

Epigenetic Profiles Distinguish Pleural Mesothelioma from Normal Pleura and Predict Lung Asbestos Burden and Clinical Outcome

Brock C. Christensen,^{1,2} E.A. Houseman,³ John J. Godleski,⁴ Carmen J. Marsit,² Jennifer L. Longacker,⁵ Cora R. Roelofs,³ Margaret R. Karagas,⁷ Margaret R. Wrensch,⁸ Ru-Fang Yeh,⁹ Heather H. Nelson,¹⁰ Joe L. Wiemels,⁹ Shichun Zheng,⁸ John K. Wiencke,⁸ Raphael Bueno,⁶ David J. Sugarbaker,⁶ and Karl T. Kelsey^{1,2}

Departments of ¹Community Health, Center for Environmental Health and Technology and ²Pathology and Laboratory Medicine, Brown University, Providence, Rhode Island; ³Department of Work Environment, University of Massachusetts Lowell, Lowell, Massachusetts; ⁴Department of Environmental Health, Harvard School of Public Health, ⁵Department of Environmental Health, Boston University School of Public Health, and ⁶Division of Thoracic Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts; ⁷Department of Community and Family Medicine, Dartmouth Medical School, Lebanon, New Hampshire; Departments of ⁸Neurological Surgery and ⁹Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California; and ¹⁰Division of Epidemiology and Community Health, Masonic Cancer Center, University of Minnesota, Minneapolis, Minnesota

Abstract

Mechanisms of action of nonmutagenic carcinogens such as asbestos remain poorly characterized. As pleural mesothelioma is known to have limited numbers of genetic mutations, we aimed to characterize the relationships among gene-locus-specific methylation alterations, disease status, asbestos burden, and survival in this rapidly fatal asbestos-associated tumor. Methylation of 1505 CpG loci associated with 803 cancer-related genes were studied in 158 pleural mesotheliomas and 18 normal pleura. After false-discovery rate correction, 969 CpG loci were independently associated with disease status ($Q < 0.05$). Classifying samples based on CpG methylation profile with a mixture model approach, methylation classes discriminated tumor from normal pleura (permutation $P < 0.0001$). In a random forests classification, the overall misclassification error rate was 3.4%, with $<1\%$ ($n = 1$) of tumors misclassified as normal ($P < 0.0001$). Among tumors, methylation class membership was significantly associated with lung tissue asbestos body burden ($P < 0.03$), and significantly predicted survival (likelihood ratio $P < 0.01$). Consistent with prior work, asbestos burden was associated with an increased risk of death (hazard ratio, 1.4; 95% confidence interval, 1.1–1.8). Our results have shown that methylation profiles powerfully differentiate diseased pleura from nontumor pleura and that asbestos burden and methylation profiles are independent predictors of mesothelioma patient survival. We have added to the growing body of evidence that cellular epigenetic dysregulation is a critical mode of action for asbestos in the induction of pleural mesothelioma. Importantly, these findings hold great promise for using epigenetic profiling in the diagnosis and prognosis of human cancers. [Cancer Res 2009;69(1):227–34]

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

Current affiliations for E.A. Houseman: Department of Community Health, Center for Environmental Health and Technology, Brown University, Providence, Rhode Island and Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts.

Requests for reprints: Karl T. Kelsey, Brown University, Box GE-5, 70 Ship Street, Providence, RI 02903. Phone: 401-863-6420; Fax: 401-863-9008; E-mail: Karl_Kelsey@Brown.edu.

©2009 American Association for Cancer Research.
doi:10.1158/0008-5472.CAN-08-2586

Introduction

A central tenet of cancer biology states that cancer is clonal, with tumors arising as the result of expansion of increasingly dysregulated cells. This insight yielded a now well-known paradigm that selective expansion of cells with a growth advantage occurs in an ordered fashion, driven primarily by genetic changes (1). This model has expanded to now include the thesis that cancers also evolve a “mutator phenotype” and become malignant as a result of somatic genetic events (2). Although this is almost certainly true of some cancers, particularly those induced by well-characterized mutagens (e.g., tobacco smoke and ionizing radiation), other known human carcinogens are not mutagenic (or are very poor mutagens) and may be less prone to induce cancers via this mechanism. Asbestos, which is known to induce mesothelioma, is an example of a nonmutagenic carcinogen. In the case of carcinogens such as asbestos, it may be that dysregulation of the somatic epigenome is equally, if not more crucial for cancer development.

Aberrant epigenetic events, including DNA hypermethylation-induced gene silencing, are well-recognized as important contributors to carcinogenesis. Methylation-associated gene silencing occurs when certain cytosines in specific clustered regions primarily located in gene promoters are hypermethylated. These regulatory CpG islands often occur in tumor suppressor genes and are thought to remain largely unmethylated in noncancerous cells. Approximately half of all human genes contain CpG islands and are, therefore, potentially subject to this type of aberrant silencing (3). Recent technologic advances allow for the simultaneous resolution of hundreds of specific, phenotypically defined cancer-related methylation events, providing a platform for the rapid epigenetic profiling of gene silencing in human tumors (4).

Malignant pleural mesothelioma is a rapidly fatal malignancy associated with asbestos exposure in ~80% of patients (5). In the United States, Great Britain, and Japan, over 5,000 cases occur annually and median survival of patients with pleural mesothelioma is <1 year (6–8). The economic burden of treating this disease and the litigation associated with asbestos exposure is estimated to exceed \$265 billion over the next four decades in the United States (9). Despite the decline in asbestos use among industrialized nations, the incidence of mesothelioma continues to increase, and it is not expected to peak until 2020, as disease latency can be as

long as 50 years (10). Importantly, asbestos is currently mined and exported throughout the world, with heavy use evident in developing nations such as China, India, and Central America (11). Asbestos-containing products are still imported to the United States, and many asbestos exposure hazards remain from earlier applications; one well-publicized example being dust from the World Trade Center towers collapse in New York City (12). A more complete understanding the molecular-genetic consequences of asbestos exposure and the mechanism of action of these mineral fibers in inducing mesothelioma is critically needed to develop more effective approaches for identifying and treating this devastating disease.

The causal link between asbestos and pleural mesothelioma has been widely accepted since 1960 (13), and the carcinogenic mechanisms of asbestos have been investigated in earnest since that time; establishing that asbestos fibers are not point mutagens but rather both clastogenic and cytotoxic *in vitro* (14, 15). Additionally, methylation-induced tumor suppressor gene silencing has been observed in recent studies of mesothelioma (16–20), leading to the hypothesis that asbestos fibers contribute to epigenetic silencing of tumor suppressor genes in this disease. Consistent with this, Tsou and colleagues (18) observed a significant association between self-reported asbestos exposure and methylation at the *MT1A*, and *MT2A* gene loci in mesotheliomas. Concomitantly, work in our laboratory, using quantitative asbestos body counts as a measure of asbestos exposure burden, revealed an association between cell cycle control tumor suppressor gene methylation and increased asbestos burden in mesothelioma (20).

To comprehensively investigate aberrant tumor-specific, phenotypically relevant methylation events in pleural mesothelioma, we profiled 158 tumors and 18 nontumorigenic parietal pleura samples for methylation at 1505 CpG dinucleotides associated with 803 cancer-related genes using the Illumina GoldenGate methylation bead array. We have definitively delineated the relationship between a comprehensive, phenotypically important CpG methylation profile and disease status, in addition to defining the tumor methylation profiles associated with patient clinical course and asbestos exposure.

Materials and Methods

Study population. Tumor material was obtained after surgical resection at Brigham and Women's Hospital through the support of the International Mesothelioma Program. Similarly, grossly nontumorigenic parietal pleura samples were taken as residual tissue during extrapleural pneumonectomy from uninvolved anatomic sites. Patients were drawn in near equivalent numbers from a pilot study conducted in 2002 ($n = 70$), and an incident case series beginning in 2005 ($n = 88$). Among identified cases, the participation rate was 85%. All patients provided informed consent under the approval of the appropriate Institutional Review Boards. All patients underwent surgical resection before other treatments. Clinical information, including histologic diagnosis, was obtained from pathology reports. Each patient was assessed for history of exposure to asbestos as well as additional demographic and environmental data by obtaining their medical and occupational history with an in-person questionnaire or interview. Additionally, we quantified asbestos bodies in samples of lung tissue from multiple sites in the resected lung (21) as previously described (22). Each tumor was pathologically examined and the amount of tumor in every sample estimated by direct microscopic evaluation and recorded as the percent tumor for that specimen. Patients were followed for survival using the National death index and last known clinical visit.

Methylation analysis. Tumor and nontumor pleural DNA was extracted from frozen tissue using the QIAamp DNA mini kit according to the manufacturer's protocol (Qiagen). DNA was modified by sodium bisulfite to convert unmethylated cytosines to uracil using the EZ DNA Methylation kit (Zymo Research) according to the manufacturer's protocol. Illumina GoldenGate methylation bead arrays were used to simultaneously interrogate 1,505 CpG loci associated with 803 cancer-related genes. The Illumina array interrogates approximately two CpGs per gene, and although sequencing methods would provide additional details, CpGs were cultivated from reports that have shown the methylation-expression relationship in large part through sequencing experiments. Bead arrays have a similar sensitivity as quantitative methylation-specific PCR and were run at the University of California at San Francisco Institute for Human Genetics, Genomics Core Facility, according to the manufacturer's protocol and as described by Bibikova and colleagues (4).

Statistical analysis. BeadStudio Methylation software from the array manufacturer Illumina was used for data set assembly. All array data points are represented by fluorescent signals from both methylated (M) and unmethylated (U) alleles, and methylation level is given by $\beta = (\max(M, 0)) / (|U| + |M| + 100)$, the average methylation (β) value is derived from the ~ 30 replicate methylation measurements and a $Cy3/Cy5$ methylated/unmethylated ratio. At each locus for each sample, the detection P value was used to determine sample performance, 3 samples (2%) with $>25\%$ of loci having a detection P value of $>1e-5$ were dropped from analysis. Similarly, CpG loci with a median detection P value of >0.05 ($n = 8, 0.5\%$), were eliminated from analysis.

Subsequent analyses were carried out using the R software (23). For exploratory and visualization purposes, hierarchical clustering was performed using R function *hclust* with Manhattan metric and average linkage. Associations between sample type, or covariates such as age or gender and methylation at individual CpG loci were tested with a generalized linear model. The β distribution of average β values was accounted for with a quasi-binomial logit link with an estimated scale variable constraining the mean between 0 and 1, in a manner similar to that described by Hsuing and colleagues (24). CpG loci where an *a priori* hypothesis existed were tested independently. In contrast, array-wide scanning for CpG loci associations with sample type or covariate used false discovery rate correction and Q values computed by the *qvalue* package in R (25).

For inference, data were clustered using a mixture model with a mixture of β distributions, and the number of classes was determined by recursively splitting the data via 2-class models, with Bayesian information criterion used at each potential split to decide whether the split was to be maintained or abandoned as described in (26). Permutation tests (running 10,000 permutations) were used to test for association with methylation class by generating a distribution of the test statistic for the null distribution for comparison to the observed distribution. For continuous variables, the permutation test was run with the Kruskal-Wallis test statistic. For categorical variables, we used a Mutual information test statistic; (27) equivalent to a likelihood ratio test comparing the saturated model for a contingency table against a model that assumes independent margins. Significant associations from permutation tests were controlled for potential confounders where appropriate using logistic regression with methylation classes and potential confounders and a likelihood ratio test of the model with and without methylation classes. For survival analyses, Cox proportional hazards models were used, and likelihood ratio tests were used to examine the significance of inclusion of the methylation classes in the models.

The R Package was also used to build classifiers with the Random Forest (RF) approach. RF is a tree-based classification algorithm similar to Classification and Regression Tree (28) and was performed on CpG average β values using RF R package version 4.5-18 by Liaw and Wiener. RF builds each individual tree by taking a bootstrap sample (sampling with replacement) of the original data, and on average, $\sim 1/3$ of the original data are not sampled (out of bag or OOB). Those sampled are used as the training set to grow the trees, and the OOB data are used as the test set. At each node of the tree, a random sample of m out of the total M variables is

Table 1. Subject gender, age, histology, and exposure for mesothelioma patients and nontumor pleural samples

| | Mesothelioma patients | Pleura donors |
|---------------------------------|-----------------------|---------------|
| Gender, <i>n</i> (%) | | |
| Female | 38 (24) | 4 (22) |
| Male | 120 (76) | 14 (78) |
| Age | | |
| Range | 30–80 | 38–77 |
| Mean (SD) | 62 (9.8) | 58 (11.3) |
| Histology, <i>n</i> (%) | | |
| Epithelioid | 116 (73) | — |
| Mixed | 37 (23) | — |
| Sarcomatoid | 5 (3) | — |
| Asbestos exposure, <i>n</i> (%) | | |
| Yes | 112 (74) | 13 (72) |
| No | 39 (26) | 5 (28) |
| Log asbestos body | | |
| Available <i>n</i> (%) | 108 (68) | — |
| Range | 0–5.5 | — |
| Mean (SD) | 2.16 (1.18) | — |

chosen, and the best split is found among the m variables. The default value for m in the RF R package is \sqrt{M} . In this analysis, we tested a range of m from half of \sqrt{M} to twice the \sqrt{M} and used the m that gave the lowest prediction error. The OOB error rate is the percentage of time the RF prediction is incorrect.

Results

To characterize the epigenetic profile of mesothelioma and nontumorigenic parietal pleura, we used the Illumina GoldenGate bead array that simultaneously interrogates 1,505 CpG sites associated with 803 cancer-related genes to generate a methylation value based on ~ 30 replicate measurements for each locus in each sample. In this study, GoldenGate arrays were used to assess methylation in 158 incident cases of mesothelioma and 18 nontumorigenic parietal pleura specimens. Exposure, demographic, and tumor characteristic data for these samples are presented in Table 1.

Array methylation data were first explored with unsupervised hierarchical clustering using Manhattan distance and average linkage for the 750 most variable autosomal CpG loci (Fig. 1). Striking differences between the epigenetic profiles of mesothelioma and nontumor pleura are observed, with almost perfect clustering of epigenetic profiles based on disease status. Next, in a univariate approach, we tested all CpG loci individually for an association between methylation and disease status, and 969 CpG loci had methylation levels that differed ($Q < 0.05$) comparing tumor and nontumor pleura after FDR correction. Of these, 727 loci associated with 493 genes had enhanced methylation in nontumor pleura, and 242 loci associated with 153 genes had more methylation in the tumors (Supplementary Table S1). Because so many loci were differentially methylated between tumor and nontumor pleura, we next applied a modified model-based form of unsupervised clustering known as mixture modeling. This approach built classes of samples based on profiles of methylation with data from all autosomal loci using a mixture of β distributions

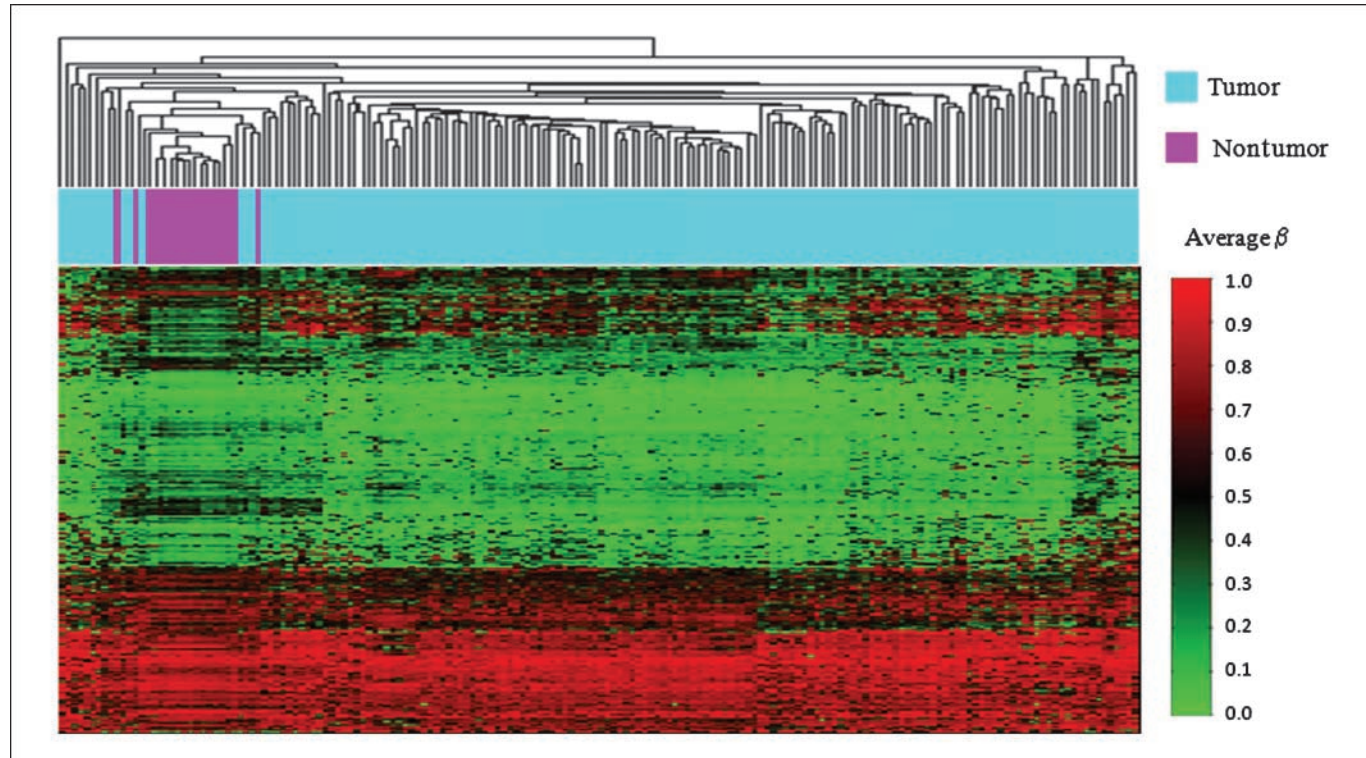


Figure 1. Unsupervised clustering of average β values in tumor and nontumor pleura. Using the R software package, normal tissue sample average β values were subjected to unsupervised hierarchical clustering based on Manhattan distance and average linkage. Each column represents a sample and each row represents a CpG locus (750 most variable autosomal loci). Above the heatmap, a tumor sample (blue) and a nontumor pleural sample (purple) are indicated. In the heatmap, average β of zero, or unmethylated (green), and average β of one, or methylated (red) are shown.

to recursively split the tumors into parsimoniously differentiated classes (29–31). All posterior class membership probabilities were numerically indistinct from 0 or 1. Applying a β mixture model to methylation data from all autosomal loci in tumors and nontumor pleura returned 11 methylation classes, their average methylation profiles, and their sample type distributions (Fig. 2). Methylation class membership was a highly significant predictor of diseased versus nondiseased tissue (permutation $P < 0.0001$). Among the 11 classes in the model, 9 classes perfectly captured only tumor or only normal, and there were 2 methylation classes containing both tumor and normal samples. To follow up, a supervised random forest classification of nontumor and tumor samples was performed. Only 1 tumor (<1%) was misclassified as a nontumor sample, and 5 nontumor samples (28%) were misclassified as tumors. The overall misclassification error rate was 3.4%, significantly lower than the expected error rate under the null hypothesis ($P < 0.0001$).

We next restricted our analyses to tumors, ($n = 158$) first applying our β mixture model approach, and Fig. 3 shows the seven methylation classes that resulted. This figure also displays the distributions of gender, histology, and asbestos body counts by methylation class. Methylation class membership was not a significant predictor of patient gender or tumor histology (data

not shown). Methylation profile class membership was not associated with the amount of tumor in the sample. However, methylation class membership significantly predicted lung asbestos body count (permutation $P < 0.04$). Because men with pleural mesothelioma have higher asbestos body counts compared with women ($P < 0.0001$; ref. 32), we controlled for gender, and methylation class membership remained a significant predictor of asbestos burden (likelihood ratio test $P < 0.03$). Based on prior published work, specific CpG loci were tested for associations between methylation and asbestos body counts; consistent with our prior data (20), tumor methylation average β values at *CDKN2A* ($P < 0.02$), *CDKN2B* ($P < 0.02$), and *RASSF1* ($P < 0.03$) were significantly and positively associated with asbestos body counts. In addition, methylation of *MT1A* [previously reported as asbestos exposure-associated by Tsou and colleagues (18)] was significantly positively associated with asbestos burden; promoter-associated CpG49 ($P < 0.04$), and exonic CpG13 ($P < 0.02$). When testing all autosomal loci for an association between methylation and asbestos burden using the *MT1A* promoter CpG 49 Q value ($Q = 0.32$) as a cutoff, there were 110 loci with an association between methylation status and asbestos burden (Supplementary Table S2). The vast majority of these 110 loci (94%) had a positive correlation between CpG methylation and asbestos body counts,

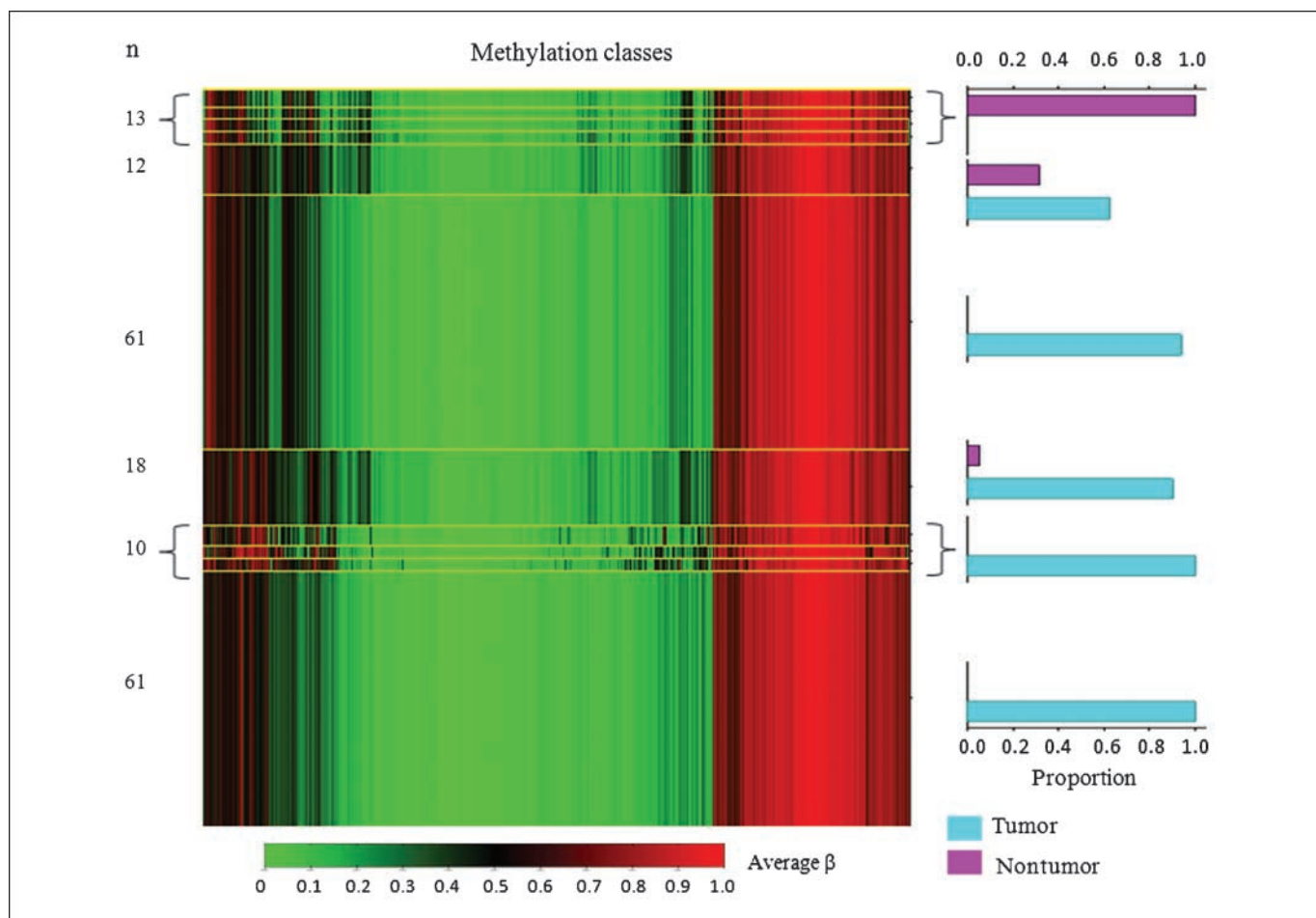


Figure 2. β mixture model of methylation profiles in mesothelioma and nontumor pleura. *Green*, methylation average β for unmethylated; *red*, methylation average β for methylated. Methylation profile classes are stacked in rows separated by yellow lines, and class height corresponds to the number of samples in each class. Class methylation at each locus is a mean of methylation for all samples within a class. *Bar charts*, the proportion of tumors and nontumor pleura samples in each class. Methylation profile classes differentiate tumor from nontumor pleura ($P < 0.0001$).

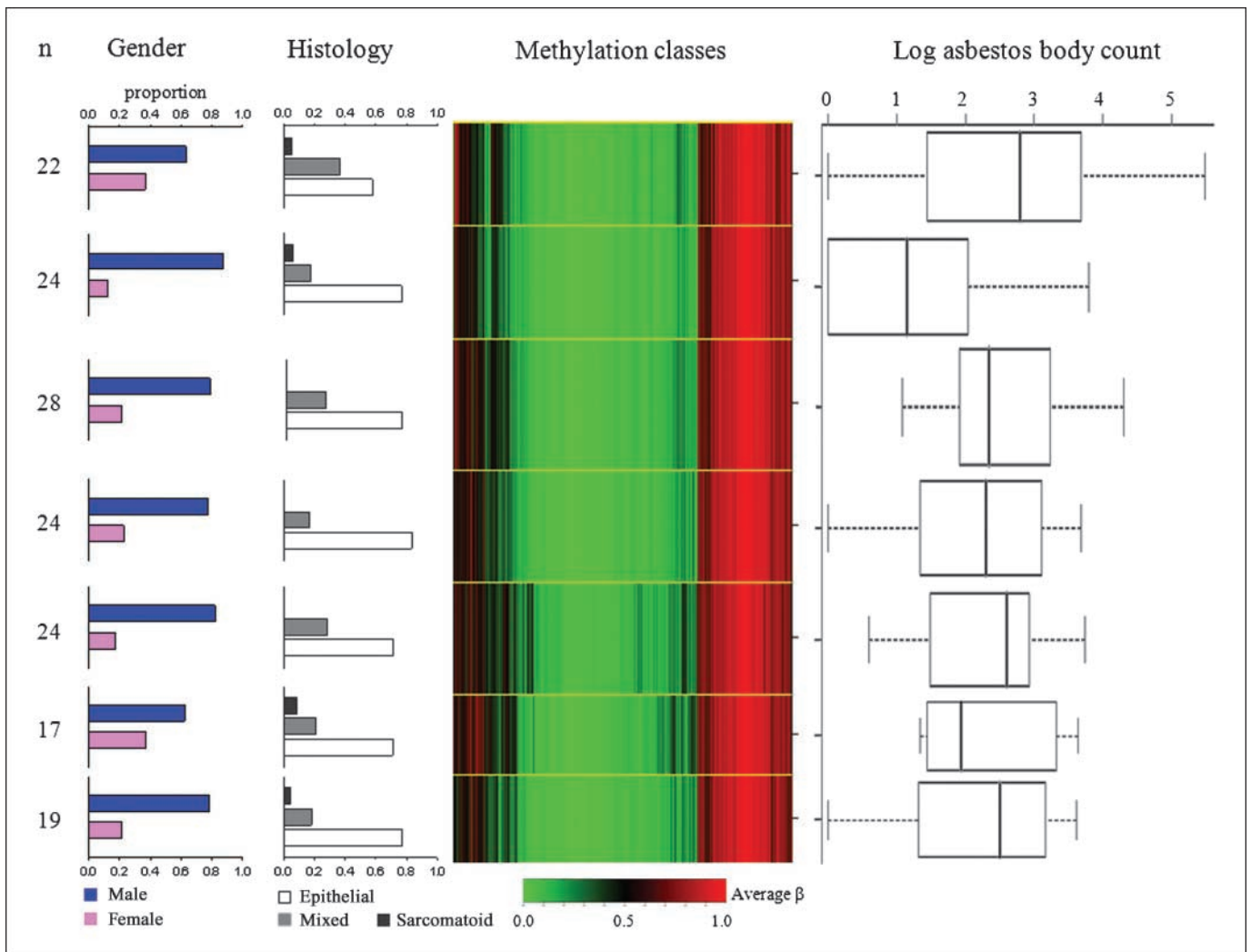


Figure 3. β Mixture model of methylation profiles in pleural mesothelioma. *Green*, methylation average β for unmethylated; *red*, methylation average β for methylated. Methylation profile classes are stacked in rows separated by yellow lines, and class height corresponds to the number of samples in each class. Class methylation at each locus is a mean of methylation for all samples within a class. *Left*, bar charts show proportions for gender and tumor histology among samples within each class. *Right*, box plots of log asbestos body counts for each class. Controlling for gender, methylation class membership predicts asbestos burden ($P < 0.03$).

indicating gene silencing was the dominant phenotype associated with asbestos associated epigenetic change.

Lastly, we examined the relationships between methylation profiles and patient outcome using Cox proportional hazards models of survival controlling for age, gender, and tumor histology. Median survival time of this population was 12.5 months with 67 months of follow-up time. In a proportional hazards model including all cases ($n = 158$), women had half the risk of death of men [hazard ratio (HR), 0.5; 95% confidence interval (CI), 0.3–0.96], and patients with mixed histology tumors were at greater risk of death compared with those with epithelial tumors (HR, 2.7; 95% CI, 1.7–4.4). Importantly, methylation class membership was also a significant predictor of patient outcome ($P < 0.01$). In particular, membership in methylation classes 4 and 7 were both independently associated with a significant 3-fold increased risk of death compared with the class with the lowest median asbestos count [95% CI (class 4), 1.4–7.0; 95% CI (class 7), 1.3–7.4; Table 2]. Where data were available ($n = 108$), and after adjustment for methylation class membership, asbestos burden was associated with a

significant 1.4-fold increased risk of death (95% CI, 1.1–1.8; Table 2). In this model, membership in methylation class 4 remained associated with a significant, nearly 3-fold increased risk of death (HR, 2.8; 95% CI, 1.1–7.1). Again, in this model including asbestos exposure, likelihood ratio tests indicate that methylation classes were significant predictors of patient outcome ($P < 0.005$).

Discussion

Exposure to asbestos is the single most important risk factor for pleural mesothelioma, and prior research has established that somatic mutations (33) and alterations in gene expression (34) are a feature of this disease. Interestingly, relatively few pathologically important mutations arise in this cancer, and there is no characteristic somatic genetic change that can be attributed to the action of asbestos (33). Furthermore, although there is consensus that gene expression (at the mRNA level) is significantly altered in mesothelioma, there is no gene expression signature representative of the action of asbestos in this disease, and there

Table 2. Gender, tumor histology, methylation profile class membership, and asbestos burden are predictors of pleural mesothelioma patient survival

| Covariate | All cases | | | Cases with asbestos burden data | | |
|----------------------------|-----------------------------------|-----------------|----------|-----------------------------------|-----------------|----------|
| | <i>n</i> (%) total <i>n</i> = 158 | HR (95% CI) | <i>P</i> | <i>n</i> (%) total <i>n</i> = 108 | HR (95% CI) | <i>P</i> |
| Age, mean (SD) | 62 (9.8) | 1.02 (1.0–1.05) | 0.09 | 61 (9.5) | 1.03 (1.0–1.1) | 0.18 |
| Gender | | | | | | |
| Male | 120 (76) | 1.0 (Reference) | — | 84 (78) | 1.0 (Reference) | — |
| Female | 38 (24) | 0.5 (0.3–0.96) | <0.04 | 24 (22) | 1.5 (0.6–3.5) | 0.38 |
| Histology | | | | | | |
| Epithelial | 109 (69) | 1.0 (Reference) | — | 74 (68) | 1.0 (Reference) | — |
| Mixed | 44 (28) | 2.7 (1.7–4.4) | <0.0001 | 31 (29) | 2.1 (1.2–3.8) | <0.02 |
| Sarcomatoid | 5 (3) | 2.8 (0.95–8.2) | 0.06 | 3 (3) | 1.2 (0.3–5.2) | 0.83 |
| Asbestos burden, mean (SD) | — | — (—) | — | 2.2 (1.2) | 1.4 (1.1–1.8) | <0.04 |
| Methylation class | | | | | | |
| 2 | 24 (15) | 1.0 (Reference) | — | 17 (16) | 1.0 (Reference) | — |
| 1 | 22 (14) | 1.4 (0.6–3.4) | 0.47 | 10 (9) | 0.5 (0.1–2.2) | 0.37 |
| 3 | 28 (18) | 0.9 (0.4–2.0) | 0.75 | 19 (18) | 0.4 (0.1–1.2) | 0.11 |
| 4 | 24 (15) | 3.1 (1.4–7.0) | <0.01 | 24 (22) | 2.8 (1.1–7.1) | <0.03 |
| 5 | 24 (15) | 1.4 (0.6–3.5) | 0.44 | 17 (16) | 0.9 (0.3–2.8) | 0.89 |
| 6 | 17 (11) | 2.0 (0.8–5.4) | 0.16 | 11 (10) | 1.2 (0.3–4.8) | 0.79 |
| 7 | 19 (12) | 3.1 (1.3–7.4) | <0.01 | 10 (9) | 1.7 (0.6–5.0) | 0.36 |

NOTE: All cases, controlled for all variables in table, model log likelihood $P < 0.01$. Cases with asbestos burden data, controlled for all variables in table, model log likelihood $P < 0.005$. Classes 1 to 7 correspond to top to bottom from Fig. 3.

remains debate about the clinical utility of mRNA expression profiling (35–37). Our work sought to definitively characterize phenotypically important alterations in the epigenome of mesothelioma. We enumerated the epigenetic status of over 800 known cancer-related genes that stably control mRNA expression; comparing normal pleura with mesotheliomas. Our findings indicate that an extremely large number of loci are epigenetically altered in mesothelioma, asbestos exposure is associated with the degree of epigenetic alteration, and that profiles of gene silencing are associated with clinical outcome. This work shows that the epigenome is a primary target of asbestos in the genesis of mesothelioma.

Precisely why epigenetic alterations are a prominent feature of pleural mesothelioma is not clear. However, it is well-known that chronic inflammation is a primary tissue response to asbestos exposure (38). Epigenetic alterations have been associated with inflammation in colon cancer, and it has been suggested that inflammation-related epigenetic alterations are common in human cancers (39). Infiltrating macrophages, neutrophils, and eosinophils have been observed at sites of fiber deposition; these cells are known to generate reactive oxygen, and nitrogen, and to induce the release of peroxidases (38, 40). Treatment of human mesothelial cells with crocidolite asbestos has been shown to increase reactive oxygen species and oxidized DNA residues such as 8-oxoguanine (41). Furthermore, it has been reported that 5-hydroxymethylcytosine can be generated by oxidation of 5-methylcytosine (42). Both 5-methylcytosine adjacent to 8-oxoguanine, and 5-hydroxymethylcytosine, have been shown to inhibit binding of methyl-CpG binding protein 2, a critical epigenetic regulator that recruits cytosine methyltransferases and histone deacetylases (43). It is also known that 5-hydroxymethylcytosine is not recognized as 5-methylcytosine by the maintenance methyltransferase DNMT1

and, hence, may lead to aberrant loss of methylation during cell replication (44). Additional base alterations occur via neutrophil and eosinophil peroxidase-derived HOCl and HOBr, which can react with DNA to form 5-chlorocytosine and 5-bromocytosine, respectively (45). These halogenated cytosines can be mistaken by DNMT1 as 5-methylcytosine during replication, thus providing a potential mechanism for asbestos-related, inflammation-induced aberrant hypermethylation (44). This is consistent with our data suggesting that asbestos burden (and the resultant inflammatory response) is associated with epigenetic alterations. Furthermore, the decades-long latency of mesothelioma would allow ample time for cellular turnover and selection of cells with altered epigenetic programs that favor survival and deregulated proliferation.

We also found that epigenetic profiles differentiate tumor from nontumor pleura ($P < 0.0001$). Similarly, a genome-wide approach profiling of gene expression at the mRNA level in mesothelioma has shown significant differential expression of over 600 genes between mesothelioma and nontumor pleura, although the precise pattern of gene expression differs in the hands of different investigators (34, 37). Epigenetic gene inactivation is inherently more stable and less prone to variability than measures of mRNA, making our approach for differentiating tumors from nontumor pleura likely more reproducible. Although our method differentiates tumors from nontumor pleura extremely well, there remained two methylation profile classes with a minor mixture of tumor and nontumor samples. It is clear that “contaminating” nontumor tissue, or samples containing a lower percent tumor cannot explain this result, as we were able to reject this based on direct observation of the amount of tumor in each specimen. Consequently, it may be that the relative epigenetic similarity between certain tumors and nontumor pleura is explained by a difference in the prevalence of somatic genetic aberrations (such as gene

deletions) among these tumors. In other words, tumors resembling nontumor pleura epigenetically may harbor significantly greater genetic alterations (relative to more epigenetically divergent tumors) that contribute to their malignant phenotype. An integrative genomics approach using common tumor sets for both epigenetic alteration profiling and genome-wide copy number alteration profiling will begin to address this question.

Hundreds of loci had different methylation levels between tumor and nontumor pleura after correcting for multiple comparisons ($Q < 0.05$). Some of the pathways and processes whose genes are differentially methylated include epigenetic regulation, cell cycle control, and inflammation, among others. For instance, *HIC1* was hypermethylated in tumors, which may prevent its transcriptional repression of the histone deacetylase *SIRT1*, and *DNMT3B* had significantly lower methylation in tumors, suggesting epigenetic dysregulation. Cell cycle control genes *WEE1* and *RASSF1* both had significantly higher methylation in tumors, suggesting uncontrolled proliferative potential, a hallmark of cancer (46). Another hypermethylated gene in tumors was *ASB4*, and a SOCS box-containing protein that inhibits c-Jun NH₂ kinases and is thought to contribute to chronic inflammation (47). Among tumors, *JAK2* had significantly decreased methylation, and concomitant silencing of *SOCS* genes may then result in increased, or constitutive STAT activation, which is a recognized inflammation-associated mechanism of tumor initiation and promotion (48). In addition, *HOXA9* was hypermethylated in tumors; *HOXA9* is an inhibitor of nuclear factor- κ B-dependent activation of leukocyte adhesion molecules intercellular adhesion molecule 1, VCAM-1, and E-selectin which recruit leukocytes, exacerbating the inflammatory response (49).

At the same time, it is tenable to posit that, although many genes harbor altered epigenetic states, only a small number of genes with epigenetic alterations are chiefly responsible for the phenotypic differences between tumor and normal tissue, and between more deadly mesothelioma compared with less deadly forms. However, if relatively few “epigenetic gatekeepers” determine phenotype, there must be a very large number of genes that are inactivated consequent to the clonal evolution of the tumor; this may be analogous to “hitchhiker” mutations that arise during tumor clonal expansion. Alternatively, a “hit” at a locus essential for epigenetic regulation, the chronic inflammatory process, or altered tumor metabolism, for example, may induce a change in overall epigenetic regulation that results in the targeting and aberrant methylation of many genes in a wholesale fashion. It is not possible to directly distinguish the alternative explanations of our data.

Among the 158 mesotheliomas studied, seven distinct methylation profile classes were identified. Previous genome-wide mRNA expression analysis of mesotheliomas was able to define two distinct subclasses of tumors that loosely correlated with histology (34). Although our epigenetic profiling data revealed more tumor subclasses, the genome-wide approach using mRNA expression as an outcome encompassed over 10 times as many loci, and used fewer tumors, both of which potentially decreased the power to resolve tumor subclasses compared with our larger sample set restricted to cancer-associated loci. We found that methylation

profiles were associated with asbestos body burden in an analysis controlling for gender ($P < 0.03$), again suggesting asbestos-associated inflammation as a mechanism driving tumors into distinct epigenetic subclasses. There were also distinct methylation profile classes composed of tumors with similar asbestos burden distributions, arguing that asbestos burden per se does not account for all of the observed epigenetic alterations in mesothelioma. However, in the context of a decades-long latency and a nonspecific, asbestos-related chronic inflammation state, alterations of different epigenetic gatekeeper genes could occur at different stages in tumorigenesis and tumor progression. Furthermore, specific genes, or gene pathways, may be more likely to be selected for, or more susceptible to asbestos-related epigenetic alterations; consistent with this, associations between CpG methylation of specific genes and asbestos previously reported by our group and others were confirmed in this study (18, 20). Here, we extended these findings to include several apoptosis-related genes such as *AATK*, *CASP2*, *CASP10*, and *PYCARD*.

Finally, overall methylation profile class membership was a significant, independent predictor of patient survival ($P < 0.01$), suggesting that epigenetic dysregulation is strongly associated with disease progression. Although general epigenetic deregulation may be highly correlated with disease prognosis, we cannot exclude epigenetic events at specific loci as responsible for predicting survival. Continued investigation into the CpG loci whose differential methylation status is associated both with membership in specific classes and prognosis is needed to elucidate the biological mechanism(s), which critically influence patient survival. Additionally, consistent with our previous finding, we confirmed that an increased asbestos burden is a significant, independent predictor of reduced survival in pleural mesothelioma (32).

In summary, our data show that epigenetic alterations are extraordinarily common in mesothelioma and discriminate the malignant phenotype from normal pleura. Epigenetic changes are also significantly associated with asbestos burden and significantly predict clinical outcome. Hence, our data suggest that phenotypically important somatic epigenetic modification is a major mode of action of asbestos in mesothelioma. Further investigation into the underlying mechanism responsible may assist in diagnosis, assessment of prognosis, and design of therapies for this rare but rapidly fatal disease.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

Received 7/10/2008; revised 9/12/2008; accepted 10/2/2008.

Grant Support: National Cancer Institute (R01 CA126939), National Institutes of Environmental Health Sciences (T32ES007155, P42ES05947), International Mesothelioma Program at Brigham and Women's Hospital (Research grant), and Mesothelioma Applied Research Foundation (Research grant).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

References

1. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990;61:759–67.
2. Loeb LA. Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res* 1991;51:3075–9.
3. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat Rev* 2002;3:415–28.
4. Bibikova M, Lin Z, Zhou L, et al. High-throughput DNA methylation profiling using universal bead arrays. *Genome Res* 2006;16:383–93.
5. Robinson BW, Lake RA. Advances in malignant mesothelioma. *N Engl J Med* 2005;353:1591–603.
6. Morinaga K, Kishimoto T, Sakatani M, Akira M,

- Yokoyama K, Sera Y. Asbestos-related lung cancer and mesothelioma in Japan. *Ind Health* 2001;39:65-74.
7. Price B, Ware A. Mesothelioma trends in the United States: an update based on Surveillance, Epidemiology, and End Results Program data for 1973 through 2003. *Am J Epidemiol* 2004;159:107-12.
 8. Roushdy-Hammady I, Siegel J, Emri S, Testa JR, Carbone M. Genetic-susceptibility factor and malignant mesothelioma in the Cappadocian region of Turkey. *Lancet* 2001;357:444-5.
 9. Bhagavatula R, Moody R, Russ J. Asbestos: a moving target. *Best's Rev* 2001;102:85-90.
 10. Peto J, Decarli A, La Vecchia C, Levi F, Negri E. The European mesothelioma epidemic. *Br J Cancer* 1999;79:666-72.
 11. Joshi TK, Gupta RK. Asbestos in developing countries: magnitude of risk and its practical implications. *Int J Occup Med Environ Health* 2004;17:179-85.
 12. Landrigan PJ, Liroy PJ, Thurston G, et al. Health and environmental consequences of the world trade center disaster. *Environ Health Perspect* 2004;112:731-9.
 13. Wagner JC, Sleggs CA, Marchand P. Diffuse pleural mesothelioma and asbestos exposure in the North Western Cape Province. *Br J Ind Med* 1960;17:260-71.
 14. Jaurand MC. Mechanisms of fiber-induced genotoxicity. *Environ Health Perspect* 1997;105 Suppl 5:1073-84.
 15. Kelsey KT, Yano E, Liber HL, Little JB. The *in vitro* genetic effects of fibrous erionite and crocidolite asbestos. *Br J Cancer* 1986;54:107-14.
 16. He B, Lee AY, Dadfarmay S, et al. Secreted frizzled-related protein 4 is silenced by hypermethylation and induces apoptosis in β -catenin-deficient human mesothelioma cells. *Cancer Res* 2005;65:743-8.
 17. Hirao T, Bueno R, Chen CJ, Gordon GJ, Heilig E, Kelsey KT. Alterations of the p16(INK4) locus in human malignant mesothelial tumors. *Carcinogenesis* 2002;23:1127-30.
 18. Tsou JA, Galler JS, Wali A, et al. DNA methylation profile of 28 potential marker loci in malignant mesothelioma. *Lung Cancer* 2007;58:220-30.
 19. Tsou JA, Shen LY, Siegmund KD, et al. Distinct DNA methylation profiles in malignant mesothelioma, lung adenocarcinoma, and non-tumor lung. *Lung Cancer* 2005;47:193-204.
 20. Christensen BC, Godleski JJ, Marsit CJ, et al. Asbestos exposure predicts cell cycle control gene promoter methylation in pleural mesothelioma. *Carcinogenesis* 2008;29:1555-9.
 21. De Vuyst P, Karjalainen A, Dumortier P, et al. 1998 ERS Task Force Report Guidelines for mineral fibre analyses in biological samples: report of the ERS working group. *Eur Respir J* 1998;11:1416-26.
 22. Churg A, Warnock ML. Correlation of quantitative asbestos body counts and occupation in urban patients. *Arch Pathol Lab Med* 1977;101:629-34.
 23. R Development CT. R: A Language and Environment for Statistical Computing. Vienna (Austria): R Foundation for Statistical Computing; 2007.
 24. Hsiung DT, Marsit CJ, Houseman EA, et al. Global DNA methylation level in whole blood as a biomarker in head and neck squamous cell carcinoma. *Cancer Epidemiol Biomarkers Prev* 2007;16:108-14.
 25. Storey J, Taylor J, Siegmund D. Strong control, conservative point estimation, and simultaneous conservative consistency of false discovery rates: A unified approach. *J Royal Stat Soc* 2004;Series B:187-205.
 26. Houseman EA, Christensen BC, Marsit CJ, et al. Model-based clustering of DNA methylation array data: a recursive-partitioning algorithm for high-dimensional data arising as a mixture of β distributions. *BMC Bioinformatics* 2008;9.
 27. Yao Y. Information-theoretic measures for knowledge discovery and data mining In: Karmeshu, editor. *Entropy Measures, Maximum Entropy Principle and Emerging Applications*: Springer; 2003. p. 115-36.
 28. Breiman L. Random Forests. *Machine Learning* 2001;45:5-32.
 29. Shen L, Toyota M, Kondo Y, et al. Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. *Proc Natl Acad Sci U S A* 2007;104:18654-9.
 30. Siegmund KD, Connor CM, Campan M, et al. DNA methylation in the human cerebral cortex is dynamically regulated throughout the life span and involves differentiated neurons. *PLoS ONE* 2007;2:e895.
 31. Siegmund KD, Laird PW, Laird-Offringa IA. A comparison of cluster analysis methods using DNA methylation data. *Bioinformatics (Oxford England)* 2004;20:1896-904.
 32. Christensen BC, Godleski JJ, Roelofs CR, et al. Asbestos burden predicts survival in pleural mesothelioma. *Environ Health Perspect* 2008;116:723-6.
 33. Sugarbaker DJ, Richards WG, Gordon GJ, et al. Transcriptome sequencing of malignant pleural mesothelioma tumors. *Proc Natl Acad Sci U S A* 2008;105:3521-6.
 34. Gordon GJ, Rockwell GN, Jensen RV, et al. Identification of novel candidate oncogenes and tumor suppressors in malignant pleural mesothelioma using large-scale transcriptional profiling. *Am J Pathol* 2005;166:1827-40.
 35. Gordon GJ, Jensen RV, Hsiao LL, et al. Using gene expression ratios to predict outcome among patients with mesothelioma. *J Natl Cancer Inst* 2003;95:598-605.
 36. Gordon GJ, Rockwell GN, Godfrey PA, et al. Validation of genomics-based prognostic tests in malignant pleural mesothelioma. *Clin Cancer Res* 2005;11:4406-14.
 37. Lopez-Rios F, Chuai S, Flores R, et al. Global gene expression profiling of pleural mesotheliomas: overexpression of aurora kinases and P16/CDKN2A deletion as prognostic factors and critical evaluation of microarray-based prognostic prediction. *Cancer Res* 2006;66:2970-9.
 38. Moalli PA, MacDonald JL, Goodglick LA, Kane AB. Acute injury and regeneration of the mesothelium in response to asbestos fibers. *Am J Pathol* 1987;128:426-45.
 39. Issa JP, Ahuja N, Toyota M, Bronner MP, Brentnall TA. Accelerated age-related CpG island methylation in ulcerative colitis. *Cancer Res* 2001;61:3573-7.
 40. Quinlan TR, Marsh JP, Janssen YM, Borm PA, Mossman BT. Oxygen radicals and asbestos-mediated disease. *Environ Health Perspect* 1994;102 Suppl 10:107-10.
 41. Chen Q, Marsh J, Ames B, Mossman B. Detection of 8-oxo-2'-deoxyguanosine, a marker of oxidative DNA damage, in culture medium from human mesothelial cells exposed to crocidolite asbestos. *Carcinogenesis* 1996;17:2525-7.
 42. Masuda T, Shinoara H, Kondo M. Reactions of hydroxyl radicals with nucleic acid bases and the related compounds in γ -irradiated aqueous solution. *J Radiat Res* 1975;16:153-61.
 43. Valinluck V, Tsai HH, Rogstad DK, Burdzy A, Bird A, Sowers LC. Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res* 2004;32:4100-8.
 44. Valinluck V, Sowers LC. Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res* 2007;67:946-50.
 45. Henderson JP, Byun J, Takeshita J, Heinecke JW. Phagocytes produce 5-chlorouracil and 5-bromouracil, two mutagenic products of myeloperoxidase, in human inflammatory tissue. *J Biol Chem* 2003;278:23522-8.
 46. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100:57-70.
 47. Li JY, Chai BX, Zhang W, Liu YQ, Ammori JB, Mulholland MW. Ankyrin repeat and SOCS box containing protein 4 (Asb-4) interacts with GPS1 (CSN1) and inhibits c-Jun NH2-terminal kinase activity. *Cell Signal* 2007;19:1185-92.
 48. Yoshimura A, Naka T, Kubo M. SOCS proteins, cytokine signalling and immune regulation. *Nat Rev Immunol* 2007;7:454-65.
 49. Trivedi CM, Patel RC, Patel CV. Homeobox gene HOXA9 inhibits nuclear factor- κ B dependent activation of endothelium. *Atherosclerosis* 2007;195:e50-60.