

Research Article

Genetic and Epigenetic Tumor Suppressor Gene Silencing Are Distinct Molecular Phenotypes Driven by Growth Promoting Mutations in Nonsmall Cell Lung Cancer

Carmen J. Marsit,¹ E. Andres Houseman,^{2,3} Heather H. Nelson,⁴ and Karl T. Kelsey^{1,2}

¹ Department of Pathology and Laboratory Medicine, Center for Environmental Health and Technology, Brown University, Providence, RI 02912, USA

² Department of Community Health, Center for Environmental Health and Technology, Brown University, Providence, RI 02912, USA

³ Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115, USA

⁴ Division of Epidemiology and Community Health, Masonic Cancer Center, University of Minnesota, Minneapolis, MN 55455, USA

Correspondence should be addressed to Karl T. Kelsey, karl_kelsey@brown.edu

Received 10 January 2008; Accepted 16 December 2008

Recommended by Cornelia Ulrich

Both genetic and epigenetic alterations characterize human nonsmall cell lung cancer (NSCLC), but the biological processes that create or select these alterations remain incompletely investigated. Our hypothesis posits that a roughly reciprocal relationship between the propensity for promoter hypermethylation and a propensity for genetic deletion leads to distinct molecular phenotypes of lung cancer. To test this hypothesis, we examined promoter hypermethylation of 17 tumor suppressor genes, as a marker of epigenetic alteration propensity, and deletion events at the 3p21 region, as a marker of genetic alteration. To model the complex biology between these somatic alterations, we utilized an item response theory model. We demonstrated that tumors exhibiting LOH at greater than 30% of informative alleles in the 3p21 region have a significantly reduced propensity for hypermethylation. At the same time, tumors with activating *KRAS* mutations showed a significantly increased propensity for hypermethylation of the loci examined, a result similar to what has been observed in colon cancer. These data suggest that NSCLCs have distinct epigenetic or genetic alteration phenotypes acting upon tumor suppressor genes and that mutation of oncogenic growth promoting genes, such as *KRAS*, is associated with the epigenetic phenotype.

Copyright © 2008 Carmen J. Marsit et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Lung cancer remains one of the most incident cancers and the leading cause of cancer death in both men and women in the United States. In 2006, there were over 162 000 deaths attributable to lung cancer in the U.S. [1]. The major cause of lung cancer is tobacco smoking, although environmental tobacco smoke, asbestos, and other environmental and industrial exposures also contribute to lung carcinogenesis [2]. Nonsmall cell lung cancer (NSCLC) is derived from the epithelial cells of the lung and bronchus and is characterized by a wide variety of molecular alterations. Included among these are activating mutations in oncogenes (such as *cMYC*, *KRAS*, *EGFR*, *CCND1*, and *BCL2*) and inactivating lesions in tumor suppressor genes [3, 4]. The inactivation of tumor

suppressor genes occurs chiefly through allele loss (both single allele loss and homozygous gene deletion) or epigenetic silencing associated with promoter hypermethylation of CpG islands in the promoters of many of these genes [5–7].

The relationship between these various genetic and epigenetic alterations has remained relatively unexplored, although it is now known that there is a surprisingly large number of genetic alterations evident in solid tumors [8]. A number of groups have reported that alterations to TP53, measured both as mutation of the gene or through altered immunohistochemical staining, are associated with greater prevalence of LOH at various loci [9–11]. There have also been links made between lung carcinogen exposures and several genetic alterations. For example, LOH of the *FHIT* gene is found more often in smokers and those exposed

to asbestos in an occupational setting [12, 13]. Work from our group and others has also suggested that carcinogen exposures may drive the type of alterations observed in specific genes, such as *CDKN2A* (encoding P16INK4A), which is more often deleted in never smoking individuals [14] but has an increasing prevalence of hypermethylation with increasing duration of tobacco smoking [15, 16]. Likewise, mutation of the *EGFR* gene occurs specifically in adenocarcinomas, and most often in women and never smoking patients [17–20]. This type of evidence reveals that distinct molecular phenotypes exist in NSCLC and that exposure, lifestyle, or a combination of these factors drive these phenotypes. Better definition of the number and character of these phenotypes may be critical for making clinical decisions about treatment course for patients, as has been evidenced in the *EGFR* mutation case [21–24]. Thus, this study was aimed at better defining the molecular phenotypes in NSCLC, closely examining the relationship between genetic and epigenetic alterations in this disease.

2. SUBJECTS AND METHODS

2.1. Study population

Eligible cases consisted of all newly diagnosed patients with resectable lung cancer who received treatment at the Massachusetts General Hospital Thoracic Surgery, Oncology, and Pulmonary Services from November 1992 through December 1996 [25]. The patients involved in these studies provided written informed consent under a protocol approved by the appropriate Institutional Review Boards. Patients with recurrent disease or nonoperable tumors were excluded. A random subset of 260 was analyzed for somatic loss (LOH) at 3p21. A subset of 185 patients of the parent study had fresh lung tumor tissue obtained for use in hypermethylation analysis. Tumor tissue was snap frozen in liquid nitrogen and stored at -80°C until processed. Demographic and epidemiological data, including all of the data on tobacco use, were gathered by interviewer review of a self-administered questionnaire completed by patients and reviewed by a single reviewer during the hospitalization for thoracic surgery.

2.2. 3p21 LOH analysis

Analysis of allelic loss at the 3p21 region at microsatellite markers *D3S1029*, *D3S3582*, *D3S3667*, *D3S3640*, *D3S1568*, and *D3S3026*, and *D3S1478* and the details of the statistical analysis and construction of the fraction allele loss (FAL) score have been previously described for this population [9, 26].

2.3. Hypermethylation analysis

DNA was extracted from fresh frozen tumor tissue using the Gentra Puregene DNA extraction kit (Gentra, Minneapolis, Minn, USA) following manufacturer's protocol. Sodium bisulfite modification of the resultant DNA was performed as previously described [27]. Briefly, DNA was denatured in

NaOH and treated with sodium bisulfite (Sigma Chemical Co., St. Louis, Miss, USA) for 16 hours. The DNA was then purified using a Wizard DNA Clean-Up Kit (Promega, Madison, Wiss, USA), treated again with NaOH, and ethanol precipitated. DNA was rehydrated in water for subsequent use in PCR amplification.

We have specifically chosen to utilize traditional methylation specific PCR (MSP) [27] for the analysis of promoter hypermethylation in these studies. We have previously examined potential biases in the sensitivity of using this assay against the relative-quantitative Taqman-based methods [28], and have seen no evidence for potential bias based on tumor quantity or tumor stage in the samples analyzed.

Sodium bisulfite modified DNA was used as the template for methylation specific PCR (MSP) as previously described in [27] using primers specific for the methylated promoters of *CDKN2A* [27], *RASSF1A* [29], *APC* [30], *PYCARD* [31], *LAMC2* [32], *SFRP1*, *SFRP2*, *SFRP4*, *SFRP5* [33], *MGMT*, *DAPK*, *RARB*, *CDH1* [34], *CDH13* [35], *MLH1* [36], *CCND2* [37], and *PRSS3* [38]. All methylation specific PCRs are optimized to detect greater than 5% methylated substrate in each sample. To control for the presence of modified DNA, primers specific to a modified region of the *ACTB* gene containing no CpG sites were used [39]. Modified circulating blood lymphocyte DNA (obtained from a control subject) and the same lymphocyte DNA completely methylated using *SssI* DNA methylase and modified by treatment with sodium bisulfite were used as the negative and positive controls, respectively, in each run.

These genes were chosen as promoter hypermethylation detected using this method has been previously shown to be correlated to transcriptional silencing of these genes, and their hypermethylation occurs in a tumor-specific pattern. We also wished to examine the silencing of tumor suppressor genes involved in a variety of cellular processes and pathways thereby not limiting the analysis to genes involved in a single pathway targeted for inactivation. The genes selected are known to be involved in processes including cell cycle control (*CDKN2A*, *RASSF1A*, *APC*, and *CCND2*), apoptosis (*DAPK*, *PYCARD*), extracellular interactions (*LAMC2*, *PRSS3*), transcriptional regulation (*RARB*), WNT signaling (*SFRP* family), cell-cell signaling (*CDH1*, *CDH13*), and DNA repair (*MGMT*, *MLH1*).

2.4. Statistical analysis

We have previously demonstrated that there are no discreet groupings of tumors by the number of genes undergoing hypermethylation [5] and we have observed that there is a great deal of correlation between methylation of the individual loci, such that simply counting the number of methylated loci in an individual tumor is statistically inappropriate. Therefore, we employed an item response theory (IRT) model [40] which has been shown to be the most appropriate latent trait technique [41] for examining this type of discrete data [42] and which allows the modeling of the propensity for methylation to be treated as a continuous variable in a regression framework.

To construct our IRT model, we used a stepwise selection combined with domain knowledge examining the effect of exposures on the latent methylation variable. Age and gender are included in the model as there have been reports of age-related methylation [43, 44], and there is a well-established difference in the prevalence of bladder cancer by gender. Histology (adenocarcinoma, squamous cell carcinoma, and others) was initially included in the model. We examined the effects of exposures which have been demonstrated to be associated with lung cancer incidence including cigarette smoking and occupational asbestos exposure. Cigarette smoking was examined as a variable comparing never, former, and current smokers as well as using measures of duration and intensity. We also examined the effect of mutation of the *KRAS* gene which has been previously reported in this population of tumors [25, 45], as mutation of this pathway is associated with a methylator phenotype in colon cancer [46]. As a marker of genetic alteration, we included the FAL score of 3p21 LOH. Variables were excluded from the model by using the AIC to determine the most parsimonious model. The final model included gender, *KRAS* mutation, and 3p21 FAL (low ≤ 0.3 versus high > 0.3) as covariates in addition to the promoter hypermethylation status of the 17 genes was examined.

We also examined the associations between propensity for hypermethylation, again characterized by the methylation latent trait, and patient survival. In a second stage analysis, we used the predictions of the methylation latent trait and employed a Cox proportional hazards model to examine the logarithm of the hazard of death as a linear function of covariates, including the methylation latent trait. All analyses were conducted in R version 2.4.1 [47], including custom software for the IRT model, available upon request.

3. RESULTS

The demographics of the study population are shown in Table 1. As expected, the majority of patients were either current or former smokers, with a mean age of approximately 67 years. Greater than 50% of the tumors were adenocarcinomas, with about 34% squamous cell carcinoma, and the remaining of the rarer histologies such as large cell. As previously noted, approximately 17% of the cases had mutation of *KRAS* at codon 12 [45], the hotspot for lung cancer, and about 34% had loss of heterozygosity (LOH) at greater than 30% of informative alleles in the 3p21 region [9].

Figure 1(a) presents the prevalence of hypermethylation of the 17 gene promoters examined in this study. The prevalence of this alteration appears highly variable by the gene examined, with some genes, such as *SFRP1* and *SFRP2* found to be hypermethylated in nearly 80% of cases, while other genes, such as *MLH1* or *APC*, rarely exhibiting hypermethylation in this series. There was moderate correlation between methylation at the individual loci, as depicted in Figure 1(b), where lighter shades represent correlation coefficients approaching 1. As we observed, this relatively high correlation between methylation events, and the broad range of prevalence of these alterations at different genes,

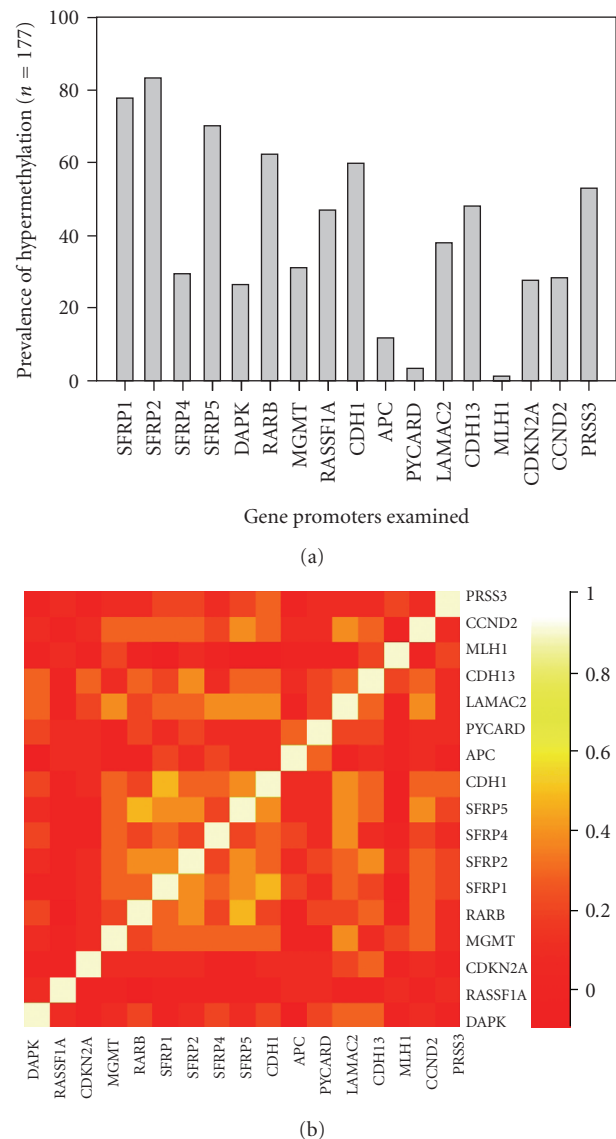


FIGURE 1: Prevalence and correlation between promoter hypermethylation at 17 genes in NSCLC. (a) Bar graph representing the prevalence of promoter hypermethylation at various gene promoters examined in 177 NSCLC tumors. (b) Heat map depicting the correlation coefficient (as depicted in the gradient bar) between individual promoter hypermethylation of 17 genes examined in NSCLC.

rather than treat these events as independent in the analysis, we employed an item response theory model to investigate the propensity for hypermethylation in these tumors. This approach takes into account this correlation and has been shown to be appropriate for similar data examined in bladder cancer [42].

The results of the item response model are listed in Table 2, which provides the coefficient values of the slopes and intercepts for each of the genes examined in this series. Although the values of the intercepts are not directly interpretable, they are provided for completeness as they are used in plotting the relationship between the latent trait and

TABLE 1: Demographics of the NSCLC case series.

Characteristic	
Age, mean (SD)	66.9(10.6)
Gender, n (prevalence)	
Female	80(0.45)
Male	97(0.55)
Histology, n (prevalence)	
Adenocarcinoma	95(0.54)
Adeno-squamous	5(0.03)
Bronchoalveolar	2(0.01)
Large cell	14(0.08)
Squamous cell	61(0.34)
Tumor TNM stage, n (prevalence)	
Stage 1	99(0.56)
Stage 2	34(0.19)
Stage 3	39(0.22)
Stage 4	5(0.03)
Smoking status, n (prevalence)	
Never	13(0.07)
Former	93(0.52)
Current	71(0.40)
KRAS codon 12 mutation, n (prevalence)	
No	147(0.83)
Yes	30(0.17)
3p21 LOH fraction allelic loss score, n (prevalence)*	
0–0.3	100(0.66)
>0.3	52(0.34)

* 3p21 LOH data not available for 25 samples.

the probability of methylation. The slope, on the other hand, indicates the strength of the relationship between the latent trait and probability of methylation of the gene. The range of values for the slope terms of these genes suggests that genes contribute differentially to the modeled underlying propensity for methylation. Comparisons with the overall prevalence demonstrate that the model is not solely driven by prevalence, as genes with relative high prevalence, such as *RASSF1A*, which are methylated in approximately 50% of cases, in fact have a nonsignificant and negative item response slope, while *CCND2* (encoding Cyclin-D2), with a prevalence of methylation of approximately 30%, has a highly significant slope of 1.3.

This model also examined the impact of exposures, demographics, tumor characteristics, and other molecular alterations on the underlying propensity for methylation. We found that only *KRAS* mutation status and fraction allelic loss score of LOH at 3p21 were significantly associated with the propensity for hypermethylation. Gender was kept in the model as it is a known confounder of *KRAS* mutation status [25]. Having mutation of *KRAS* leads to an approximately 0.5 unit *increase* in latent trait mean (where latent trait is scaled to have unit standard deviation) ($P < .05$). On the other hand, having a 3p21 LOH fraction allele loss score of >0.3, indicating that greater than 30% of informative loci

examined in the region demonstrated LOH, was associated with a statistically significant *reduction* by 0.5 unit in the methylation latent trait mean ($P < .02$). We found no significant association between overall patient survival and the methylation latent trait in a second stage analysis.

4. DISCUSSION

The existence of a CpG island methylator phenotype (CIMP) has been demonstrated in a number of tumor types, most conclusively in colorectal, and gastric cancers [44, 48–50]. In colon cancer, the CIMP phenotype is associated with genetic mutation of the *BRAF* gene, and it has also been suggested that this phenotype, particularly through its associated methylation of the *MLH1* gene, is responsible for mismatch repair deficiency and thus the microsatellite instability observed in a subset of colorectal cancers [46]. Mutation of the *KRAS* gene has also been associated with CIMP in colorectal cancer [48, 50]. Our recent work has also suggested that tumors from a number of other sites, including the lung, may exhibit differences in their underlying propensity for hypermethylation [5]. Assessment of the methylator phenotype in lung cancer is complicated by a lack of understanding as to which genes should be assessed to determine this phenotype, as those used in colon

TABLE 2: Item response model for NSCLC promoter hypermethylation and covariates.

	Est	SE	P
<i>Item response intercepts</i>			
<i>DAPK</i>	−1.10	0.22	<.0001
<i>RASSF1A</i>	−0.20	0.17	.24
<i>CDKN2A</i>	−0.84	0.18	<.0001
<i>MGMT</i>	−0.66	0.27	.02
<i>RARB</i>	0.82	0.28	.004
<i>SFRP1</i>	2.24	0.42	<.0001
<i>SFRP2</i>	2.80	0.51	<.0001
<i>SFRP4</i>	−0.86	0.28	.002
<i>SFRP5</i>	1.67	0.42	.0001
<i>CDH1</i>	0.88	0.36	.01
<i>APC</i>	−1.83	0.25	<.0001
<i>PYCARD</i>	−2.99	0.42	<.0001
<i>LAMAC2</i>	−0.38	0.40	.35
<i>CDH13</i>	0.03	0.24	.90
<i>MLH1</i>	−3.77	0.57	<.0001
<i>CCND2</i>	−0.90	0.32	.006
<i>PRSS3</i>	0.20	0.19	.30
<i>Item response slopes</i>			
<i>DAPK</i>	0.50	0.24	.04
<i>RASSF1</i>	−0.16	0.17	.33
<i>CDKN2A</i>	0.11	0.19	.56
<i>MGMT</i>	1.02	0.26	.0001
<i>RARB</i>	0.98	0.24	<.0001
<i>SFRP1</i>	1.70	0.29	<.0001
<i>SFRP2</i>	1.68	0.35	<.0001
<i>SFRP4</i>	0.99	0.26	.0002
<i>SFRP5</i>	1.67	0.27	<.0001
<i>CDH1</i>	1.43	0.28	<.0001
<i>APC</i>	0.15	0.22	.50
<i>PYCARD</i>	0.61	0.25	.01
<i>LAMAC2</i>	1.82	0.27	<.0001
<i>CDH13</i>	0.84	0.24	.0005
<i>MLH1</i>	0.14	0.40	.72
<i>CCND2</i>	1.32	0.28	<.0001
<i>PRSS3</i>	0.43	0.19	.02
<i>Latent trait mean effects</i>			
<i>Gender</i>			
Male	Referent		
Female	−0.34	0.22	.11
<i>KRAS codon 12 mutation</i>			
No mutation	Referent		
Mutation	0.54	0.28	.05
<i>3p21 LOH FAL score</i>			
0–0.3	Referent		
>0.3	−0.56	0.24	.02

cancers may not be appropriate. Thus, we have sought to more thoroughly examine what factors may be driving this

propensity, and if there are specific molecular phenotypes extant in NSCLC.

As we previously demonstrated, there are no discreet groupings of tumors by the number of genes undergoing hypermethylation [5], a finding that has led some to question the existence of a methylator phenotype [51]. We have also shown (Figure 1(b)) that there is a great deal of correlation between methylation of the individual loci, such that simply counting the number of methylated loci in an individual tumor is statistically inappropriate. Therefore, in order to employ appropriate statistical methodologies as well as avoid biased separation of tumors into classes based on arbitrary counts, we employed an item response theory model which has been shown to be the most appropriate latent trait technique for examining this type of discrete data [42] and which allows the propensity for methylation to be treated as a continuous variable in a regression framework.

Using this approach, we observe differences in the contribution to the underlying methylation latent trait by the different loci examined, an observation that would be lost if the methylation events were simply counted and thus considered equal. This suggests that some genes, such as *SFRP2*, *SFRP5*, and *CCND2*, may be more informative (due to their significant and greater item response slopes) than genes with nonsignificant or small item response slopes, such as *RASSF1A* or *CDKN2A*. This is of interest as *CDKN2A* (encoding P16INK4A) is often considered as a marker of CIMP in colorectal cancer [49], but in NSCLC, silencing of this gene appears less informative for predicting the overall propensity for methylation. *CDKN2A* has also been shown to be one of the earliest genes identified to become hypermethylated during lung carcinogenesis, with detection possible in epithelial cells from smokers prior to lung cancer diagnosis [52]. Thus, alteration of this gene may be more important in the early stages of carcinogenesis and may be common in tumors irrespective of their overall propensity for methylation.

We observed that tumors exhibiting more extensive LOH at the 3p21 region showed a lower propensity for methylation. This is particularly interesting, as it is often thought that LOH and methylation may be occurring together since 2 hits are needed to inactivate tumor suppressors. Further, the tumor suppressor *RASSF1A*, which is located at 3p21, was included in the panel of genes in our model, suggesting that the deletion phenotype and methylation phenotype are perhaps even less correlated than we report here. Our results, from this examination, our previous examination of LOH at 3p21 [9], and our work showing p16 deletion and methylation to be reciprocal [14], suggest that allele loss in the 3p21 region, beyond identifying the site of a specific NSCLC-specific tumor suppressor gene, is a more general marker of a propensity for genetic inactivation of tumor suppressor genes via allelic imbalance or loss. Our data demonstrates that tumors predisposed to this type of genetic alteration are less likely to undergo epigenetic alterations and vice versa. That is, our result suggests that tumor suppressor gene silencing in NSCLCs arises either through a preponderance of allele loss events or epigenetic silencing events, occurring in a roughly dichotomous fashion.

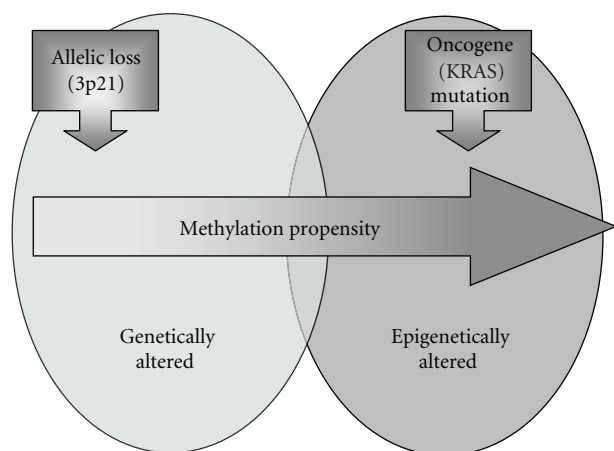


FIGURE 2: Model of the molecular phenotypes of NSCLC driven by promoter hypermethylation propensity. The model depicts the inverse relationship between epigenetic and genetic alterations and its positive association between epigenetic propensity and growth promoting oncogene activation.

Our model also demonstrates that *KRAS* mutation is associated with a greater propensity for methylation, a result similar to that reported in colon cancer [48, 50]. Figure 2 presents a model describing this overall relationship. The mechanism that underlies the connection between *KRAS* mutation and increased methylation propensity remains unclear; indeed, the same can be said for the established connection between *BRAF* mutation and CIMP, although it is known that these two oncogenes operate in the same cell signaling pathway [48, 53]. Our study as well as previous studies cannot absolutely discern which event is occurring first, mutation of these genes or establishment of greater promoter methylation, but one can speculate that the oncogenic activation of *KRAS* may drive cell division, thus increasing the possibility for alteration of promoter methylation profiles. At the same time, silencing of genes involved in DNA damage or repair through promoter hypermethylation may allow for the propagation of mutation of these oncogenes. Approaches employing model systems with oncogenic mutation of these growth promoting genes may help to shed light on the timing of these events and help to answer this question.

A limitation of this study was the use of a panel of selected loci and a limited regional examination of deletion. Thus, our results may be indicative of gene-specific selection events and not of the broader phenotypes that we have proposed. To more conclusively examine, this requires genomic-level approaches be applied for examination of the epigenetic and genetic character within the same tumor, and new technologies are becoming available which can allow for this type of examination on clinical specimens. As these technologies develop, it will be critical to employ appropriate and rigorous statistical methods, such as those used here, to analyze the data to allow for an understanding of the biology driving these alterations. Our results, though, do provide an

impetus to test these hypotheses using these more genome-wide approaches.

More broadly, these results suggest the existence of distinct molecular pathways of tumor suppressor gene inactivation in sporadic tumors. Classifying tumors based on these pathways may lead to improved understanding of the etiology of these diseases, as it may improve the overall classification of disease for genetic association studies. For example, one may posit that individuals with polymorphisms in, perhaps, genes involved in DNA repair (particularly recombinational repair) that lead to reduced repair capacity, may be at higher risk for the genetically altered subclass of tumors. This improved subclassification also holds tremendous clinical utility as it may help to define patients' response to specific chemotherapeutic regimens. In this series, we did not observe any relationship between the methylation latent trait and overall patient survival, but we could not examine specific subgroups based on treatment, which could be critical for understanding this relationship.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health Grants ES05974, ES007373, and CA100679 and Flight Attendants Medical Research Institute Young Clinical Scientist Award (C. J. Marsit).

REFERENCES

- [1] A. Jemal, R. Siegel, E. Ward, et al., "Cancer statistics, 2006," *CA: A Cancer Journal for Clinicians*, vol. 56, no. 2, pp. 106–130, 2006.
- [2] D. Schottenfeld, "Epidemiology of lung cancer," in *Lung Cancer: Principles and Practice*, H. I. Pass, J. B. Mitchell, D. H. Johnson, and A. T. Turrissi, Eds., pp. 305–321, Lippincott-Raven, Philadelphia, Pa, USA, 1996.
- [3] W. D. Travis, J. Linder, and B. Mackay, "Classification, histology, cytology, and electron microscopy," in *Lung Cancer: Principles and Practice*, H. I. Pass, J. B. Mitchell, D. H. Johnson, and A. T. Turrissi, Eds., pp. 361–395, Lippincott-Raven, Philadelphia, Pa, USA, 1996.
- [4] J. D. Minna, J. A. Roth, and A. F. Gazdar, "Focus on lung cancer," *Cancer Cell*, vol. 1, no. 1, pp. 49–52, 2002.
- [5] C. J. Marsit, E. A. Houseman, B. C. Christensen, et al., "Examination of a CpG island methylator phenotype and implications of methylation profiles in solid tumors," *Cancer Research*, vol. 66, no. 21, pp. 10621–10629, 2006.
- [6] S. A. Belinsky, "Silencing of genes by promoter hypermethylation: key event in rodent and human lung cancer," *Carcinogenesis*, vol. 26, no. 9, pp. 1481–1487, 2005.
- [7] J. D. Minna, K. Fong, S. Zöchbauer-Müller, and A. F. Gazdar, "Molecular pathogenesis of lung cancer and potential translational applications," *The Cancer Journal*, vol. 8, supplement 1, pp. S41–S46, 2002.
- [8] T. Sjöblom, S. Jones, L. D. Wood, et al., "The consensus coding sequences of human breast and colorectal cancers," *Science*, vol. 314, no. 5797, pp. 268–274, 2006.
- [9] C. J. Marsit, M. Hasegawa, T. Hirao, et al., "Loss of heterozygosity of chromosome 3p21 is associated with mutant *TP53* and better patient survival in non-small-cell lung cancer," *Cancer Research*, vol. 64, no. 23, pp. 8702–8707, 2004.

- [10] C. J. Marsit, D.-H. Kim, M. Liu, et al., "Hypermethylation of *RASSF1A* and *BLU* tumor suppressor genes in non-small cell lung cancer: implications for tobacco smoking during adolescence," *International Journal of Cancer*, vol. 114, no. 2, pp. 219–223, 2005.
- [11] S. Zienolddiny, D. Ryberg, M. O. Arab, V. Skaug, and A. Haugen, "Loss of heterozygosity is related to *p53* mutations and smoking in lung cancer," *British Journal of Cancer*, vol. 84, no. 2, pp. 226–231, 2001.
- [12] H. H. Nelson, J. K. Wiencke, L. Gunn, J. C. Wain, D. C. Christiani, and K. T. Kelsey, "Chromosome 3p14 alterations in lung cancer: evidence that *FHIT* exon deletion is a target of tobacco carcinogens and asbestos," *Cancer Research*, vol. 58, no. 9, pp. 1804–1807, 1998.
- [13] G. Sozzi, L. Sard, L. De Gregorio, et al., "Association between cigarette smoking and *FHIT* gene alterations in lung cancer," *Cancer Research*, vol. 57, no. 11, pp. 2121–2123, 1997.
- [14] K. S. Kraunz, H. H. Nelson, M. Lemos, J. J. Godleski, J. K. Wiencke, and K. T. Kelsey, "Homozygous deletion of *p16^{INK4a}* and tobacco carcinogen exposure in nonsmall cell lung cancer," *International Journal of Cancer*, vol. 118, no. 6, pp. 1364–1369, 2006.
- [15] S. Toyooka, R. Maruyama, K. O. Toyooka, et al., "Smoke exposure, histologic type and geography-related differences in the methylation profiles of non-small cell lung cancer," *International Journal of Cancer*, vol. 103, no. 2, pp. 153–160, 2003.
- [16] D.-H. Kim, H. H. Nelson, J. K. Wiencke, et al., "*p16^{INK4a}* and histology-specific methylation of CpG islands by exposure to tobacco smoke in non-small cell lung cancer," *Cancer Research*, vol. 61, no. 8, pp. 3419–3424, 2001.
- [17] K. Matsuo, H. Ito, Y. Yatabe, et al., "Risk factors differ for non-small-cell lung cancers with and without *EGFR* mutation: assessment of smoking and sex by a case-control study in Japanese," *Cancer Science*, vol. 98, no. 1, pp. 96–101, 2007.
- [18] I. Y. S. Tam, L. P. Chung, W. S. Suen, et al., "Distinct epidermal growth factor receptor and *KRAS* mutation patterns in non-small cell lung cancer patients with different tobacco exposure and clinicopathologic features," *Clinical Cancer Research*, vol. 12, no. 5, pp. 1647–1653, 2006.
- [19] S. Toyooka, M. Tokumo, H. Shigematsu, et al., "Mutational and epigenetic evidence for independent pathways for lung adenocarcinomas arising in smokers and never smokers," *Cancer Research*, vol. 66, no. 3, pp. 1371–1375, 2006.
- [20] M. Tokumo, S. Toyooka, K. Kiura, et al., "The relationship between epidermal growth factor receptor mutations and clinicopathologic features in non-small cell lung cancers," *Clinical Cancer Research*, vol. 11, no. 3, pp. 1167–1173, 2005.
- [21] L. V. Sequist, V. A. Joshi, P. A. Jänne, et al., "Response to treatment and survival of patients with non-small cell lung cancer undergoing somatic *EGFR* mutation testing," *The Oncologist*, vol. 12, no. 1, pp. 90–98, 2007.
- [22] D. W. Bell, T. J. Lynch, S. M. Haserlat, et al., "Epidermal growth factor receptor mutations and gene amplification in non-small-cell lung cancer: molecular analysis of the IDEAL/INTACT gefitinib trials," *Journal of Clinical Oncology*, vol. 23, no. 31, pp. 8081–8092, 2005.
- [23] Y. Tomizawa, H. Iijima, N. Sunaga, et al., "Clinicopathologic significance of the mutations of the *epidermal growth factor receptor* gene in patients with non-small cell lung cancer," *Clinical Cancer Research*, vol. 11, no. 19, part 1, pp. 6816–6822, 2005.
- [24] T. Mitsudomi, T. Kosaka, H. Endoh, et al., "Mutations of the epidermal growth factor receptor gene predict prolonged survival after gefitinib treatment in patients with non-small-cell lung cancer with postoperative recurrence," *Journal of Clinical Oncology*, vol. 23, no. 11, pp. 2513–2520, 2005.
- [25] H. H. Nelson, D. C. Christiani, E. J. Mark, J. K. Wiencke, J. C. Wain, and K. T. Kelsey, "Implications and prognostic value of *K-ras* mutation for early-stage lung cancer in women," *Journal of the National Cancer Institute*, vol. 91, no. 23, pp. 2032–2038, 1999.
- [26] T. Hirao, H. H. Nelson, T. D. S. Ashok, et al., "Tobacco smoke-induced DNA damage and an early age of smoking initiation induce chromosome loss at 3p21 in lung cancer," *Cancer Research*, vol. 61, no. 2, pp. 612–615, 2001.
- [27] J. G. Herman, J. R. Graff, S. Myöhänen, B. D. Nelkin, and S. B. Baylin, "Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 18, pp. 9821–9826, 1996.
- [28] C. J. Marsit, M. R. Karagas, A. Andrew, et al., "Epigenetic inactivation of *SFRP* genes and TP53 alteration act jointly as markers of invasive bladder cancer," *Cancer Research*, vol. 65, no. 16, pp. 7081–7085, 2005.
- [29] M. van Engeland, G. M. J. M. Roemen, M. Brink, et al., "*K-ras* mutations and *RASSF1A* promoter methylation in colorectal cancer," *Oncogene*, vol. 21, no. 23, pp. 3792–3795, 2002.
- [30] A. K. Virmani, A. Rath, U. G. Sathyanarayana, et al., "Aberrant methylation of the *adenomatous polyposis coli* (*APC*) gene promoter 1A in breast and lung carcinomas," *Clinical Cancer Research*, vol. 7, no. 7, pp. 1998–2004, 2001.
- [31] A. Virmani, A. Rath, K. Sugio, et al., "Aberrant methylation of *TMS1* in small cell, non small cell lung cancer and breast cancer," *International Journal of Cancer*, vol. 106, no. 2, pp. 198–204, 2003.
- [32] U. G. Sathyanarayana, R. Maruyama, A. Padar, et al., "Molecular detection of noninvasive and invasive bladder tumor tissues and exfoliated cells by aberrant promoter methylation of laminin-5 encoding genes," *Cancer Research*, vol. 64, no. 4, pp. 1425–1430, 2004.
- [33] H. Suzuki, D. N. Watkins, K.-W. Jair, et al., "Epigenetic inactivation of *SFRP* genes allows constitutive WNT signaling in colorectal cancer," *Nature Genetics*, vol. 36, no. 4, pp. 417–422, 2004.
- [34] S. Zöchbauer-Müller, K. M. Fong, A. K. Virmani, J. Geradts, A. F. Gazdar, and J. D. Minna, "Aberrant promoter methylation of multiple genes in non-small cell lung cancers," *Cancer Research*, vol. 61, no. 1, pp. 249–255, 2001.
- [35] S. Toyooka, H. I. Pass, N. Shivapurkar, et al., "Aberrant methylation and simian virus 40 tag sequences in malignant mesothelioma," *Cancer Research*, vol. 61, no. 15, pp. 5727–5730, 2001.
- [36] A. O.-O. Chan, R. R. Broaddus, P. S. Houlihan, J.-P. J. Issa, S. R. Hamilton, and A. Rashid, "CpG island methylation in aberrant crypt foci of the colorectum," *The American Journal of Pathology*, vol. 160, no. 5, pp. 1823–1830, 2002.
- [37] E. Evron, C. B. Umbricht, D. Korz, et al., "Loss of cyclin D2 expression in the majority of breast cancers is associated with promoter hypermethylation," *Cancer Research*, vol. 61, no. 6, pp. 2782–2787, 2001.
- [38] K. Yamashita, K. Mimori, H. Inoue, M. Mori, and D. Sidransky, "A tumor-suppressive role for trypsin in human cancer progression," *Cancer Research*, vol. 63, no. 20, pp. 6575–6578, 2003.
- [39] C. A. Eads, K. D. Danenberg, K. Kawakami, et al., "Methyl-Light: a high-throughput assay to measure DNA methylation," *Nucleic Acids Research*, vol. 28, no. 8, article e32, pp. 1–8, 2000.

- [40] R. K. Hambleton and L. L. Cook, "Latent trait models and their use in analysis of educational testing data," *Journal of Educational Measurement*, vol. 14, no. 2, pp. 75–96, 1977.
- [41] B. N. Sánchez, E. Budtz-Jørgensen, L. M. Ryan, and H. Hu, "Structural equation models: a review with applications to environmental epidemiology," *Journal of the American Statistical Association*, vol. 100, no. 472, pp. 1443–1455, 2005.
- [42] E. A. Houseman, C. J. Marsit, M. Karagas, and L. M. Ryan, "Penalized item response theory models: application to epigenetic alterations in bladder cancer," *Biometrics*, vol. 63, no. 4, pp. 1269–1277, 2007.
- [43] T. Takahashi, H. Shigematsu, N. Shivapurkar, et al., "Aberrant promoter methylation of multiple genes during multistep pathogenesis of colorectal cancers," *International Journal of Cancer*, vol. 118, no. 4, pp. 924–931, 2006.
- [44] M. Toyota, N. Ahuja, H. Suzuki, et al., "Aberrant methylation in gastric cancer associated with the CpG island methylator phenotype," *Cancer Research*, vol. 59, no. 21, pp. 5438–5442, 1999.
- [45] H. H. Nelson, D. C. Christiani, J. K. Wiencke, E. J. Mark, J. C. Wain, and K. T. Kelsey, "K-ras mutation and occupational asbestos exposure in lung adenocarcinoma: asbestos-related cancer without asbestosis," *Cancer Research*, vol. 59, no. 18, pp. 4570–4573, 1999.
- [46] D. J. Weisenberger, K. D. Siegmund, M. Campan, et al., "CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with *BRAF* mutation in colorectal cancer," *Nature Genetics*, vol. 38, no. 7, pp. 787–793, 2006.
- [47] R Development Core Team, "R: a language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria, 2005.
- [48] W. S. Samowitz, H. Albertsen, J. Herrick, et al., "Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer," *Gastroenterology*, vol. 129, no. 3, pp. 837–845, 2005.
- [49] M. Toyota, N. Ahuja, M. Ohe-Toyota, J. G. Herman, S. B. Baylin, and J.-P. J. Issa, "CpG island methylator phenotype in colorectal cancer," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 15, pp. 8681–8686, 1999.
- [50] M. Toyota, M. Ohe-Toyota, N. Ahuja, and J.-P. J. Issa, "Distinct genetic profiles in colorectal tumors with or without the CpG island methylator phenotype," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 2, pp. 710–715, 2000.
- [51] K. Yamashita, T. Dai, Y. Dai, F. Yamamoto, and M. Perucho, "Genetics supersedes epigenetics in colon cancer phenotype," *Cancer Cell*, vol. 4, no. 2, pp. 121–131, 2003.
- [52] S. A. Belinsky, W. A. Palmisano, F. D. Gilliland, et al., "Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers," *Cancer Research*, vol. 62, no. 8, pp. 2370–2377, 2002.
- [53] H. Rajagopalan, A. Bardelli, C. Lengauer, K. W. Kinzler, B. Vogelstein, and V. E. Velculescu, "Tumorigenesis: *RAF/RAS* oncogenes and mismatch-repair status," *Nature*, vol. 418, no. 6901, p. 934, 2002.