protein biosynthesis
Ribosomal protein L24	Rpl24	641145	Protein biosynthesis
Ribosomal protein L27a	Rpl27a	640012	Protein biosynthesis
Ribosomal protein L29	Rpl29	583177	Protein biosynthesis
Ribosomal protein S11	Rps11	640585	Protein biosynthesis
Ribosomal protein S3a	Rps3a	639738	Protein biosynthesis
ATP-binding cassette, sub-family E (OABP), member 1	Abce1	583718	Protein folding

Table 1. (Continued)

Gene name	Symbol	Clone ID	Biological function
DnaJ (Hsp40) homolog, subfamily A, member 2	Dnaja2	1225247	Protein folding
Tubulin tyrosine ligase-like family, member 4	Ttll4	582821	Protein modification
Calpain 2	Capn2	583248	Proteolysis
Cathepsin L	Cts 1	640032	Proteolysis
Insulin degrading enzyme	Ide	574241	Proteolysis
Protein-kinase, interferon-inducible double stranded RNA dependent inhibitor	Prkrir	640886	Regulation of cell proliferation
Annexin A1	Anxa1	640284	Regulation of cell proliferation
Chloride channel, nucleotide-sensitive, 1A	Clns1A	639949	Regulation of cell volume
Eukaryotic translation initiation factor 4E binding protein 2	Eif4ebp2	582913	Regulation of protein biosynthesis
Furin	Furin	583759	Regulation of protein biosynthesis
Activator of basal transcription	Abce1	583686	Regulation of transcription
Ankyrin repeat domain 10	Ankrd10	583740	Regulation of transcription
Aryl hydrocarbon receptor nuclear translocator-like	Arntl	582931	Regulation of transcription
Ets variant gene 3	Etv3	583667	Regulation of transcription
General transcription factor IIB	Gtf2b	640038	Regulation of transcription
H6 homeo box 3	Hmx3	583663	Regulation of transcription
Homeodomain interacting protein kinase 1	Hipk1	583858	Regulation of transcription
Human immunodeficiency virus type I enhancer binding protein 2	Hivep2	1282235	Regulation of transcription
Nuclear receptor interacting protein 1	Nrip1	582841	Regulation of transcription
Insulin induced gene 2	Insig2	640096	Response to sterol depletion
Stress-induced phosphoprotein 1	Stip1	583886	Response to stress
LPS-responsive beige-like anchor	Lrba	573828	Signal transduction
Protein phosphatase 2, regulatory subunit B (B56), gamma isoform	Ppp2r5c	640201	Signal transduction
Putative homeodomain transcription factor 1	Phtf1	583390	Transcriptional control
Single-stranded DNA binding protein 2	Ssbp2	641018	Transcriptional control
SRY-box containing gene 4	Sox4	582937	Transcriptional control
SRY-box containing gene 4	Sox4	639914	Transcriptional control
Trans-acting transcription factor 1	Sp1	583862	Transcriptional control
Transcription elongation factor A, 3	Tcea3	573147	Transcriptional control
Transcriptional adaptor 2	Tada2l	583605	Transcriptional control
Transducin (beta)-like 1X-linked receptor 1	Tbl1xr1	640061	Transcriptional control
Tripartite motif protein 24	Trim24	576305	Transcriptional control
Zinc fingers and homeoboxes protein 1	Zhx1	640278	Transcriptional control
RIKEN cDNA 4930566A11 gene	4930566A11Rik	583433	Transcriptional repressor
ARP1 actin-related protein 1 homolog A	Actr 1A	583567	Transport
ATPase, H+ transporting, lysosomal V0 subunit E	Atp6vOe	583287	Transport
Calcium channel, voltage-dependent, beta 3 subunit	Cacnb3	583697	Transport
Exportin, tRNA	Xpot	639640	Transport
Mitogen-activated protein kinase 8 interacting protein 3	Mapk8ip3	576539	Transport
Mitogen-activated protein kinase 8 interacting protein 3	Mapk8ip3	576539	Transport
Nuclear RNA export factor 1 homolog	Nxf1	574499	Transport
RAB3D, member RAS oncogene family	Rab3d	575981	Transport
Solute carrier family 25, member 16	Slc25a16	640182	Transport
Synaptophysin-like protein	Sypl	640468	Transport
Translocase of outer mitochondrial membrane 40 homolog	Tomm40	583739	Transport
Ubiquinol-cytochrome c reductase, Rieske iron-sulfur polypeptide 1	Uqcrfs1	640752	Transport
Solute carrier family 11,member 1	SLC11A1	140625	Transport
RAN binding protein 9	Ranbp9	583408	Transport

[B56], gamma isoform, accession number NM_012023), which participates in the signal transduction cascade. These two genes participate in the regulation of the cell cycle. We also identified *Cdk4* (cyclin-dependent kinase 4, accession number NM_009870), another gene involved in the regulation of the cell cycle by cyclin-dependent protein kinase activity. Finally we detected the induction of the *Ada* gene (adenosine deaminase, accession number NM_007398.3). The protein encoded by this gene is important in the development and function of the immune system in humans and mice (18).

Among the repressed genes we have highlighted *Mre11a* (meiotic recombination 11 homolog A, accession number NM_018736.2) and *Parp3* (poly ADP-ribose polymerase family, member 3, accession number NM_145619.2), which encode proteins involved in the response to DNA damage. The *Phtf1* transcription factor (accession number NM_013629), which encodes an enzyme

involved in protein amino acid phosphorylation, was also repressed in lymphocytes following tumor cell injection.

These induced and repressed genes represent candidate genes for the detection of cancer by a gene profiling assay. These genes may be useful in further investigations using an *in vivo* carcinogenesis model (e.g., colon carcinoma induced by azoxymethane and dextran sodium sulfate treatment of mice) (19). Although an orthotopic model of cancer could be more physiologically relevant, it features intrinsic experimental limitations that would have made the analysis presented in this paper difficult. Specifically, the identification of specific animals bearing tumor cells before the appearance of a tumor mass is difficult in an orthotopic model. As this study aimed to examine the expression profile of lymphocytes before a tumor could be detected, we used a system in which the number of tumor cells could be precisely controlled.

The statistically significant shift in the gene expression profile of peripheral lymphocytes shortly after the injection of cancer cells validates the development of an assay that uses microarray-based gene expression profiling of peripheral blood immune cells as a method of early detection of cancer. Moreover, these results serve as a first step in the development of a study to examine the peripheral lymphocytes of patients with cancer and the development of a data bank with the expression signatures corresponding to the different cancers. This could be useful as a reference data bank of lymphocyte gene expression signatures in response to cancer.

The flow cytometry and real-time PCR were performed in the laboratories of Dr. Célio Lopes Silva and Dr. Zilá Luz Paulino Simões, respectively, at the University of São Paulo campus of Ribeirão Preto, Brazil. The cDNA clones used for the microarray preparation were kindly provided by Dr. Catherine Nguyen from the INSERM (Institut National de la Santé et de la Recherche Médicale) U928, Marseille, France. This work is a part of the Ph.D. thesis of MMCM, which was awarded in 2008 by the CAPES (Coordenação de Aperfeiçoamento Pessoal de Nível Superior)/ Ministry of Education, Brazil.

1. Staudt LM, Brown PO. Genomic view of the immune system. Ann Rev Immunol 18:829–859, 2000.
2. Shaffer AL, Rosenvald A, Hurt EM, Giltane JM, Lam LT, Pickreal OK, Staudt LM. Signatures of the immune response. Immunity 15:37–85, 2001.
3. Alizadeh A, Staudt LM. Genomic-scale gene expression profile of normal and malignant immune cells. Curr Opin Immunol 12:219–225, 2000.
4. Pardoll D. Does the immune system see tumors as foreign or self? Ann Rev Immunol 21:807–839, 2003.
5. Marques MM, Junta CM, Cardoso RS, Mello SS, Sakamoto-Hojo ET, Donadi EA, Passos GA. Changes in the gene expression profiling of the thymus in response to fibrosarcoma growth. Mol Cell Biochem 276:81–88, 2005.
6. Pereira E, Tamia-Ferreira MC, Cardoso RS, Mello SS, Sakamoto-Hojo, ET, Passos GA, Donadi EA. Immunosuppressive therapy modulates T lymphocyte gene expression in patients with systemic lupus erythematosus. Immunology 113:99–105, 2004.
7. Trevisan GL, Rassi DM, Baião AM, Sandrin-Garcia P, Mello SS, Tamia-Ferreira MC, Junta CM, Fachin AL, Marques MM, Sakamoto-Hojo ET, Donadi EA, Passos GA. Using cDNA microarrays to identify human $CD19^{+}$ B cell gene products (ESTs) originated from systemic lupus erythematosus susceptibility loci. Autoimmun Rev 5:319–323, 2006.
8. Kovary K, Armelin MC, Armelin HA. Ha-Ras-1 oncogene dosage differentially effects BALB/3T3 cell's growth factor requirement and tumorigenicity. Oncogene Res 4:55–64, 1989.
9. Kovary K, Armelin MC, Armelin HA. Glucocorticoid dexamethazone reversibly complements RJ-Ras oncogene to transform mouse embryo BALB-3T3 cells. J Cell Biochem 41:171–177, 1989.
10. Tusher VG, Tibishirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 98:5116–5121, 2001.
11. Zar JH. Biostatistical Analysis (4th ed.). New Jersey: Prentice Hall, 1999.
12. Hedge P, Qi K, Abernathy C, Gay S, Dharap R, Gaspard JE, Hughes E, Lee NS, Quackenbush JA. Concise guide to cDNA microarray analysis. BioTechniques 29:548–562, 2000.
13. Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J. TM4: a free, open-source system for microarray data management and analysis. BioTechniques 34:374–378, 2003.
14. Sharma P, Sahn NS, Tibishirani R, Skaane P, Urdal P, Berghagen H, Jensen M, Kristiansen L, Moen C, Sharma P, Zaka A, Arnes J, Sauer T, Akslen LA, Schlichting E, Børresen-Dale AL, Lönneborg A. Early detection of breast cancer based on gene-expression patterns in peripheral blood cells. Breast Cancer Res 7:R634–R644, 2005.
15. Bertucci F, Nasser V, Granjeaud S, Eisinger F, Adelaïde J, Tagett R, Loriod B, Giaconia A, Benziane A, Devilard E, Jacquemier J, Viens P, Nguyen C, Birnbaum D, Houlgatte R. Gene expression profiles of poor-prognosis primary breast cancer correlate with survival. Human Mol Genet 11:863–872, 2002.
16. Whitney AR, Diehn M, Popper SJ, Alizadeh AA, Boldrick JC, Relman DA, Brown PO. Individuality and variation in gene expression patterns in human blood. Proc Natl Acad Sci U S A 100:1896–1901, 2003.
17. Faca VM, Song KS, Wang H, Zhang Q, Krasnoselsky AL, Newcomb LF, Plentz RR, Gurumurthy S, Redston MS, Pitteri SJ, Pereira-Faca SR, Ireton RC, Katayama H, Glukhova V, Phanstiel D, Brenner DE, Anderson MA, Misek D, Scholler N, Urban ND, Barnett MJ, Edelstein C, Goodman GE, Thornquist MD, McIntosh MW, DePinho RA, Bardeesy N, Hanash SM. A mouse to human search for plasma proteome changes associated with pancreatic tumor development. PLoS Med 5:0953–0967, 2008.
18. Aldrich MB, Blackburn MR, Kellems RE. The importance of adenosine deaminase for lymphocyte development and function. Biochem Biophys Res Comm 272:311–315, 2000.
19. Tanaka T, Kohno H, Suzuki H,Yamada Y, Sugie S, Mori H. A novel inflammation-related mouse colon carcinogenesis model induced by azoxymethane and dextran sodium sulfate. Cancer Sci 94: 965–973, 2003.