

## A Molecular Diagnostic Test for Distinguishing Lung Adenocarcinoma from Malignant Mesothelioma Using Cells Collected from Pleural Effusions

Andrew J. Holloway,<sup>1,2</sup> Dileepa S. Diyagama,<sup>1</sup> Ken Opeskin,<sup>3,4</sup> Jenette Creaney,<sup>5</sup> Bruce W.S. Robinson,<sup>5,6</sup> Richard A. Lake,<sup>5,6</sup> and David D.L. Bowtell<sup>1,2</sup>

**Abstract Purpose:** Patients with malignant mesothelioma or adenocarcinoma of the lung often present with respiratory complications associated with a malignant pleural effusion. Distinguishing between these malignancies is frequently problematic, as many of the clinical, cytologic, and histologic features of the diseases overlap. Following cytologic analysis of pleural effusions, subsequent confirmatory tissue biopsies involve increased patient morbidity and expense. We have therefore designed a gene expression – based test to classify the primary tumor causing a malignant pleural effusion, using cells collected from the effusion itself.

**Experimental Design:** We have used microarray data for 190 lung adenocarcinomas and 33 malignant mesotheliomas to identify genes differentially expressed between the two diseases. Genes expressed in normal mesothelial cells were removed, allowing the development of a PCR-based test to measure the expression of genes that discriminate between mesothelioma and lung adenocarcinoma from cytology specimens.

**Results:** Applying an real-time PCR – based assay involving 17 genes to 13 independent samples from biopsy-proven malignant mesothelioma and lung adenocarcinomas resulted in the correct identification of all samples.

**Conclusions:** We have developed a test that is able to distinguish between lung adenocarcinoma and mesothelioma in cells collected from pleural effusions.

Approximately 90% of patients with malignant pleural mesothelioma (1) and 15% of patients with lung adenocarcinoma (2) present with a malignant pleural effusion (MPE). Clinically differentiating between the most common, epithelioid form of mesothelioma and lung adenocarcinoma is often problematic, as many of the features of the two diseases overlap, the most significant being chest pain, shortness of breath, and fatigue in older adults (3). A correct diagnosis is important, as the treatment regimes for the two diseases are distinct (4–7),

particularly with respect to the recent demonstration of efficacy of pemetrexed in malignant mesothelioma (8, 9). There may also be medicolegal implications regarding environmental exposure to asbestos if a diagnosis of mesothelioma is made (10).

Multiple approaches have been used to make the differential diagnosis between lung adenocarcinoma and mesothelioma using cells or tissue. Currently, immunohistochemistry and other stains are commonly used for diagnosis. In particular, immunohistochemistry using antibodies to factors such as calretinin, carcinoembryonic antigen, cytokeratins, and thyroid transcription factor, are commonly used (11–15). In general, whereas histologic examination with immunohistochemistry using a panel of antibodies usually allows accurate diagnosis (16, 17), there remains some doubt regarding the best markers to use, and there is always some subjectivity regarding histologic interpretation.

The most common initial approach to the differential diagnosis of pleural samples is cytology (18). Several cytologic features are reported to be useful in differentiating mesothelioma from lung adenocarcinoma, but no one set of features is pathognomic (19). Furthermore, the consistency of observation of some cytologic features is dependent on the mode of collection and processing of cell specimens (20), and there is a subjective element in the observation and interpretation of diagnostic features. Overall, the sensitivity of cytology for the diagnosis of malignant mesothelioma has been reported to vary from 4% to 77% (18), implying that the method has limited utility in clinical settings. Other markers to distinguish mesothelioma from lung adenocarcinoma in pleural effusions

**Authors' Affiliations:** <sup>1</sup>Ian Potter Centre for Cancer Genomics and Predictive Medicine, Peter MacCallum Cancer Centre, East Melbourne, Departments of <sup>2</sup>Biochemistry and Molecular Biology and <sup>3</sup>Pathology, University of Melbourne, Parkville, <sup>4</sup>Department of Anatomical Pathology, St. Vincent's Hospital, Fitzroy, Victoria, <sup>5</sup>School of Medicine and Pharmacology, University of Western Australia, Crawley, and <sup>6</sup>Western Australian Institute of Medical Research, Queen Elizabeth II Medical Centre, Nedlands, Western Australia, Australia  
Received 4/26/06; revised 6/20/06; accepted 6/29/06.

**Grant support:** Australian National Health and Medical Research Council project grant (350466) to A.J. Holloway, B.W.S. Robinson, and R.A. Lake.

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.

**Note:** Supplementary data for this article are available at Clinical Cancer Research Online (<http://clincancerres.aacrjournals.org/>).

**Requests for reprints:** Andrew J. Holloway, Ian Potter Centre for Cancer Genomics and Predictive Medicine, Peter MacCallum Cancer Centre, Locked Bag 1, A'Beckett Street, Melbourne, Victoria, Australia 8006. Phone: 61-3-9656-1789; Fax: 61-3-9656-1414; E-mail: [andrew.holloway@petermac.org](mailto:andrew.holloway@petermac.org).

© 2006 American Association for Cancer Research.

doi:10.1158/1078-0432.CCR-06-1027

have been suggested, such as the specific accumulation of hyaluronic acid in mesothelioma cells (21), or CDKN2A loss (22), although this genetic lesion may be absent in as many as 30% of mesotheliomas (23). Typically, therefore, both cytologic examination of pleural fluid and subsequent histologic examination of pleural biopsy material are used for diagnosis.

Recently, there has been a growing interest in the use of gene expression profiling for diagnosis and prognostication in many cancers. We have previously shown that tumor type can be readily identified in cases of cancers of unknown primary site using expression profiling (24). More specifically, gene expression profiling has been shown to distinguish mesothelioma from lung adenocarcinoma in samples of solid tumors. Gordon et al. (25) reported that by using PCR to measure the expression of four genes, it was possible to correctly classify tumors as being mesothelioma or lung adenocarcinoma. In preliminary work, we found that this test did well in solid tumor specimens (75% accuracy; Supplementary Fig. S1), but the utility of this test for evaluating pleural effusions is unknown. Given that mesothelial cells contaminate MPEs derived from both lung adenocarcinoma and mesothelioma, it is possible that a classifier developed on

solid tissue samples of these diseases may not work effectively with corresponding pleural effusions.

Here, we describe an expression-based classifier designed to differentiate adenocarcinoma of the lung and mesothelioma, through the analysis of MPE. By specifically avoiding genes that are expressed in normal mesothelial cells, we were able to develop a quantitative real-time PCR (RT-PCR)-based assay that reliably discriminates between lung adenocarcinoma and mesothelioma using fresh or formalin-fixed paraffin-embedded samples from MPEs. We show the utility of this test in a series of patients presenting with a MPE, which was followed by subsequent biopsy to confirm the diagnosis. The assay represents the first expression-based classifier to facilitate the differential diagnosis of lung adenocarcinoma and mesothelioma by making use of MPE. This may potentially avoid the need for tissue biopsy in the future.

### Materials and Methods

**Microarray data.** Microarray data for lung adenocarcinoma and mesothelioma was previously described by Gordon et al. (25), and is

**Table 1.** Genes differentially expressed between samples of lung adenocarcinoma and mesothelioma and not expressed in normal mesothelial cells

Gene (GenBank accession no.)	Primers (5' > 3')	Product size (bp)	Affymetrix fold change (log <sub>2</sub> )	Real-time fold change (log <sub>2</sub> )
<i>SFTP3</i> (NM_000542)	FWD: TTTGTGGAGCAGCACACG REV: TACACTGGAGAGGGCTGGAC	115	7.67	10.7
<i>CEACAM6</i> (M18728)	FWD: ATGAAGAAGCAACCGGACAG REV: ACAGCATCCTTGTCTCCAC	97	6.64	10.4
<i>WFDC2</i> (X63187)	FWD: CCGACAACCTCAAGTGCTG REV: CGAGCTGGGGAAAGTTAATG	105	4.61	3.59
<i>AGR2</i> (AF038451)	FWD: CGAACCTGCAGATACAGCTC REV: GACAGACAGAAGGGCTTGGA	105	4.48	7.40
<i>CLDN7</i> (AJ011497)	FWD: CAAAATGTACGACTCGGTGCT REV: CACAAACATGGCCAGGAAG	100	4.07	4.33
<i>CLDN3</i> (NM_001306)	FWD: AACATCATCAGTCGCAGAA REV: AGCAGCGAGTCGTACACCTT	98	4.07	5.33
<i>ERBB3</i> (NM_001982)	FWD: GATGGGAACCTTGAGATTG REV: CACGAGGACATAGCCTGTCA	91	3.80	4.41
<i>FLJ12443</i> (AF052162)	FWD: CCTCGCACTTGTGCGATC REV: ATGGCCTTCAGCAGGAAGT	96	3.16	2.32
<i>XMP</i> (U52100)	FWD: CTGCAGCCTTGCTGTTTCAT REV: GTGCAATTCGTGTTGTTGGT	103	2.67	3.73
<i>LAD</i> (U42408)	FWD: CTTACAGACCTCCGACTGTG REV: GCTGTGTGGGTGAGGACATA	102	2.36	1.90
<i>SPINT1</i> (BT007425)	FWD: CTCACGAGGGAAGTGTACCG REV: GGTTCCCTGGGGTTGTACCTT	93	1.67	1.59
<i>GAS6</i> (NM_000820)	FWD: CCTCTCTGTGGCACTGGT REV: CGCAGACCTTGATCTCCATT	112	-2.27	-1.06
<i>SEMA3C</i> (AB000220)	FWD: CAGCCCAAGCAAGAGTTTA REV: TCTAAAGGATGGTGGGAAAGG	89	-4.12	-2.50
<i>KIBRA</i> (AB020676)	FWD: AAGCGGTGGAAAAGGAC REV: GGCTTCCTCCAGTTCTCTCA	114	-2.41	-1.09
<i>GFPT2</i> (NM_005110)	FWD: AGGCCGTGGAATTCTTCTTT REV: GAGTTTCCCATCAGCCACTG	110	-3.47	-2.86
<i>S1-5 (EFEMP1)</i> (NM_004105)	FWD: GCTAAGCAGTGACAGGCTCA REV: TTCCAGGTTTCATTGACACA	93	-4.12	-3.37
<i>RALDH2</i> (AB015226)	FWD: GTAGTGGGGAGTCCCTTTGA REV: CCACACCACTCTGGATGAGTT	97	-4.53	-5.75

NOTE: The primers used for quantitative RT-PCR analysis and the expected product sizes are shown. Log<sub>2</sub>-transformed fold change in median expression in lung adenocarcinoma relative to malignant mesothelioma for both Affymetrix U95A and quantitative RT-PCR are given. Gene expression levels were normalized relative to GAPDH probesets (Affymetrix data) or to quantitative RT-PCR controls.

available at <http://www.chestsurg.org/publications/2002-microarray.aspx>. Briefly, Affymetrix U95A GeneChip data was provided for 190 lung adenocarcinomas, 20 carcinoid tumors of the lung, 6 small cell carcinomas of the lung, 21 squamous cell carcinomas of the lung, 17 samples of normal lung tissue, and 33 mesotheliomas. For this study, a subset comprising the 190 lung adenocarcinomas and 33 mesotheliomas was used. We generated additional Affymetrix U133 plus 2.0 data for 7 samples of normal mesothelial cells using standard Affymetrix protocols with 3 µg of total RNA starting material. Data was extracted and processed from GCOS (GeneChip Operating Software, Affymetrix) CHP files using Genespring (Agilent Technologies, Foster City, CA). Data was further processed by transforming (resetting) data points that are <0.01 to 0.01. A per-gene normalization was used across all samples.

**Isolation and expression profiling of normal human mesothelial cells.** Normal mesothelial cells were recovered from pericardial fluid obtained from patients undergoing intrathoracic surgery with no evidence of malignancy. Cells were washed in RPMI 1640 supplemented with L-glutamine and containing 15% FCS, then cultured in this medium further supplemented with insulin/transferrin/selenium A (Invitrogen, Carlsbad, CA), 10 µmol/L of 2-mercaptoethanol (Sigma, Castlehill, New South Wales, Australia), 20 mmol/L of HEPES (Sigma), and 400 µg/L of hydrocortisone (Sigma) and antibiotics [120 mg/L of benzylpenicillin (CSL, Parkville, Victoria, Australia), 4 mg/L of gentamicin (Invitrogen), and 2.5 mg/L of Amphotericin (Sigma)]. The cells were incubated in a humidified 5% CO<sub>2</sub> atmosphere at 37°C, and the medium was changed after the first day and every third day. When confluent, the cells were washed in PBS and treated with 0.02% trypsin in 0.007% EDTA (Sigma) solution for 5 minutes at 37°C to detach them from the plastic. Cells were cultured for a maximum of three passages, that is, until ~2 × 10<sup>7</sup> cells were available to enable parallel RNA extraction and immunostaining. The cells were then washed with complete medium and counted. RNA was extracted from cells as previously described (24). Only normal mesothelial cell cultures which had >95% of cells staining positive for calretinin (DakoCytomation, Glostrup, Denmark) were used for studies.

**Gene filtering.** To identify genes not expressed in normal mesothelial cells, ~21,000 genes with absent expression, as determined from CHP files from seven samples of normal human cultured mesothelial cells profiled as described above, were selected. U133 plus 2.0 probe-set gene identifiers were converted to U95A v2 probe-set identifiers using the "translate between array platforms" function in GeneSpring. Approximately 25,000 probe-set identifiers were found for U95Av2 GeneChips. Using only these genes as a starting list, we identified genes differentially expressed between mesothelioma and lung adenocarcinoma using a signal to noise metric. Briefly, genes were ranked according to a signal to noise metric as determined by the equation:

$$S_g = (\mu_{\text{class 1}} - \mu_{\text{class 2}}) / (\sigma_{\text{class 1}} + \sigma_{\text{class 2}})$$

where, for each gene,  $\mu_{\text{class 1}}$  represents the mean value and  $\sigma_{\text{class 1}}$  represents the SD for that gene in samples from class 1. Seventeen genes were selected for further analysis by quantitative RT-PCR.

**Tissue samples.** Tissues were sourced from the Peter MacCallum Cancer Centre, St. Vincent's Hospital, Melbourne or PathWest, Perth, Western Australia. Patient consent and relevant Human Research Ethics Committee approval were obtained consistent with Australian National Health and Medical Research Council National Statement on Ethical Conduct in Research Involving Humans (1999). Fresh-frozen solid tumor or cytology samples (lung adenocarcinoma or mesothelioma) were collected from patients with an unequivocal diagnosis based on a combination of confirmed primary site, and review of histology and immunohistochemical staining. Formalin-fixed paraffin-embedded specimens were collected from patients with an initial pleural effusion sample with an either definitive diagnosis of mesothelioma or lung adenocarcinoma and a later tissue biopsy sample available for definitive diagnosis. The cases were centrally reviewed to confirm the original diagnosis.

**Quantitative RT-PCR analysis of gene expression.** Primers were designed for putative discriminating genes using the program Primer 3

([http://frodo.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)). Wherever possible, primers were designed to span a large intron (>500 bp) as determined using the University of California, Santa Cruz genome browser (<http://genome.ucsc.edu/>). Product size was set to between 85 and 115 bp, with an optimal value of 100 bp. Optimal  $T_m$  was set to 60°C, with a range from 58.5°C to 61.5°C. Primers for the genes used in the final analysis are shown in Table 1. Total RNA was isolated from fresh tissue or cell specimens, or from formalin-fixed paraffin-embedded tissue, using methods previously described (24). One microgram of total RNA was converted to cDNA using poly(T) priming and Moloney murine leukemia virus point mutant reverse transcriptase kit (Promega, Madison, WI). Ten nanograms of cDNA was used as template in RT-PCR reactions. Reactions were set up in 384-well microtiter plates by a Biomek2000 robot (Beckman, Fullerton, CA), and run in an ABI7900HT (Applied Biosystems, Foster City, CA) instrument according to the manufacturer's standard SYBR green assay. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was included as a normalizing control. Primer sequences were forward: TCAACGACCACCTTTGCAAGCTCA and reverse: GCTGGTGGTCCAGGGTCTTACT. Fold change values relative to GAPDH ( $X$ ), were calculated using the standard  $\delta Ct$  formula (24) as follows:

$$X = 2^{-\Delta Ct}, \text{ where } \Delta Ct = Ct(\text{target}) - Ct(\text{GAPDH})$$

**Bioinformatic analysis.** Data was clustered using the program Cluster and visualized using the program TreeView (26). Class prediction was carried out using GeneSpring. Briefly, quantitative RT-PCR gene expression data from fresh-frozen tumor specimens of lung adenocarcinoma ( $n = 13$ ) and mesothelioma ( $n = 12$ ) was used to train a support vector machine algorithm. To achieve the optimum output, we refined the algorithm by adjusting the gene selection method, the number of genes used, and Kernel function variables. The optimum output (100% correct in leave-one-out cross-validation of training set) was obtained when we used Fisher's exact test using all 17 genes and a Radial basis Kernel function. The algorithm was then applied to the test set, quantitative RT-PCR gene expression data derived from

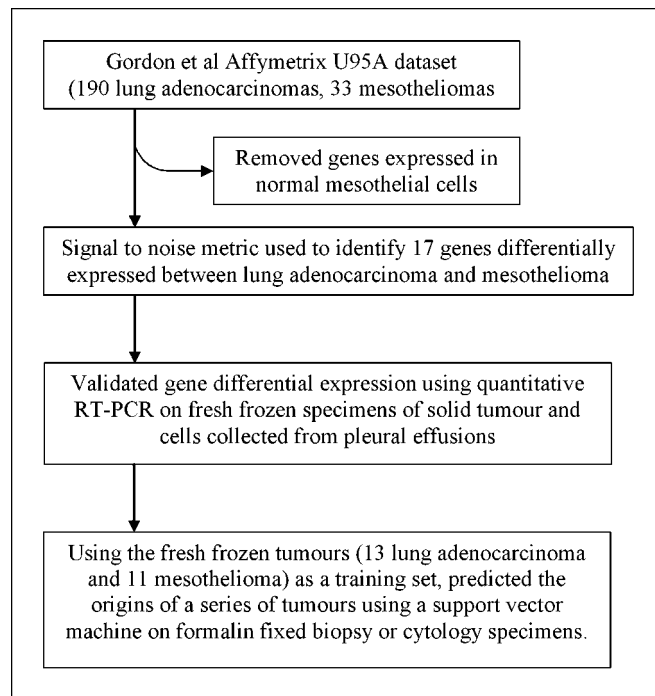


Fig. 1. Flowchart describing the experimental flow in designing a test with applicability to cells collected from pleural effusions.

formalin-fixed paraffin-embedded cytology or biopsy specimens of lung adenocarcinoma ( $n = 13$ ) or malignant mesothelioma ( $n = 18$ ).

## Results

We initially used the gene set defined by Gordon et al. (25) to diagnose the origin of the tumor in patients with either lung adenocarcinoma or mesothelioma and achieved a classification rate equivalent to that published when solid tissue biopsies were tested (75%; Supplementary Fig. S1). We were, however, unable to accurately determine the nature of the primary tumor using MPE from lung adenocarcinoma and mesothelioma, perhaps because of the presence of nonmalignant mesothelial cells in pleural effusions. We therefore sought to specifically develop a test which could classify MPE.

A schematic description of our approach to assay development and validation is provided in Fig. 1. To identify genes not expressed in normal mesothelial cells for inclusion in our gene selection algorithm, we profiled a series of seven samples of primary cultures of normal human mesothelial cells using Affymetrix U133 plus 2.0 GeneChips. Using GC-RMA, we identified ~21,000 probesets, which were below the limit of reliable detection in all samples. To integrate this data with the Gordon

Affymetrix U95v2 data set (25), we matched the U133 plus 2.0 probeset identifiers to U95v2 identifiers. This yielded a list of ~25,000 Affymetrix U95v2 probesets with expression below the limit of detection in normal mesothelial cells. We then ranked this gene list by differential expression between lung adenocarcinoma and mesothelioma according to a signal to noise metric, and identified a list of 17 differentially expressed genes.

To validate the restricted expression of the 17 genes, we designed primer pairs and did quantitative RT-PCR on RNA derived from an independent series of fresh-frozen solid tissue biopsy specimens of lung adenocarcinoma ( $n = 13$ ), mesothelioma ( $n = 12$ ), and normal mesothelial cells ( $n = 7$ ). Gene expression levels were normalized to GAPDH. As expected, there was a significant difference in the expression of the majority of the target genes between the two malignancies (15 of 17,  $P < 0.05$ ; GAS6 and KIBRA did not reach significance, but were retained for the support vector machine analysis, see below), and no gene showed substantial expression in normal mesothelial cells (Fig. 2). We normalized the expression of the genes using the expression of GAPDH in the Affymetrix data set, and found a very high degree of correlation between the Affymetrix data and the RT-PCR data for the 17 genes (Pearson correlation 0.96; Table 1).

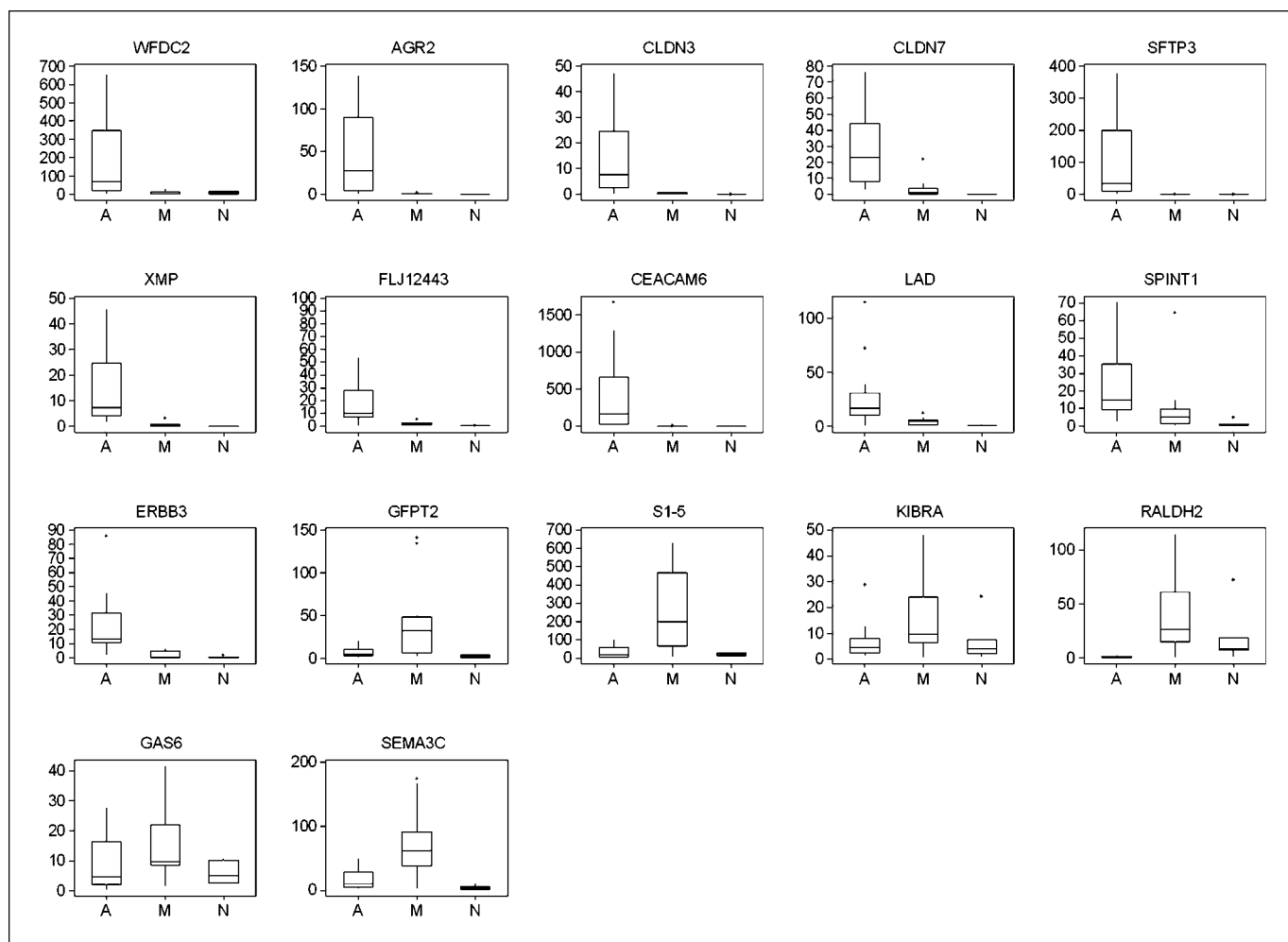


Fig. 2. Boxplots showing expression of selected genes, in lung adenocarcinoma, mesothelioma, and cultured normal mesothelial cells. Gene expression values, relative to GAPDH, are shown for lung adenocarcinoma (A), mesothelioma (M), and normal mesothelial cells (N).

Clustering of tumors using GAPDH normalized Affymetrix data for the 17 genes yielded a near-perfect segregation of the two tumor classes (Fig. 3A), indicating, as expected, that the genes were useful in segregating the two tumor classes. The genes were similarly effective in discriminating between fresh-frozen solid tumor samples in a validation set (Fig. 3B). To investigate whether the genes were useful in discriminating between pleural effusions caused by either lung adenocarcinoma or malignant mesothelioma, we did quantitative RT-PCR analysis on a series of fresh-frozen specimens of cells collected from pleural effusions, from either lung adenocarcinoma ( $n = 5$ ), or mesothelioma ( $n = 11$ ), unrelated to the solid tumors above. Although the clustering did not show perfect segregation of the two classes (one lung adenocarcinoma grouped with the mesothelioma samples; Fig. 3C), differential expression of the genes was essentially recapitulated in the cytology specimens.

To show the application of a test using this gene set in a clinical context, we selected a series of additional patients from whom formalin-fixed paraffin-embedded matched cytology and biopsy specimens were available. We prepared RNA from tissue sections, and did quantitative RT-PCR on cDNA in a manner identical to the samples previously assayed. Using a support vector machine trained using the fresh-frozen samples of solid tumors of known origin described above, we predicted the origin of both the cytology and biopsy specimens in our series of patients (Table 2). A result was obtained for all cytology specimens, and all biopsy specimens except one in which no usable RNA was obtained (patient 5305). In some patients, multiple cytology and biopsy specimens were

**Table 2.** Validation of diagnostic test on a series of archival, paraffin-embedded specimens of tumor biopsy or pleural effusion cell pellet from patients with lung adenocarcinoma or mesothelioma

Patient no.	Pleural effusion cell pellet	Biopsy specimen
Patients with lung adenocarcinoma		
5799	Lung adenocarcinoma	Lung adenocarcinoma (2)
6775	Lung adenocarcinoma (2)	Lung adenocarcinoma
9143	Lung adenocarcinoma	Lung adenocarcinoma (3)
5305	Lung adenocarcinoma	no RNA
6702	Lung adenocarcinoma	Lung adenocarcinoma
Patients with mesothelioma		
5450	Mesothelioma	Mesothelioma
6854	Mesothelioma	Mesothelioma
D033	Mesothelioma	Mesothelioma
E071	Mesothelioma	Mesothelioma
F037	Mesothelioma	Mesothelioma
F404	Mesothelioma	Mesothelioma
G431	Mesothelioma (2)	Mesothelioma
L209	Mesothelioma (2)	Mesothelioma

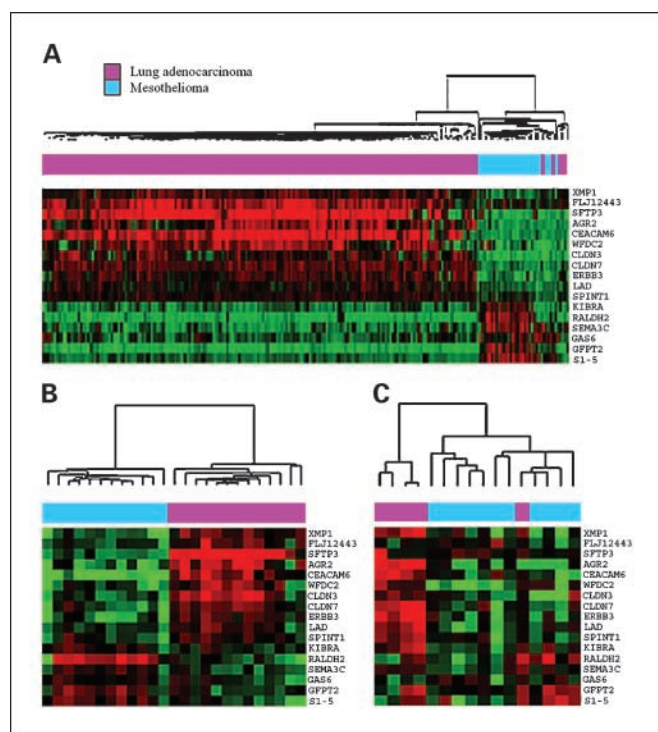
NOTE: Numbers in parentheses indicate the number of independent samples tested in patients from whom multiple samplings were available.

available, and the test was consistent in predicting the source of the underlying malignancy in all samples. When the quantitative RT-PCR result was compared with the definitive diagnosis made from pathology review of cytology, histology and immunohistochemistry, and clinical history, the test was 100% accurate in identifying the malignancy (Supplementary Table S1). Importantly, all cytology specimens yielded a correct result, indicating that the test would have identified the type of malignancy in these patients prior to biopsy of tumor material.

## Discussion

The classification of the origin and prognosis of tumors using gene expression-based tests is an attractive adjunct to current approaches (27). One of the areas in which this type of information is likely to be useful is in making differential diagnoses between tumors that are clinically and histologically similar, for example, MPE from lung adenocarcinoma and mesothelioma. It is generally accepted that tumor samples contain more than one cell type, and unless steps such as microdissection are taken, expression analysis represents the sum of the genes expressed in all cell types (28). Thus, if the distinction between two tumors is restricted to a minority subset of the total cells in a sample, it is reasonable to assume that these differences may be missed unless an approach is taken to minimize the contribution of the major nontumor subset of cells.

A simple comparison of gene expression differences between solid tissue samples of mesothelioma and lung adenocarcinoma would be expected to identify genes expressed predominantly in the mesothelial cell lineage. If so, the large number of normal nonmalignant mesothelial cells present in many MPEs, irrespective of the nature of the primary tumor, would be expected to interfere with accuracy of such a classifier. Although we have not investigated whether this accounted for our failure to accurately



**Fig. 3.** A, hierarchical cluster analysis of U95A Affymetrix gene expression data set using 17 genes selected with differential expression between mesothelioma and lung adenocarcinoma. B, cluster analysis of quantitative RT-PCR data obtained from fresh-frozen mesothelioma and lung adenocarcinoma samples for the same genes as shown in (A). C, cluster analysis of quantitative RT-PCR analysis of fresh-frozen samples of MPE taken from patients with biopsy-proven mesothelioma or lung adenocarcinoma. All data are normalized to GAPDH expression.

classify MPE when we used the genes defined by Gordon et al. (25), we note that both calretinin and VAC- $\beta$  are abundantly expressed in our samples of normal mesothelial cells.

We therefore used a subtractive approach to remove genes that are expressed in mesothelial cells, whereas focusing on genes that are likely to be differentially expressed between lung adenocarcinoma and mesothelioma. The genes selected were differentially expressed between the two malignancies in question, but they did not necessarily have expression patterns restricted solely to the tumor type with which the marker was associated, implying that lung adenocarcinoma and mesothelioma share the expression of some genes with other malignancies or normal tissues. Several of the genes we have identified as discriminating markers have been previously shown to be expressed in malignancies of various types, for example *CEACAM6* (29), *claudin 3* and *claudin 7* (30, 31), *LAD-1* (32), *AGR2* (33, 34), and *ERBB3* (35, 36). Other genes have previously been shown to have tissue-restricted expression, and have not been shown to be expressed in either of these malignancies, for example *SFTP3* (37), which has expression restricted to the adult lung. Only one gene, *claudin 7*, was also found in the classifier built by Gordon et al. (25).

We believe that excluding genes expressed in normal mesothelial cells is crucial to designing a test for use in specimens of pleural effusion. We are aware, however, that our gene selection method was probably compromised by this requirement for the following reasons: first, we do not anticipate that the relatively small sample of normal mesothelial cells we profiled following *in vitro* culture will necessarily express all genes expressed in mesothelial cells (normal and reactive) in a MPE *in vivo*, and second, given the large similarity of gene expression in normal mesothelial cells and mesothelioma, it was difficult to identify genes with expression restricted to just the tumor, which were also differentially expressed with lung adenocarcinoma. Nevertheless, we have been able to identify a gene set differentially transcribed by lung adenocarcinoma and mesothelioma tumor cells for the purposes of making a differential diagnosis in these two malignancies. The usual differential diagnosis of malignant tumor in pleural fluid is lung adenocarcinoma and mesothelioma. Uncommonly, MPE can

be generated from other malignancies, such as metastatic breast and ovarian cancer, lymphoma, and other histologic subtypes of lung cancer. Clearly, evaluation of the test for these less common causes of MPE would be desirable, as would an evaluation of the test on a larger number of cases of malignant mesothelioma and adenocarcinoma.

We have shown that the test we designed can be done on formalin-fixed cytology specimens. There are several advantages in accessing cytology specimens to make definitive diagnoses of tumor origins. First, a MPE is a very common mode of presentation of patients with either lung adenocarcinoma or mesothelioma (38), and cells collected from MPE represent an easily collected source of material from which to make a diagnosis. Furthermore, the test we describe can be used on formalin-fixed paraffin-embedded specimens, making it applicable within the context of standard pathology specimen handling. We have also shown that our test is accurate in tumor biopsy specimens. Therefore, in those cases with inaccessible pleural effusion, analysis of a pleural biopsy should allow a distinction between mesothelioma and adenocarcinoma of the lung.

In conclusion, we present a robust test which is able to discriminate between adenocarcinoma of the lung and malignant mesothelioma using cells collected from MPE, even after formalin fixation and paraffin embedding. Pleural fluid can be obtained more easily and with significantly less patient morbidity than a solid tissue biopsy of tumor material. We therefore suggest that a molecular analysis of cells collected from a MPE may be useful in the differential diagnosis of lung adenocarcinoma and malignant mesothelioma, particularly as quantitative RT-PCR becomes a more widely practiced technique in pathology laboratories.

### Acknowledgments

We are grateful to Dr. Amanda Segal, PathWest, for expert advice and assistance with this study; Michelle Murphy, Tumour Immunology Group at UWA, for sharing unpublished data; and to members of the Ian Potter Centre for Cancer Genomics and Predictive Medicine at the Peter MacCallum Cancer Centre, particularly Drs. Alex Boussioutas, Izhak Haviv, Andreas Möller, David Thomas, Anna Tinker, and Richard Tothill for helpful comments on this work. We wish especially to thank all patients who donated samples to this study.

### References

1. Nguyen GK, Akin MR, Villanueva RR, Slatnik J. Cytopathology of malignant mesothelioma of the pleura in fine-needle aspiration biopsy. *Diagn Cytopathol* 1999; 21:253–9.
2. Hung TL, Chen FF, Liu JM, et al. Clinical evaluation of HER-2/neu protein in malignant pleural effusion-associated lung adenocarcinoma and as a tumor marker in pleural effusion diagnosis. *Clin Cancer Res* 2003;9:2605–12.
3. Tang P, Vatsia SK, Teichberg S, Kahn E. Pulmonary adenocarcinoma simulating malignant mesothelioma. *Arch Pathol Lab Med* 2001;125:1598–600.
4. Pass HI. Malignant pleural mesothelioma: surgical roles and novel therapies. *Clin Lung Cancer* 2001;3: 102–17.
5. Onn A, Tsuboi M, Thatcher N. Treatment of non-small-cell lung cancer: a perspective on the recent advances and the experience with gefitinib. *Br J Cancer* 2004; 91 Suppl 2:S11–7.
6. Robinson BW, Lake RA. Advances in malignant mesothelioma. *N Engl J Med* 2005;353:1591–603.
7. Neragi-Miandoab S. Multimodality approach in management of malignant pleural mesothelioma. *Eur J Cardiothorac Surg* 2006;29:14–9.
8. Kerr C. Pemetrexed combination improves mesothelioma survival. *Lancet Oncol* 2005;6:548.
9. Ceresoli GL, Zucali PA, Favaretto AG, et al. Phase II study of pemetrexed plus carboplatin in malignant pleural mesothelioma. *J Clin Oncol* 2006;24:1443–8.
10. Cugell DW, Kamp DW. Asbestos and the pleura: a review. *Chest* 2004;125:1103–17.
11. Ordonez NG. The immunohistochemical diagnosis of mesothelioma: a comparative study of epithelioid mesothelioma and lung adenocarcinoma. *Am J Surg Pathol* 2003;27:1031–51.
12. Ordonez NG. Value of mesothelin immunostaining in the diagnosis of mesothelioma. *Mod Pathol* 2003;16: 192–7.
13. Ordonez NG. Role of immunohistochemistry in differentiating epithelial mesothelioma from adenocarcinoma. Review and update. *Am J Clin Pathol* 1999; 112:75–89.
14. Abutaily AS, Addis BJ, Roche WR. Immunohistochemistry in the distinction between malignant mesothelioma and pulmonary adenocarcinoma: a critical evaluation of new antibodies. *J Clin Pathol* 2002;55: 662–8.
15. Hecht JL, Pinkus JL, Pinkus GS. Monoclonal antibody MOC-31 reactivity as a marker for adenocarcinoma in cytologic preparations. *Cancer* 2005;108:56–9.
16. King JE, Thatcher N, Pickering CA, Hasleton PS. Sensitivity and specificity of immunohistochemical markers used in the diagnosis of epithelioid mesothelioma: a detailed systematic analysis using published data. *Histopathology* 2006;48:223–32.
17. Hammar SP. Macroscopic, histologic, histochemical, immunohistochemical, and ultrastructural features of mesothelioma. *Ultrastruct Pathol* 2006;30:3–17.
18. Renshaw AA, Dean BR, Antman KH, Sugarbaker DJ, Cibas ES. The role of cytologic evaluation of pleural fluid in the diagnosis of malignant mesothelioma. *Chest* 1997;111:106–9.
19. Whitaker D. Invited review. The cytology of malignant mesothelioma. *Cytopathology* 2000;11:139–51.
20. Ylagan LR, Zhai J. The value of ThinPrep and cytospin preparation in pleural effusion cytological

- diagnosis of mesothelioma and adenocarcinoma. *Diagn Cytopathol* 2005;32:137–44.
21. Afify AM, Stern R, Michael CW. Differentiation of mesothelioma from adenocarcinoma in serous effusions: the role of hyaluronic acid and CD44 localization. *Diagn Cytopathol* 2005;32:145–50.
  22. Illei PB, Ladanyi M, Rusch VW, Zakowski MF. The use of CDKN2A deletion as a diagnostic marker for malignant mesothelioma in body cavity effusions. *Cancer* 2003;99:51–6.
  23. Illei PB, Rusch VW, Zakowski MF, Ladanyi M. Homozygous deletion of CDKN2A and codeletion of the methylthioadenosine phosphorylase gene in the majority of pleural mesotheliomas. *Clin Cancer Res* 2003;9:2108–13.
  24. Tothill RW, Kowalczyk A, Rischin D, et al. An expression-based site of origin diagnostic method designed for clinical application to cancer of unknown origin. *Cancer Res* 2005;65:4031–40.
  25. Gordon GJ, Jensen RV, Hsiao LL, et al. Translation of microarray data into clinically relevant cancer diagnostic tests using gene expression ratios in lung cancer and mesothelioma. *Cancer Res* 2002;62:4963–7.
  26. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 1998;95:14863–8.
  27. Ramaswamy S, Golub TR. DNA microarrays in clinical oncology. *J Clin Oncol* 2002;20:1932–41.
  28. Todd R, Lingen MW, Kuo WP. Gene expression profiling using laser capture microdissection. *Expert Rev Mol Diagn* 2002;2:497–507.
  29. Blumenthal RD, Hansen HJ, Goldenberg DM. Inhibition of adhesion, invasion, and metastasis by antibodies targeting CEACAM6 (NCA-90) and CEACAM5 (carcinoembryonic antigen). *Cancer Res* 2005;65:8809–17.
  30. Morin PJ. Claudin proteins in human cancer: promising new targets for diagnosis and therapy. *Cancer Res* 2005;65:9603–6.
  31. Swisshelm K, Macek R, Kubbies M. Role of claudins in tumorigenesis. *Adv Drug Deliv Rev* 2005;57:919–28.
  32. Santin AD, Zhan F, Bellone S, et al. Gene expression profiles in primary ovarian serous papillary tumors and normal ovarian epithelium: identification of candidate molecular markers for ovarian cancer diagnosis and therapy. *Int J Cancer* 2004;112:14–25.
  33. Smirnov DA, Zweitzig DR, Foulk BW, et al. Global gene expression profiling of circulating tumor cells. *Cancer Res* 2005;65:4993–7.
  34. Liu D, Rudland PS, Sibson DR, Platt-Higgins A, Barraclough R. Human homologue of cement gland protein, a novel metastasis inducer associated with breast carcinomas. *Cancer Res* 2005;65:3796–805.
  35. Holbro T, Civenni G, Hynes NE. The ErbB receptors and their role in cancer progression. *Exp Cell Res* 2003;284:99–110.
  36. Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* 2001;2:127–37.
  37. Vamvakopoulos NC, Modi WS, Floros J. Mapping the human pulmonary surfactant-associated protein B gene (SFTP3) to chromosome 2p12-→p11.2. *Cytogenet Cell Genet* 1995;68:8–10.
  38. DiBonito L, Falconieri G, Colautti I, Bonifacio D, Dudine S. The positive pleural effusion. A retrospective study of cytopathologic diagnoses with autopsy confirmation. *Acta Cytol* 1992;36:329–32.