

# A Prognostic Test for Adenocarcinoma of the Lung From Gene Expression Profiling Data<sup>1</sup>

Gavin J. Gordon, William G. Richards,  
David J. Sugarbaker, Michael T. Jaklitsch, and  
Raphael Bueno<sup>2</sup>

Division of Thoracic Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115

## Abstract

Until recently, it has been impossible to determine which patients with resected stage I lung cancer are among the 30% who will die of metastatic cancer within 5 years of surgery. Bioinformatics tools applied to lung cancer expression profiling data have identified prognostic genes that have been used to develop predictor models, but thus far, these models have not been incorporated into routine clinical use because of their inherent complexity and requirement for relatively large numbers of genes. We have used ratios of gene expression to overcome these limitations. Here, we evaluate the ability of this technique to identify patients with stage I lung adenocarcinoma at risk for recurrence. We derived candidate ratio-based tests from analysis of 36 stage I lung adenocarcinoma samples using previously published expression profiling data. Eleven of these tests were identified for additional study and assessed for classification accuracy in an independent set of 60 stage I lung adenocarcinoma samples. We then evaluated the optimal ratio-based test in the independent samples using Kaplan-Meier survival analysis. Finally, we examined the ability of this test to predict outcome in a set of 97 stage I breast adenocarcinoma. We found that subsets of the independent lung cancer samples predicted to be associated with either good or poor outcome using the optimal ratio-based test differed significantly ( $P = 0.0056$ ) in terms of survival with a classification accuracy of 74% ( $P = 0.0043$ , Fisher's exact test). When this test was applied to stage II and III lung cancers, most specimens were classified as poor outcome cancers. Interestingly, we found that the same test significantly ( $P = 0.0417$ ) predicted recurrence of stage I breast tumors, suggesting that at least some of the marker genes we identified may have generalized prognostic

significance for adenocarcinoma. Our results provide additional evidence that expression ratios are highly accurate in predicting cancer recurrence and may be used in a simple test to predict response to surgical therapy in early-stage lung adenocarcinoma.

## Introduction

Lung cancer is the most common cause of cancer mortality in the United States for both men and women. In 2001, it is estimated that ~169,500 men and women were diagnosed in the United States with lung cancer and that 157,400 died from this malignancy.<sup>3</sup> Approximately 80% of lung cancer patients have NSCLC<sup>4</sup>, a histological category of primary lung cancer that includes adenocarcinoma, the most common form of NSCLC, as well as squamous cell carcinoma and large cell carcinoma. All patients diagnosed with NSCLC, regardless of subtype, are evaluated in a uniform manner and offered treatment based on their stage at presentation. Up to 34,000 lung cancer patients in the United States present with stage I or II disease and are amenable to effective surgical treatment (1). These patients with early-stage NSCLC enjoy up to 70% 5-year survival after surgery alone. Patients with stage IIIA NSCLC usually undergo neoadjuvant chemotherapy followed by surgery or radiation therapy. Patients with stage IIIB or IV lung cancer are usually only candidates for palliative chemotherapy and radiation therapy (2).

A prognostic test for early-stage lung adenocarcinoma is likely to improve survival by identifying patients who are more likely to recur and may, therefore, benefit from adjuvant or neoadjuvant therapy. A number of prognostic markers and pathological characteristics of resected lung cancer have been reported to correlate with outcome. These include the degree of differentiation, tumor size, lymphovascular invasion, and the expression of certain markers such as p53 and K-ras (3). However, none have been validated for routine clinical use.

Gene expression profiling using microarrays has been used to successfully predict disease related outcome in multiple cancers (4–11). Recently, several groups have published the results of gene profiling with microarrays of large cohorts of lung adenocarcinoma tissues linked to clinical outcome data (11–13). These studies propose prognostic models based on analysis of adenocarcinoma samples for which unique genetic profiles are correlated with treatment-related outcome. Unfortunately, these models are difficult to assess clinically because they rely on measuring expression levels of a relatively large numbers of genes using costly data acquisition platforms (*i.e.*, microarrays) and sophisticated algorithms/software. In addition, clinical use of the models is hindered by the inability to

Received 1/17/03; revised 4/30/03; accepted 5/16/03.

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.

<sup>1</sup> This work was partly funded by grants to R. B. from the Brigham Surgical Group Foundation, The Milton Fund of Harvard Medical School, and the National Cancer Institute (CA-098501) and by a grant to G. J. G. from the Cancer Research and Prevention Foundation.

<sup>2</sup> To whom requests for reprints should be addressed, at Brigham and Women's Hospital, Division of Thoracic Surgery, 75 Francis Street, Boston, MA 02115. Phone: (617) 732-8148, Fax: (617) 582-6171; E-mail: rbueno@partners.org.

<sup>3</sup> Internet address: <http://www.cancer.org>.

<sup>4</sup> The abbreviations used are: NSCLC, non-small cell lung cancer; DFS, disease-free survival; FNA, fine needle aspiration.

Table 1 Lung cancer prognostic genes

Accession no.	<i>P</i>	Ratio <sup>a</sup>	Description (LocusLink Symbol)
Expressed at relatively higher levels in good outcome tumors			
MI2529	0.014	2.1	Apolipoprotein E (APOE)
X00437	0.015	2.8	T cell receptor $\beta$ (TRB@)
D87436	0.034	2.6	Lipin 2 (LPIN2)
Expressed at relatively higher levels in poor outcome tumors			
K03195	$8.5 \times 10^{-5}$	0.49	Solute carrier family 2, member 1 (SLC2A1)
X65614	$7.5 \times 10^{-4}$	0.26	S100 calcium-binding protein P (S100P)
X70040	0.0012	0.49	Macrophage-stimulating 1 receptor (MST1R)
X93036	0.0042	0.33	FXD domain-containing ion transport regulator 3 (FXD3)

<sup>a</sup> Average expression level in good outcome samples/average expression level in poor outcome samples.

analyze a sample independently without reference to other samples. However, these studies do support the hypothesis that in addition to stage, specific features of the tumors' gene expression may be used to predict, with high accuracy, the clinical outcome of the patient in response to a specific therapy.

We recently described a method for translating expression profiling data into clinically relevant tests using ratios of gene expression (14). We have successfully used this method to develop tests for the differential diagnosis of lung adenocarcinoma and mesothelioma (14), the prognosis of mesothelioma (15), and the diagnosis of prostate cancer.<sup>5</sup> Development of specific tests is based on an initial supervised comparison of gene expression data between two groups that differ with respect to a chosen clinical characteristic. Although fundamentally similar to linear discriminant analysis (16), expression ratio-based models are statistically robust and offer several clear advantages that better facilitate the transition to clinical use (15). This method can be used to create clinical tests that are platform independent and require only small amounts of RNA and, by extension, tumor tissue. Here, we report the application of this method in developing a prognostic test for patients with resected stage I lung adenocarcinoma. The tissue specificity of the test was also assessed by applying it to expression profiling data of stage I breast adenocarcinoma.

## Materials and Methods

**Gene Expression Profiling Data.** Microarray data for lung adenocarcinoma tissues was obtained from two sources. Gene expression data for the training set of stage I lung adenocarcinoma samples were obtained using Affymetrix high-density oligonucleotide microarrays (HuFL or 6800 chip) with probe sets representing ~7000 genes (11). Gene expression data for the test set of samples were obtained using Affymetrix high-density oligonucleotide microarrays (U95A chip) with probe sets representing ~12,000 genes (12). Gene expression data for resected stage I (lymph node-negative) breast cancer tissues were obtained from a single source using microarrays containing ~25,000 genes (10).

**Data and Statistical Analysis.**<sup>6</sup> The selection of predictor genes for use in expression ratio-based prognosis was performed essentially as described (14, 15). We identified prognostic genes that were discriminatory between two subsets of training set samples: those from patients considered to be cured from cancer by surgery (*i.e.*, good outcome, survival >48 months,  $n = 25$ ) and those from patients who died from

recurrent cancer (*i.e.*, poor outcome, survival < 48 months,  $n = 11$ ). Using a two-sided Student's (parametric) *t* test for pairwise comparisons of average gene expression levels, we first searched all of the genes represented on the microarray for those with a statistically significant  $\geq 2$ -fold difference in average expression levels between good outcome and poor outcome tumors. To minimize the effects of background noise, the list of distinguishing genes was additionally refined by requiring that the mean expression level be >600 in at least one of the two sample subsets. We chose for final analysis those genes that fit the filtering criteria and were also represented on the expression profiling platform of the test set of samples. For the analysis of test set samples, we again defined good ( $n = 30$ ) and poor ( $n = 16$ ) outcome as survival >48 and <48 months, respectively. For determination of classification accuracy only, we also designated as poor prognosis those patients still alive but with recurrent cancer within 48 months of surgery because patient status data were available for test set samples. Of the seven prognostic genes identified in the training set (Table 1), a total of two were represented by multiple Affymetrix probe sets on the expression-profiling platform of the test set (LocusLink ID): *APOE* and *TRB@*. Uninformative expression data were removed from consideration by excluding those probe sets for which average expression levels of all test set samples were <50 and that were not called "Present" for a majority of samples. Two probe sets were excluded (32795\_at, and 31449\_at), and the remainder were averaged for each of the two genes to give a final expression value for each gene in all samples. Data from three highly accurate gene expression ratios were combined by calculating the geometric mean,  $(R_1 R_2 R_3)^{1/3}$ , where  $R_i$  represents a single ratio value. This is equivalent to the average of  $[\log_2(R_1), \log_2(R_2), \log_2(R_3)]$  and has the effect of giving equal weight to ratio fold-changes of identical magnitude but opposite direction. Kaplan-Meier analysis was used to estimate survival where only known death events were uncensored. The log-rank test was used to statistically assess differences among multiple survival curves. The classification accuracy of the model in the test set of samples was assessed using Fisher's exact test (*i.e.*,  $2 \times 2$  contingency table). All differences were determined to be statistically significant if  $P < 0.05$ . All calculations and statistical comparisons were generated using S-Plus (17).

## Results

**Identification of Prognostic Molecular Markers in Lung Adenocarcinoma and Prediction of Outcome Using Gene Expression Ratios.** Twenty-three candidate prognostic genes were found to fit the filtering criteria in this analysis. Three of these were expressed at relatively higher levels in good out-

<sup>5</sup> R. Bueno, K. R. Loughlin, M. H. Powell, G. J. Gordon. (2003) A diagnostic test for prostate cancer from gene expression profiling data. *J Urology*. In Press.

<sup>6</sup> Additional information and detailed methods for all analyses can be found at our websites: <http://www.chesturg.org> and <http://www.generatios.com>.

Table 2 Accuracy of all ratio combinations in classifying training set samples<sup>a</sup>

	SLC2A1	S100P	MST1R	FXYD3
APOE	67% (24/36)	83% (30/36)	69% (25/36)	78% (28/36)
TRB@	78% (28/36)	64% (23/36)	81% (29/36)	69% (25/36)
LPIN2	89% (32/36)	58% (21/36)	81% (29/36)	75% (27/36)

<sup>a</sup> Seven diagnostic genes (identified in a training set of samples as described in the text) were used to calculate a total of 12 possible expression ratios (column/row intersection). The accuracy of each ratio in classifying training set samples was examined, and predictions are stated as the fraction diagnosed correctly.

come samples, and 20 were expressed at relatively higher levels in poor outcome samples. For additional study, we chose the four genes most significantly overexpressed in poor outcome samples and all three of the genes expressed at relatively higher levels in good outcome samples (Table 1). Using these genes, we next determined whether expression ratios could accurately classify the 36 samples used to train the model. We calculated a total of 12 possible expression ratios/sample by dividing the expression value of each of the three genes expressed at relatively higher levels in good outcome samples (*i.e.*, *APOE*, *TRB@*, and *LPIN2*) by the expression value of each of the 4 genes expressed at relatively higher levels in poor outcome samples (*i.e.*, *SLC2A1*, *S100P*, *MST1R*, and *FXYD3*). Samples with ratio values > 1 were predicted to be good outcome, and those with ratio values < 1 were predicted to be poor outcome. The overall accuracy of the 12 expression ratios varied widely (average = 74%, range, 58–89%, Table 2). To incorporate the predictive power of multiple prognostic genes (*i.e.*, ratios), we calculated the geometric mean (see “Materials and Methods”) for all 220 possible nonredundant three-ratio combinations and found that we could easily identify training samples with accuracy that met or exceeded that of any of the gene pair ratios when used alone. We chose for additional study a small number of these three-ratio tests (5%, 11 of 220, Table 3) that were ≥90% accurate overall and also ≥90% accurate within each subset of training set samples (*i.e.*, good and poor outcome). There were clearly multiple three-ratio combinations that were highly accurate in classifying training set samples. One of these tests (*APOE/S100P*, *LPIN2/SLC2A1*, and *LPIN2/MST1R*) used the three most accurate single ratios that individually identified >80% of the training set samples and for this reason was predicted to ultimately represent the optimal test.

**Verification of Expression Level Ratios as Outcome Predictors.** Next, we tested the ability of these 11 highly accurate expression ratio combinations to predict outcome in a separate cohort of 60 stage I lung adenocarcinoma tumor samples for which expression profiling data and associated clinical data were available (12). This dataset (12) was also used by the investigators that generated the current training set data (11) to independently validate a 50-gene prognostic model developed in their laboratory. Similar to these investigators (11), we initially excluded from analysis those samples with <40% tumor cell content and/or samples with mixed histology (*e.g.*, adenosquamous) to ensure equal comparability of samples, thus leaving 46 samples (*i.e.*, the test set) in the primary analysis. The Kaplan-Meier estimated 4-year survival of this cohort (66%, Fig. 1A) was consistent with expected survival for resected stage I lung cancer. We then used the expression values for all seven prognostic genes to calculate the 11 most accurate three-ratio combinations in these 46 samples. Samples with combined scores >1 and <1 were predicted to be associated with good and poor outcome, respectively. The classification

Table 3 Most accurate three-ratio combinations in the training set

All 220 possible 3-ratio combinations were filtered for accuracy according to the criteria in the text. The 11 most accurate three-ratio combinations in classifying training set samples are presented as the fraction diagnosed correctly.

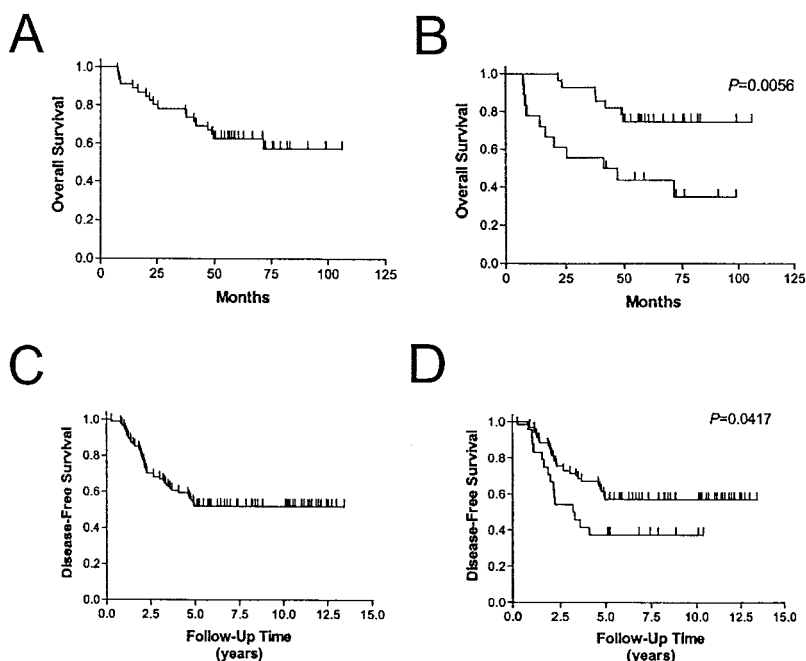
Three-ratio test <sup>a</sup>	Overall accuracy (n = 36)	Good prognosis accuracy (n = 25)	Poor prognosis accuracy (n = 11)
1/4, 3/4, 3/5	97%	100%	91%
1/4, 3/5, 3/6	94%	96%	91%
1/5, 3/4, 3/6	94%	96%	91%
1/6, 3/4, 3/5	94%	96%	91%
1/4, 2/4, 3/5	92%	92%	91%
1/4, 2/5, 3/4	92%	92%	91%
1/5, 2/4, 3/4	92%	92%	91%
2/4, 3/4, 3/7	92%	92%	91%
2/4, 3/6, 3/7	92%	92%	91%
2/6, 3/4, 3/7	92%	92%	91%
2/7, 3/4, 3/6	92%	92%	91%

<sup>a</sup> Ranked by overall accuracy then good prognosis accuracy. Genes (from Table 1) are represented by numbers for simplicity: 1-APOE; 2-TRB@; 3-LPIN2; 4-SLC2A1; 5-S100P; 6-MST1R; and 7-FXYD3.

accuracy for the 11 ratio-based tests are listed in Table 4 (average = 66%, range, 61–74%). The three most accurate tests used the same five genes (*APOE*, *S100P*, *LPIN2*, *SLC2A1*, and *MST1R*) in various different combinations and were equally proficient at classifying test set samples. Using predictions made from the test using the most accurate single ratios from the training set (*APOE/S100P*, *LPIN2/SLC2A1*, and *LPIN2/MST1R*), we could accurately (74%, 34 of 46) and significantly ( $P = 0.0043$ , Fisher’s exact test) predict the identity of the test set samples in this analysis. This test was 77% (23 of 30) and 69% (11 of 16) accurate at classifying good and poor outcome samples, respectively. To take into account censor status, we performed Kaplan-Meier survival analysis using predictions made by the same three-ratio test. We found that the estimated survival associated with good and poor prognosis subjects identified using expression ratios were significantly different ( $P = 0.0056$ , log-rank test; Fig. 1B). A total of 29 samples were assigned to the good outcome subset (median survival not reached) and 17 to the poor outcome subset (median survival = 44 months).

We also examined the optimal three ratio combination (*APOE/S100P*, *LPIN2/SLC2A1*, and *LPIN2/MST1R*) in stage II ( $n = 24$ ) and stage III ( $n = 10$ ) test set specimens (12). The same selection criteria were applied to these samples (*i.e.*, samples with <40% tumor cell content or mixed histology were excluded) leaving 15 stage II samples and 10 stage III samples in the analysis. We found that 87% (13 of 15) of stage II tumors and 100% (10 of 10) stage III tumors were predicted to be in the poor prognosis subset regardless of ultimate outcome. Unlike the stage I patients who were treated by surgery alone, many of the stage II and III patients received chemotherapy and/or radiotherapy either before or after surgery and, for this reason, are not appropriate for use in survival analysis.

**Sensitivity of Expression Ratios as Outcome Predictors in Lung Adenocarcinoma.** We originally identified prognostic genes from samples containing >70% tumor cellularity (11). To test outcome predictor models using these genes in an independent group of samples, we initially only considered test set samples with relatively high tumor cell content to ensure equal comparability between training and test sample sets. However, it would be beneficial for any clinical test to demonstrate adequate sensitivity when samples with relatively low



**Fig. 1.** Kaplan-Meier survival and time-to-relapse predictions for stage I lung and breast adenocarcinoma patients, respectively. **A**, overall survival for all 46 stage I lung adenocarcinoma patients from which the test set was chosen. The estimated 4-year survival for entire cohort was 66%. **B**, overall survival in the test set of samples for good outcome (*top line*, median survival not reached) and poor outcome (*bottom line*, median survival = 44 months) subsets as defined by the optimal three-ratio prognostic test. This three-ratio test uses expression data from five genes and significantly ( $P = 0.0056$ ) predicts outcome in an independent set of 46 stage I lung adenocarcinoma samples. **C**, DFS for 97 stage I breast adenocarcinoma patients. **D**, DFS in these 97 samples for the subsets predicted by the optimal three-ratio test developed in lung cancer to have a more favorable outcome (*top line*, median DFS not reached) and a less favorable outcome (*bottom line*, median DFS = 3.3 months). This three-ratio test uses expression data from five genes and significantly ( $P = 0.0417$ ) predicts outcome in an independent set of 97 stage I breast adenocarcinoma samples. Hash marks indicate censored data (*i.e.*, patients still alive at follow-up for **A** and **B** and patients without cancer recurrence at follow-up for **C** and **D**).

**Table 4** Most accurate three-ratio combinations in the test set

The 11 most accurate three-ratio combinations identified in the training set of samples were examined in test set samples. Ratio combinations are presented as the fraction diagnosed correctly and are listed in order of accuracy in the training set (from Table 3).

Three-ratio test <sup>a</sup>	Overall accuracy (n = 46)	Good prognosis accuracy (n = 30)	Poor prognosis accuracy (n = 16)
1/4, 3/4, 3/5	72%	73%	69%
1/4, 3/5, 3/6	74%	77%	69%
1/5, 3/4, 3/6	74%	77%	69%
1/6, 3/4, 3/5	74%	77%	69%
1/4, 2/4, 3/5	61%	83%	19%
1/4, 2/5, 3/4	61%	83%	19%
1/5, 2/4, 3/4	61%	83%	19%
2/4, 3/4, 3/7	65%	93%	13%
2/4, 3/6, 3/7	63%	90%	13%
2/6, 3/4, 3/7	63%	90%	13%
2/7, 3/4, 3/6	63%	90%	13%

<sup>a</sup> Genes (from Table 1) are represented by numbers for simplicity: 1-APOE; 2-TRB@; 3-LPIN2; 4-SLC2A1; 5-S100P; 6-MSTR1; and 7-FXYD3.

tumor cell content are assayed. Therefore, we determined, using the optimal three-ratio test from above, the classification accuracy in the 14 stage I test set samples with low tumor cell content (<40%, minimum 10%) originally excluded from consideration. This test resulted in 64% (9 of 14) of the samples correctly identified with 67% (6 of 9) and 60% (3 of 5) classification accuracy in the good and poor prognosis subsets, respectively. Although classification accuracy was moderately better than that expected by chance alone, our predictions were not statistically significant for this small cohort ( $P = 0.58$ , Fisher's exact test).

**Prognosis of Other Adenocarcinomas Using Expression Ratios Developed in Lung as Outcome Predictors.** We next examined whether the prognostic test was specific for lung adenocarcinoma or could be applicable to other types of ade-

nocarcinoma. We hypothesized that adenocarcinomas of both lung and breast origin would exhibit some degree of overlap in prognosis-related gene expression profiles despite originating from different tissue types. To test this hypothesis, we performed Kaplan-Meier time-to-recurrence survival analysis in breast cancer using predictions made by the optimal three-ratio test developed in lung cancer (*APOE/S100P*, *LPIN2/SLC2A1*, and *LPIN2/MSTR1*). For this study, we obtained expression profiling data for a total of 97 breast cancer samples originally used to develop a microarray-based predictor model that stratified patient samples into prognostic groups based on cancer recurrence (10). The Kaplan-Meier estimated median DFS was not reached for these 97 samples (Fig. 1C). We used the breast cancer microarray data to calculate the geometric means of the three ratios and used the value to assign each patient to a good or poor prognosis subset. We then compared these groups using Kaplan-Meier survival analysis and discovered that predictions made using the three-ratio lung cancer prognostic test were able to produce significantly different ( $P = 0.0417$ , log-rank test; Fig. 1D) DFS curves in breast cancer. The classification accuracy of this model was 60% (58 of 97) and was determined by comparing ratio-based predictions to the known patient status pertaining to cancer recurrence (10). The median DFS was not reached for the subset predicted to have a more favorable outcome and was 3.3 years for the subset predicted to have a less favorable outcome.

## Discussion

In this article, we describe a prognostic test for stage I adenocarcinoma of the lung that could easily be extended to the clinical setting and is amenable to widespread analysis and validation by multiple investigators. Despite fundamental differences in experimental design, there are some similarities in the composition of both our training and test sample sets and those of Beer *et al.* (11) who developed an algorithm that uses 50 genes to predict outcome in lung cancer. Although classification accuracy was not reported in the Beer *et al.* study (11),

Kaplan-Meier survival analysis by these investigators resulted in significantly different curves ( $P = 0.006$ , Fig. 3H). Our study also produced significantly different survival curves with a similar  $P$  ( $P = 0.0056$ , Fig. 1B) but used 45 fewer genes in a manner amenable to reverse transcriptase-PCR data acquisition and analysis of individual samples.

The observed classification accuracy of 74% for the optimal ratio combination is encouraging within the context of this limited proof-of-principle study because gene discovery conditions were constrained by the number of sequences represented on the expression profiling platform of the training set. We are currently evaluating more comprehensive microarray platforms for the identification of differentially expressed genes. Ultimately, the accuracy of an expression ratio-based prognostic algorithm will be substantially enhanced by combining the ratio test results with proven clinical prognostic variables in addition to patient-specific parameters. These latter factors would not necessarily be reflected in the expression signature and should add independent prognostic information to improve the accuracy of the ratio-based test.

The issue of identifying the minimal number of genes necessary for maximal accuracy deserves consideration. Our experience has shown that multiple ratio tests incorporating three ratios (*i.e.*, up to six genes) are a good starting point for optimization. In the current study we examined all combinations of four ratios as well and, not surprisingly, found multiple four-ratio tests that were highly accurate in the training set. However, none of these tests were as accurate as the most accurate three-ratio tests in classifying test set samples (data not shown), suggesting that a three-ratio test may be optimal for ratio-based prognosis under these circumstances. In fact, increasing the number of genes included in a ratio-based model will ultimately prove detrimental because additional predictor genes are added with progressively less discriminating power as reflected in the increasing  $P$ s obtained during the initial supervised analysis.

Our study may also indicate that accurate classification of lung cancer prognostic subsets using the gene ratio technique requires tumor samples with relatively high tumor cell content because we found that successful classification of low tumor cell content samples is only moderately better (64%) than that expected by chance alone. These findings are consistent with previous studies using alternative classification methods (11). However, our results are not conclusive given the relatively small cohort size of low tumor cell content samples and the fact that all 12 classification errors in stage I samples with >40% tumor content were from samples with >60% tumor cell content and 9 of 12 (75%) were from samples with >70% tumor cell content.

Published studies seeking to identify new prognostic molecular markers in cancer have largely focused on comparing samples within a single tissue type. The ability of a prognostic test developed and validated in lung adenocarcinoma to significantly predict recurrence in breast adenocarcinoma is intriguing, but the potential implications of this finding are presently unknown. We did not observe any obvious trends in the expression levels of individual prognostic genes in misclassified breast cancer samples and are currently examining whether ratio-based tests created using breast cancer profiling data are likewise amenable to predicting prognosis in lung cancer. Still, these results support the hypothesis that survival-related similarities exist in the global expression patterns of both adenocarcinomas and specifically that some of these can be reflected in a simple test using five genes. Direct support for this idea can be found in the current study; the *SI00P* gene we independently

discovered in lung adenocarcinoma is also a known marker of tumor progression in breast cancer (18, 19). Furthermore, pathologists have previously described the degree of differentiation, lymphovascular invasion, and mucin production as independent prognostic markers within adenocarcinomas from a variety of tissue origins. Although the optimal predictor genes for any tumor may be tissue specific, discovery of prognostic markers suitable for use in both adenocarcinomas will reveal fundamental similarities in gene expression patterns within adenocarcinomas in general and perhaps lead to the discovery of potential therapeutic targets.

A total of five genes comprised the most accurate three-ratio test in this study: *APOE*; *SI00P*; *SLC2A1*; *LPIN2*; and *MSTR*. Of these, only two (*SI00P* and *SLC2A1*) were listed among the top survival genes in the initial analysis of the training dataset (11), suggesting that multiple genes can be used to predict outcome in lung cancer, the choice of which is based on the particular model used in the analysis. Nevertheless, four of the five genes in our study have clearly documented diagnostic and prognostic implications in other cancers. *APOE*, a lipid homeostasis protein, is highly expressed in ovarian carcinoma (20, 21). Similar to the current study, *SI00P* is also preferentially expressed in prostate tumors with relatively worse prognosis (*i.e.*, hormone refractory tumors; Ref. 22). Abnormal expression of *SLC2A1* (alias *GLUT1*) has been observed in a number of epithelial malignancies, including ovarian cancer, where protein expression levels are directly proportional to tumor aggressiveness (23), consistent with our observation that higher levels of this gene are associated with worse prognosis in lung cancer. Finally, overexpression of *MSTR1* (alias *RON*, a member of the *MET* proto-oncogene family) has been found to cause the formation of lung tumors with atypical phenotypes *in vivo* (24).

In addition to dividing the stage I tumors in the test set into two subsets with statistically different survival, the ratio test also assigned the majority of stage II and all of the stage III specimens to the poor outcome subset. This could suggest that the genetic profile of stage I cancer likely to recur is similar to that of more advanced tumors with respect to the prognostic genes identified in this study. Because the cause of death from lung cancer is usually metastatic disease and patients with stage II and III lung cancer have progressively higher incidence of metastatic disease, it is reasonable to hypothesize that the test developed herein measures inherent tumor metastatic potential. This suggestion challenges the idea that distant tumors arise from relatively rare cells within the primary tumor that have metastatic potential and is supported by a recent study by Ramaswamy *et al.* (25) in multiple tumor types. One other explanation for the assignment of stage II and all of the stage III specimens to the poor outcome subset is that the selected ratio-based test reflects the inherent presence of micrometastatic nodal disease because stage II and III cancer by definition have involved lymph nodes. Ratio-based prognosis within stage II and stage III patients was not possible because too few samples were assigned to the good outcome group and because different types of additional treatments were given to many of these patients.

The identification of patients at higher risk for recurrence may not immediately improve survival in the absence of effective therapeutic options. To date, however, adjuvant chemotherapy has not proven effective in controlling metastatic disease after resection of the primary tumor. This issue is currently being examined in a cooperative group trial randomizing stage IB patients to adjuvant chemotherapy *versus* observation (CALGB 9633). One potentially effective clinical approach

would be ratio-based testing of tumor specimens obtained by FNA before any intervention. Percutaneous *trans*-thoracic FNA of lung nodules is a safe and well-accepted diagnostic technique that has been applied to lesions as small as 8–10 mm (26). Stage I patients predicted to have poor prognosis after FNA biopsy, for example, could be offered participation in clinical trials of neoadjuvant therapy using protocols proven useful in patients with stage IIIA lung cancer (27). In fact, we have ongoing studies to determine the suitability of FNA biopsy material for analysis using gene expression ratios.<sup>7</sup>

## References

- Mountain, C. F. Revisions in the international system for staging lung cancer. *Chest*, *111*: 1710–1717, 1997.
- Fraire, A. Pathology of lung cancer. In: J. Aisner, R. Arriagada, M. Green, N. Martini, and M. Perry (eds.), *Comprehensive Textbook of Thoracic Oncology*, pp. 245–275. Baltimore, MD: Williams & Wilkins, 1996.
- Kwiatkowski, D. J., Harpole, D. H. J., Godleski, J., Herndon, J. E., Shieh, D. B., Richards, W., Blanco, R., Xu, H. J., Strauss, G. M., and Sugarbaker, D. J. Molecular-pathologic substaging in 244 stage I non-small cell lung cancer patients: clinical implications. *J. Clin. Oncol.*, *16*: 2468–2477, 1998.
- Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science (Wash. DC)*, *286*: 531–537, 1999.
- Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, O., Pergamenschikov, A., Williams, C., Zhu, S. X., Lonning, P. E., Borresen-Dale, A.-L., Brown, P. O., and Botstein, D. Molecular portraits of human breast tumours. *Nature (Lond.)*, *406*: 747–752, 2000.
- Hedenfalk, I., Duggan, D., Chen, Y., Radmacher, M., Bittner, M., Simon, R., Meltzer, P., Gusterson, B., Esteller, M., Kallioniemi, O.-P., Wilfond, B., Borg, A., and Trent, J. Gene expression profiles in hereditary breast cancer. *N. Engl. J. Med.*, *344*: 539–548, 2001.
- Khan, J., Wei, J. S., Ringner, M., Saal, L. H., Ladanyi, M., Westermann, F., Berthold, F., Schwab, M., Antonescu, C. R., Peterson, C., and Meltzer, P. S. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nat. Med.*, *7*: 673–679, 2001.
- Welsh, J. B., Sapinoso, L. M., Su, A. I., Kern, S. G., Wang-Rodriguez, J., Moskaluk, C. A., Frierson, H. F., and Hampton, G. M. Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. *Cancer Res.*, *61*: 5974–5978, 2001.
- Dhanasekaran, S. M., Barrette, T. R., Ghosh, D., Shah, R., Varambally, S., Kurachi, K., Pienta, K. J., Rubin, M. A., and Chinnaiyan, A. M. Delineation of prognostic biomarkers in prostate cancer. *Nature (Lond.)*, *412*: 822–826, 2001.
- van 't Veer, L. J., Dai, H., van de Vijver, M. J., He, Y. D., Hart, A. A. M., Mao, M., Peterse, H. L., van der Kooy, K., Marton, M. J., Witteveen, A. T., Schreiber, G. J., Kerkhoven, R. M., Roberts, C., Linsley, P. S., Bernards, R., and Friend, S. Gene expression profiling predicts clinical outcome of breast cancer. *Nature (Lond.)*, *415*: 530–536, 2002.
- Beer, D. G., Kardia, S. L. R., Huang, C.-C., Giordana, T. J., Levin, A. M., Misek, D. E., Lin, L., Chen, G., Gharib, T. G., Thomas, D. G., Lizyness, M. L., Kuick, R., Hayasaka, S., Taylor, J. M. G., Iannettoni, M. D., Orringer, M. B., and Hanash, S. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat. Med.*, *8*: 816–824, 2002.
- Bhattacharjee, A., Richards, W. G., Staunton, J., Li, C., Monti, S., Vasa, P., Ladd, C., Beheshti, J., Bueno, R., Gillette, M., Loda, M., Weber, G., Mark, E. J., Lander, E. S., Wong, W., Johnson, B. E., Golub, T. R., Sugarbaker, D. J., and Meyerson, M. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma sub-classes. *Proc. Natl. Acad. Sci. USA*, *98*: 13790–13795, 2001.
- Garber, M. E., Troyanskaya, O. G., Schluens, K., Petersen, S., Thaesler, Z., Pacyna-Gengelbach, M., van de Rijn, M., Rosen, G. D., Perou, C. M., Whyte, R. I., Altman, R. B., Brown, P. O., Botstein, D., and Petersen, I. Diversity of gene expression in adenocarcinoma of the lung. *Proc. Natl. Acad. Sci. USA*, *99*: 13784–13789, 2002.
- Gordon, G. J., Jensen, R. V., Hsiao, L.-L., Gullans, S. R., Blumenstock, J. E., Ramaswami, S., Richards, W. G., Sugarbaker, D. J., and Bueno, R. Translation of microarray data into clinically relevant cancer diagnostic tests using gene expression ratios in lung cancer and mesothelioma. *Cancer Res.*, *62*: 4963–4967, 2002.
- Gordon, G. J., Hsiao, L.-L., Jensen, R. V., Gullans, S. R., Blumenstock, J. E., Richards, W. G., Sugarbaker, D. J., and Bueno, R. Using gene expression ratios to predict outcome among patients with mesothelioma. *J. Natl. Cancer Inst. (Bethesda)*, *95*: 598–6058, 2003.
- Dudoit, S., Fridlyand, J., and Speed, T. P. Comparison of discrimination methods for the classification of tumors using gene expression data. *J. Am. Stat. Assoc.*, *97*: 77–87, 2002.
- Venables, W. N., and Ripley, B. D. *Modern Applied Statistics with S-Plus*. New York, NY: Springer, 1997.
- Russo, J., Hu, Y. F., Silva, I. D., and Russo, I. H. Cancer risk related to mammary gland structure and development. *Microsc. Res. Tech.*, *52*: 204–223, 2001.
- Guerreiro Da Silva, I. D., Hu, Y. F., Russo, I. H., Ao, X., Salicioni, A. M., Yang, X., and Russo, J. S100P calcium-binding protein overexpression is associated with immortalization of human breast epithelial cells *in vitro* and early stages of breast cancer development *in vivo*. *Int. J. Oncol.*, *16*: 231–240, 2000.
- Hough, C. D., Sherman-Baust, C. A., Pizer, E. S., Montz, F. J., Im, D. D., Rosenshein, N. B., Cho, K. R., Riggins, G. J., and Morin, P. J. Large-scale serial analysis of gene expression reveals genes differentially expressed in ovarian cancer. *Cancer Res.*, *60*: 6281–6287, 2000.
- Hough, C. D., Cho, K. R., Zonderman, A. B., Schwartz, D. R., and Morin, P. J. Coordinately up-regulated genes in ovarian cancer. *Cancer Res.*, *61*: 3869–3876, 2001.
- Mousses, S., Bubendorf, L., Wagner, U., Hostetter, G., Kononen, J., Cornelison, R., Goldberger, N., Elkahoulou, A. G., Willi, N., Koivisto, P., Ferhle, W., Raffeld, M., Sauter, G., and Kallioniemi, O. P. Clinical validation of candidate genes associated with prostate cancer progression in the CWR22 model system using tissue microarrays. *Cancer Res.*, *62*: 1256–1260, 2002.
- Kalir, T., Wang, B. Y., Goldfischer, M., Haber, R. S., Reder, I., Demopoulos, R., Cohen, C. J., and Burstein, D. E. Immunohistochemical staining of GLUT1 in benign, borderline, and malignant ovarian epithelia. *Cancer (Phila.)*, *94*: 1078–1082, 2002.
- Chen, Y. Q., Zhou, Y. Q., Fisher, J. H., and Wang, M. H. Targeted expression of the receptor tyrosine kinase RON in distal lung epithelial cells results in multiple tumor formation: oncogenic potential of RON *in vivo*. *Oncogene*, *21*: 6382–6386, 2002.
- Ramaswamy, S., Ross, K. N., Lander, E. S., and Golub, T. A molecular signature of metastasis in primary solid tumors. *Nat. Genet.*, *33*: 49–54, 2002.
- Li, H., Boisselle, P., Shepard, J., Trotman-Dickenson, B., and McLoud, T. Diagnostic accuracy and safety of CT-guided percutaneous needle aspiration biopsy of the lung: comparison of small and large pulmonary nodules. *Am. J. Roentgenol.*, *167*: 105–109, 1996.
- Bueno, R., Richards, W. G., Swanson, S. J., Jaklitsch, M. T., Lukanich, J., Mentzer, S. J., and Sugarbaker, D. J. Nodal stage after induction therapy for stage IIIA lung cancer determines patient survival. *Ann. Thorac. Surg.*, *70*: 1826–1831, 2000.
- Bueno, R., Deters, L. A., Nitz, M. D., Lieberman, B. C., Gordon, G. J., Differential diagnosis of solitary lung nodules using gene expression ratios, submitted for publication.

<sup>7</sup> R. Bueno, L. A. Deters, M. D. Nitz, B. C. Lieberman, G. J. Gordon. Differential diagnosis of solitary lung nodules using gene expression ratios, submitted for publication.