VOLUME 31 · NUMBER 32 · NOVEMBER 10 2013

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

A Prognostic DNA Methylation Signature for Stage I Non–Small-Cell Lung Cancer

Juan Sandoval, Jesus Mendez-Gonzalez, Ernest Nadal, Guoan Chen, F. Javier Carmona, Sergi Sayols, Sebastian Moran, Holger Heyn, Miguel Vizoso, Antonio Gomez, Montse Sanchez-Cespedes, Yassen Assenov, Fabian Müller, Christoph Bock, Miquel Taron, Josefina Mora, Lucia A. Muscarella, Triantafillos Liloglou, Michael Davies, Marina Pollan, Maria J. Pajares, Wenceslao Torre, Luis M. Montuenga, Elisabeth Brambilla, John K. Field, Luca Roz, Marco Lo Iacono, Giorgio V. Scagliotti, Rafael Rosell, David G. Beer, and Manel Esteller

Author affiliations appear at the end of this article.

Published online ahead of print at www.jco.org on September 30, 2013.

Supported by Grant No. HEALTH-F2-2010-258677-CURELUNG Project from the European Community's Seventh Framework Programme (FP7/2007-2013); Grants No. PI10/02992, PI10/00166, RD06/0020/0066, and RD06/0020/0062 from the Institute of Health Carlos III; the Cellex Foundation; and the Roy Castel Lung Cancer Foundation.

J.S. and J.M.-G. contributed equally to this article.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Manel Esteller, MD, PhD, Cancer Epigenetics and Biology Program, 3rd Floor, Hospital Duran i Reynals, Av. Gran Via 199-203, 08908 L'Hospitalet, Barcelona, Catalonia, Spain; e-mail: mesteller@idibell.cat.

0732-183X/13/3132w-4140w/$20.00

DOI: 10.1200/JCO.2012.48.5516

ABSTRACT

Purpose
Non–small-cell lung cancer (NSCLC) is a tumor in which only small improvements in clinical outcome have been achieved. The issue is critical for stage I patients for whom there are no available biomarkers that indicate which high-risk patients should receive adjuvant chemotherapy. We aimed to find DNA methylation markers that could be helpful in this regard.

Patients and Methods
A DNA methylation microarray that analyzes 450,000 CpG sites was used to study tumoral DNA obtained from 444 patients with NSCLC that included 237 stage I tumors. The prognostic DNA methylation markers were validated by a single-methylation pyrosequencing assay in an independent cohort of 143 patients with stage I NSCLC.

Results
Unsupervised clustering of the 10,000 most variable DNA methylation sites in the discovery cohort identified patients with high-risk stage I NSCLC who had shorter relapse-free survival (RFS; hazard ratio [HR], 2.35; 95% CI, 1.29 to 4.28; $P = .004$). The study in the validation cohort of the significant methylated sites from the discovery cohort found that hypermethylation of five genes was significantly associated with shorter RFS in stage I NSCLC: *HIST1H4F*, *PCDHGB6*, *NPBWR1*, *ALX1*, and *HOXA9*. A signature based on the number of hypermethylated events distinguished patients with high- and low-risk stage I NSCLC (HR, 3.24; 95% CI, 1.61 to 6.54; $P = .001$).

Conclusion
The DNA methylation signature of NSCLC affects the outcome of stage I patients, and it can be practically determined by user-friendly polymerase chain reaction assays. The analysis of the best DNA methylation biomarkers improved prognostic accuracy beyond standard staging.

J Clin Oncol 31:4140-4147. © 2013 by American Society of Clinical Oncology

INTRODUCTION

Non–small-cell lung cancer (NSCLC) is the leading cause of cancer-related death.[1] The poor prognosis of patients with NSCLC is associated with several factors, among which are late disease diagnosis and the small number of effective drugs. The absence of validated prognostic biomarkers could also be relevant, because even patients with stage I NSCLC who undergo potentially curative surgical resection are at high risk of dying from recurrent disease, with a 5-year relapse rate of 35% to 50%.[1] Although adjuvant platinum-based chemotherapy is beneficial in more advanced resected disease, in which most of the patients have a high risk of recurrence,[2-6] it has failed to show a survival benefit for patients at stage I.[7,8] One explanation for these negative data in the early stages could be the lack of biologic factors predicting their recurrence and the fact that, in the absence of useful biomarkers, all stage I NSCLCs are pooled, making it more difficult to draw meaningful clinical conclusions.

In the search for new potential biomarkers of human cancer, the hypermethylation of the CpG island sequences located in the promoter regions of tumor suppressor genes are gaining prominence.[9-11] We wondered whether DNA methylation markers could also be used to provide a prognostic snapshot of lung tumors. Herein, we have obtained DNA methylation signatures associated with shorter relapse-free survival (RFS) in stage I NSCLCs that could be useful in the design of clinical trials for

adjuvant chemotherapy in the expanding population of those diagnosed with early-stage lung cancer.

PATIENTS AND METHODS

Study Design and Patient Population

Patients were eligible to enter the study as part of either discovery or validation cohorts if they underwent surgical resection of NSCLC in any of the international participating institutions. Patients treated with neoadjuvant therapy and/or patients with large cell-carcinoma were not included in the study. The clinical characteristics of the NSCLC surgical samples obtained are provided in Table 1. Descriptors of the patients by site of origin and for each single case are included in the Data Supplement. Tumors were collected by surgical resection from patients who provided consent and under approval by the institutional review boards. The median clinical follow-up was 7.2 years. Follow-up was performed by using radiographic imaging (chest x-ray and computed tomography scans), and time of recurrence was noted. In addition, 25 histologically normal lung tissue counterparts without any histologic evidence of malignancy were also analyzed (Data Supplement). The NSCLC tumor samples were studied in a consecutive manner as they arrived at the centralized DNA methylation facility and passed the technical quality checks.

Procedures

The DNA methylation status of 450,000 CpG sites was established by using the Infinium 450K Methylation Array.[12,13] The methylation score of

Table 1. Clinical Characteristics of the Discovery and Validation Cohorts

Characteristic	Discovery Cohort (n = 444)		Subset From Discovery Cohort for RFS Analysis (n = 198)*		Subset From Discovery Cohort for RFS Analysis (stage I) (n = 147)*		Validation Cohort (stage I) (n = 143)†	
	No.	%	No.	%	No.	%	No.	%
Age, years								
Median	65		65.5		65.9		63.7	
Range	35-90		38-85		38-85		32-78	
Sex								
Male	254	57	107	54	78	53	126	88
Female	190	43	91	46	69	47	17	12
Smoking history								
Current or former smoker	334	75	169	85	127	86	122	85
Nonsmoker	47	11	25	13	17	12	10	7
Unknown	63	14	4	2	3	2	11	8
Disease stage								
I	237	53	147	74	147	100	143	100
II	94	21	22	11	0	0	0	0
III	102	23	26	13	0	0	0	0
IV	11	3	3	2	0	0	0	0
Tumor type								
Adenocarcinoma	322	73	155	78	118	80	79	55
Squamous cell carcinoma	122	27	43	22	29	20	64	45
Thoracic surgery practice								
Lobectomy	396	90	172	86	132	90	117	82
Pneumonectomy	23	5	13	7	3	2	3	2
Segmentectomy	24	5	13	7	12	8	2	1
Unknown	1	0	0	0	0	0	21	15
Adjuvant treatment								
None	211	48	198	100	147	100	143	100
Chemotherapy	24	5	0	0	0	0	0	0
Chemotherapy plus radiotherapy	12	3	0	0	0	0	0	0
Radiotherapy	27	6	0	0	0	0	0	0
Unknown	170	38	0	0	0	0	0	0
Recurrence								
Yes	161	36	98	49	53	36	51	36
No	150	34	100	51	94	64	92	64
Unknown	133	30	0	0	0	0	0	0
RFS, months								
Average	46.7		50.8		60.7		42.3	
Range	0.6-224		0.6-224		0.6-224		2.6-130	
Origin of the samples								
Europe	291	66	100	51	68	46	142	99
United States	153	34	98	49	79	54	1	1

Abbreviation: RFS, relapse-free survival.
*Patients from the discovery cohort who had undergone resection of non–small-cell lung cancer and did not receive adjuvant chemotherapy before relapse.
†All patients from the validation cohort had undergone resection of non–small-cell lung cancer and did not receive adjuvant chemotherapy before relapse.

each CpG is represented as a β value. Samples were clustered in an unsupervised manner by using the 10,000 most variable β values for CpG methylation according to the standard deviation for the CpG sites located in promoter regions by hierarchical clustering using the complete method for agglomerating the Manhattan distances (Data Supplement). DNA methylation microarray data are available from the National Center for Biotechnology Information's Gene Expression Omnibus.[14] Pyrosequencing analyses to determine CpG methylation status were conducted as previously described.[15]

Statistical Analysis

Assay results were compared with patient outcomes in a double-blind manner. Median follow-up duration was calculated according to the inverse Kaplan-Meier method. Differences in distributions between groups were examined by the χ^2 test. The Kaplan-Meier method was used to estimate RFS, and differences among the groups were analyzed with the log-rank test. Hazard ratios (HRs) from univariate Cox regression analysis were used to determine the association of clinicopathologic features with relapse. Multivariate Cox proportional hazards regression was used to evaluate independent prognostic factors associated with RFS.

RESULTS

Characteristics of Patients in the Discovery Cohort

Clinical characteristics of the 444 patients in the discovery cohort are listed in Table 1. Descriptors of the patients by site of origin and for each single case are shown in the Data Supplement. The clinicopathologic characteristics of the lung tumors studied were related to the site of origin (United States *v* Europe).[16-18]

DNA Methylation Profiles Identify Two Groups With Different RFS Rates

We first evaluated a genome-wide DNA methylation profile of the original cohort of 444 patients with lung tumors, which included two NSCLC subtypes (adenocarcinoma and squamous cell carcinomas) by using a previously validated 450,000 CpG methylation microarray.[12,13] In addition, 25 histologically normal lung tissue counterparts without any histologic evidence of malignancy were also analyzed (Data Supplement).

The analyses of CpG methylation β values from the DNA methylation microarray within all primary NSCLCs (n = 444) and histologically normal tissues (n = 25) identified 10,000 promoter CpGs with the most variable CpG methylation levels (Data Supplement). These 10,000 top-ranked CpG sites were plotted in an unsupervised manner in the 444 primary NSCLCs (Fig 1A). The hierarchical clustering distinguished two main types of tumors that accounted for 70 (16%; group A) and 374 (84%; group B) patients. The χ^2 tests showed a significantly higher proportion of the adenocarcinoma histologic type in group A (χ^2 test $P = .02$), but no other significant differences in the distribution of the tumors according to stage, sex, or smoking history between group A and group B were observed (Data Supplement).

We investigated whether these two DNA methylation groups had any effect on the RFS of these patients. We analyzed the subset of patients who had undergone resection of NSCLC and had not received adjuvant chemotherapy before relapse, because of the possible confounding effect of chemotherapy on the RFS. Overall survival was not selected as an end point for the study because it could be affected by subsequent therapies received at relapse. Overall, 198 patients with NSCLC met the criteria for inclusion in the RFS cohort. Most importantly, these group A patients with NSCLC had a significantly shorter RFS, as shown in the Kaplan-Meier survival analysis (log-rank test $P < .001$; Fig 1B) and in the univariate (HR, 2.45; $P < .001$) and multivariate (HR, 2.40; $P < .001$) Cox regression analyses of stage, histology, smoking history, age, and sex (Data Supplement). In reference to histology, the unsupervised clustering analysis of either adenocarcinomas or squamous cell carcinomas also identified a group associated with shorter RFS (HR, 2.47; $P = .002$ and HR 4.93; $P = .001$, respectively; Data Supplement).

We wanted to extend these observations to identify those NSCLC tumors that, despite their low stage, are prone to recurrence. The selection of these patients is critical because approximately 30% to 40% of patients with stage I NSCLC die of recurrent disease.[19-21] To address this, the profile of the aforementioned 10,000 promoter CpGs, which had already shown their prognostic value throughout all NSCLC stages, was plotted in an unsupervised manner in the 237 patients with stage I NSCLC (Fig 1C). Hierarchical clustering distinguished two main types of tumors, accounting for 63 (27%; group 1) and 174 (73%; group 2) patients. The χ^2 tests revealed no significant differences in the distribution of the tumors in the two groups by sex, smoking history, and histologic type (Data Supplement). Among the 237 patients with stage I NSCLC, 147 met the described criteria for inclusion in the RFS cohort. The ineligible patients (n = 90) did not show a higher recurrence rate (χ^2 test $P = .12$). Group 1 identified patients with high-risk stage I NSCLC that had lower RFS, as revealed by the Kaplan-Meier survival analysis (log-rank test $P = .03$; Fig 1D) and in the univariate (HR, 1.85; $P = .037$) and multivariate (HR, 2.35; $P = .004$) Cox regressions of histology, smoking history, age, and sex (Data Supplement). In reference to histology, the unsupervised clustering analysis of the adenocarcinomas in stage I also identified a group associated with shorter RFS (HR, 2.94; $P = .003$; Data Supplement), and a trend was observed for squamous cell carcinomas (HR, 2.55; $P = .09$; Data Supplement). We also performed a Cox analysis that included smoking pack-years as a covariate. We categorized pack-years[22] as less than 30 or $\geq$ 30. The inclusion of the pack-year data value did not change the significant association of group 1 tumors with shorter RFS (HR, 2.3; $P = .007$). For all the patients with stage I NSCLC (because stage IA and IB have different outcomes), we also added this particular feature (according to the sixth revision of the TNM classification criteria) to the Cox regression multivariate analysis and group 1 remained significantly associated with shorter RFS (HR, 2.12; $P = .018$). The inclusion of tumor size within stage I (also an indicator of poor prognosis in NSCLC) in the Cox analysis did not alter the significant association of group 1 tumors with shorter RFS (HR, 2.02; $P = .05$). The reclassification of the stage I tumors according to the seventh revision of the TNM classification criteria also confirmed that group 1 patients remained significantly associated with shorter RFS (HR, 2.14; $P = .05$).

Identification of Candidate Genes as DNA Methylation Biomarkers of Shorter RFS in the Discovery Cohort of Stage I NSCLC

The identification of a DNA methylation signature for stage I NSCLC that predicts early recurrence might be useful, but the finding of a smaller panel of DNA methylation biomarkers could simplify the process. To achieve this goal, we developed an integrative approach to rank the CpG sites that, according to their methylation status (β values), were best at discriminating the 444 NSCLC samples from the 25 histologically normal lung tissue

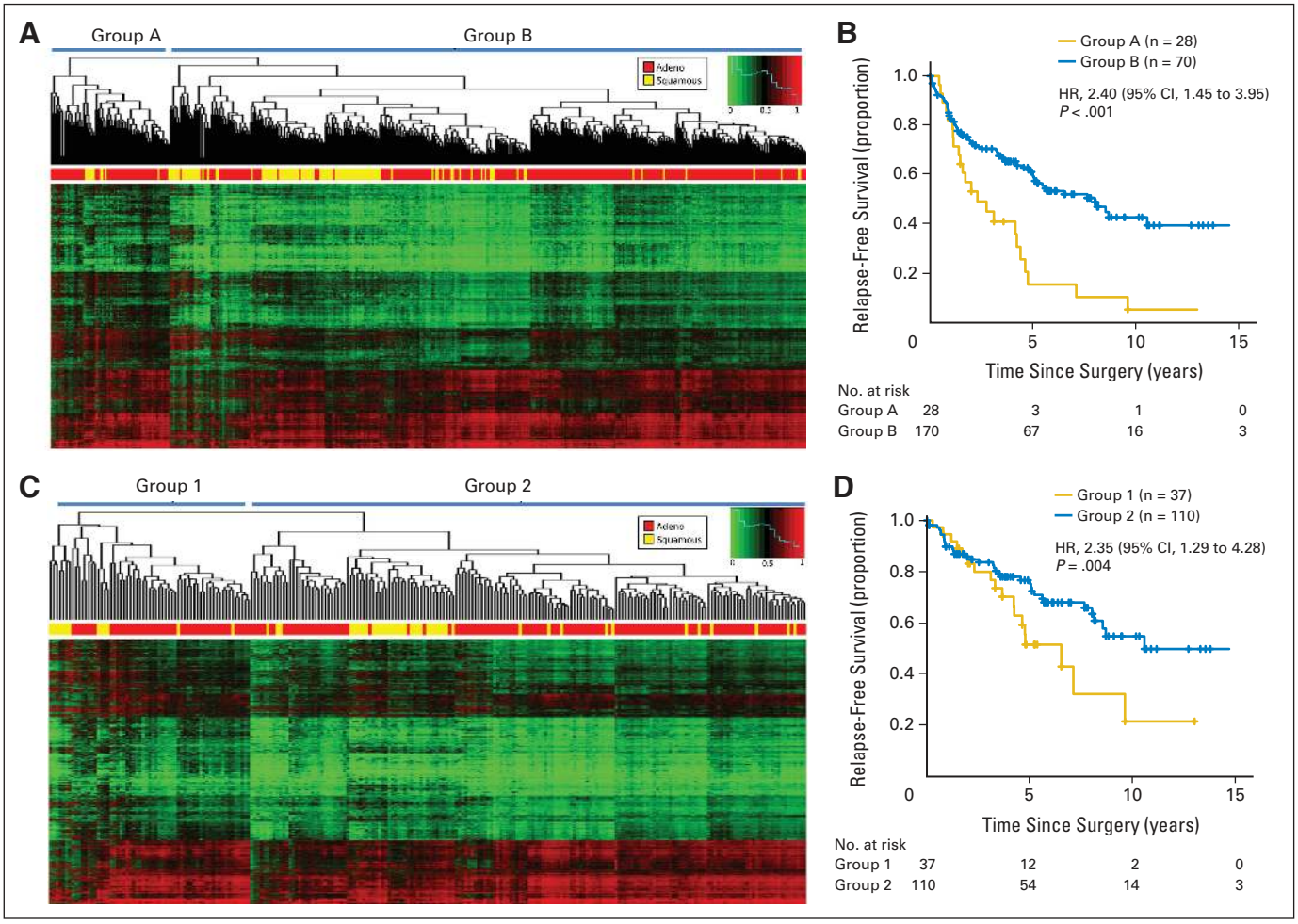

Fig 1. DNA methylation signatures associated with recurrence-free survival (RFS) in non–small-cell lung cancer samples. (A) Unsupervised hierarchical clustering and heat map associated with the methylation profile (according to the color scale shown) of the 444 primary tumor specimens based on the 10,000 most variable promoter β values. Two different histologic subtypes are represented: adenocarcinoma (Adeno; red) and squamous cell carcinoma (yellow). Each column represents an individual patient and each row an individual CpG. (B) Kaplan-Meier analysis for RFS among the 198 patients with RFS information according to the two groups obtained from the clustering. The *P* value corresponds to the hazard ratio (HR) adjusted by multivariate regression (including age, sex, smoking history, stage, and histologic type). (C) Unsupervised clustering and the heat map of the subset of 237 patients with stage I non–small-cell lung cancer. (D) Kaplan-Meier estimates for RFS among the subset of 147 patients with RFS information according to the two groups obtained in the clustering. The *P* value reflects the HR adjusted as in the analysis in (B).

samples. This analysis identified 338 highly ranked CpG sites (Data Supplement). From these, we focused on the CpGs located in regulatory regions: promoter CpG islands[9-11] and shores.[23,24] We found that 150 of the 338 CpG sites were located in the described regions. All of these 150 CpG sites were present in the 10,000 CpG sites used in the clustering. CpG hypermethylation of these 150 sites was significantly enriched in group A versus group B (*t* test $P < .001$) and in group 1 versus group 2 (*t* test $P < .001$), supporting their potential prognostic value. Thus, we tested the methylation value of each of these 150 CpG sites for RFS in the 147 stage I tumors by Kaplan-Meier survival and multivariate Cox regression. We identified 54 CpGs corresponding to 42 genes that were significantly associated with shorter RFS at a 10% false discovery rate (Data Supplement). Our data mining approach can be complemented by others and, in this regard, the promoter CpG sites of other methylation markers in lung cancer[25] did not pass the criteria used. However, we confirmed that *CDH13* and *RASSF1A* hypermethylation was associated with shorter RFS (HR, 3.47; $P = .01$ and HR, 2.17; $P = .02$, respectively) in the 147 stage I tumors.

Validation of Candidate Genes as DNA Methylation Biomarkers of Shorter RFS in an Independent Cohort of Stage I NSCLC

Once we had identified 42 genes with CpG promoter methylation that influenced RFS in our initial discovery cohort of 147 stage I tumors, we sought to validate these single DNA methylation markers in an additional cohort of 143 patients with stage I NSCLC (Table 1). Descriptors of the patients by site of origin and for each single sample are shown in the Data Supplement. All these new NSCLC samples were obtained from patients who had undergone a resection and did not receive adjuvant chemotherapy. The validation cohort, in comparison to the discovery set, was significantly enriched in European samples and, thus, in affected men and squamous cell carcinomas.[16,17] The methylation levels at the described CpG sites were analyzed by pyrosequencing[15] to test a more affordable large-scale approach. Methylation value by pyrosequencing was obtained from the average of each of the CpG dinucleotides included in the sequence analyzed (Data Supplement). Because the DNA material was limited, we selected the top 10 genes (Data Supplement) with an HR of more than

2 at a 10% false discovery rate (Data Supplement). By histology, four (80%) of the top five candidates in the adenocarcinoma set were also present in the overall 10-gene candidate list (Data Supplement).

Of these 10 candidate DNA methylation biomarkers associated with recurrence in the discovery cohort by using the DNA methylation microarray, five (50%) were significantly associated with recurrence ($P < .05$) in the validation cohort of 143 stage I NSCLC samples analyzed by pyrosequencing. These were the genes histone cluster1 H4F (*HIST1H4F*; HR, 3.55; $P < .001$), protocadherin gamma subfamily B6 (*PCDHGB6*; HR, 2.95; $P = .002$), neuropeptide B/W receptor 1

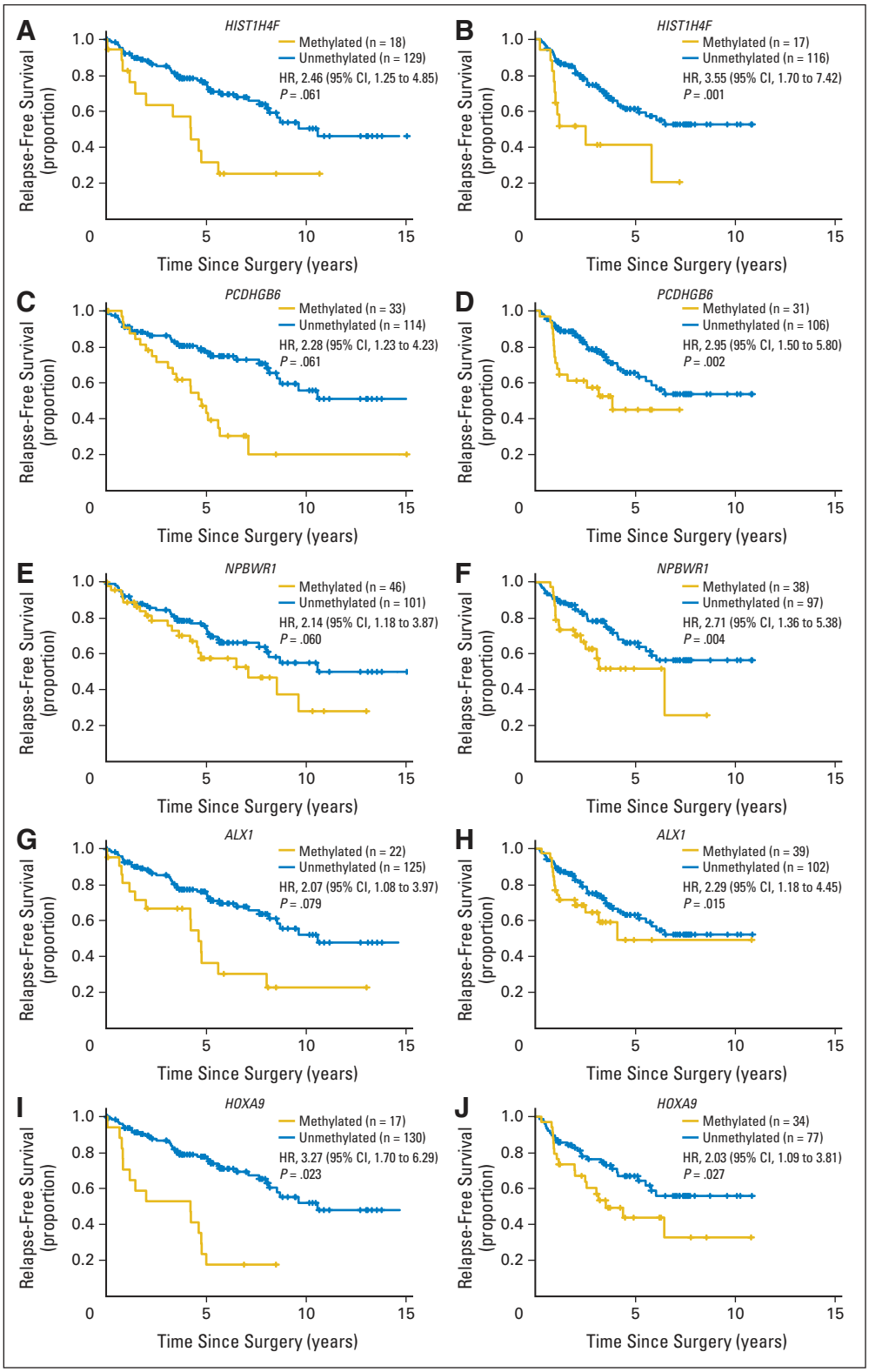

Fig 2. Kaplan-Meier estimates of relapse-free survival in patients with stage I non–small-cell lung cancer. Patients by methylation status of the five validated genes. (A, C, E, G, I) Kaplan-Meier estimates for relapse-free survival of the final five validated genes in the subset of 147 patients in stage I from the discovery cohort. Methylation status was determined by the Infinium 450k Methylation Array. (B, D, F, H, J) Corresponding Kaplan-Meier estimates for the same genes in the 143 patients in stage I included in the validation cohort. In this case, methylation status was determined by pyrosequencing. The *P* values correspond to hazard ratios (HRs) adjusted by multivariate regression (including age, sex, smoking history, and histologic type).

(*NPBWR1*; HR, 2.71; $P = .004$), ALX homeobox protein 1 (*ALX1*; HR, 2.29; $P = .015$), and homeobox A9 (*HOXA9*; HR, 2.03; $P = .027$; Fig 2 and Data Supplement). In addition, three other genes (30%) showed a trend toward significance (*OTX2*; HR, 1.82; $P = .11$; *TRIM58*; HR, 1.57; $P = .14$; and *TRH*; HR, 4.23; $P = .17$; Data Supplement). The pyrosequencing values for the five significant genes (*HIST1H4F*, *PCDHGB6*, *NPBWR1*, *ALX1*, and *HOXA9*) in all studied samples, histologically normal tissues (n = 25), and primary NSCLC (n = 143) are provided in the Data Supplement.

We also observed a greater risk of shorter RFS, according to Kaplan-Meier plots, when stage I NSCLCs harbored a large number of the five statistically significant hypermethylated markers (*HIST1H4F*, *PCDHGB6*, *NPBWR1*, *ALX1*, and *HOXA9*). To obtain the most useful methylation signature, we chose the cutoff of zero to one versus two or more hypermethylated markers, because it was the best one in resembling the percentage of expected recurrences.[1,19-21] The described methylation signature divides the patients with stage I tumors into two arms: patients with zero to one methylated markers that show longer RFS and those with two or more hypermethylated genes that were associated with a higher risk of poor RFS by Kaplan-Meier estimates (Fig 3A). The heavily hypermethylated group identified patients with high-risk stage I NSCLC who had shorter RFS, as shown by the Kaplan-Meier survival analysis (log-rank test $P = .010$; Fig 3A) and the univariate (HR, 2.26; $P = .012$) and multivariate (HR, 3.24; $P = .001$) Cox regressions (Data Supplement). The identified methylation signature remained significantly associated with shorter RFS in the Cox regression multivariate analysis, even when

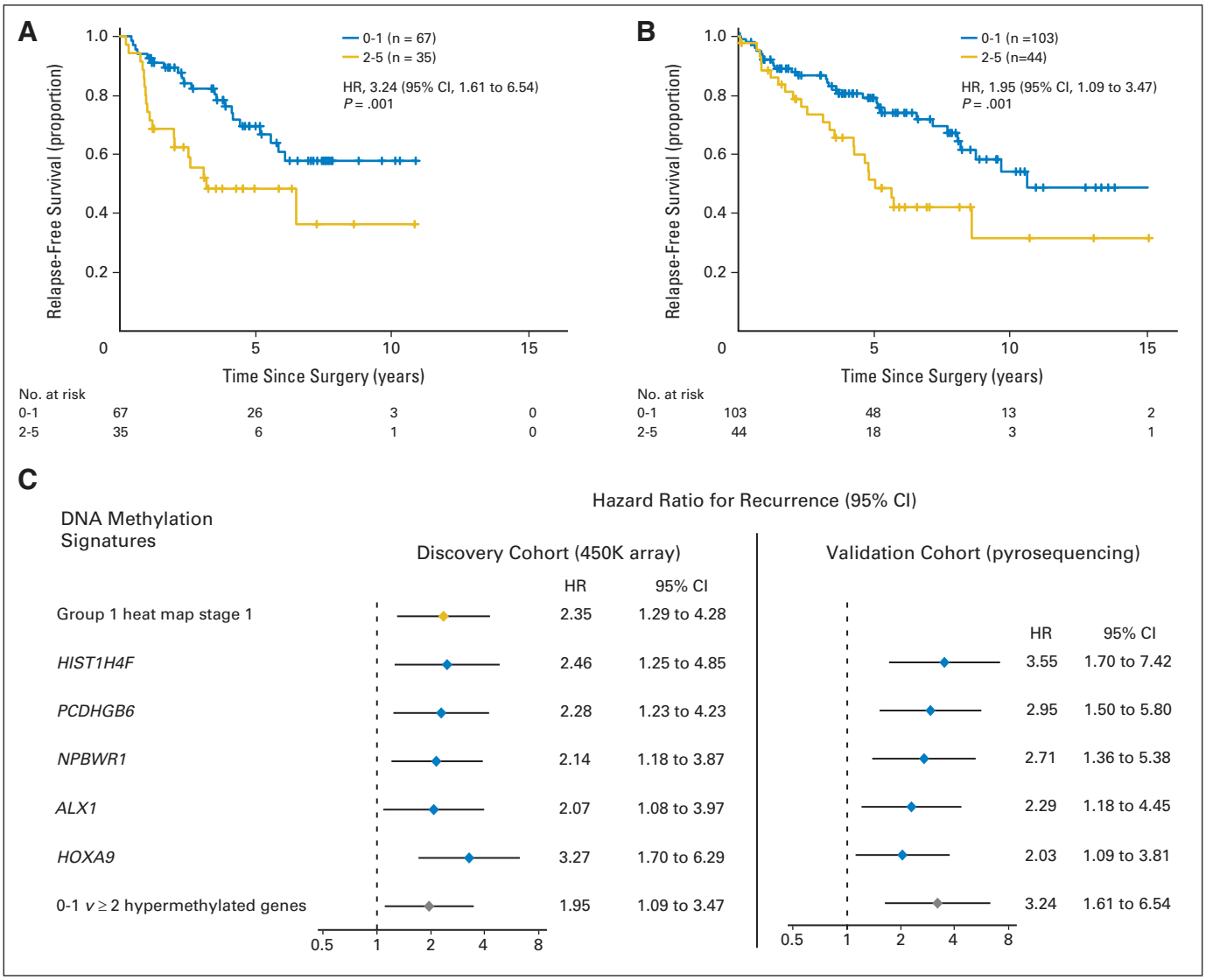

Fig 3. (A-B) Kaplan-Meier estimates of relapse-free survival by number of methylated genes and (C) forest plot with hazard ratios for recurrence in stage I non–small-cell lung cancer. (A, B) In each panel, patients are grouped into methylation low (blue) or methylation high (gold) groups according to the number of methylated genes (zero to one *v* two to five) from the five-gene signature (including *HIST1H4F*, *NPBWR1*, *PCDHGB6*, *ALX1*, and *HOXA9*). (A) Patients from the validation cohort analyzed by pyrosequencing. (B) Patients from the discovery cohort analyzed by the DNA methylation microarray. The *P* values correspond to hazard ratios (HRs) adjusted by multivariate regression (including age, sex, smoking history, and histologic type). (C) The forest plot shows the multivariate Cox regression for the various DNA methylation classifiers of patients with stage I non–small-cell lung cancer. Data for the group 1 heat map in stage I were obtained from the discovery cohort with the 450K array. Data from each of the five significant genes and the zero to one $v \geq$ two hypermethylated genes signature model were obtained from both the discovery cohort with the 450K array and the validation cohort by pyrosequencing. The prognostic value for each gene or signature was adjusted for age, sex, smoking history, and histologic type.

stage I tumors were subdivided into IA and IB according to the sixth revision of the TNM classification (HR, 3.09; $P = .002$). The reclassification of the stage I tumors according to the seventh revision of the TNM classification criteria also confirmed the relevance of the enriched hypermethylation group for shorter RFS in 103 original stage I tumors for which all the necessary clinicopathologic information was available (HR, 2.89; $P = .010$). The inclusion of tumor size in the Cox analysis within stage I did not alter the significant association of tumors with two or more methylated markers with shorter RFS (HR, 2.88; $P = .011$). Because 80% of recurrences of stage I NSCLC occur within 3 years of surgery,[19] we also calculated how many patients relapsed in this period. We observed that 48% (95% CI, 39.8% to 56.4%) of patients from the enriched methylated group relapsed, but only 18% (95% CI, 16.1% to 19.5%) of those in the low methylated group (zero to one methylated markers). Finally, as expected, the prognostic zero to one versus two or more hypermethylated genes signature obtained by pyrosequencing in the validation cohort was also observed in the 147 stage I NSCLCs from the discovery cohort studied by the DNA methylation microarray (HR, 1.95; $P = .023$; Fig 3B). Overall, we have identified DNA methylation classifiers that, at a different level of resolution, are potential prognostic biomarkers of shorter RFS in stage I NSCLC (Fig 3C).

DISCUSSION

One challenge in lung cancer management is that, despite complete surgical resection, patients with early-stage NSCLC are at considerable risk of recurrence and death. For this reason, we studied samples from stage I patients with the aim of identifying candidate DNA methylation biomarkers that can distinguish between patients at low risk of relapse and those at high risk for whom adjuvant treatment could be prescribed. Our results have distinguished two prognostic groups of stage I NSCLC at two levels of resolution by using a DNA methylation microarray profile that includes 10,000 CpG sites and by obtaining a methylation signature based on five genes derived from Cox regression models that could simplify the decision-making process.

Our study represents the largest cohort of primary NSCLCs studied for high-resolution DNA methylation analyses with a clinical orientation, a complement to the genomics[26] and expression[27] data. Other genomic approaches with lower resolution have determined DNA methylation profiles in NSCLCs,[15,28-32] although they have not focused on stage I. Candidate gene approaches have also suggested DNA methylation markers that are linked with prognosis in NSCLC.[25,33] An example is provided by the suggested association between *p16INK4a, CDH13, RASSF1A,* and *APC* hypermethylation and early recurrence in stage I lung cancer.[25] In addition, some CpG methylation events may be associated with better prognosis,[34] and their identification will require further analyses. It is also noteworthy to explain that useful hypermethylated markers to add to those characterized herein are possible and can be obtained from further mining of our publically available DNA methylation data.

Among the genes in our five-gene methylation signature, *HOXA9* hypermethylation has been described in lung tumorigenesis.[29,35,36] Although the association of *HOXA9* methylation with RFS was not assessed, *HOXA9* hypermethylation relates to poor prognosis in other tumor types.[37,38] For the other genes, *PCDHGB6* and *NPBWR1* hypermethylation occur in breast[39] and prostate[40] cancer, respectively, and both are associated with poor prognosis. Although our analysis was not aimed at finding markers of chemoresponse, the observed CpG hypermethylation of a particular histone gene in the high-risk group (*HIST1H4F*) warrants further research because a small subset of patients with NSCLC are sensitive to histone deacetylase inhibitors.[41] The identified patients with lung cancer whose high risk is associated with the described DNA methylation markers might also be a candidate group to receive DNA demethylating agents.[42]

The introduction of new therapies in NSCLC, such as epidermal growth factor receptor and anaplastic lymphoma kinase inhibitors, is a promising avenue for improving the outcome of these patients, but the target population is small. Although surgery remains the reference treatment in stage I NSCLCs, recurrence of the disease still occurs. Adjuvant platinum-based chemotherapy is beneficial in stage II and IIIa NSCLC.[2-6] Most studies have failed to show a survival benefit for adjuvant chemotherapy in stage I,[7,8] although a trend was observed for stage IB.[7,8] However, no molecular biomarkers were investigated in those trials. If we could identify stage I NSCLCs associated with shorter RFS, we could design stage-specific clinical trials in which a benefit of adjuvant therapies could accrue to the high-risk population. The DNA methylation markers identified herein, once they have been externally validated, could be useful for generating treatment guidelines for early-stage lung tumors.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) and/or an author's immediate family member(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a "U" are those for which no compensation was received; those relationships marked with a "C" were compensated. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.
Employment or Leadership Position: None **Consultant or Advisory Role:** None **Stock Ownership:** None **Honoraria:** John K. Field, Chinese Cancer Association 2013, 14th European Lung Cancer Congress 2013, 5th Asia Pacific Lung Cancer Conference 2012 **Research Funding:** None **Expert Testimony:** None **Patents:** None **Other Remuneration:** None

AUTHOR CONTRIBUTIONS

Conception and design: Juan Sandoval, Jesus Mendez-Gonzalez, Manel Esteller
Provision of study materials or patients: Ernest Nadal, Montse Sanchez-Cespedes, Josefina Mora, Lucia A. Muscarella, Marina Pollan, Luis Montuenga, Elisabeth Brambilla, John K. Field, Luca Roz, Giorgio V. Scagliotti, David Beer
Collection and assembly of data: Juan Sandoval, Jesus Mendez-Gonzalez, Ernest Nadal, Guoan Chen, F. Javier Carmona, Sergi Sayols, Sebastian Moran, Holger Heyn, Miguel Vizoso, Antonio Gomez, Montse Sanchez-Cespedes, Christoph Bock, Miquel Taron, Josefina Mora, Lucia A. Muscarella, Triantafillos Liloglou, Michael Davies, Marina Pollan, Maria J. Pajares, Wenceslao Torre, Luis M. Montuenga, Elisabeth Brambilla, John K. Field, Luca Roz, Marco Lo Iacono, Giorgio V. Scagliotti, Rafael Rosell, David G. Beer, Manel Esteller
Data analysis and interpretation: Juan Sandoval, Jesus Mendez-Gonzalez, Antonio Gomez, Yassen Assenov, Fabian Müller, Christoph Bock, Manel Esteller
Manuscript writing: All authors
Final approval of manuscript: All authors

REFERENCES

1. Siegel R, Naishadham D, Jemal A: Cancer statistics, 2012. CA Cancer J Clin 62:10-29, 2012
2. Scagliotti GV, Fossati R, Torri V, et al: Randomized study of adjuvant chemotherapy for completely resected stage I, II, or IIIA non-small-cell lung cancer. J Natl Cancer Inst 95:1453-1461, 2003
3. Waller D, Peake MD, Stephens RJ, et al: Chemotherapy for patients with non-small cell lung cancer: The surgical setting of the Big Lung Trial. Eur J Cardiothorac Surg 26:173-182, 2004
4. Winton T, Livingston R, Johnson D, et al: Vinorelbine plus cisplatin vs. observation in resected non-small-cell lung cancer. N Engl J Med 352:2589-2597, 2005
5. Douillard JY, Rosell R, De Lena M, et al: Adjuvant vinorelbine plus cisplatin versus observation in patients with completely resected stage IB-IIIA non-small-cell lung cancer (Adjuvant Navelbine International Trialist Association [ANITA]): A randomised controlled trial. Lancet Oncol 7:719-727, 2006
6. Arriagada R, Dunant A, Pignon JP, et al: Long-term results of the international adjuvant lung cancer trial evaluating adjuvant cisplatin-based chemotherapy in resected lung cancer. J Clin Oncol 28:35-42, 2010
7. Pignon JP, Tribodet H, Scagliotti GV, et al: Lung adjuvant cisplatin evaluation: A pooled analysis by the LACE Collaborative Group. J Clin Oncol 26:3552-3559, 2008
8. Strauss GM, Herndon JE 2nd, Maddaus MA, et al: Adjuvant paclitaxel plus carboplatin compared with observation in stage IB non-small-cell lung cancer: CALGB 9633 with the Cancer and Leukemia Group B, Radiation Therapy Oncology Group, and North Central Cancer Treatment Group Study Groups. J Clin Oncol 26:5043-5051, 2008
9. Herman JG, Baylin SB: Gene silencing in cancer in association with promoter hypermethylation. N Engl J Med 349:2042-2054, 2003
10. Esteller M: Epigenetics in cancer. N Engl J Med 358:1148-1159, 2008
11. Baylin SB, Jones PA: A decade of exploring the cancer epigenome: Biological and translational implications. Nat Rev Cancer 11:726-734, 2011
12. Sandoval J, Heyn H, Moran S, et al: Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. Epigenetics 6:692-702, 2011
13. Bibikova M, Barnes B, Tsan C, et al: High density DNA methylation array with single CpG site resolution. Genomics 98:288-295, 2011
14. National Center for Biotechnology Information: Gene Expression Omnibus. http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE39279
15. Fernandez AF, Assenov Y, Martin-Subero JI, et al: A DNA methylation fingerprint of 1628 human samples. Genome Res 22:407-419, 2012
16. Travis WD, Brambilla E, Muller-Hermlink HK, et al: Pathology and genetics of tumours of the lung, pleura, thymus and heart, in, World Health Organization (eds): Classification of Tumours, Lyon, France, IARC Press, 2004
17. Curado MP, Edwards B, Shin HR, et al: Cancer Incidence in Five Continents, Vol IX. Lyon, France, International Agency for Research on Cancer (IARC) Scientific Publication 160, 2007
18. Janssen-Heijnen ML, Coebergh JW: The changing epidemiology of lung cancer in Europe. Lung Cancer 41:245-258, 2003
19. Martini N, Bains MS, Burt ME, et al: Incidence of local recurrence and second primary tumors in resected stage I lung cancer. J Thorac Cardiovasc Surg 109:120-129, 1995
20. Mountain CF: Revisions in the International System for Staging Lung Cancer. Chest 111:1710-1717, 1997
21. Hoffman PC, Mauer AM, Vokes EE. Lung cancer. Lancet 355:479-485, 2000
22. Fujisawa T, Iizasa T, Saitoh Y, et al: Smoking before surgery predicts poor long-term survival in patients with stage I non-small-cell lung carcinomas. J Clin Oncol 17:2086-2091, 1999
23. Irizarry RA, Ladd-Acosta C, Wen B, et al: The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 41:178-186, 2009
24. Hansen KD, Timp W, Bravo HC, et al: Increased methylation variation in epigenetic domains across cancer types. Nat Genet 43:768-775, 2011
25. Brock MV, Hooker CM, Ota-Machida E, et al: DNA methylation markers and early recurrence in stage I lung cancer. N Engl J Med 358:1118-1128, 2008
26. Cancer Genome Atlas Research Network: Comprehensive genomic characterization of squamous cell lung cancers. Nature 489:519-525, 2012
27. Kratz JR, He J, Van Den Eeden SK, et al: A practical molecular assay to predict survival in resected non-squamous, non-small-cell lung cancer: Development and international validation studies. Lancet 379:823-832, 2012
28. Brena RM, Morrison C, Liyanarachchi S, et al: Aberrant DNA methylation of OLIG1, a novel prognostic factor in non-small cell lung cancer. PLoS Med 4:e108, 2007
29. Rauch T, Wang Z, Zhang X, et al: Homeobox gene methylation in lung cancer studied by genome-wide analysis with a microarray-based methylated CpG island recovery assay. Proc Natl Acad Sci U S A 104:5527-5532, 2007
30. Christensen BC, Marsit CJ, Houseman EA, et al: Differentiation of lung adenocarcinoma, pleural mesothelioma, and nonmalignant pulmonary tissues using DNA methylation profiles. Cancer Res 69:6315-6321, 2009
31. Selamat SA, Chung BS, Girard L, et al: Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression. Genome Res 22:1197-1211, 2012
32. Heller G, Babinsky VN, Ziegler B, et al: Genome-wide CpG island methylation analyses in non-small cell lung cancer patients. Carcinogenesis 34:513-521, 2013
33. Toyooka S, Suzuki M, Maruyama R, et al: The relationship between aberrant methylation and survival in non-small-cell lung cancers. Br J Cancer 91:771-774, 2004
34. Claus R, Lucas DM, Stilgenbauer S, et al: Quantitative DNA methylation analysis identifies a single CpG dinucleotide important for ZAP-70 expression and predictive of prognosis in chronic lymphocytic leukemia. J Clin Oncol 30:2483-2491, 2012
35. Son JW, Jeong KJ, Jean WS, et al: Genome-wide combination profiling of DNA copy number and methylation for deciphering biomarkers in non-small cell lung cancer patients. Cancer Lett 311:29-37, 2011
36. Nelson HH, Marsit CJ, Christensen BC, et al: Key epigenetic changes associated with lung cancer development: Results from dense methylation array profiling. Epigenetics 7:559-566, 2012
37. Alaminos M, Davalos V, Cheung NK, et al: Clustering of gene hypermethylation associated with clinical risk groups in neuroblastoma. J Natl Cancer Inst 96:1208-1219, 2004
38. Di Vinci A, Casciano I, Marasco E, et al: Quantitative methylation analysis of HOXA3, 7, 9, and 10 genes in glioma: Association with tumor WHO grade and clinical outcome. J Cancer Res Clin Oncol 138:35-47, 2012
39. Miyamoto K, Fukutomi T, Akashi-Tanaka S, et al: Identification of 20 genes aberrantly methylated in human breast cancers. Int J Cancer 116:407-414, 2005
40. Cottrell S, Jung K, Kristiansen G, et al: Discovery and validation of 3 novel DNA methylation markers of prostate cancer prognosis. J Urol 177:1753-1758, 2007
41. Ramalingam SS, Maitland ML, Frankel P, et al: Carboplatin and paclitaxel in combination with either vorinostat or placebo for first-line therapy of advanced non-small-cell lung cancer. J Clin Oncol 28:56-62, 2010
42. Juergens RA, Wrangle J, Vendetti FP, et al: Combination epigenetic therapy has efficacy in patients with refractory advanced non-small cell lung cancer. Cancer Discov 1:598-607, 2011

Affiliations

Juan Sandoval, Jesus Mendez-Gonzalez, F. Javier Carmona, Sergi Sayols, Sebastian Moran, Holger Heyn, Miguel Vizoso, Antonio Gomez, Montse Sanchez-Cespedes, and Manel Esteller, Bellvitge Biomedical Research Institute; Josefina Mora, Hospital de la Santa Creu i Sant Pau; Manel Esteller, University of Barcelona and Institucio Catalana de Recerca i Estudis Avançats, Barcelona; Miquel Taron and Rafael Rosell, Catalan Institute of Oncology, Badalona, Catalonia; Marina Pollan, Instituto de Salud Carlos III, Madrid; Maria J. Pajares and Luis M. Montuenga, University of Navarra; Wenceslao Torre, Clínica University de Navarra, Pamplona, Spain; Ernest Nadal, Guoan Chen, and David G. Beer, University of Michigan Medical School, Ann Arbor, MI; Yassen Assenov and Fabian Müller, Max Planck Institute, Saarbrücken, Germany; Christoph Bock, Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria; Lucia A. Muscarella, Istituto Di Ricovero e Cura a Carattere Scientifico (IRCCS) Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy; Triantafillos Liloglou, Michael Davies, and John K. Field, The University of Liverpool Cancer Research Centre, Liverpool, United Kingdom; Elisabeth Brambilla, Centre Hospitalier Universitaire A Michallon, Grenoble, France; Luca Roz, IRCCS Foundation National Cancer Institute, Milan; Marco Lo Iacono and Giorgio V. Scagliotti, University of Torino, Orbassano (Torino), Italy.

Acknowledgment

We thank Diana García, Carles Arribas, Vito Michele Fazio, Annamaria la Torre, and Ruben Pio for their support.