Translation of Microarray Data into Clinically Relevant Cancer Diagnostic Tests Using Gene Expression Ratios in Lung Cancer and Mesothelioma[1]

Gavin J. Gordon, Roderick V. Jensen, Li-Li Hsiao, Steven R. Gullans, Joshua E. Blumenstock, Sridhar Ramaswamy, William G. Richards, David J. Sugarbaker, and Raphael Bueno[2]

Division of Thoracic Surgery [G. J. G., W. G. R., D. J. S., R. B.], and Renal Division, Department of Medicine, [L-L. H.], Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115; Department of Physics, Wesleyan University, Middletown, Connecticut 06457 [R. V. J., J. E. B.]; Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Cambridge, Massachusetts 02139 [S. R. G.]; Department of Adult Oncology, Dana-Farber Cancer Institute [S. R.], Harvard Medical School, Boston, Massachusetts 02115; and Whitehead Institute/Massachusetts Institute of Technology Center for Genome Research, Cambridge, Massachusetts 02139 [S. R.]

ABSTRACT

The pathological distinction between malignant pleural mesothelioma (MPM) and adenocarcinoma (ADCA) of the lung can be cumbersome using established methods. We propose that a simple technique, based on the expression levels of a small number of genes, can be useful in the early and accurate diagnosis of MPM and lung cancer. This method is designed to accurately distinguish between genetically disparate tissues using gene expression ratios and rationally chosen thresholds. Here we have tested the fidelity of ratio-based diagnosis in differentiating between MPM and lung cancer in 181 tissue samples (31 MPM and 150 ADCA). A training set of 32 samples (16 MPM and 16 ADCA) was used to identify pairs of genes with highly significant, inversely correlated expression levels to form a total of 15 diagnostic ratios using expression profiling data. Any single ratio of the 15 examined was at least 90% accurate in predicting diagnosis for the remaining 149 samples (*e.g.,* test set). We then examined (in the test set) the accuracy of multiple ratios combined to form a simple diagnostic tool. Using two and three expression ratios, we found that the differential diagnoses of MPM and lung ADCA were 95% and 99% accurate, respectively. We propose that using gene expression ratios is an accurate and inexpensive technique with direct clinical applicability for distinguishing between MPM and lung cancer. Furthermore, we provide evidence suggesting that this technique can be equally accurate in other clinical scenarios.

INTRODUCTION

MPM[3] is a highly lethal pleural malignancy (1). Patients with MPM frequently present with a malignant unilateral pleural effusion or pleural thickening. However, ADCA that is metastatic to the pleura of lung or other origin is a far more common etiology for patients presenting with a unilateral pleural effusion (1). The ultimate treatment strategies depend on the correct pathological diagnosis. Early MPM is best treated with extrapleural pneumonectomy followed by chemoradiation, whereas metastatic lung cancer is treated with chemotherapy alone (2). Not infrequently, distinguishing MPM from ADCA of the lung is challenging from both clinical and pathological perspectives (3). Fluid cytology is diagnostic in only 33% of the cases (2, 4), and sufficient additional tissue from an open surgical biopsy is often required for immunohistochemistry and cytogenetic analysis (1).

Current bioinformatics tools recently applied to microarray data have shown utility in predicting both cancer diagnosis (5) and outcome (6). Although highly accurate, their widespread clinical relevance and applicability are unresolved. The minimum number of predictor genes is not known, and the discrimination function can vary (for the same genes) based on the location and protocol used for sample preparation (5). Profiling with microarray requires relatively large quantities of RNA making the process inappropriate for certain applications. Also, it has yet to be determined whether these approaches can use relatively low-cost and widely available data acquisition platforms such as RT-PCR and still retain significant predictive capabilities. Finally, the major limitation in translating microarray profiling to patient care is that this approach cannot currently be used to diagnose individual samples independently and without comparison with a predictor model generated from samples the data of which were acquired on the same platform.

In this study, we have explored an alternative approach using gene expression measurements to predict clinical parameters in cancer. Specifically, we have explored the feasibility of a simple, inexpensive test with widespread applicability that uses ratios of gene expression levels and rationally chosen thresholds to accurately distinguish between genetically disparate tissues. This approach circumvents many of the problems that prevent the penetration of expression profiling research into the clinical setting. We found that expression ratio-based diagnosis of MPM and lung cancer was similarly accurate compared with standard statistical methods of class discrimination such as linear discrimination analysis (7) and similar models (5) while addressing many of their deficiencies.

MATERIALS AND METHODS

Tumor Tissues. A combined total of 245 discarded MPM and lung ADCA surgical specimens were freshly collected (and snap-frozen) from patients who underwent surgery at BWH between 1993 and 2001. Lung ADCA tumors consisted of both primary malignancies and metastatic ADCAs of breast and colon origin. All of the MPM samples used in these studies contained relatively pure tumor (>50% tumor cells in a high power field examined in a section adjacent to the tissue used). Linked clinical and pathological data were obtained for all of the patients who contributed tumor specimens and were rendered anonymous to protect patient confidentiality. Studies using human tissues were approved by and conducted in accordance with the policies of the Institutional Review Board at BWH.

Microarray Experiments. Total RNA (7 μg) was prepared from whole tumor blocks using Trizol Reagent (Invitrogen Life Technologies, Inc. Carlsbad, CA) and processed as described previously (8–10). cRNA was hybridized to human U95A oligonucleotide probe arrays (Affymetrix, Santa Clara, CA) using a protocol described previously (10). Data from 64 of 245 samples were discarded after visual inspection of hybridization data revealed obvious scanning artifacts, leaving a total of 31 MPM samples and 150 ADCA samples (139 patient tumors and 11 duplicates). Microarrays for all of the ADCA samples and 12 MPM samples were processed at the Dana-Farber Cancer Institute and the Whitehead Institute. The remaining 19 MPM samples were processed separately at BWH. Microarray data for the ADCA samples have been previously published (11). Bhattacharjee *et al*. (11) used microarray data from ADCAs used in this study in combination with additional samples, but not samples of MPM, to identify distinct subclasses within ADCA of the lung and to search for prognostic markers. However, their study did not provide any comparison of gene expression between ADCA and MPM.

Fig. 1. Tumor diagnosis using expression ratios. In *A*, patterns of relative expression levels for the eight genes selected from the training set can be extended to the remaining samples. Relative expression levels increase from low to high per legend. *B*, graphic depiction of the magnitude and direction, in all 149 samples comprising the test set, of the value for two independent ratios (*calretinin*:*Claudin-7* and *VAC-β*: *TACSTD1*) chosen for further study. *C*, the eight individual samples (represented by *colored bars*) that were misdiagnosed using one ratio or the other from *B* (*blue bars*, misdiagnosed MPM samples; *red bars*, misdiagnosed ADCA samples).

Real-Time Quantitative RT-PCR. Total RNA (2 μg) was reverse-transcribed into cDNA using Taq-Man Reverse Transcription reagents (Applied Biosystems, Foster City, CA) and quantified using all recommended controls for SYBR Green-based detection. Primers amplifying portions of claudin-7, VAC-β, TACSTD1, and calretinin cDNA (synthesized by Invitrogen Life Technologies, Inc.) had the following sequences (forward and reverse, respectively): claudin-7, 5′-GTTCCTGTCCTGGGAATGAG-3′ and 5′-AAGGAGATC-CCAGGTCACAC-3′; VAC-β, 5′-CCAGCCTTTCGGTCTTCTAT-3′ and 5′-CTGGAGGAAGTTGGGAAGAG-3′; TACSTD1 (5′-AGCAGCTTGAA-ACTGGCTTT-3′ and 5′-AACGATGGAGTCCAAGTTCTG-3′; calretinin 5′-AGGACCTGGAGATTGTGCTC-3′, 5-GAGTCTGGGTAGACGCATCA-3′.

Data Analysis. Gene expression levels were appropriately scaled to facilitate comparison of data from arrays hybridized at different times and/or using multiple scanners.[4] When the "average difference" was negative (*i.e.,* negligible expression level), the absolute value was used. A two-tailed Student's *t* test was used to compare the log(gene expression levels) for all of the 12,600 genes on the microarray between samples from a training set consisting of 16 MPM and 16 ADCA samples. All of the differences in the mean log(expression levels) between the samples in the two groups in the training set were determined to be statistically significant if $P < 2 \times 10^{-6}$. Kaplan-Meier survival analysis was performed and the difference between multiple survival curves was assessed using the log-rank test. All of the statistical comparisons (including linear discrimination analysis) were performed using S-PLUS (12). To generate the graphical representations of relative gene expression levels, all of the expression levels were first normalized within samples by setting the average (mean) to zero and the SD to 1. Scaled levels were assigned RGB values (representing 20 shades) for colorimetric display as a spectrum representing relative gene expression levels.

RESULTS

Identification of Diagnostic Molecular Markers. We searched all of the genes represented on the microarray for those with a highly

significant difference ($P < 2 \times 10^{-6}$; ≥8-fold) in average expression levels between both tumor types in the training set of 16 ADCA and 16 MPM samples. For further analysis, we chose the eight genes with the most statistically significant differences and a mean expression level >600 in at least one of the two training sample sets (gene name, GenBank accession no.): *calretinin*, X56667 ($P = 8 \times 10^{-12}$); *VAC-β*, X16662, ($P = 8 \times 10^{-13}$); *TACSTD1*, M93036 ($P = 6 \times 10^{-12}$); *claudin-7*, AJ011497 ($P = 2 \times 10^{-9}$); *TITF-1*, U43203 ($P = 10^{-9}$); *MRC OX-2 antigen*, X05323 ($P = 5 \times 10^{-13}$); *PTGIS*, D83402 ($P = 10^{-10}$); and *hypothetical protein KIAA0977*, AB023194 ($P = 9 \times 10^{-11}$). Five of these genes were expressed at relatively higher levels in MPM tumors (*calretinin*, *VAC-β*, *MRC OX-2*, *PTGIS*, and *KIAA0977*) and three were expressed at relatively higher levels in ADCA tumors (*TACSTD1*, *claudin-7*, and *TITF-1*). We then investigated whether expression patterns of these genes extended to all of the samples (Fig. 1*A*).

Diagnostic Accuracy of Gene Expression Ratios. Using the eight genes identified in the initial training set, we calculated 15 expression ratios per sample by dividing the expression value of each of the five genes expressed at relatively higher levels in MPM by the expression value of each of the three genes expressed at relatively higher levels in ADCA. Then, we tested the diagnostic accuracy of these ratios in the 149 remaining samples not included in the training set (*i.e.,* 15 MPM and 134 ADCA). Samples with ratio values >1 were called MPM and those with ratio values <1 were called ADCA. We found that these ratios could be used to correctly distinguish between ADCA and MPM tumors with a high degree of accuracy (Table 1).

To incorporate data from multiple ratios, we then randomly chose a pair of independent ratios (*calretinin*:*claudin-7* and *VAC-β*:*TAC-STD1*) and examined their predictive accuracy in the test set. Each ratio (*calretinin*:*claudin-7* and *VAC-β*:*TACSTD1*) was 97% (145 of 149) accurate with four errors (Figs. 1, *B* and *C*). Thus, a total of eight samples were incorrectly diagnosed using either ratio. However, these

Table 1 *Accuracy of all ratio combinations in predicting tumor diagnosis in test set*

Eight candidate diagnostic genes were identified in a training set of samples as described in "Materials and Methods." A total of 15 possible expression ratios (column/row intersection) were calculated in which both genes used to form the ratio possessed inversely correlated expression levels in both tumor types. The accuracy of each ratio in predicting diagnosis was examined in the 149 remaining tumor specimens not included in the training set (15 MPM and 134 ADCA). Predictions are stated as the fraction diagnosed correctly.

	Claudin-7	TACSTD1	TITF-1
Calretinin	97% (145/149)	98% (146/149)	91% (136/149)
VAC-β	97% (144/149)	97% (145/149)	94% (140/149)
MRC OX-2	97% (145/149)	97% (145/149)	95% (142/149)
KIAA097	97% (145/149)	95% (142/149)	94% (140/149)
PTGIS	97% (145/149)	97% (144/149)	96% (143/149)

two ratios were in disagreement for all eight of the incorrectly diagnosed samples (Fig. 1*C*). When the diagnostic call of both ratios is combined, the final analysis results in 95% (141 of 149) of tumors correctly diagnosed with zero errors and eight no-calls. No-calls were conservatively made for samples when both ratios did not return the same diagnosis (Fig. 1*C*). To predict a diagnosis for the eight no-calls, we randomly chose an additional ratio (*MRC OX-2:TITF-1*, Table 1). The addition of a third ratio established a majority diagnosis for the eight samples that could not previously be determined using only two ratios. Using all three ratios (*i.e.,* six genes), 99% (148 of 149) of tumors were correctly diagnosed; seven no-calls were resolved and one sample was incorrectly diagnosed.

Comparison with Linear Discrimination Analysis. Standard statistical methods of class discrimination (7), such as linear discrimination analysis, can also be used to achieve similar results for these three pairs of genes. We first determined a linear combination of measured expression levels for each pair of genes that provided maximal discrimination between the two sets of tumor samples in the training set. When applied to the test set samples, the linear discrimination functions for the (*calretinin, claudin-7*), (*VAC-β, TACSTD1*), and (*MRC OX-2, TITF-1*) pairs each gave six, five, and four misclassifications, respectively. However, only one sample was incorrectly diagnosed in all three tests combined. In fact, the same errant sample was identified in the application of both the three ratio tests and the three linear discriminant tests. This sample was originally obtained from a patient with the clinical and pathological diagnosis of ADCA. This specimen was annotated by a pathologist reviewing frozen sections of all specimens before RNA preparation as having unusual histological features raising suspicion of a "germ cell tumor or sarcoma."

Verification of Microarray Data and Validation of Ratio-based Diagnosis. We used real-time quantitative RT-PCR (*a*) to confirm gene expression levels of diagnostic molecular markers identified in microarray-based analysis; and (*b*) to demonstrate that ratio-based diagnosis of MPM and lung cancer is equally accurate using data obtained from another platform. We randomly chose 12 tumor samples each of MPM and ADCA from those used in microarray analysis and then calculated expression ratios for *calretinin:claudin-7* and *VAC-β:TACSTD1*. Expression ratios correctly diagnosed 96% (23 of 24) of samples, with zero errors and one no-call (Fig. 2).

We have also explored the usefulness of expression ratios in predicting clinical parameters under more challenging circumstances, *i.e.,* when predictor genes have substantially higher *P*s and smaller fold-change differences in average expression levels. In this analysis, we used previously published microarray data (6) for a set of 60 medulloblastoma tumors with linked clinical data (Dataset "C") to create a ratio-based test designed to predict patient outcome after treatment. Of these 60 samples, 39 and 21 originated from patients classified as "treatment responders" and "treatment failures," respec-

tively. We used a training set composed of 20 randomly chosen samples (10 responders and 10 failures) to identify predictor genes. A total of 10 genes fit our filtering criteria ($P < 0.05$, >2-fold change in average expression levels, and at least one mean >200), and we chose the most significant three genes expressed at relatively higher levels in each group for further analysis (gene name, GenBank accession no.) as follows: *histone 2A*, M37583 ($P = 0.012$); *GTPase rho C*, L25081 ($P = 0.026$); *protein gene product 9.5*, X04741 ($P = 0.046$); *neurofilament-66*, S78296 ($P = 0.0025$); *sulfonylurea receptor*, U63455 ($P = 0.0067$); and cell surface protein *HCAR*, U90716 ($P = 0.030$). *Histone 2A*, *GTPase rho C*, and *protein gene product 9.5* were expressed at relatively higher levels in treatment failures, whereas *neurofilament-66*, *sulfonylurea receptor*, and *HCAR* were expressed at relatively higher levels in treatment responders. Using the previously stated diagnostic criteria, we calculated a total of nine possible expression ratios using data from these six genes and examined their predictive accuracy in the remaining samples (*i.e.,* the test set, $n = 29$ responders and $n = 11$ failures). The accuracy of these nine ratios in classifying test set samples varied greatly (average, 60%; range, 43–70%). Because individual ratios use data from only two genes, we next determined whether combining multiple ratios (*i.e.,* genes) resulted in increased classification accuracy. To give equal weight to ratios with identical magnitude but opposite direction, we calculated the geometric mean for all nine ratios and found that this value was nearly as accurate (68%, 27 of 40) in classifying test set samples as the most accurate individual ratio (70%, 28 of 40). The accuracy rate of this model (68%, 27 of 40) is somewhat lower than that reported by Pomeroy *et al.* (78%, 47 of 60) for an optimal eight-gene and *k*-nearest neighbor (*k*-NN) model (6). However, the difference is likely attributable to sample size considerations. (These investigators used all 60 samples to train multiple predictor models ranging from 2 to 21 genes and cross-validated the optimal 8-gene model using statistical techniques.) Finally, we performed Kaplan-Meier survival analysis using predictions made from the nine-ratio geometric mean value. We found that a nine-ratio (six-gene) model could significantly ($P = 0.0079$, log-rank test) predict patient outcome after treatment in the test set of samples (Fig. 3). There was no overlap in the list of genes comprising our model and the eight-gene *k*-NN model of Pomeroy *et al.* (6), which suggested that multiple genes are present in this malignancy that have similar predictive capability.

DISCUSSION

Accurate diagnosis of cancer (or any disease) is the first critical step in choosing appropriate treatment that will hopefully result in the best
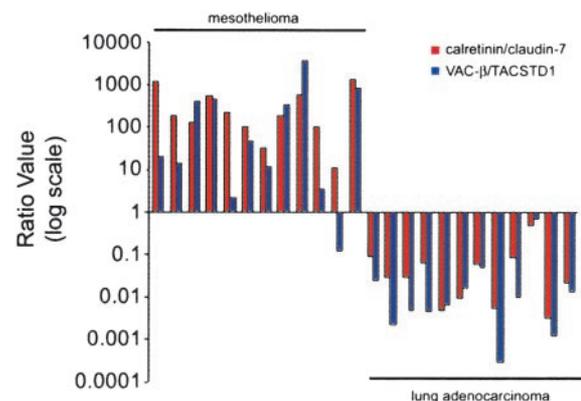
Fig. 2. Validation of microarray data and ratio-based diagnosis. Quantitative RT-PCR was used to obtain ratio values for 12 MPM and 12 ADCA tumors. In this case, the two ratios correctly identified 23 of 24 samples with 1 no-call.
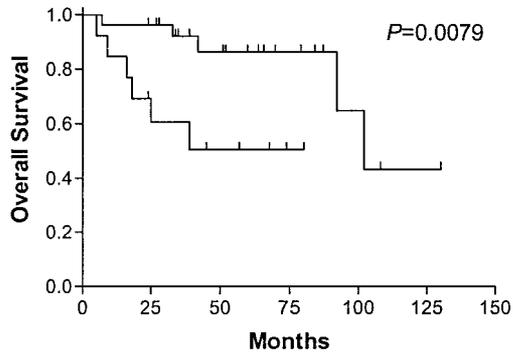
Fig. 3. Kaplan-Meier survival predictions for medulloblastoma patients. Overall survival for patients predicted to be treatment responders (*top line*) and treatment failures (*bottom line*) using a six-gene (nine-ratio) model in a true test set of samples (*n* = 40). *Hash marks*, censored data.

possible outcome. We propose that the ratio-based method described herein that uses expression levels of carefully chosen genes can be a simple, inexpensive, and highly accurate means to distinguish MPM from ADCA of the lung and that this method is applicable to many other clinical scenarios. We have also shown that multiple highly accurate ratios can be combined to form a simple diagnostic tool using the ratio direction ("majority rules" approach, *e.g.,* MPM and lung cancer diagnosis) or the ratio magnitude (calculation of the geometric mean, *e.g.,* prediction of outcome in medulloblastoma). The gene expression ratio method, by virtue of the fact that it is a ratio (*a*) negates the need for a third reference gene when determining expression levels, (*b*) is independent of the platform used for data acquisition, (*c*) requires only small quantities of RNA (as little as 10 pg using RT-PCR), (*d*) does not require the coupling of transcription to translation for chosen genes, and (*e*) permits analysis of individual samples without reference to additional "training samples" the data for which were acquired on the same platform. For these reasons, expression ratios are more likely to find immediate use in clinical settings because they confer several advantages compared with other equally accurate techniques, such as linear discriminant analysis.

The small *P*s and large fold-differences in average expression levels between genes used in expression ratio-based diagnosis of MPM and lung cancer are not surprising given that both tumor types have different cell types of origin. It is important to note that we have not determined in the present study the exact magnitude and consistency by which gene expression needs to differ between any two groups to allow the usage of a simple ratio test. In other clinical scenarios, the differences in gene expression patterns between groups to be distinguished may be more subtle, thus necessitating a relaxed filtering criteria in choosing potential predictor genes. Even in these cases, simple ratios can still be a highly accurate means of predicting clinical parameters. We have also found that expression ratios are useful in predicting the outcome after therapy in MPM, using genes with considerably higher *P*s and lower fold-differences in average expression levels than those used in the present study.[5] In the present study, we have used previously published microarray data (6) to identify a small number of predictor genes that were able to significantly predict outcome after therapy in medulloblastoma in a true test set of samples using simple expression ratios. Nevertheless, in some cases, larger numbers of genes (and perhaps sophisticated software) and/or initial expression profiling of a larger number of specimens for the training set may be required to achieve acceptable predictive power.

The selection of diagnostic genes for MPM and lung cancer was based solely on our stated criteria. Nevertheless, many of the molecular markers with the lowest *P*s and greatest difference in average expression levels have notable cancer relevance and/or are known to have tissue-specific expression patterns. *Calretinin* (13, 14) and *TITF-1* (15, 16) are part of several immunohistochemical panels currently used in the diagnosis of MPM and lung cancer. *Claudin* family members are expressed in various cancers (17, 18), and *TACSDT1* (also known as *TROP1*) is a recently described marker for carcinoma cells and, as a cell surface receptor protein, has been postulated to play a role in the growth regulation of tumor cells (19, 20). The discovery of diagnostic gene ratios is likely to make possible future clinical tests to definitively diagnose MPM and ADCA using smaller tissue specimens and perhaps pleural effusions. In this way, the need for diagnostic surgery in many of these patients may be eliminated.

The expression ratio technique represents a substantial improvement over past efforts to translate the strengths of expression profiling into simple tests with clinical relevancy. Many bioinformatics tools under development and testing are quite complex and/or rely upon data from large numbers of "training samples" to establish a diagnosis for unknown samples. The end result is that the practical use of microarray data remains beyond the scope of many scientists and clinicians. Similarly, no comprehensive method has been proposed to translate the results of tumor profiling to the analysis of individual tissues. As a consequence, no simple yet effective clinical applications have resulted from microarray research. The expression ratio technique represents a powerful use of microarray data that can be easily adapted and extended to routine clinical application without the need for additional sophisticated analysis.

ACKNOWLEDGMENTS

REFERENCES

1. Aisner, J. Diagnosis, staging, and natural history of pleural mesothelioma. *In:* J. Aisner, R. Arriagada, M. R. Green, N. Martini, and M. C. Perry (eds.), Comprehensive Textbook of Thoracic Oncology, pp. 799–785. Baltimore: Williams and Wilkins, 1996.
2. Pass, H. Malignant pleural mesothelioma: surgical roles and novel therapies. Clin. Lung Cancer, *3:* 102–117, 2001.
3. Ordonez, N. G. The immunohistochemical diagnosis of epithelial mesothelioma. Hum. Pathol., *30:* 313–323, 1999.
4. Nguyen, G-K., Akin, M-R. M., Villanueva, R. R., and Slatnik, J. Cytopathology of malignant mesothelioma of the pleura in fine-needle aspiration biopsy. Diagn. Cytopathol., *21:* 253–259, 1999.
5. 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 Landers, E. S. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science (Wash. DC), *286:* 531–537, 1999.
6. Pomeroy, S. L., Tamayo, P., Gaasenbeek, M., Sturla, L. M., Angelo, M., McLaughlin, M. E., Kim, J. Y. H., Goumnerova, L. C., Black, P. M., Lau, C., *et al.* Prediction of central nervous system embryonal tumor outcome based on gene expression. Nature (Lond.), *415:* 436–442, 2002.
7. 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.
8. Wang, K., Gan, L., Jeffery, E., Gayle, M., Gown, A. M., Skelly, M., Nelson, P. S., Ng, W. V., Schummer, M., Hood, L., and Mulligan, J. Monitoring gene expression profile changes in ovarian carcinomas using cDNA microarrays. Gene (Amst.), *229:* 101–108, 1999.

9. Warrington, J. A., Nair, A., Hahadevappa, M., and Tsyganskaya, M. Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. Physiol. Genomics, *2:* 143–147, 2000.

10. O'Dell, S. D., Bujac, S. R., Miller, G. J., and Day, I. N. Associations of IGF2 ApaI RFLP and INS VNTR class I allele size with obesity. Eur. J. Hum. Genet., *7:* 565–576, 1999.

11. 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.

12. Venables, W. N., and Riley, B. D. Modern Applied Statistics with S-Plus. New York: Springer, 1997.

13. Lozano, M. D., Panizo, A., Toledo, G. R., Sola, J. J., and Pardo-Mindan, J. Immunocytochemistry in the differential diagnosis of serous effusions: a comparative evaluation of eight monoclonal antibodies in Papanicolaou stained smears. Cancer (Phila.), *93:* 68–72, 2001.

14. Sato, S., Okamoto, S., Ito, K., Konno, R., and Yajima, A. Differential diagnosis of mesothelial and ovarian cancer cells in ascites by immunocytochemistry using Ber-EP4 and calretinin. Acta Cytol., *44:* 485–488, 2000.

15. Di Loreto, C., Puglisi, F., Di Lauro, V., Damante, G., and Beltrami, C. A. TTF-1 protein expression in malignant pleural mesotheliomas and adenocarcinomas of the lung. Cancer Lett., *124:* 73–78, 1998.

16. Ordonez, N. G. The value of antibodies 44–36A. SM3, HBME-1, and thrombomodulin in differentiating epithelial pleural mesothelioma from lung adenocarcinoma, Am. J. Surg. Pathol., *21:* 1399–1408, 1997.

17. Michl, P., Buchholz, M., Rolke, M., Kunsch, S., Lohr, M., McClane, B., Tsukita, S., Leder, G., Adler, G., and Gress, T. M. Claudin-4: a new target for pancreatic cancer treatment using *Clostridium perfringens* enterotoxin. Gastroenterology, *121:* 678–684, 2001.

18. 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 (Phila.), *60:* 6281–6287, 2000.

19. Alberti, S., Nutini, M., and Herzenberg, L. A. DNA methylation prevents the amplification of *TROP1*, a tumor-associated cell surface antigen gene. Proc. Natl. Acad. Sci. USA, *91:* 5833–5837, 1994.

20. Calabrese, G., Crescenzi, C., Morizio, E., Palka, G., Guerra, E., and Alberti, S. Assignment of TACSTD1 (alias TROP1. M4S1) to human chromosome 2p21 and the refinement of mapping of TACSTD2 (alias TROP2, M1S1) to human chromosome 1p32 by *in situ* hybridization, Cytogenet. Cell Genet., *92:* 164–165, 2001.