Global Gene Expression Profiling of Circulating Endothelial Cells in Patients with Metastatic Carcinomas

Denis A. Smirnov, Bradley W. Foulk, Gerald V. Doyle, Mark C. Connelly, Leon W.M.M. Terstappen, and S. Mark O’Hara

Immunicon Corp., Huntingdon Valley, Pennsylvania

Abstract

Increased numbers of endothelial cells are observed in peripheral blood of cancer patients. These circulating endothelial cells (CECs) may contribute to the formation of blood vessels in the tumor or reflect vascular damage caused by treatment or tumor growth. Characterization of these cells may aid in the understanding of the angiogenic process and may provide biomarkers for treatment efficacy of angiogenesis inhibitors. To identify markers typical for CECs in cancer patients, we assessed global gene expression profiles of CD146 immunomagnetically enriched CECs from healthy donors and patients with metastatic breast, colorectal, prostate, lung, and renal cancer. From the generated gene profiles, a list of 61 marker genes for CEC detection was generated, and their expression was measured by real-time quantitative PCR in blood samples from 81 metastatic cancer patients and 85 healthy donors that were immunomagnetically enriched for CECs. A set of 34 genes, among which novel CEC-associated genes, such as THBD, BST1, TIE1, POSTN1, SELE, SORT1, and DTR, were identified that were expressed at higher levels in cancer patients compared with healthy donors. Expression of the VWF, DTR, CDH5, TIE, and IGFBP7 genes were found to discriminate between cancer patients and “healthy” donors with a receiver operating characteristic curve accuracy of 0.93. Assessment of the expression of these genes may provide biomarkers to evaluate treatment efficacy. (Cancer Res 2006; 66(6): 2918-22)

Introduction

Approximately 1.2 trillion endothelial cells in an average adult cover the internal surface of blood and lymphatic vessels and actively participate in many physiologic events, such as vascular permeability, inflammation, hemostasis, and angiogenesis (1). Consequently, endothelial cells are involved in multiple clinical pathologic conditions, including cancer, cardiovascular diseases, autoimmune diseases, and infectious diseases (1, 2). Endothelial cells can be found in peripheral blood, the site of origin, and precise role of these circulating endothelial cells (CECs) are not yet understood (3). Increased numbers of CECs have been observed in peripheral blood of patients with a variety of disorders, such as myocardial infarction, infectious vasculitis, kidney transplant rejection, and cancer (3, 4). Enumeration and characterization of CECs may offer a unique opportunity to study the vasculature and improve our understanding of a variety of disease processes. In cancer, in particular, detection and characterization of CECs may be useful for monitoring clinical treatment efficacy of antiangiogenesis therapies, such as bevacizumab, SU6668, and others (4, 5).

Reverse transcription-PCR (RT-PCR) is widely used for detection of circulating tumor cells in the peripheral blood (6, 7). To date, however, RT-PCR is used infrequently to study endothelial cells in peripheral blood, which is likely due to the lack of suitable and tested markers (4, 8). Our objective here was to identify markers for monitoring CECs in peripheral blood samples. We profiled gene expression in CECs enriched from seven cancer patients with high number of detected CECs and three “healthy” donors using microarrays, identified candidate marker genes, and validated their performance in a broader group of cancer patients and healthy donors using quantitative RT-PCR. As a result of these experiments, we identified a set of marker genes that can be potentially used to monitor CECs in peripheral blood by quantitative RT-PCR.

Materials and Methods

Patient samples, immunomagnetic sample preparation, and CEC enumeration. Blood from cancer patients and healthy volunteers was drawn into a 10-mL EDTA-containing vacutainer tubes (Becton Dickinson, Franklin Lakes, NJ) for RNA extraction and a 10-mL CellSave preservative tube (Immunicon Corp. Huntingdon Valley, PA) for CEC enumeration. Samples were kept at room temperature and processed within 36 hours after collection. Information about all samples used in this study is presented in Supplementary Tables S1A and S2A online. All participants signed an Institutional Review Board–approved informed consent form before study participation. The healthy individuals used for comparison with the patients had no known illness or fever at the time of draw and had no known history of malignant disease. The CellTracks AutoPrep system and the CellSpotter Analyzer system (Immunicon) were used to enrich and enumerate endothelial cell from peripheral blood. Four milliliters of collected blood were used for CEC enumeration and RNA extraction. Briefly, endothelial cells were immunomagnetically captured using ferrofluids coated with CD146 antibodies. The enriched cells were then labeled with the nuclear dye 4′,6-diamidino-2-phenylindole (DAPI), CD105 antibodies conjugated to phycoerythrin (CD105-PE), and the pan-leukocyte antibody CD45 conjugated to allophycocyanin (CD45-APC). The CD146-enriched, fluorescently labeled cells were identified as CECs when the cells exhibited the DAPI+/CD105+/CD45-APC/CD45-PE phenotype (3). For RNA, endothelial cells were immunomagnetically captured using ferrofluids coated with CD146 antibodies while the fluorescent labeling was omitted.

RNA isolation, target preparation, microarray hybridization, and microarray data analysis. For gene expression studies following immunomagnetic enrichment, CECs were lysed by adding 100 μL of Trizol reagent (Invitrogen, Carlsbad, CA). RNA from all samples was isolated using Trizol reagent, DNase I treated, and Trizol repurified. Ten nanograms of total RNA from the CEC-enriched fraction from each of the seven cancer patients with high CEC counts and three “healthy” donors were used to prepare biotinylated hybridization targets with Affymetrix’s eukaryotic...
small sample target labeling assay, version II.1 Biotinylated target cRNA was then hybridized to an Affymetrix Focus array containing >8,500 verified human sequences according to manufacturer's instructions, and gene expression data were obtained using the Affymetrix Microarray Analysis Suite, version 5.0. A global scaling normalization procedure to normalize the expression data to the target value of 150 was done. More detailed information about the microarray experiments is available online (submission no. MIAMEXPress E-MEXP-476).2

Multigene quantitative real-time RT-PCR analysis. To narrow down the list of candidate genes for the real-time RT-PCR verification studies, we focused on the genes with high ratios between median expression intensity value from cancer patients and "healthy" donors and minimal expression in leukocytes based on data published in the Cancer Gene Anatomy Project SAGE database.3 We also measured expression levels of a few well-characterized endothelial markers whose expression was not detectable using Affymetrix chips (e.g., CDH5) or that showed minimal differences between CEC-enriched samples from cancer patients and healthy donors (e.g., VCAM1, CD34, and SEL). Expression of the selected candidate marker genes was evaluated in a separate set of metastatic cancer patients and a control group of healthy volunteers. To ensure that a sufficient amount of cDNA was available for multigene analysis, the RNA extracted from the CEC-enriched fraction of each blood sample was subjected to one round of amplification using the MessageAmp aRNA kit (Ambion, Austin, TX) according to the manufacturer's instructions. A total of 25 ng of the resulting aRNA was reverse transcribed. cDNA was diluted 30-fold with distilled water, and then a volume of 10 µL of the cDNA samples was used in each RT-PCR reaction. Where possible, primer sequences that amplified a product of about 100 bp within 300 bases of the 3'-end of the transcript were selected (see Supplementary Table S2B). Quantitative RT-PCR was done using the SYBR Green PCR Master Mix and an ABI Prism 7000 Detection System (Applied Biosystems, Foster City, CA). Gene expression levels were determined using a standard calibration curve prepared from gene-specific RT-PCR products with known concentrations. Gene expression levels between samples were normalized using the expression levels of the ribosomal protein RPS27A gene, which was shown recently to exhibit the least amount of variability in expression levels among different tissues (9).

Statistical analysis of gene expression data. To allow logarithmic transformation of the negative observations (zero transcripts detected) in the quantitative real-time PCR data, a constant c = 2 was added to all measured number of transcripts for all genes in all clinical samples (10). To present expression data (obtained by microarrays or by quantitative RT-PCR) for each gene in each profiled sample, log 2–transformed ratios between expression values in tested samples and the median expression value between cancer patients and healthy donors were calculated. Heat maps were constructed using the Multiple Experiment Viewer software (11). Gene expression levels in the CTC-enriched fraction of the blood samples from the set of metastatic cancers and the normal donors were compared using the Mann-Whitney test to identify genes with significantly different expression levels (see Supplementary Table S2C and D). Potential diagnostic value of each gene and their combination was estimated using an area under the receiver operating characteristic curve (AUC ROC).

Results

Generation of global expression profiles for CECs from patients with metastatic cancers. An integrated sample preparation system was used to enrich CECs from 4 mL of blood
using magnetic nanoparticles conjugated to monoclonal antibodies against CD146 expressed on endothelial cells and a subset of activated T lymphocytes. Despite a 10,000-fold enrichment, CECs are still outnumbered by carried over leukocytes and the subset of activated T lymphocytes. From previous spike-in experiments, it was estimated that a considerable number of genes expressed in circulating cells could be identified with the Affymetrix GeneChip system if ~50 to ≥100 captured cells were present in a background of ~1,000 to 10,000 leukocytes (data not shown). Based on CEC enumeration information, we selected CEC-enriched samples from seven cancer patients with high number of CECs (renal cancer, 102 CECs; prostate cancer, 87 CECs; lung cancer, 88 CECs; colorectal cancer, 557 CECs; breast cancer, 81 CECs; breast cancer, 299 CECs; breast cancer, 118 CECs; Supplementary Table S1A). To identify potential markers for detection and characterization of carcinoma associated CECs, gene expression profiles from CEC-enriched samples from abovementioned seven cancer patients were compared with CEC-enriched samples from three “healthy” donors (containing 1, 43, and 48 detected CECs; Supplementary Table S1A).

After a global scaling procedure was used to normalize the expression data between experiments, we selected ~160 genes as potential markers for detection of CECs in peripheral blood (Fig. 1; Supplementary Table S1A-B). We concentrated on the genes with high ratios between median expression values in cancer samples and median expression values in “healthy” donors (Fig. 1A). Among the genes more abundant in CEC-enriched samples from cancer patients were a number of genes known to be associated with endothelial function. This list includes tissue inhibitor of metalloproteinase (TIMP2), a protein involved in regulation of angiogenesis; thrombomodulin (THBD), an endothelial cell surface glycoprotein; endoglin (ENG or CD105), a membrane glycoprotein primarily associated with human vascular endothelium; adrenomedullin (ADM), a potent vasodilator and a hypotensive agent; vascular endothelial growth factor (VEGF), a well-known mitogen for vascular endothelial cells; and CD146, a gene that encodes endothelial cell adhesion protein that was used to capture CECs from peripheral blood (3, 12–15). These findings indicate that gene expression profiling of CECs can be a useful tool for identification of potential marker genes for CEC detection.

Verification of CEC-specific expression of the candidate genes by quantitative real-time RT-PCR. It was crucial to confirm that the global expression profiles generated from CECs enriched from peripheral blood of the seven cancer patients are reflective of CEC gene expression signatures in a larger population of patients. Therefore, expression of the candidate genes selected from the microarray analyses was measured in the CEC-enriched blood fractions from 81 metastatic colorectal cancer patients and a control group of 55 apparently healthy normal donors using quantitative real-time RT-PCR (see Supplementary Table S2A). Consistent with previous reports (4), we observed statistically significant difference (P < 0.05, Mann-Whitney test) in frequencies of CECs in peripheral blood of patients with metastatic cancer and “healthy” volunteers. In blood samples from 55 “healthy” donors, 2 to 378 CECs/4 mL were present (mean = 29, median = 13), whereas in 62 samples from metastatic cancer patients, CECs ranged from 1 to 1,560 (mean = 70, median = 26).

2 http://www.ebi.ac.uk/miamexpress/. 

Figure 2. Results of CEC expression profiling in set of samples from cancer patients and healthy donors by quantitative RT-PCR. A, fold differences between median (median ratio) and mean (mean ratio) expression values between samples from cancer patients and healthy donors for each gene are presented in the form of a heat map (black to yellow scale). B, expression data for each gene in each profiled sample are presented in the form of a heat map (green to red scale) of log 2-transformed ratios between expression values in tested samples and the median expression value between cancer patients and healthy donors. C, values of AUC ROC for each gene.
For quantitative RT-PCR studies, 61 genes were selected. In addition to genes selected from the microarray experiments, we also tested a small number of well-characterized endothelial markers whose expression was not detectable using Affymetrix chips (e.g., CDH5) or that showed minimal differences between CEC-enriched samples from cancer patients and healthy donors (e.g., VCAM1, CD34, and SEL1; Supplementary Table S1B). Among the 61 genes tested by quantitative RT-PCR, transcripts for 34 genes were more abundant (P < 0.01; Mann-Whitney test) in CEC-enriched samples from 81 metastatic cancer patients than in 55 “healthy” donors (Supplementary Table S2A). Figure 2 depicts expression data for these 34 genes that show greater number of transcripts detected in CEC-enriched samples from patients with metastatic cancer when compared with CEC-enriched samples from “healthy” control donors. Among the genes that are more abundant in CEC-enriched samples from cancer patients are several with well-characterized endothelial functions. This list includes von Willebrand factor (VWF), THBD, TIMP2, protein receptor tyrosine kinase (TIE), and its family member TIE2 (TEK; refs. 12, 16). Among the genes with less studied roles in endothelial cells are peroxidin (POSTN), osteonectin (SPARC), diphtheria toxin receptor (DTR or HB-EGF), neuregulin 1 (NRG1), bone marrow stromal cell antigen 1 (BST1), sortilin (SORT1), and others. Interestingly, recent reports identified several of these genes, including POSTN, SPARC, DTR, and NRG1, as potent promoters of angiogenesis and enhancers of tumor growth (17–20). These observations suggest that differences in gene expression observed between CEC-enriched samples from cancer patients and “healthy” donors may be associated with active processes of angiogenesis and tumor growth in cancer patients.

Among the cancer samples that were used to validate expression of potential markers by quantitative real-time PCR, there were 12 samples from colorectal cancer, 22 samples from lung cancer, and 29 samples from breast cancer patients (Supplementary Table S2A). When compared expression pattern for the 34 genes that were more abundant in CEC-enriched samples from cancer patients than in “healthy” donors, we observed no statistically significant difference in gene expression profiles between patients with colorectal, lung, and breast cancer (Kruskal-Wallis test, \( \alpha = 0.01 \); Supplementary Table S2E).

Discrimination between patients with metastatic cancers and “healthy” donors using real-time RT-PCR gene expression data. It was reported that the levels of CECs correlate well with the degree of tumor angiogenesis or the response to angiogenic therapy (5). Moreover, recent reports also indicate that high plasma concentrations of endothelial cell–derived proteins (e.g., vWF and VCAM1) correlate with advanced diseases and significantly poor prognosis of patients with metastatic carcinomas (21, 22). These observations suggest that quantitative real-time RT-PCR analysis may also have clinical use for the characterization of the CECs.

To estimate potential use of identified marker genes to discriminate between cancer patients and “healthy” donors, we calculated the AUC ROC values for all the 34 genes that were more abundant in CEC-enriched samples from cancer patients than in “healthy” donors (Fig. 2C). AUC ROC values are commonly used as a summary measure of diagnostic accuracy (23). AUC ROC values for measured genes varied from the lowest of 0.64 (for TNS) to the highest of 0.88 (for VWF; Fig. 2C; Supplementary Table S2D) and exceeded for most genes the AUC ROC value of 0.65 calculated strictly based on the number of CECs detected in tested clinical samples. From a number of gene expression studies, it is clear that the use of multiple markers usually improves the accuracy (24). Indeed, a simple combination of genes with the highest AUC ROC values (VWF, DTR, CDH5, TIE, and IGFBP7) was estimated to have an AUC ROC value of 0.93, which is considered to be an excellent classification accuracy (25). This suggests that just a few markers can be used to develop a highly accurate RT-PCR based test to characterize CECs.

Discussion

Little is known about molecular characteristics of the CECs that are detected in peripheral blood of cancer patients. To identify a set of genes that can be used for detection and characterization of CECs in peripheral blood, we first compared gene expression profiles from CEC-enriched samples from the blood of patients with metastatic breast, colorectal, prostate, lung, and renal cancers with CEC-enriched samples from three “healthy” donors. Expression of the 61 candidate genes was then measured by quantitative real-time PCR in the CEC-enriched blood fractions from 81 metastatic cancer patients and a control group of 55 apparently healthy normal donors. Transcripts for the 34 genes were significantly more abundant in CEC-enriched samples from metastatic cancer patients than in “healthy” donors. Among these genes several have well-known association with endothelial function (e.g., VWF, TIE, and CDH5). Others genes (e.g., POSTN, SPARC, DTR, and NGR1) were shown to be involved in angiogenesis and tumor growth. These results indicate that the differences between CEC-enriched samples from cancer patients and “healthy” donors may be indicative of the active processes of angiogenesis and tumor growth in cancer patients. We also identified several genes whose role in vascular function and angiogenesis is yet to be elucidated. This list includes BST1, SORT1, TNS, and others. Our analysis of ROC curves suggest that quantitative real-time RT-PCR analysis of samples immunomagnetically enriched for CECs can be used to accurately detect and characterize endothelial cells found in the blood of patients with metastatic cancer and potentially other diseases. We believe that global expression profiles of CECs may provide insights that could improve our understanding of endothelial function in cancer and could lead to the development of both novel noninvasive diagnostic tools as well as novel therapeutic targets.

Acknowledgments

Received 11/8/2005; revised 12/22/2005; accepted 1/20/2006.

Grant support: Immunicon Corp.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

References


Global Gene Expression Profiling of Circulating Endothelial Cells in Patients with Metastatic Carcinomas

Denis A. Smirnov, Bradley W. Foulk, Gerald V. Doyle, et al.


**Updated version**  
Access the most recent version of this article at:  
http://cancerres.aacrjournals.org/content/66/6/2918

**Supplementary Material**  
Access the most recent supplemental material at:  
http://cancerres.aacrjournals.org/content/suppl/2007/10/04/66.6.2918.DC1

**Cited articles**  
This article cites 25 articles, 12 of which you can access for free at:  
http://cancerres.aacrjournals.org/content/66/6/2918.full#ref-list-1

**Citing articles**  
This article has been cited by 5 HighWire-hosted articles. Access the articles at:  
http://cancerres.aacrjournals.org/content/66/6/2918.full#related-urls

**E-mail alerts**  
Sign up to receive free email-alerts related to this article or journal.

**Reprints and Subscriptions**  
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

**Permissions**  
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.