A Macrophage Gene Expression Signature Defines a Field Effect in the Lung Tumor Microenvironment

Robert S. Stearman,1 Lori Dwyer-Nield,2 Michael C. Grady,1 Alvin M. Malkinson,2 and Mark W. Geraci1

Departments of Medicine/Pulmonary Sciences and Critical Care Medicine, and Pharmaceutical Sciences, University of Colorado Health Sciences Center, Denver, Colorado

Abstract

One area of intensive investigation is to understand complex cellular and signaling interactions in the tumor microenvironment. Using a novel, although straightforward, microarray approach, we defined a gene expression signature from the lung tumor microenvironment in the murine A/J-urethane model of human lung adenocarcinoma. The tumor microenvironment is reflected by the composition of the cell types present and alterations in mRNA levels, resulting in a “Field Effect” around the tumor. The genes composing the Field Effect expression signature include proteases and their inhibitors, inflammation markers, and immune signaling molecules. By several criteria, the Field Effect expression signature can be attributed to the macrophage lineage, suggesting a qualitative change in the expression pattern of tumor-associated macrophages (TAM) observed in lung tumors. The protein expression levels for a number of Field Effect genes were verified by Western blot analysis of lung homogenates, and for their expression in macrophages and parenchymal cells outside of the tumors by immunohistochemistry. In addition, the Field Effect expression signature was used to classify bronchoalveolar lavage (BAL) cells from tumor-bearing or age-matched control mice. Using a variety of statistical measures, the Field Effect expression signature correctly classified the BAL cells >94% of the time. Finally, the protein levels for several Field Effect genes were higher in cell-free BAL fluid, indicating they may be secreted by the TAMs. This work suggests that TAMs generate a unique gene expression signature within the tumor microenvironment, and this signature could potentially be used for identifying lung cancer from BAL cells and/or fluid. [Cancer Res 2008;68(1):34–43]

Introduction

Lung cancer continues to be the number one cause of cancer mortality in the United States, with >210,000 new cases and 160,000 deaths estimated for 2007 (1). Tobacco smoke is the most widespread risk factor accounting for lung cancer in 85% to 90% of cases (2, 3). Environmental risks include occupational exposures to asbestos, uranium, and other mining activities, second-hand tobacco smoke, and household radon. In a study of Iowa women with (cases) or without (controls) lung cancer, the greatest risk factors were current or ex-smoker status (adjusted odds ratio, 13.5), and an elevated odds ratio was present even 25 years after smoking cessation (4). Recently released data (1) covering 1975 to 2002, showed a relatively constant 5-year survival rate of ~15% after lung cancer diagnosis, independent of race or gender, with a higher 5-year survival rate (49%) if early stage disease is discovered before detectable lymph node or metastasis involvement. In contrast, colorectal cancer had an overall 5-year survival rate of 64% over the same study period. Although lung cancer is the most common cancer affecting both sexes, it had the third lowest 5-year survival rate, exceeding only liver and pancreatic cancers, which account for ~50,000 cases per year combined.

Survival studies indicate that early detection of lung cancer as stage I disease is the most significant factor in providing the highest 5-year survival rate (1, 5, 6). Unfortunately, there are no effective low-cost options available for general annual screening of the tobacco smoking population, estimated at 20% to 25% of the adults in the United States (1). Two recent studies examining the utility of spiral computed tomography as an annual screening method had conflicting results, although they used different clinical outcome measures (7, 8). Both studies did agree that spiral computed tomography does increase early detection and treatment of lung cancer in the easily identifiable at-risk population (a smoking history and/or occupational exposures; age >40 years). The significance of improved survival rate versus decreased mortality must await the results from two randomized controlled trials currently under way by a Dutch-Belgian collaboration and the National Cancer Institute.

One promise of gene expression microarray studies in lung cancer has been to develop novel approaches to diagnostics for disease classification and patient stratification into high probability treatment outcome groups (9–12). This overall strategy has found considerable success in the treatment of breast cancer, moving from research tool to clinical diagnostics, to identify best treatment outcome groups (9–12). This overall strategy has found considerable success in the treatment of breast cancer, moving from research tool to clinical diagnostics, to identify best treatment outcome and recurrence using a 70-gene panel (MammaPrint; ref. 13). Recent reports suggest a similar approach is feasible using lung cancer tissue to identify high- and low-risk groups for recurrence and survival (14–16), although these approaches do not address early detection methodologies.

As a first step toward identifying biomarkers of early lung cancer, we used the well-characterized A/J mouse-urethane model of human adenocarcinoma (12, 17). We reasoned that histologically appearing, normal lung tissue from urethane-treated mice would have an altered gene expression profile because of proximity to the tumor microenvironment, when compared with normal lung tissue from age-matched untreated animals. The altered gene expression could be derived from lung parenchymal cells and/or immune cells infiltrating and responding to the tumor. In lung cancer, the tumor microenvironment has been extensively studied, especially with

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Requests for reprints: Robert S. Stearman, Box C272, Room BB 3B10, 4200 East Ninth Avenue, Denver, CO 80262. Phone: 303-315-2317; Fax: 303-315-5632; E-mail: Robert.Stearman@uchsc.edu.

©2008 American Association for Cancer Research.
doi:10.1158/0008-5472.CAN-07-0988
regard to a wide variety of chemokine (C-C motif) signaling pathways after activation of nuclear factor-κB (18–20). Chemotactic factors produced by tumors recruit blood-derived monocytes, and within the tumor microenvironment, they mature to tumor-associated macrophages (TAMs; refs. 21–23).

TAMs play a central role in many tumor-stromal interactions including angiogenesis, extracellular matrix remodeling, and invasion/metastasis (21, 24, 25). The degree of macrophage infiltration into human non–small cell lung cancers is correlated with the microvessel counts and inversely correlated with survival (21). In addition, alveolar-derived macrophages isolated by bronchoalveolar lavage (BAL) from lung cancer patients have altered in vitro characteristics (26). In transgenic mouse models, elevating the number of pulmonary macrophages overexpressing FGF10 in type 2 and Clara cells induced spontaneous development of lung tumors in the absence of exogenous carcinogens (27). Finally, chromosomal sites containing genes that regulate macrophage behavior map to lung cancer susceptibility sites (28). TAMs receive signals from diverse cell types within the tumor microenvironment and release signals that affect a variety of different cell types, including epithelial and endothelial cells, fibroblasts, and lymphocytes (24, 29). TAMs alter epithelial cell phenotype by stimulating proliferation, inhibiting apoptosis, and/or changing their morphology/differentiation status so that initiated epithelial cells adapt a more motile and less sessile configuration (30). Macrophages are phenotypically heterogeneous, and extremes of this continuum are called M1 (classic activation) and M2 (alternative activation; refs. 23, 29). M1 macrophages are stimulated by IFNγ and microbial products (lipopolysaccharide), are bactericidal, and express inducible nitric oxide synthase (iNOS). The M2 state is induced by the cytokines interleukin (IL)-4 and IL-13, produces polyamines, lymphocyte suppression, and arginase expression. In A/J mice, alveolar macrophage phenotype changes from M0 (expressing lymphocyte suppression, and arginase expression) to M1 (iNOS expressing) in mice with carcinomas (22).

In this study, we compared the gene expression profiles of normal-appearing lung tissue adjacent to tumors with age-matched, untreated controls, and we identified a set of 46 genes specifically altered by the tumor microenvironment. Supporting the observed biology of TAMs, these genes are highly expressed in the monocytic-macrophage cell lineage. The increased protein expression for a number of these candidates was verified by immunohistochemistry of lung cancer tissue. Importantly, these genes had strong predictive value when used to classify expression data from BAL-derived macrophages from urethane-treated A/J mice in an independent study. These results suggest that gene expression information contained within samples from outside the tumor itself (surrogate tissues) are informative for predicting tumor status, and that TAMs play a central role in lung tumorigenesis.

Materials and Methods

Expression levels of specific genes verified by quantitative real-time PCR or Western blot analyses. The following genes were verified by the methods described below and are listed here using their standardized gene symbol notation3 in alphabetical order for reference: acid phosphatase 5 (Acp5; tartrate resistant), C-C motif ligand 6 (Ccl6), CD68 antigen (Cd68), chitinase 3-like 1 (Chi3l1), cathespins D (Ctsd), cathespins K (Ctsk), cathespins Z (Ctsz), coagulation factor VII (F7), lipoprotein lipase (Lpl), leucine-rich α-2-glycoprotein 1 (Lrg1), lymphocyte antigen 75 (Lyt75), nephroblastoma overexpressed gene (Nov), prostat glandin-endoperoxide synthase 1 [Ptgs1; cyclooxygenase (COX)1], signal-regulatory protein α (previously designated Ptpn1; Sirpα), secreted phosphoprotein 1 (osteopontin; Spp1).

Murine model of human lung cancer. As previously reported (12), male A/J mice (6–8 weeks old; The Jackson Laboratory) were given a single 1-mg/g i.p. injection of urethane or saline vehicle (age-matched untreated controls). At two time points (early, 24–26 weeks; late, 42 weeks) after urethane or saline treatment, mice were sacrificed by lethal pentobarbital injection, and their lungs excised for examination and molecular characterization. Tumors were dissected from neighboring tissue using a dissection microscope, and the adjacent tissue was placed in RNAlater (Ambion) for microarray analysis or in 20 mmol/L HEPS (pH 7.4), 2 mmol/L EDTA, 2 mmol/L EGTA, 1 mmol/L DTT, 10% glycerol, 5 μg/mL aprotinin, 10 μmol/L leupeptin, and 1 mmol/L 4-(2-aminoethoxy) benzenesulfon fluoride for protein analysis. Lungs from age-matched untreated controls underwent similar processing.

RNA isolation and microarray analysis. RNA was isolated from lung tissues or BAL derived cells (below) using Trizol Reagent (Invitrogen) and RNeasy kits (Qiagen), followed by quality control measures using UV spectra characteristics (NanoDrop) and Bioanalyzer (Agilent) for size and integrity of the total RNA (12). Total RNA (2–5 μg from lung tissues or 0.25–1 μg from BAL cells) was converted to fragmented biotin-labeled cRNA by the Affymetrix kit reagents according to the manufacturer (Affymetrix). Labeled cRNAs were hybridized to MG-U74Av2 (lung tissue samples) or MOE430 2.0 (BAL samples) Affymetrix microarrays were washed, developed, and scanned using Fluidics and Scanner Workstations. The lung tissue microarray dataset consisted of four possible different samples: normal lung tissues or tissues adjacent to tumor from two different time points, 24 to 26 (early) or 42 (late) weeks after urethane injection. The number of replicate mice used for lung tissue microarrays was 5 early normal, 8 early adjacent, 7 late normal, and 7 late adjacent. Data quality characteristics of the MG-U74Av2 microarrays were as previously reported (12). The adjacent tissue datasets were deposited at Gene Expression Omnibus3 (31) under the series accession GSE2514 (12). Microarrays completed from BAL-derived cells were from 7 control and 11 urethane-treated A/J mice. Two control BAL RNA samples were combined (C6_C13 BAL) to give sufficient total RNA for labeling, and 1 control sample (C36; indicated with an asterisk) was later identified as from a mouse with a spontaneous lung tumor. The MOE430 2.0 microarrays had an average background intensity of 77 (SD = 64), and average percent present called 43% (SD = 5.5%). Affymetrix probe IDs from the various microarrays used in this analysis were annotated and cross-referenced using NetAffx (32). The new microarray datasets described in this work, including.cel and .exp files, have been deposited at the Gene Expression Omnibus under the series accession GSE7269. Processing and statistical analysis of the microarray datasets was completed using BRB-ArrayTools (v3.4.0a; 6/2006) freely available from Dr. Rich Simon (33).5 The BRB-ArrayTools suite makes extensive use of parametric and permutation analyses, as well as estimation of significance by False Discovery Rate. Classifiers are crossvalidated by a “leave one out” strategy by several different methods. More limited sets of probeIDs intensity data were analyzed using Excel XP and GraphPad Prism 4. Cluster and heatmap diagrams were generated in BRB-ArrayTools, which implements the algorithms according to Eisen et al. (34), using centered genes, centered correlations, and average linkage settings.

Measurement of RNA levels by quantitative real-time PCR. All reagent kits, primer/probe sets, and the GeneAmp 5700 Instrument were from Applied Biosystems, Inc. and used according to the manufacturer’s instructions. Briefly, 100 ng of total RNA was converted to cDNA using the

---


5 http://louis.nicl.gov/BRB-ArrayTools.html
High Capacity cDNA Reverse Transcription kit in 100 μL reaction and diluted 2-fold with water. Five microliters of cDNA was used per primer/probe (1× concentration) set reaction in TaqMan Universal PCR MasterMix (20 μL final volume) and run for 40 cycles under standard conditions. Mouse β-actin endogenous control primer/probe set was used on all samples and a plasmid DNA dilution standard curve validated amplification efficiencies (r = 0.99; >5 replicates). C\textsubscript{T} values are averages of RNA derived from 5 age-matched or 12 urethane-treated A/J mouse BAL cells, each sample run in duplicate. ΔC\textsubscript{T} value of each gene was calculated relative to its β-actin (lower ΔC\textsubscript{T} values indicate more abundant transcripts) and fold increase observed in urethane treated versus controls as 2^(-ΔΔC\textsubscript{T}).

**Protein biomarker analysis and immunohistochemistry.** Protein extract preparation, Western methodology, and antibodies used are described in the Supplementary Method. For immunohistochemistry, A/J mouse lungs were perfused by instilling saline through the pulmonary artery to clear the blood from the lungs. Lungs were inflated with 10% buffered formalin before formalin fixation. Fixed lungs were paraffin embedded and 4 μm sections were cut and affixed to glass microscope slides. Tissue sections were rehydrated, endogenous peroxidases quenched in 1% H\textsubscript{2}O\textsubscript{2} in methanol, and subjected to antigen retrieval in 100 mmol/L sodium citrate buffer in a steamer. After blocking in PBS containing 10% of the appropriate blocking serum, sections were incubated with primary antibody (Supplementary Method) for 90 min in a humidified chamber at 37°C. Proteins were visualized using Vectastain kits according to the manufacturer’s instructions (Vector Laboratories) and diaminobenzidine.

**BAL.** The tracheas were cannulated with an 18-gauge angiocatheter, and 1 mL PBS-0.6 mMol/L EDTA was instilled and aspirated back into the syringe as described previously (35). For cell isolation, lungs were rinsed 5 times with 1 mL PBS, the rinses combined and centrifuged at 500 × g for 10 min. The rinses were then aspirated off, leaving ~500 μL per sample, and the cells transferred to a 1.5 mL microfuge tube, pelleted at 16,000 × g for 1 min, and directly lysed in 400 μL Trizol Reagent (Invitrogen) for RNA purification. For protein analysis, BAL fluid (~600 μL/mouse) was centrifuged at 16,000 × g to remove cells before being concentrated in a Microcon (Millipore) concentrator with a 3 kDa cutoff to remove cells before being concentrated in a Microcon (Millipore) concentrator with a 3 kDa cutoff to ~50 μL/sup. Protein concentration was determined and samples were diluted into protein-loading buffer. Insufficient protein amounts of BAL samples from vehicle-treated mice did not allow comparable SDS-PAGE analyses, so only the samples from 24- and 42-week urethane-treated animals were analyzed for Ctsd, Ctsz, Chi3l1, Spp1, and Sirpα content. Four separate replicate mice were used to produce BAL fluid protein samples. The BAL protein yield from the 24-week urethane-treated mice was substantially lower than that from the 42-week urethane-treated mice. Immunoblot analysis was performed as described above, except that blots were stripped with 1% Tris-HCl (pH 6.7), 0.8% β-mercaptoethanol, and 2% SDS at 50°C for 30 min before reuse.

**Results**

We and others have used microarray technology to reveal expression differences between lung cancer and its surrounding tissue, with the goals of tumor classification, therapeutic responsiveness, recurrence, and murine modeling of this disease state (10-12, 36, 37). As an extension of the analysis of the A/J-urethane murine model of human lung cancer (12), our approach here was to ascertain the influence of the tumor microenvironment on gene expression in adjacent tissue by comparison to age-, sex-, and environmental exposure-matched normal control A/J mice (Supplementary Fig. S1). Given the extensive literature on the interrelationships between the roles of inflammation processes, immune and stromal cells, and tumorigenesis (19, 21, 28), we hypothesized that a novel gene expression signature would be found in what is histologically appearing normal lung tissue as a result of the influence of the tumor microenvironment (herein called the “Field Effect”). The tumor microenvironment could produce a specific gene expression signature due to immune cell infiltration, endothelial cell alterations (e.g., angiogenesis), and/or the variety of cellular responsiveness to signaling pathways. In addition, a Field Effect signature may aid in biomarker discovery from a more distantly removed surrogate tissue source (38) such as BAL.

**Developing an expression signature of the tumor microenvironment.** Our microarray dataset consisted of normal or adjacent to tumor A/J mouse lung tissues from two different time points, 24 to 26 or 42 weeks after urethane injection (early versus late time points). Using a four-way classifier algorithm in BRB-ArrayTools (testing for significance in four categories: normal, adjacent, early, and late), a total of 204 probeIDs were identified with parametric P values of ≤0.0001 (see supplementary Excel Spreadsheet for complete signal intensity and annotation data). On the MG-U74Av2 microarray, with ~12,000 probeIDs, <2 probeIDs would be expected by random chance at this level of significance. As shown in a cluster dendrogram (Fig. 1A), the 204 probeID classifiers generate a tree diagram with 2 main branches [early (yellow) versus late (black) samples; Age subdivided by proximity to lung tumors [normal (green) versus adjacent (blue); Field]. Two criteria were jointly used to designate whether a probeID was classifying samples based on their proximity to lung tumor (Field) or because the animals in this study were used at two time points (Age). First, on a probeID basis, was the Student’s t test for significant differences in the log\textsubscript{2} signal intensities with a P value of <0.001, when comparing both potential pairs of samples (normal versus adjacent or early versus late). Second, which t test comparison was smaller (i.e., more statistically significant), the Field or Age comparison. Using both comparisons, the 204 probeIDs were divided into 48 Field probeIDs (24%; see Table 1 for listing of Field Effect genes), 139 Age probeIDs (68%), and 17 not clearly separated by these criteria into either category (8%). The Field probeIDs had significant absolute signal intensities (>64 arbitrary units) and each signal intensity distribution of the probeID was separable by sample type (Supplementary Fig. S2). If the Field Effect genes are reflective of the tumor microenvironment, the magnitude of the difference in signal intensities (adjacent-normal) should increase as the tumor becomes more aggressive at the later time point. A ΔSignal Intensity graph (Supplementary Fig. S3) shows that 25 of the 43 up-regulated genes in the tumor microenvironment were significantly increased, whereas 3 of the 5 down-regulated genes were significantly decreased at the late time point.

The Field and Age t test comparisons gave unambiguous categorization, as the ratio of t tests (e.g., Field t test P value/Age t test P value) were clearly different (median Field probeIDs t test ratio, 30,000; median Age probeIDs t test ratio, 78,000). To visualize the significance of these two sets of probeIDs, supervised clustering of the microarray datasets clearly identified the samples appropriately [Fig. 1B (Field) and C (Age)]. The expression level heatmap of the Field probeIDs shows the majority of genes (90%) are up-regulated (red) in the adjacent tissues, suggesting the presence of a novel gene expression signature (Fig. 1D).

**Assessment of cell lineage characteristics within the tumor microenvironment.** Two in silico analyses were done to ascertain if the Field genes can be attributed to a likely candidate for cellular origin. The murine Field probeIDs were converted to their appropriate human ortholog probeIDs for querying the SymAtlas database of normal human tissue expression (39). Each probeID was scored for its expression level by tissue type and marked as
“overexpression” if the signal intensity of the tissue was significantly higher than its median level (Supplementary Fig. S4). A tissue was scored as overexpressing a gene when the signal intensity was greater than the median of all tissues plus thrice the SD of the measured signal intensities. The Field probeIDs were highly expressed in whole blood, specifically peripheral blood CD14+ monocytes and bone marrow CD33+ myeloid cells, the precursors to tissue matured macrophages. As a further test, the murine probeID orthologs of the human immune cell type specific genes, characterized by Du et al. (40), were used for supervised clustering of the A/J normal and adjacent microarray datasets (Supplementary Fig. S5). In no case were the major immune cell gene clusters able to distinguish the normal from adjacent samples. These results suggest a unique myeloid—monocyte—macrophage lineage for the cells found in the adjacent lung tissue as responsible for the gene expression signature.

Figure 1. Microarray analysis of the A/J-urethane lung cancer model identifies Field Effect classifiers. A, RNAs from murine A/J-urethane lung tissues from either adjacent (A or A42) or age-matched normal controls (N or N42) were analyzed using MG-U74Av2 microarrays (Affymetrix) and the data were classified using a four-way comparison for age (24 or 42 wk after urethane treatment) or source of tissue (adjacent or normal control). The resulting 204 probeID classifier was used for supervised clustering of the datasets and displayed in a dendrogram format. Color coding bars are used to indicate tissue source (green, normal; blue, adjacent) and age (yellow, early, 24 wk postinjection; black, late, 42 wk postinjection). The datasets are distributed into two main branches by age (early versus late), and then each branch subdivided by location of tissue (adjacent versus normal). B, the microarray datasets were analyzed by supervised clustering using the Field Effect probeIDs (n=48; adjacent versus normal) identified from the four-way comparison classifier (n=204). The data are displayed in a dendrogram format with color coding as in Fig. 1A. Two main branches clearly divide the microarray datasets by adjacent and normal lung tissues. C, the microarray datasets were analyzed by supervised clustering using the Age Effect probeIDs (n=139; early, 24 wk postinjection; late, 42 wk postinjection) identified from the four-way comparison classifier (n=204). The data are displayed in a dendrogram format with color coding as in Fig. 1A. Two main branches are found, clearly dividing the microarray datasets by early and late lung tissues. D, gene expression heatmap displays the degree of regulation for the Field Effect probeIDs with the probeIDs down-regulated in the adjacent tissue listed as the top five rows followed by the up-regulated probeIDs. Maximum color saturation was set at ~3-fold change in gene expression. The data are displayed above the heatmap as a dendrogram with the color coding as in Fig. 1A.
### Table 1. Field effect genes

<table>
<thead>
<tr>
<th>ProbeID</th>
<th>Description</th>
<th>Gene symbol</th>
<th>Fold change (adjacent/normal)</th>
<th>Parametric P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>160564_at</td>
<td>Lipocalin 2</td>
<td>Lcn2</td>
<td>4.26</td>
<td>1.57E−05</td>
</tr>
<tr>
<td>97420_at</td>
<td>Leucine-rich α-2-glycoprotein 1</td>
<td>Lrg1</td>
<td>3.65</td>
<td>2.00E−07</td>
</tr>
<tr>
<td>98467_at</td>
<td>Inter α-trypsin inhibitor, heavy chain 4</td>
<td>Itih4</td>
<td>3.19</td>
<td>7.10E−06</td>
</tr>
<tr>
<td>101912_at</td>
<td>RIKEN cDNA 1100001G20 gene</td>
<td>1100001G20Rik</td>
<td>2.98</td>
<td>&lt;1E−07</td>
</tr>
<tr>
<td>160406_at</td>
<td>Cathepsin K</td>
<td>Ctsk</td>
<td>2.61</td>
<td>1.30E−06</td>
</tr>
<tr>
<td>92849_at</td>
<td>Chemokine (C-C motif) ligand 6</td>
<td>Ccl6</td>
<td>2.53</td>
<td>9.20E−06</td>
</tr>
<tr>
<td>100507_at</td>
<td>Nephroblastoma-overexpressed gene</td>
<td>Nov</td>
<td>2.36</td>
<td>2.00E−07</td>
</tr>
<tr>
<td>97519_at</td>
<td>Secreted phosphoprotein 1</td>
<td>Spp1</td>
<td>2.16</td>
<td>2.70E−06</td>
</tr>
<tr>
<td>104308_at</td>
<td>Integrin α X</td>
<td>Itgax</td>
<td>2.11</td>
<td>1.36E−05</td>
</tr>
<tr>
<td>96787_at</td>
<td>Serine (or cysteine) peptidase inhibitor, clade A (α-1 antiprotease, antitrypsin), member 10</td>
<td>Serpina10</td>
<td>2.07</td>
<td>8.99E−05</td>
</tr>
<tr>
<td>95957_at</td>
<td>Prostaglandin-endoperoxide synthase 1</td>
<td>Pigs1</td>
<td>2.07</td>
<td>6.40E−06</td>
</tr>
<tr>
<td>92918_at</td>
<td>Coagulation factor VII</td>
<td>F7</td>
<td>2.03</td>
<td>3.20E−06</td>
</tr>
<tr>
<td>98859_at</td>
<td>Acid phosphatase 5, tartrate resistant</td>
<td>Acp5</td>
<td>1.97</td>
<td>2.86E−05</td>
</tr>
<tr>
<td>162198_f_at</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>103258_at</td>
<td>Lymphocyte antigen 75</td>
<td>Ly7y</td>
<td>1.94</td>
<td>&lt;1E−07</td>
</tr>
<tr>
<td>95784_at</td>
<td>Paired-lg–like receptor A1</td>
<td>Pir1</td>
<td>1.86</td>
<td>2.23E−05</td>
</tr>
<tr>
<td>93354_at</td>
<td>Apolipoprotein C-1</td>
<td>Apoc1</td>
<td>1.86</td>
<td>1.70E−05</td>
</tr>
<tr>
<td>103016_s_at</td>
<td>CD68 antigen</td>
<td>C6d8</td>
<td>1.82</td>
<td>5.58E−05</td>
</tr>
<tr>
<td>97413_at</td>
<td>RIKEN cDNA 1600029D21 gene</td>
<td>1600029D21Rik</td>
<td>1.78</td>
<td>3.00E−07</td>
</tr>
<tr>
<td>92762_at</td>
<td>C-type lectin domain family 4, member a2</td>
<td>Clec4a2</td>
<td>1.75</td>
<td>9.60E−06</td>
</tr>
<tr>
<td>92633_at</td>
<td>Cathepsin Z</td>
<td>Ctsz</td>
<td>1.74</td>
<td>7.91E−05</td>
</tr>
<tr>
<td>161895_s_at</td>
<td>cDNA sequence BC054822</td>
<td>BC054822</td>
<td>1.69</td>
<td>1.23E−05</td>
</tr>
<tr>
<td>99071_at</td>
<td>Macrophage-expressed gene 1</td>
<td>Mpeg1</td>
<td>1.67</td>
<td>7.03E−05</td>
</tr>
<tr>
<td>100607_at</td>
<td>Phospholipase D family, member 3</td>
<td>Pld3</td>
<td>1.66</td>
<td>6.31E−05</td>
</tr>
<tr>
<td>99952_at</td>
<td>Chitinase 3-like 1</td>
<td>Chi3l1</td>
<td>1.64</td>
<td>1.00E−04</td>
</tr>
<tr>
<td>93611_at</td>
<td>Lipoprotein lipase</td>
<td>Lpl</td>
<td>1.58</td>
<td>7.92E−05</td>
</tr>
<tr>
<td>93818_at</td>
<td>Cathepsin D</td>
<td>Ctsd</td>
<td>1.58</td>
<td>8.00E−07</td>
</tr>
<tr>
<td>161458_at</td>
<td>Coagulation factor VII</td>
<td>F7</td>
<td>1.54</td>
<td>2.21E−05</td>
</tr>
<tr>
<td>94747_at</td>
<td>Colony-stimulating factor 2 receptor, β 1, low affinity (granulocyte-macrophage)</td>
<td>Cat2rb1</td>
<td>1.47</td>
<td>2.11E−05</td>
</tr>
<tr>
<td>103070_at</td>
<td>Protein tyrosine phosphatase, nonreceptor type substrate 1</td>
<td>Sirpa</td>
<td>1.46</td>
<td>4.57E−05</td>
</tr>
<tr>
<td>98543_at</td>
<td>Cathepsin S</td>
<td>Cts</td>
<td>1.44</td>
<td>1.50E−05</td>
</tr>
<tr>
<td>101851_at</td>
<td>CD200 antigen</td>
<td>Cd200</td>
<td>1.43</td>
<td>1.10E−05</td>
</tr>
<tr>
<td>161819_f_at</td>
<td>Lyosomal-associated protein transmembrane 5</td>
<td>Laptm5</td>
<td>1.43</td>
<td>3.77E−05</td>
</tr>
<tr>
<td>104300_at</td>
<td>IQ motif containing GTPase-activating protein 1 (lqgap1), mRNA</td>
<td>Iqgap1</td>
<td>1.40</td>
<td>8.00E−05</td>
</tr>
<tr>
<td>94056_at</td>
<td>Stearoyl-CoA desaturase 1</td>
<td>Scd1</td>
<td>1.37</td>
<td>7.60E−06</td>
</tr>
<tr>
<td>103210_at</td>
<td>Colony-stimulating factor 2 receptor, β 2, low affinity (granulocyte-macrophage)</td>
<td>Csfrb2</td>
<td>1.35</td>
<td>1.07E−05</td>
</tr>
<tr>
<td>94713_at</td>
<td>Myosin Vlla</td>
<td>Myo7a</td>
<td>1.33</td>
<td>2.40E−06</td>
</tr>
<tr>
<td>92232_at</td>
<td>Suppressor of cytokine signaling 3</td>
<td>Socs3</td>
<td>1.31</td>
<td>9.02E−06</td>
</tr>
<tr>
<td>101389_at</td>
<td>Scavenger receptor class B, member 2</td>
<td>Scarb2</td>
<td>1.31</td>
<td>1.29E−05</td>
</tr>
<tr>
<td>96657_at</td>
<td>Spermindine/sperrnine N1-acetyl transferase 1</td>
<td>Sat1</td>
<td>1.27</td>
<td>3.49E−05</td>
</tr>
<tr>
<td>103584_at</td>
<td>RIKEN cDNA 5830471E12 gene</td>
<td>5830471E12Rik</td>
<td>1.27</td>
<td>2.62E−05</td>
</tr>
<tr>
<td>100581_at</td>
<td>Cystatin B</td>
<td>Catb</td>
<td>1.26</td>
<td>8.57E−05</td>
</tr>
<tr>
<td>100611_at</td>
<td>Lysozyme</td>
<td>Lysz</td>
<td>1.16</td>
<td>1.47E−05</td>
</tr>
<tr>
<td>101585_at</td>
<td>Progesterone receptor membrane component 1</td>
<td>Pgrmc1</td>
<td>0.84</td>
<td>7.20E−06</td>
</tr>
<tr>
<td>102871_at</td>
<td>Eph receptor B6</td>
<td>Ephb6</td>
<td>0.75</td>
<td>4.58E−05</td>
</tr>
<tr>
<td>96812_at</td>
<td>Smoothened homologue (Drosophila)</td>
<td>Smo</td>
<td>0.75</td>
<td>1.30E−06</td>
</tr>
<tr>
<td>92448_s_at</td>
<td>Glial cell line–derived neurotrophic factor family receptor α 2</td>
<td>Gfna2</td>
<td>0.61</td>
<td>4.89E−05</td>
</tr>
<tr>
<td>92449_at</td>
<td>Glial cell line–derived neurotrophic factor family receptor α 2</td>
<td>Gfna2</td>
<td>0.52</td>
<td>2.01E−05</td>
</tr>
</tbody>
</table>

NOTE: The gene descriptions, symbols, fold change (adjacent/normal), and estimated P value from the parametric statistical test in BRB-ArrayTools classifier algorithm are given for each probeID identified as a Field gene. The 48 probeIDs represent 46 different genes, with the majority of the genes (91%) being up-regulated (higher in the adjacent tissue than normal control). Previously published work (41) independently supports increased protein level of Ptgs1 (COX1), and our studies reported here validate increased protein expression of Ctsd, Ctsz, Chi3l1, Spp1, and Sirpa in adjacent lung tissue.
Validation of increased protein levels of the Field genes in the tumor microenvironment. The protein level of five different Field genes (Ctsd, Ctsz, Chi3l1, Spp1, and Sirpa) was verified by Western blot analysis of normal and adjacent A/J lung homogenates (Fig. 2). In each case, the protein levels were higher in the adjacent tissues compared with the controls. The expression of Ctsz, Chi3l1, Spp1, and Sirpa were significantly increased (P < 0.05) as determined by densitometry measurements of the exposed films.

Field Effect classifiers correctly predict BAL samples from tumor-bearing mice. If the Field genes are highly expressed in the vicinity of lung cancer due to infiltrating macrophages, we reasoned that the same genes would be useful in classifying microarray datasets derived from the cells recovered by BAL. As previously found (22), urethane-treated mice have a significant increase in total cells recovered by BAL compared with controls (3.7-fold; P < 0.0001). However, the cell type distribution was indistinguishable between urethane-treated and control mice, with macrophages the major cell type recovered (~95%) followed by lymphocytes (~5%) and neutrophils (~1%). RNA was prepared from cells recovered by BAL from 6 normal untreated controls and 11 urethane-treated A/J mice. These datasets were used to generate a BAL-specific classifier probeID lists (168 probeIDs; P < 0.0001), and a supervised cluster dendrogram using this list is shown in Fig. 4A. Each BAL sample is closely grouped in either tumor-bearing (T; red) or control (C; green) branches of the tree diagram with only sample C36 misidentified (magenta asterisk). On inspection of the formalin-fixed lungs from this mouse, it clearly had developed a spontaneous tumor, something the A/J strain is known to do as it ages (17, 42), suggesting why it was more closely grouped with the BAL samples from treated mice. The same BAL datasets were clustered with supervision using the independently derived Field probeIDs (49 probeIDs; Fig. 4B) producing a similar degree of accuracy as the BAL-derived classifiers. Using the prediction algorithm within BRB-ArrayTools, the Field probeIDs correctly predicted 94% to 100% of the BAL samples using a variety of statistical approaches (compound covariate, nearest neighbors, nearest centroid, and support vector machine; ref. 33). Interestingly, the Field gene and BAL-derived classifiers had only three genes in common: Aep5, Lrg1, and the expressed sequence tag (EST) 1100001G20Rik (Fig. 4C). Although the majority of the Field genes were not selected as BAL classifiers, most had significant absolute signal intensities (>64 arbitrary units) and nonoverlapping signal intensity distributions in the BAL microarray datasets (Supplementary Fig. S7).

The BAL samples were correctly classified using the Field Effect genes (derived from lung tissue samples), although the majority of these genes did not overlap with the BAL-specific classifiers. For this reason, we verified the BAL microarray signal intensities for a number of different Field genes by quantitative real-time PCR (qRT-PCR) analysis (Fig. 5A and B). Theoretically, the Field gene classifiers could give accurate predictions of the BAL samples yet have opposite expression levels (e.g., up-regulated Field genes are down-regulated in BAL RNA). The mRNA levels for the 11 genes tested were significantly up-regulated in BAL cells from treated animals compared with control BAL cells, showing 1.5- to 21-fold increases (P ≤ 0.01). The qRT-PCR results further support the
in silico assessment (Supplementary Figs. S4 and S5) that the Field Effect signature is attributable to macrophages because 95% of the recovered cells in BAL are macrophages. In addition, Western blot analysis showed increased protein levels in cell-free BAL fluid for the same proteins measured in lung homogenates (Fig. 5C). This suggests that macrophages infiltrating the tumor microenvironment have a unique Field gene expression signature reflected in both the mRNAs of the BAL cells and the secreted proteins in BAL of urethane-treated A/J mice.

Discussion

As with many cardiovascular diseases, lung cancer risk can easily be assigned to a specific patient population, with the overwhelming risk factors being cigarette smoke inhalation and environmental exposures (1). Unlike cardiovascular diseases, there are few clinically valid methods for identifying individuals with early stage disease, when it is most curable by surgical resection and chemotherapy. The controversial use of spiral computed tomography in asymptomatic smokers (7, 8) also suffers from high monetary cost and relatively poor rate of new patient identification during the extensive follow-up period. An alternative approach uses laser-induced fluorescence endoscope bronchoscopy to localize early stage bronchial lesions (43). Although screening studies continue to show the survival advantage of early diagnosis, no strategy other than direct biopsy exists for the evaluation of peripheral nodules. Our studies were designed to investigate the gene expression signature of the lung tumor microenvironment and evaluate the utility of BAL expression profiling.

In our previous work, we presented evidence using a microarray approach that the murine A/J mouse-urethane system is an excellent animal model of human adenocarcinoma due to the high degree of conservation in the altered gene expression patterns due to the cancer (12). We have now extended the study of the

![Image](image.png)

**Figure 3.** Immunohistochemical analysis of Field Effect proteins in fixed lung tissues. Formalin-fixed lung tissues from normal (42-wk untreated controls) and urethane-treated (24 or 42 wk after injection) A/J mice were stained for immunohistochemistry analyses of five different Field gene products (Ctsd, Ctsz, Chi3l1, Spp1, and Sirpa). Colored arrows, pointing toward representative cells or locations, show cell type identifications. Intensity of the brownish gray staining was generally highest in 42-wk urethane-treated animals. Twenty-four–week untreated animals were similar to 42-wk normal controls (not shown). Magnification bar, 40 μm.
remodeling (cathepsins D, K, S, and Z; proteinase inhibitors cystatin B and α1-antitrypsin; and integrin αX), macrophage markers (CD68 and macrophage-expressed gene 1), and cellular immune functions (Ly75; Cc66; CD200; colony-stimulating factor 2 receptor, β1, and low affinity (granulocyte-macrophage); suppressor of cytokine signaling 3; and signal-regulatory protein-α). In addition, PtgS1 was found up-regulated within the Field genes, a well-characterized marker for the role of inflammation in cancer (12, 35).

A number of recent reports have identified macrophage infiltration into the lung tumor microenvironment as an important consequence of multiple signaling pathways, including C-C motif and angiogenesis pathways (19, 21, 24). Our data argue strongly for the infiltration of a specific macrophage phenotype into the lung tumor microenvironment. These macrophages produce a gene expression signature (Field Effect genes) attributed to their presence in the tumor microenvironment, perhaps due to functional differences from the normally resident lung macrophages. The “strength” of the Field genes was shown by their predictive abilities in correctly classifying cells recovered by saline BAL of urethane or control A/J mouse lungs. In addition, a number of secreted proteins, predicted to be overexpressed with in the tumor microenvironment, were increased in cell-free BAL fluid. Because there is a high degree of similarity between the A/J-urethane model and human lung adenocarcinoma, we speculate that the Field genes could have predictive value in analyzing cells and/or fluid from BAL of cigarette-smoking patients. Ideally, a unique signature found by gene expression or proteomic analysis will be identified in high-risk asymptomatic individuals with early stage lung cancer.

The influence of immune cell infiltration on metastatic potential and patient survival has been examined in patients with colorectal cancer (44, 45). Instead of using a microarray approach, large-scale flow cytometry (48 CD markers and 10 surface markers/receptors) and qRT-PCR measurements (20 genes) were used to create a panel of cellular immune markers. Differing from our lung cancer results, their studies focused on the adaptive immune response, specifically the effector T cells, which were important in colorectal cancer, rather than inflammatory or immunosuppressive molecules. They found that T-cell markers of migration, activation, and differentiation were increased in colorectal tumors without signs of metastatic invasion. The density, type, and location of immune cells within the colorectal tumors, gave a better predictor of patient survival than the current histologic methods to stage their cancers and was reproduced in two additional patient populations.

The tumor microenvironment is clearly complex in terms of cell types within the milieu, regulation of signaling pathways, and tissue remodeling. In selected cases, potential novel therapeutic targets have been identified by analysis of the tumor microenvironment. Strikingly, the single largest functional group within the lung cancer Field genes is proteases (cathepsins D, K, S, and Z; proteinase inhibitors and specific cathepsin gene knockouts caused decreased pancreatic tumor formation, impaired angiogenesis, and lowered metastatic potential. Although cathepsins are typically lysosomal,
in pancreatic cancer, they are found extracellularly at the interface between normal and tumor cells, implicating them in extracellular matrix remodeling and angiogenesis. We found elevated protein levels for cathepsins D and S in BAL fluid from urethane-treated A/J mice, suggesting these proteases are also present in the extracellular environment of lung adenocarcinoma.

In summary, many studies suggest that secreted chemotactic and differentiating factors affect TAMs in the tumor microenvironment.

**Figure 5.** Validation of increased expression of Field Effect genes in BAL. A, mRNA levels for 11 different Field genes were determined using qRT-PCR with primer/probe pairs. \( \Delta C_t \) values were measured relative to \( \beta \)-actin as the endogenous control mRNA in BAL from 5 normal and 11 treated animals. Smaller \( \Delta C_t \) values indicate higher expression relative to \( \beta \)-actin in the BAL samples. Fold increase (tumor BAL/control BAL) was calculated by the \( \Delta \Delta C_t \) with the \( t \) test \( P \) value of <0.02 in all cases. Columns, mean; bars, the SE. B, the protein levels for five different Field genes (Ctsd, Ctsz, Chi3l1, Spp1, and Sirpa) were determined by Western blot analysis of cell-free BAL fluid. In each case, the protein levels were higher in the BAL fluid from treated animals compared with the controls. The expression of Ctsd, Chi3l1, and Spp1 was significantly increased (\( P < 0.05 \)) as determined by densitometry measurements of the exposed films.
genases and cell-free BAL fluid. Four different cathepsins were correctly classified BAL-derived cells from urethane-treated or normal tissue. The Field genes identified in TAMs were able to distinguish macrophages in tumors from those in normal lung. TAMs have a phenotype that promotes tumor growth and metastasis. The use of TAM genes in clinical trials of lung cancer may be a promising therapeutic approach.

**Acknowledgments**

Received 3/19/2007; revised 8/7/2007; accepted 10/8/2007.

**Grant support:** FAMRI Clinical Innovation Award (R.S., Stearns), R01 HL 72340 (M.W. Geraci), R01 CA 96131 (M.A. Malkinson and M.W. Geraci), R01 CA 83897 (A.M. Malkinson), and P50 CA 81887 (M.W. Geraci).

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**


A Macrophage Gene Expression Signature Defines a Field Effect in the Lung Tumor Microenvironment

Robert S. Stearman, Lori Dwyer-Nield, Michael C. Grady, et al.


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/68/1/34

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2007/12/27/68.1.34.DC1

Cited articles
This article cites 49 articles, 16 of which you can access for free at:
http://cancerres.aacrjournals.org/content/68/1/34.full.html#ref-list-1

Citing articles
This article has been cited by 7 HighWire-hosted articles. Access the articles at:
/content/68/1/34.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.