Disease-associated Expression Profiles in Peripheral Blood Mononuclear Cells from Patients with Advanced Renal Cell Carcinoma


ABSTRACT

Expression profiling has demonstrated that transcriptomes of primary malignancies differ from those in normal tissue. It is unknown, however, whether there exist “surrogate” transcriptional markers in peripheral blood mononuclear cells (PBMCs) of patients with solid tumors. We identified transcripts expressed differentially between PBMCs from renal cell carcinoma patients and normal subjects, some of which appear to reflect specific immune responses of circulating cells. We also identified small sets of predictor genes distinguishing PBMCs from renal cell carcinoma patients and normal volunteers with high accuracy. The present findings have important implications for diagnosis and future clinical pharmacogenomic studies of antitumor therapies.

INTRODUCTION

RCC1 comprises the majority of all cases of kidney cancer and is one of the 10 most common cancers in industrialized countries (1). The 5-year survival rate for advanced RCC is less than 5% (2). RCC is usually detected by imaging methods, and 30% of apparently nonmetastatic patients undergo resection after surgery and eventually die of disease (3). Recent expression profiling studies have demonstrated that the transcriptional profiles of primary malignancies are radically altered from the transcriptional profiles of the corresponding normal tissue (for a review see Ref. 4). Specific microarray studies examining RCC tumor transcriptional profiles in detail (5) have identified many classes of genes altered between normal kidney tissue and primary RCC tumors. Despite the progress in expression profiling of primary malignant tissues, it is currently unknown whether in the context of RCC or any other active solid tumor burden there exist correspondingly distinct markers of gene expression in the PBMCs of affected individuals. In the present study, global expression profiles of PBMCs from RCC patients were compared with PBMC profiles from normal volunteers using oligonucleotide arrays for the purpose of identifying surrogate transcriptional markers of disease in the blood of RCC patients.

MATERIALS AND METHODS

Clinical Parameters and Demographics of Patients and Normal Volunteers. PBMCs were isolated from peripheral blood of 20 normal volunteers (12 females and 8 males) and 45 RCC patients (18 females and 27 males) participating in a Phase II study. Consent for the pharmacogenomic portion of the clinical study was received, and the project was approved by the local Institutional Review Boards at the participating clinical sites. The RCC tumors were classified at each site as conventional (clear cell) carcinomas (24), granular (1), papillary (3), or mixed subtypes (7). Ten tumors were classified as unknown. RCC patients were primarily of Caucasian descent (44 Caucasians and 1 African-American) and had a mean age of 58 years (range, 40–78). Normal volunteers were of exclusively Caucasian descent, with a mean age of 42 years (range, 29–58).

PBMC Preparation, Isolation of RNA, and Hybridization of Targets to Microarrays. PBMCs from individuals were isolated from whole blood samples (8 ml) collected into cell purification tubes according to the standard procedure. All normal and RCC blood samples were shipped or stored overnight before processing. Total RNA was isolated from PBMC pellets using the RNeasy mini kit (Qiagen, Valencia, CA), and labeled probe for oligonucleotide arrays was prepared using a modification of the procedure described by Lockhart et al. (6). Labeled probes were hybridized to oligonucleotide arrays comprising over 12,600 human sequences (HgU95A, Affymetrix), according to the Affymetrix Expression Analysis Technical Manual (Affymetrix). Gene Expression Data Reduction. Data analysis and absent/present call determination were performed on raw fluorescent intensity values using GENECHIP 3.2 software (Affymetrix). “Present” calls were calculated by GENECHIP 3.2 software by estimating whether a transcript is detected in a sample based on the strength of the signal of the gene compared with background. The “average difference” values for each transcript were normalized to “frequency” values using the scaled frequency normalization method (7), in which the average differences for 11 control cRNAs with known abundance spiked into each hybridization solution were used to generate a global calibration curve. This calibration was then used to convert average difference values for all transcripts to frequency estimates, stated in units of parts per million ranging from 1:300,000 (~3 pm) to 1:1,000,000 (1.000 pm). Statistical and Clustering Analyses. Unsupervised hierarchical clustering of genes and/or arrays on the basis of similarity of their expression profiles was performed using the procedure of Eisen et al. (8). Nearest neighbor network and supervised prediction were performed using Genecluster version 2.0,4 which has been described previously (9). For hierarchical clustering and nearest neighbor analysis, data were log transformed and normalized to have a mean value of zero and a variance of one. To identify the disease-associated transcripts, a Student’s t test was used to compare normal PBMC expression profiles to renal carcinoma PBMC profiles.

Additional Samples from the GeneLogic GX2000 Bioexpress Database and Fold Change Analysis. Expression profiles measured on HgU95 chips of renal carcinoma biopsies (n = 47) and nonmalignant normal kidney tissues (n = 60), WBCs from nondiseased volunteers (n = 4) and WBCs from non-RCC end-stage renal failure patients (n = 9), or unstimulated CD4 T cells in culture (n = 3) and anti-CD3/anti-CD28-stimulated CD4 T cells in culture (n = 3) were accessed from the GX2000 BioExpress database (GeneLogic, Gaithersburg, MD). Data were processed in Affymetrix Micro Array Suite 4 and then normalized using the GeneLogic normalization algorithm. Fold changes were calculated in the GX2000 Fold Change analysis tool, which uses a geometric mean to calculate average changes in the expression of a gene between groups of samples.

Received 3/26/03; revised 6/10/03; accepted 6/20/03.

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. 1Present address: Millennium Pharmaceuticals, 75 Sidney Street, Cambridge, MA 02139. 2To whom requests for reprints should be addressed, at Wyeth Research, 1 Burtt Road, Andover, MA 01810. Phone: (978) 247-1156; Fax: (978) 247-1133; E-mail: mburczynski@wyeth.com. 3The abbreviations used are: RCC, renal cell carcinoma; PBMC, peripheral blood mononuclear cell; ppm, parts per million.

RESULTS

Identification of a RCC Disease Gene Classification Set in Peripheral Blood. Expression profiling analysis of the 20 normal PBMC RNA samples and 45 RCC PBMC RNA samples revealed that of the 12,626 genes on the HgU95A chip, 5,249 genes met the initial criteria for further analysis (at least one present call and at least one frequency >10 ppm). On average, 4,023 transcripts were detected as “present” in the 45 RCC PBMCs, whereas 4,254 expressed transcripts were detected as “present” in the 20 normal PBMCs.

An initial unsupervised cluster analysis approach, which hierarchically groups samples and genes based on correlation coefficients (8), was performed using the 5249 genes passing the main filtering criteria (Fig. 1A). The dendrogram describing sample relationships grouped the majority of RCC PBMCs (42 of 45) into a single RCC-specific cluster, whereas expression patterns of normal PBMCs and a small subset of RCC PBMCs (3 of 45) formed a separate cluster (Fig. 1B). A fold change analysis and Student’s t test (two-tailed distribution; two-sample unequal variance) identified transcripts defined as differentially expressed between RCC PBMCs and normal PBMCs. In total, transcript levels of 184 genes differed, on average, by 2-fold or greater between normal and RCC PBMCs, with an unadjusted P below 0.001 in a t test between the groups. Of these, 132 transcripts were expressed in at least 15% of the PBMC samples (present in 10 or more of the 65 profiles) and are presented in Table 1.

Fig. 1. Global expression analysis of PBMCs from RCC patients and normal volunteers. Total RNA obtained from PBMCs of 45 RCC patients and PBMCs from 20 normal patients were analyzed on oligonucleotide arrays containing more than 12,000 full-length human genes. In total, 65 samples were analyzed on individual arrays. In no case were samples pooled. A, unsupervised hierarchical cluster analysis of normal and RCC PBMCs using all expressed genes present in at least one sample and possessing a frequency of greater than 10 ppm in at least one sample (5,249 genes total). Red indicates genes that are elevated relative to the average expression value across all experiments, and green indicates genes that are decreased relative to the average expression value. B, a dendrogram of sample relatedness using all expressed gene expression values is shown. RCC patient PBMC expression profiles are denoted by yellow bars, and normal volunteer PBMC expression profiles are indicated by gray bars.
## Table 1: One hundred thirty-two disease-associated transcripts in RCC PBMCs

Transcripts are sorted by P value in the Student’s t test (t test P). Fold changes are reported as RCC average frequency/normal average frequency, in which a number equal to or greater than 2.0 represents a transcript induced in RCC PBMCs. The number of samples possessing levels greater than 10 ppm are presented for each transcript (No. freq >10), and the number of samples in which the transcript was considered “present” are also indicated (No. present). The last column indicates whether the transcript was detected in other settings or model systems, as appropriate.

<table>
<thead>
<tr>
<th>GenBank accession no.</th>
<th>Gene annotation</th>
<th>Unigene ID</th>
<th>Fold change</th>
<th>t test P</th>
<th>No. present</th>
<th>No. freq &gt;10</th>
<th>Common with...</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF025535</td>
<td>Leukocyte immunoglobulin-like receptor, subfamily B member 3</td>
<td>Hs.105928</td>
<td>2.3</td>
<td>1.7959E-07</td>
<td>64</td>
<td>62</td>
<td>Activated T cells</td>
</tr>
<tr>
<td>AF032886</td>
<td>Forkhead box O3A</td>
<td>Hs.14845</td>
<td>2.9</td>
<td>1.824E-07</td>
<td>65</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>U05770</td>
<td>Annexin A5</td>
<td>Hs.30071</td>
<td>2.1</td>
<td>1.9983E-07</td>
<td>64</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>AF021111</td>
<td>Homo sapiens clone 23953 mRNA sequence</td>
<td>Hs.21334</td>
<td>2.2</td>
<td>2.806E-07</td>
<td>26</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>M64925</td>
<td>Membrane protein, palmitoylated 1 (55kD)</td>
<td>Hs.1861</td>
<td>3.4</td>
<td>2.921E-07</td>
<td>65</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>M10267</td>
<td>Tropomyosin 1 (o)</td>
<td>Hs.77899</td>
<td>2.1</td>
<td>3.018E-07</td>
<td>22</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>M62986</td>
<td>Human lipocortin (LIP)</td>
<td>Hs.32485</td>
<td>2.3</td>
<td>3.217E-07</td>
<td>17</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>D86961</td>
<td>Lipoma HMGIC fusion partner-like 2</td>
<td>Hs.79299</td>
<td>3.7</td>
<td>3.542E-07</td>
<td>35</td>
<td>12</td>
<td>Renal failure</td>
</tr>
<tr>
<td>AA187563</td>
<td>ESTs, weakly similar to 38kDa splicing factor [H. sapiens]</td>
<td>Hs.283976</td>
<td>4.0</td>
<td>1.1431E-06</td>
<td>29</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>D32143</td>
<td>Biliverdin reductase B (flavin reductase (NADPH))</td>
<td>Hs.76289</td>
<td>3.8</td>
<td>1.1763E-06</td>
<td>65</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>X12491</td>
<td>Cathepsin L</td>
<td>Hs.78056</td>
<td>35.0</td>
<td>1.02E-06</td>
<td>64</td>
<td>64</td>
<td>Activated T cells</td>
</tr>
<tr>
<td>AL080235</td>
<td>Homo sapiens mRNA; cDNA DKFZp434F152 (from clone DKFZp434F152)</td>
<td>Hs.323383</td>
<td>6.8</td>
<td>7.0412E-07</td>
<td>64</td>
<td>65</td>
<td>Renal failure</td>
</tr>
<tr>
<td>D78793</td>
<td>Superoxide dismutase 2, mitochondrial</td>
<td>Hs.37283</td>
<td>2.5</td>
<td>3.955E-06</td>
<td>45</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>AF241797</td>
<td>Three prime repair exomuclease 1</td>
<td>Hs.278408</td>
<td>3.2</td>
<td>4.0173E-06</td>
<td>23</td>
<td>15</td>
<td>Activated T cells</td>
</tr>
<tr>
<td>H12458</td>
<td>Y12403.x1 Soares placenta Nh2HP Homo sapiens cDNA clone IMAGE: 148517 3' similar to WT6</td>
<td>Hs.283976</td>
<td>4.0</td>
<td>1.1411E-06</td>
<td>29</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>S78798</td>
<td>Cluster incl 578798: 1-phosphatidylinositol-4-phosphate 5-kinase isozyme C [human; PBLs, mRNA, 1.83 nt]</td>
<td>Hs.108966</td>
<td>2.5</td>
<td>4.1232E-06</td>
<td>50</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>M94856</td>
<td>Fatty acid binding protein 5 (psoriasis-associated)</td>
<td>Hs.323383</td>
<td>6.8</td>
<td>7.0412E-07</td>
<td>64</td>
<td>65</td>
<td>Renal failure</td>
</tr>
<tr>
<td>J05070</td>
<td>Matrix metalloproteinase 9 (gelatinase B, 92kD gelatinase, 92kD type IV collagenase)</td>
<td>Hs.32485</td>
<td>2.3</td>
<td>3.217E-07</td>
<td>17</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>M25915</td>
<td>Clusterin (complement lysis inhibitor, SP-40, testosterone-repressed prostate message 2, apolipoprotein J)</td>
<td>Hs.75106</td>
<td>2.2</td>
<td>5.682E-06</td>
<td>65</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>D7071</td>
<td>Nucleoporin 214kD (CAIN)</td>
<td>Hs.170285</td>
<td>2.2</td>
<td>6.273E-06</td>
<td>26</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>D30784</td>
<td>Epidergulin</td>
<td>Hs.115263</td>
<td>3.1</td>
<td>5.6944E-06</td>
<td>44</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>X79355</td>
<td>Tubulin, β polypeptide</td>
<td>Hs.89719</td>
<td>2.6</td>
<td>5.8172E-06</td>
<td>40</td>
<td>42</td>
<td>Renal failure and activated T cells</td>
</tr>
<tr>
<td>D14689</td>
<td>Neocardin 214kD (CAIN)</td>
<td>Hs.170285</td>
<td>2.2</td>
<td>6.273E-06</td>
<td>19</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>AL031230</td>
<td>Homo sapiens mRNA; cDNA DKFZp564A132 (from clone DKFZp564A132)</td>
<td>Hs.284205</td>
<td>2.2</td>
<td>6.7601E-06</td>
<td>50</td>
<td>18</td>
<td>Renal failure</td>
</tr>
<tr>
<td>Z32684</td>
<td>Kerb blood group precursor (McLeod phenotype)</td>
<td>Hs.78919</td>
<td>3.0</td>
<td>7.4352E-06</td>
<td>60</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>X12496</td>
<td>Glycoprotein C (Bell blood group)</td>
<td>Hs.81994</td>
<td>2.3</td>
<td>8.0744E-06</td>
<td>65</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>U16386</td>
<td>Human putative cyclin G1 interacting protein mRNA, partial sequence</td>
<td>Hs.92374</td>
<td>2.4</td>
<td>8.1446E-06</td>
<td>40</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>U43774</td>
<td>Fc fragment of IgA, receptor for</td>
<td>Hs.193122</td>
<td>2.4</td>
<td>9.3169E-06</td>
<td>27</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>M35999</td>
<td>Integrin, β 3 (platelet glycoprotein IIIa, antigen CD61)</td>
<td>Hs.87149</td>
<td>2.6</td>
<td>1.0112E-05</td>
<td>48</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>L07648</td>
<td>MAX-interacting protein 1</td>
<td>Hs.116830</td>
<td>3.1</td>
<td>1.054E-05</td>
<td>65</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>M24069</td>
<td>Cold shock domain protein A</td>
<td>Hs.198726</td>
<td>2.2</td>
<td>1.0541E-05</td>
<td>65</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>AFO061034</td>
<td>Tumor necrosis factor α-inducible protein with leucine zipper domains; Huntingtin interacting protein L</td>
<td>Hs.278898</td>
<td>2.5</td>
<td>1.228E-05</td>
<td>60</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>U29091</td>
<td>Selenium binding protein 1</td>
<td>Hs.334841</td>
<td>6.9</td>
<td>1.246E-05</td>
<td>58</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>U61111</td>
<td>Protein phosphatase 1, regulatory (inhibitor) subunit 2</td>
<td>Hs.129751</td>
<td>2.0</td>
<td>1.282E-05</td>
<td>11</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>M27819</td>
<td>Interleukin 17 receptor</td>
<td>Hs.185920</td>
<td>2.3</td>
<td>1.282E-05</td>
<td>46</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>MLN097</td>
<td>Interleukin 1 receptor accessory protein</td>
<td>Hs.173936</td>
<td>2.1</td>
<td>2.471E-06</td>
<td>44</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>L07648</td>
<td>MAX-interacting protein 1</td>
<td>Hs.116830</td>
<td>2.7</td>
<td>1.9613E-06</td>
<td>65</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>M19267</td>
<td>Tropomyosin 1 (o)</td>
<td>Hs.77899</td>
<td>2.3</td>
<td>1.9998E-06</td>
<td>47</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>GenBank accession no.</td>
<td>Gene annotation</td>
<td>Unigene ID</td>
<td>Fold change</td>
<td>t test p</td>
<td>No. present</td>
<td>No. freq &gt;10</td>
<td>Common with:</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------------------------------------------</td>
<td>------------</td>
<td>-------------</td>
<td>----------</td>
<td>-------------</td>
<td>---------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>J04102</td>
<td>V-ets avian erythroblastosis virus E26 oncogene homolog 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A8007943</td>
<td>KIAA0474 gene product</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z33115</td>
<td>BCL2-like 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AF001461</td>
<td>Core promoter element binding protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Activated T cells</td>
</tr>
<tr>
<td>D14874</td>
<td>Adrenomedullin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M34480</td>
<td>Integrin, α 2β (platelet glycoprotein Ib of IbIIa complex, antigen CD41B)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U97067</td>
<td>Cluster incl U97067: homo sapiens α-catenin-like protein mRNA, complete cds.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M26683</td>
<td>Small inducible cytokine A2 (monocyte chemotactic protein 1, homologous to mouse Sig-je)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA527880</td>
<td>NADH dehydrogenase (ubiquinone 1 beta subcomplex, 7 (18kD, B18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X72308</td>
<td>Small inducible cytokine A7 (monocyte chemotactic protein 3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M63835</td>
<td>Fc fragment of IgG, high affinity 1a, receptor for (CD64)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U28389</td>
<td>Erythrocyte membrane protein band 4-9 (dematin)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U10249</td>
<td>Epithelial protein up-regulated in carcinoma, membrane associated protein 17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L45242</td>
<td>RalA binding protein 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U63631</td>
<td>Solute carrier family 6 (neurotransmitter transporter, creatine), member 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J02973</td>
<td>Thrombomodulin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AF141349</td>
<td>Tubulin, β</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Activated T cells</td>
</tr>
<tr>
<td>D63940</td>
<td>MAX-interacting protein 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U61836</td>
<td>Human putative cyclin G1 interacting protein mRNA, partial sequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AI349593</td>
<td>Hemoglobin, epsilon 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AF141349</td>
<td>Tubulin, β</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Activated T cells</td>
</tr>
<tr>
<td>L66065</td>
<td>MAX dimerization protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AF065338</td>
<td>TSPAN5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z35491</td>
<td>BCL2-associated ananogene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB023211</td>
<td>Peptidyl arginine deiminase, type II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M27492</td>
<td>Interleukin 1 receptor, type 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Activated T cells</td>
</tr>
<tr>
<td>X00737</td>
<td>Nucleotide phosphorylase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N74607</td>
<td>ESTs, Highly similar to AQUAPORIN 3 [H. sapiens]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X17644</td>
<td>G1 to S phase transition 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Activated T cells</td>
</tr>
<tr>
<td>AF156570</td>
<td>Ganglioside expression factor 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Activated T cells</td>
</tr>
<tr>
<td>V00505</td>
<td>Hemoglobin, Δ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Renal failure</td>
</tr>
<tr>
<td>X90999</td>
<td>Hydroxyacetyl glutathione hydrolase; glyoxalase 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D38583</td>
<td>S100 calcium-binding protein A11 (calgizzardin)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Activated T cells</td>
</tr>
<tr>
<td>X79535</td>
<td>Tubulin, β polypeptide</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AI679353</td>
<td>EST</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U76248</td>
<td>Seven in absentia (Drosophila) homolog 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA131149</td>
<td>S100 calcium-binding protein B18 (calgranulin A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Activated T cells</td>
</tr>
<tr>
<td>AB018293</td>
<td>KIAA0750 gene product</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K00635</td>
<td>V-fos FBJ murine osteosarcoma viral oncogene homolog</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Activated T cells</td>
</tr>
<tr>
<td>U00630</td>
<td>Plasminogen activator inhibitor, type II (arginine-serpin)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U89066</td>
<td>Pyridoxal (pyridoxine, vitamin B6) kinase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AL049250</td>
<td>Homo sapiens mRNA; cDNA DKFp564D113 (from clone DKFzp564D113)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M36820</td>
<td>GRO2 oncogene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X58851</td>
<td>Myosin, light polypeptide 4, alcali; atrial, embryonic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S77763</td>
<td>Nuclear factor (erythroid-derived, 25kD, 43kD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M60298</td>
<td>Erythrocyte membrane protein band 4-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X90857</td>
<td>Conserved gene telomeric to alpha globin cluster</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K02401</td>
<td>Chorionic somatomammetropsin hormone 1 (placental lactogen)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M25329</td>
<td>Selectin P (granule membrane protein 140kD, antigen CD62)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X14787</td>
<td>Thrombospondin 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D83664</td>
<td>S100 calcium-binding protein A12 (calgranulin C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M38690</td>
<td>TSPAN9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T77016</td>
<td>Tropomodulin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y97324</td>
<td>Adipose differentiation-related protein, adipophillin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* na = not available.  
* EST = expressed sequence tag.
Probing the Molecular Basis for RCC Disease Transcripts in Peripheral Blood. We next compared the differentially expressed transcripts in RCC PBMCs to the 10 transcripts most strongly up-regulated in RCC tumors (n = 47) relative to normal kidney tissue (n = 60), using profiles downloaded from the Bioexpress Database (GeneLogic). The RCC tumor-specific transcripts that possessed the highest average fold differences in expression between RCC tumor tissue and normal kidney were unchanged between normal and RCC PBMCs (data not shown), suggesting that shed RCC tumor cells did not contribute significantly to the disease-associated transcripts identified in PBMCs isolated from RCC patients.

In a second approach, we compared the differentially expressed genes in RCC PBMCs with genes differentially expressed between unstimulated CD4+ T cells (n = 3 normal donors) and CD4+ T cells (n = 3 normal donors) stimulated ex vivo with anti-CD3 and anti-CD28 in culture. Stimulated CD4+ T cells possessed 14 transcripts that were greater than 2-fold changed in the same direction (induced or repressed) as the disease-associated transcripts in RCC PBMCs (Table 1).

In a third approach, we compared the differentially expressed genes in RCC PBMCs with genes differentially expressed between PBMCs from non-RCC end-stage renal failure patients (n = 9 individuals) and PBMCs from normal volunteers (n = 4 individuals). Of these, nine transcripts differentially expressed in PBMCs from renal failure patients were also disease-associated transcripts in RCC PBMCs (Table 1). Thus, our marker gene list from PBMCs of RCC patients contains a subset of markers commonly involved in immune responses measured ex vivo (CD4+ T-cell activation) and in responses of circulating leukocytes to renal dysfunction observed in vivo.

Classification of RCC and Normal Status Using Patterns of Expression in Peripheral Blood Cells. We next sought to apply our results by determining the ability of a minimal gene set(s) to classify RCC versus normal status using expression patterns in the peripheral blood. To initially build and subsequently train the classifiers, 70% of the RCC PBMCs (n = 31) and 70% of the normal PBMCs (n = 14) were selected randomly and used as the training set. We used the Genecluster default correlation metric (9) to identify genes with expression levels most highly correlated with the classification vector characteristic of the training set. All 5249 genes meeting the main filter criteria were screened using this approach.

Prediction was also performed in Genecluster using the weighted voting method. In this method, the expression level of each gene in the classifier set contributes to an overall vote on the classification of the sample (10). The prediction strength is a combined variable that indicates the support for one class or the other and can vary between 0 (narrow margin of victory) and 1 (wide margin of victory) in favor of the predicted class.

Predictor sets containing between 2 and 20 genes were evaluated by leave-one-out cross-validation to identify the predictor set with the highest accuracy for classification of the samples in the training set (Fig. 2A). An eight-gene classifier set containing the top genes up-regulated in RCC and the four top genes down-regulated in RCC was found to yield the highest cross-validation prediction accuracy on the training set. The relative expression levels for this eight-gene classifier set were shown in Fig. 2B, and the individual prediction confidence scores for each sample in the training set using this eight-gene classifier set are presented, in the same order, in Fig. 2C. For illustrative purposes, we assigned a positive sign to the prediction strengths resulting in votes for RCC and a negative sign to prediction strengths resulting in votes for normal PBMCs.

Finally, we classified the remaining test set of samples using the eight-gene classifier. On the test set, the predicted class matched the true class in all cases, although for 1 of the 19 test samples the prediction strength was negligible (Fig. 2D). These studies demonstrate the feasibility of predicting RCC versus normal status using expression patterns found in a limited number of gene transcripts in mononuclear cells from peripheral blood.

DISCUSSION

One of the objectives of clinical pharmacogenomic studies is to determine whether accessible (often nontarget) tissues such as peripheral blood can be used to identify surrogate markers of disease in vivo. Overall, clinical pharmacogenomics is a relatively new discipline that seeks to mine transcriptional expression profiles in surrogate tissues from human patients for the multifold purpose of identifying three main types of diagnostics: (a) markers of disease; (b) markers of drug exposure; and (c) markers of drug efficacy. The successful identification of each type of diagnostic is predicated on several implicit assumptions, respectively: (a) that surrogate tissues will exhibit sufficiently robust and reproducible alterations in transcription between diseased and normal states; (b) that pharmacological intervention will modify any transcriptional alterations observed in surrogate tissues in the disease state; and (c) that the transcriptional profiles of surrogate tissues in disease states before or during therapy will be related to whether the patient will respond to pharmacological intervention. In the field of oncology, it is, therefore, of interest to determine whether transcriptional profiles from the clinically accessible compartment of peripheral blood in patients with solid tumors will provide disease-associated transcripts that can be monitored in future pharmacogenomic studies.

In the present study we identified 132 genes with transcripts detected as present in at least 15% of the samples that varied greater than 2-fold between RCC and normal PBMCs with significance at the 99.9% confidence level or greater. Among the most significantly elevated transcripts (P < 10^-7) in RCC PBMCs were several inflammatory-related genes, including Toll-like receptor 2, galectin-3, interleukin 1 receptor antagonist, and aquaporin-9, a water channel implicated in leukocyte migration. The unchanged levels of many other cytokines, chemokines, and their respective receptors between normal and RCC PBMCs suggested that a specific, rather than global, activation of PBMCs constituted an important part of the disease signature in RCC peripheral blood.

It is interesting to note that the vast majority of the transcripts detected as significantly changed in PBMCs from RCC patients also possessed the highest variability across the RCC PBMC profiles. This finding indicates that although the levels of these transcripts were significantly distinct from levels in normal PBMCs, there was relative heterogeneity of expression of these transcripts across RCC patients. This observation is similar to the recent finding by Whitney et al. (11) that variation in normal PBMC expression is lower than the variation in patients with diffuse large B-cell lymphomas and chronic lymphocytic leukemias. The results here demonstrate that this may also be the case for diseases of solid tumors. It will be of great interest to determine whether any of these disease-associated, yet highly variable, transcripts in RCC PBMCs will be correlated with any clinical categories of response, once clinical indices of outcome and follow-up are satisfactorily measured in these patients.

Our comparisons clearly identified transcripts, the expression levels of which were altered in PBMCs from RCC patients relative to PBMCs of disease-free volunteers. However, it was unknown whether PBMC profiles of patients with RCC tumors of different origin might be further distinct. In a subsequent analysis, we used a multiclassification approach to determine whether transcripts in PBMCs from RCC patients were significantly distinct between patients with renal tumors of different cellular origin (clear cell versus mixed versus...
papillary). Significance testing by random permutation revealed that no transcripts were significantly correlated ($P < 0.05$) in PBMCs of patients with specific tumor types within the RCC population (data not shown). Therefore, in this small study, expression profiles in the surrogate tissue PBMCs were not sufficiently distinct to allow classification of RCC tumor type on the basis of PBMC transcriptional profiles alone.

Previous studies have used high-sensitivity arrays or RT-PCR in the peripheral blood to diagnose tumor status (12, 13), based on the detection of transcripts derived from metastatic cells in the circulation. Although this was unlikely to contribute to the disease-associated gene set in these studies, we nonetheless formally explored this possibility and other potential biological bases for the disease associated gene set in PBMCs of RCC patients. None of the most highly induced transcripts in RCC tumors were detected in PBMCs from RCC patients, consistent with the hypothesis that shed tumor cells did not contribute to the disease-associated transcripts observed in RCC PBMCs. We also compared RCC PBMC profiles measured in vivo with those of ex vivo-activated CD4$^+$ T cells in culture. RCC has been characterized as an immunogenic tumor similar to melanoma (14, 15). Comparison of profiles measured in CD4$^+$ T cells stimulated ex vivo with anti-CD3/anti-CD28 antibodies to profiles measured in PBMCs isolated from RCC patients identified 14 transcripts commonly induced or repressed 2-fold or more. The genes commonly regulated in RCC PBMCs and in activated CD4$^+$ T cells often showed quite significant induction ex vivo. These results support a hypothesis that the expression of at least a subset of the disease-associated genes observed in RCC PBMCs may result from an activation of circulating T cells and/or other leukocytes in response to the presence of the tumor. In a final comparison, we also identified another set of nine genes that were commonly regulated in PBMCs from advanced RCC patients and in PBMCs from patients with end-stage renal failure.
Because the patients in this trial possessed adequate renal function (serum creatinine <1.5 X upper limit of normal or a calculated creatinine clearance >60 ml/min) at the time of trial entry, we anticipated that very few, if any, alterations in PBMC profiles would be due to renal dysfunction. It is, however, possible that the regulation of this small subset of disease-associated transcripts detected in RCC PBMCs could be due to alterations in leukocyte expression profiles in response to early (as yet undetectable) renal dysfunction onset in the RCC patients.

Ongoing studies in our laboratory have demonstrated that PBMCs from RCC patients can be accurately distinguished not only from PBMCs of normal volunteers but also from PBMCs of patients with other types of solid tumors (prostate and head and neck cancer; data not shown). Using a multiclass approach, we have predicted RCC status using profiles in PBMCs with moderately high accuracy (70%), comparable with the overall accuracy achieved by Ramaswamy et al. (16) across a large database of primary tumor biopsy profiles. If these preliminary results are confirmed in a larger patient population and across multiple tumor types in individuals with nonadvanced cases of disease, it is possible that expression profiles in PBMCs could ultimately be used for diagnostic purposes.

As clinical pharmacogenomic analyses gain acceptance and become more commonplace in clinical trials, it is increasingly evident that microarrays will eventually be used as diagnostic devices. Recent consideration has been given to the use of microarrays as medical devices (17). One of the important issues will be to establish a rigorous and numerically-based method for reporting expression “pattern” results from a diagnostic assay and how an associated reference range for that pattern will be calculated and reported (18). We are currently using the weighted voting method described herein to collapse expression pattern results from many genes into a single numerical confidence score (9). One important advantage of this method is that it reports a prediction strength score, indicative of the confidence in the prediction for each patient. A confidence threshold can, thus, be established to optimize the accuracy of prediction and minimize the incidence of both false positive and false negative results. In the future, average confidence scores collected for the accumulating pool of correctly diagnosed patients and correctly nondiagnosed disease-free individuals could be calculated, and a reference range of values, for the particular predictive gene set diagnostic in question, could be reported. Alternatively, a reference disease-free RNA standard could be run alongside the clinical RNA sample in question, although the requirements and the source of such a standard remain to be defined.

In summary, the present study has established that there appear to be disease-associated genes in the PBMCs of patients with RCC. On the basis of these data, it is possible that because PBMCs circulate throughout the bodily tissues, their expression profiles may serve as a sensitive indicator and physiological monitor of disease and health. If additional analyses bear these findings out, the genes identified here represent the foundation on which to build disease-specific gene sets that can be used as part of a molecular diagnosis of disease using peripheral blood. We suggest that global expression profiling of the peripheral blood will identify gene sets of limited size that may ultimately be developed into clinical assays. Additional research and larger patient populations will ultimately determine the exact identities of the transcripts in circulating leukocytes with the greatest predictive power in clinical diagnosis, and establish the limits and caveats associated with the ability to predict/distinguish each disease type using expression profiles in peripheral blood.

ACKNOWLEDGMENTS

We thank the many patients and volunteers who donated clinical samples for these studies. Expert technical assistance was provided by Christine Reilly. We also thank Andrew Hill, Fred Immerrman, William Mounts, Andrew Strahs, and Maryann Whitley for many thoughtful discussions and assistance in data analysis and experimental design throughout these studies.

REFERENCES

Disease-associated Expression Profiles in Peripheral Blood Mononuclear Cells from Patients with Advanced Renal Cell Carcinoma


Updated version Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/63/18/6069

Cited articles This article cites 15 articles, 7 of which you can access for free at:
http://cancerres.aacrjournals.org/content/63/18/6069.full.html#ref-list-1

Citing articles This article has been cited by 11 HighWire-hosted articles. Access the articles at:
/content/63/18/6069.full.html#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.