Table 4. Correlation between DNA methylation levels of hallmark genes for Clusters I, II and III and the clinicopathological parameters in the validation cohort (Continued)

(A) Hallmark genes for Cluster I

		DNA methylation level in non-cancerous lung tissue (N) samples[2] (mean ± SD)											
		Pleural anthracosis			Emphysematic change			Atypical adenomatous hyperplasia			Tumor anthracosis		
Target ID [1]	Gene symbol	G1	G2-3	*p*-Value[3]	Negative	Positive	*p*-Value[3]	Absence	Presence	*p*-Value[3]	Negative	Positive	*p*-Value[3]
cg11733245	IL2RA	−0.037 ± 0.042	−0.035 ± 0.051	9.29×10^{-1}	−0.023 ± 0.038	−0.047 ± 0.060	1.10×10^{-1}	−0.035 ± 0.051	−0.013 ± 0.023	1.09×10^{-1}	−0.009 ± 0.035	−0.047 ± 0.050	$\underline{7.07\times10^{-3}}$
cg22325572	CD53	−0.024 ± 0.083	−0.028 ± 0.058	9.23×10^{-1}	−0.017 ± 0.051	−0.040 ± 0.067	1.90×10^{-1}	−0.029 ± 0.060	−0.008 ± 0.057	4.68×10^{-1}	0.003 ± 0.057	−0.041 ± 0.057	$\underline{2.96\times10^{-2}}$
cg15691199	CEBPE	−0.029 ± 0.080	−0.018 ± 0.064	7.94×10^{-1}	−0.016 ± 0.050	−0.022 ± 0.079	7.79×10^{-1}	−0.022 ± 0.066	0.015 ± 0.021	$\underline{1.52\times10^{-2}}$	−0.002 ± 0.061	−0.028 ± 0.067	2.28×10^{-1}
cg16927606	U2AF1L4	−0.012 ± 0.079	−0.004 ± 0.054	8.25×10^{-1}	−0.005 ± 0.048	−0.006 ± 0.065	9.54×10^{-1}	−0.010 ± 0.057	0.032 ± 0.025	$\underline{1.48\times10^{-2}}$	0.009 ± 0.051	−0.012 ± 0.057	2.31×10^{-1}
cg16240480	EDARADD	−0.040 ± 0.103	−0.047 ± 0.072	8.91×10^{-1}	−0.032 ± 0.071	−0.062 ± 0.075	1.61×10^{-1}	−0.048 ± 0.074	−0.017 ± 0.070	3.86×10^{-1}	0.001 ± 0.049	−0.066 ± 0.076	$\underline{1.24\times10^{-3}}$
cg05596756	FAM113B	−0.010 ± 0.085	−0.010 ± 0.062	9.92×10^{-1}	0.000 ± 0.052	−0.022 ± 0.073	2.22×10^{-1}	−0.014 ± 0.064	0.026 ± 0.025	$\underline{2.14\times10^{-2}}$	0.011 ± 0.057	−0.020 ± 0.065	1.30×10^{-1}
cg08040471	C17orf62	0.000 ± 0.073	−0.014 ± 0.066	7.11×10^{-1}	−0.002 ± 0.051	−0.024 ± 0.080	2.80×10^{-1}	−0.015 ± 0.068	0.013 ± 0.029	1.21×10^{-1}	0.019 ± 0.053	−0.027 ± 0.066	$\underline{2.03\times10^{-2}}$
cg20622019	ADA	−0.021 ± 0.073	−0.033 ± 0.068	7.39×10^{-1}	−0.028 ± 0.069	−0.035 ± 0.066	7.31×10^{-1}	−0.036 ± 0.069	0.008 ± 0.021	$\underline{6.34\times10^{-3}}$	0.002 ± 0.051	−0.047 ± 0.068	$\underline{1.14\times10^{-2}}$
cg05109049	EVI2B	−0.057 ± 0.106	−0.041 ± 0.090	7.56×10^{-1}	−0.023 ± 0.093	−0.062 ± 0.084	1.32×10^{-1}	−0.044 ± 0.090	−0.009 ± 0.102	5.07×10^{-1}	−0.002 ± 0.054	−0.058 ± 0.099	$\underline{1.72\times10^{-2}}$
cg07973967	CD79B	−0.034 ± 0.105	−0.028 ± 0.075	9.16×10^{-1}	−0.017 ± 0.060	−0.042 ± 0.094	2.83×10^{-1}	−0.031 ± 0.080	0.007 ± 0.021	$\underline{2.06\times10^{-2}}$	−0.007 ± 0.067	−0.039 ± 0.081	1.69×10^{-1}

(B) Hallmark genes for Cluster II and III

		DNA methylation level in non-cancerous lung tissue (N) samples[5] (mean ± SD)									
		Lymphatic invasion			Nodal status				Pathological Tumor-Node-Metastasis stage			
Target ID [4]	Gene symbol	Negative	Positive	*p*-Value[6]	N0	N1	N2-3	*p*-Value[6]	IA-IB	IIA-IIB	IIIA-IV	*p*-Value[6]
cg26606064	*EI24*	0.018 ± 0.098	0.098 ± 0.113	$\underline{3.01\times10^{-2}}$	0.004 ± 0.098	0.010 ± 0.094	0.118 ± 0.091	$\underline{1.27\times10^{-3}}$	0.009 ± 0.102	−0.008 ± 0.082	0.118 ± 0.091	$\underline{1.14\times10^{-3}}$
cg17872476	*VTI1A*	−0.035 ± 0.134	−0.126 ± 0.137	$\underline{4.32\times10^{-2}}$	−0.023 ± 0.132	0.022 ± 0.127	−0.149 ± 0.121	$\underline{5.69\times10^{-3}}$	−0.020 ± 0.116	−0.016 ± 0.166	−0.149 ± 0.121	$\underline{6.70\times10^{-3}}$
cg21063899	*SCEL*	−0.043 ± 0.093	−0.109 ± 0.110	6.08×10^{-2}	−0.024 ± 0.082	−0.024 ± 0.090	−0.141 ± 0.096	$\underline{2.81\times10^{-4}}$	−0.021 ± 0.089	−0.033 ± 0.062	−0.141 ± 0.096	$\underline{2.60\times10^{-4}}$
cg14074641	*ABCC12*	−0.017 ± 0.118	−0.083 ± 0.117	8.60×10^{-2}	0.005 ± 0.104	−0.006 ± 0.112	−0.118 ± 0.114	$\underline{2.09\times10^{-3}}$	0.000 ± 0.111	0.014 ± 0.084	−0.118 ± 0.114	$\underline{2.00\times10^{-3}}$

[1]Probe ID for the Infinium HumanMethylation27 Bead Array.
[2]$\Delta\beta_{N\text{-}averageC}$.
[3]*p* values (Welch's *t*-test) of <0.05 are underlined.
[4]Probe ID for the Infinium HumanMethylation27 Bead Array.
[5]$\Delta\beta_{N\text{-}averageC}$.
[6]*p* values (Welch's *t*-test) of <0.05 are underlined.

strengthened in, T samples (Table 1). These findings are compatible with the "field cancerization" concept in the lung.[29] In our previous study using the Infinium assay, we proved that DNA methylation alterations in N samples resulted in silencing of tumor-related genes in tumorous tissue.[15] However, the correlation between the results of the Infinium assay in N samples and carcinogenetic factors was not examined in detail.

In this epigenetic clustering of patients with LADCs based on DNA methylation profiles in N samples, many of the patients belonging to Cluster I were heavy smokers. In fact, pleural anthracosis, which mainly reflects the long-term cumulative effects of cigarette smoking, was marked in the lungs of patients belonging to Cluster I. Smoking is known to be a cause of COPD. In fact, many patients in Cluster I actually suffered from obstructive ventilation impairment, and histological findings compatible with emphysema and lung fibrosis were observed in their N samples. Moreover, recurrent inflammation is generally associated with COPD,[30] and histological findings compatible with respiratory bronchiolitis[20,21] were actually observed in the lungs of patients belonging to Cluster I. Inflammation is known to be one of the major causes of DNA methylation alterations in precancerous conditions in various organs, such as chronic hepatitis[16,17] and chronic pancreatitis.[31,32] Taken together, the data suggest that the DNA methylation profiles characterizing Cluster I may be established in lung tissue through the long-term cumulative effects of cigarette smoking *via* chronic inflammation under the conditions of COPD. Unlike the previous study, which revealed aberrant DNA methylation of several tumor-related genes in lung cancers themselves of patients with COPD,[33] this study demonstrated for the first time the presence of distinct DNA methylation profiles related to COPD in N samples, based on genome-wide analysis.

The majority of patients belonging to Cluster II were non-smokers, especially young females. DNA methylation profiles characterizing Cluster II may reflect the carcinogenetic pathway that is unrelated to cigarette smoking. Mutation of the *EGFR* gene is well known to be a driver of LADCs in young female non-smokers, especially in Asia.[34] However, Cluster II included LADCs without *EGFR* gene mutations in non-smokers (data not shown), indicating that DNA methylation profiles in Cluster II were not entirely induced by *EGFR* mutation.

Although many of the patients belonging to Cluster III were smokers, the average number of cigarettes smoked per day × year index was lower in Cluster III than in Cluster I. In fact severe pleural anthracosis was not so frequently evident in the lungs of patients belonging to Cluster III. In addition, the incidence of emphysematous change and fibrosis was lower in the adjacent lung tissue of patients in Cluster III than in that of patients in Cluster I. DNA methylation profiles in Cluster III did not develop from a background of chronic inflammation in COPD, but may have developed rapidly before the long-term effects of cigarette smoking had accumulated in the adjacent lung tissue (possibly through more direct effects of carcinogens related or unrelated to cigarette smoking). However, to evaluate more precisely the effects of smoking on DNA methylation profiles, detailed DNA methylation analysis should be performed using purified epithelial cells, such as those from the airway epithelium.

Distinct DNA methylation profiles seem to be established in the non-cancerous lung during the carcinogenetic pathway *via* inflammation in COPD in heavy smokers (Fig. 2*a*), the carcinogenetic pathway unrelated to cigarette smoking (Fig. 2*b*), and the carcinogenetic pathway that occurs not *via* COPD but possibly *via* more direct effects of carcinogens (Fig. 2*c*). Each pathway may have distinct target genes as hallmarks for Clusters I, II and III (Table 3 and Supporting Information Table S3). Among 120 hallmark genes for Clusters I, II and III, 119 (one exception, *ABCC12*, being shared between Clusters II and III) showed ordered differences of DNA methylation from C to N, and then to T samples of the relevant cluster ($p < 0.05$, Jonckheere–Terpstra trend test, Table 3 and Supporting Information Table S3), indicating that a distinct DNA methylation profile in N samples of each cluster is inherited during progression to Ts.

A proportion of genes described in Table 3 and Supporting Information Table S3 may be simple hallmarks of each cluster (simple target genes of each carcinogenetic pathway). However, at least a proportion of DNA methylation alterations occurring during each carcinogenetic pathway actually result in altered expression of target genes, and may participate in establishment of the clinicopathological characteristics of LADCs in each cluster. The DNA methylation profiles in Cluster I may participate in the generation of locally invasive LADCs, which have a large diameter, a progressed T stage, a high histological grade, frequent pleural invasion and tumor anthracosis. DNA methylation profiles in Cluster II may participate in the generation of clinicopathologically less aggressive LADCs with a favorable outcome. DNA methylation profiles in Cluster III may participate in the generation of the most aggressive LADCs showing frequent lymphatic vessel invasion, blood vessel invasion, a high N stage, a high TNM stage and a poor outcome.

Table 3 includes homeobox genes, such as *IRX2* and *HOXD8*, a gene that has been implicated in cell migration, *SPARCL1*, and genes that have been implicated in apoptosis, such as *RGS5* and *EI24*. *IRX2* is known to participate in early lung development in mouse embryos.[35] *HOXD8* is known to be methylated and/or down-regulated in human malignancies, especially in metastatic, rather than in primary lesions.[36,37] *SPARCL1* is an extracellular matrix glycoprotein known to be correlated with cancer invasion.[38,39] *RGS5* is a member of the family of molecules regulating G protein signaling, and stimulates hypoxia-inducible apoptosis.[40] Positive correlations between *RGS5* expression and both tumor differentiation and a favorable outcome have been reported.[41,42] *EI24* is induced by *p53*, suppresses cell growth and induces apoptosis.[43] Reduced expression associated with DNA methylation of *IRX2, HOXD8, SPARCL1, RGS5* and *EI24* in our cohort of LADCs has been confirmed using expression microarray (data not shown). It is

feasible that these target genes of each carcinogenetic pathway participate in determining the clinicopathological characteristics of LADCs in each cluster.

In the validation cohort, the DNA methylation status of hallmark genes identified in N samples of Cluster I was significantly correlated with pleural anthracosis, which reflects the long-term cumulative effects of smoking, and COPD (pulmonary emphysema) in the adjacent lung and tumor anthracosis, which reflect active cancer–stromal interaction in LADCs. The DNA methylation status of the hallmark gene identified in N samples of Cluster II was significantly correlated with lower aggressiveness (low N stage and low TNM stage) of LADCs in the validation cohort. The DNA methylation status of hallmark genes identified in N samples of Cluster III was significantly correlated with aggressiveness of LADCs, such as lymph vessel invasion, a high N stage and a high TNM stage, in the validation cohort. Thus, correlations between distinct DNA methylation profiles in N samples and both carcinogenetic background factors in the adjacent lung tissue and clinicopathological characteristics of LADCs were confirmed in the validation cohort (Table 4).

Receiver operating characteristic curve (ROC) analysis was performed for N samples in the learning cohort, and the thresholds of the representative hallmark genes described in Table 4 were set so that they were nearest to the top left corner of the ROC. Using these thresholds, the sensitivity, specificity and accuracy for prediction of lymphatic vessel involvement, lymph node metastasis, TNM stage and patient outcome (recurrence and death) were calculated in both the learning and validation cohorts (Supporting Information Table S4). Even though Supporting Information Table S4 suggests that the aggressiveness of tumors developing in the same individual patients and patient outcome may be predictable on the basis of DNA methylation status in N samples, further examinations will be needed to set strict criteria for maximal sensitivity, specificity and accuracy.

Although bulk tissue comprising several cell lineages, for a large number of C, N and T samples, was examined in this study, it would be preferable to examine the DNA methylation status of purified cells. Therefore, the DNA methylation status of the representative gene *CASP8* (Infinium probe ID: cg26799474), included in Table 1B, was compared between cancer cells and normal peripheral airway epithelial cells obtained by microdissection from formalin-fixed, paraffin-embedded tissues of representative patients with LADCs and patients without primary lung cancers, respectively, using pyrosequencing. The DNA methylation levels in T samples (0.279 ± 0.184) were significantly lower than those in C samples (0.689 ± 0.042) by Infinium assay ($p = 3.64 \times 10^{-4}$). Such a significant difference was reproduced upon comparison with microdissected normal airway epithelium: pyrosequencing showed that the DNA methylation levels in microdissected cancer cells (0.273 ± 0.313) were significantly lower than those in microdissected normal airway epithelial cells (0.765 ± 0.104) ($p = 2.74 \times 10^{-3}$).

Differences in DNA methylation levels among different cell lineages, such as epithelial and stromal components, are also an important issue. Cancer cells and their stromal cells, such as cancer-associated fibroblasts, were again collected separately by microdissection from formalin-fixed, paraffin-embedded tissues from representative patients with LADCs. The DNA methylation levels of representative genes described in Table 1B were evaluated quantitatively by pyrosequencing. In one of the examined genes (*CASP8* [Infinium probe ID: cg26799474]), the DNA methylation statuses of cancer cells (0.273 ± 0.313) and stromal cells (0.219 ± 0.094) were almost equal, indicating that both may be affected by carcinogenetic factors. For the other examined gene (*LHX1* [Infinium probe ID: cg22660578]), the DNA methylation statuses of cancer cells (0.096 ± 0.141) and stromal cells (0.538 ± 0.486) differed from each other, probably reflecting differences in susceptibility to the effects of carcinogens, or differences in cell lineage.

In summary, DNA methylation profiles reflecting carcinogenetic background factors, such as smoking, inflammation and COPD, appear to be established in adjacent lung tissue in patients with LADCs. Such DNA methylation profiles in adjacent lung tissue may play a role in determining the aggressiveness of tumors developing in the same individual patients, and thus patient outcome.

Acknowledgement

T. Sato is an awardee of a research resident fellowship from the Foundation for Promotion of Cancer Research in Japan.

References

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin* 2013;63:11–30.
2. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers—a different disease. *Nat Rev Cancer* 2007;7:778–90.
3. Baylin SB, Jones PA. A decade of exploring the cancer epigenome—biological and translational implications. *Nat Rev Cancer* 2011;11:726–34.
4. Heller G, Zielinski CC, Zochbauer-Muller S. Lung cancer: from single-gene methylation to methylome profiling. *Cancer Metastasis Rev* 2010;29: 95–107.
5. Arai E, Kanai Y. DNA methylation profiles in precancerous tissue and cancers: carcinogenetic risk estimation and prognostication based on DNA methylation status. *Epigenomics* 2010;2: 467–81.
6. Kanai Y. Genome-wide DNA methylation profiles in precancerous conditions and cancers. *Cancer Sci* 2010;101:36–45.
7. Arai E, Chiku S, Mori T, et al. Single-CpG-resolution methylome analysis identifies clinicopathologically aggressive CpG island methylator phenotype clear cell renal cell carcinomas. *Carcinogenesis* 2012;33:1487–93.
8. Zöchbauer-Müller S, Lam S, Toyooka S, et al. Aberrant methylation of multiple genes in the upper aerodigestive tract epithelium of heavy smokers. *Int J Cancer* 2003;107: 612–6.
9. Eguchi K, Kanai Y, Kobayashi K, et al. DNA hypermethylation at the D17S5 locus in non-small cell lung cancers: its association with smoking history. *Cancer Res* 1997;57: 4913–5.
10. Lamy A, Sesboüé R, Bourguignon J, et al. Aberrant methylation of the CDKN2a/p16INK4a gene promoter region in preinvasive bronchial lesions: a prospective study in high-risk patients without invasive cancer. *Int J Cancer* 2002;100: 189–93.
11. Belinsky SA, Palmisano WA, Gilliland FD, et al. Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. *Cancer Res* 2002;62:2370–7.

12. Bibikova M, Le J, Barnes B, et al. Genome-wide DNA methylation profiling using Infinium(R) assay. *Epigenomics* 2009;1:177–200.
13. Selamat SA, Chung BS, Girard L, et al. Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression. *Genome Res* 2012;22:1197–211.
14. Lockwood WW, Wilson IM, Coe BP, et al. Divergent genomic and epigenomic landscapes of lung cancer subtypes underscore the selection of different oncogenic pathways during tumor development. *PLoS One* 2012;7:e37775.
15. Sato T, Arai E, Kohno T, et al. DNA methylation profiles at precancerous stages associated with recurrence of lung adenocarcinoma. *PLoS One* 2013;8:e59444.
16. Nagashio R, Arai E, Ojima H, et al. Carcinogenetic risk estimation based on quantification of DNA methylation levels in liver tissue at the precancerous stage. *Int J Cancer* 2011;129:1170–9.
17. Arai E, Ushijima S, Gotoh M, et al. Genome-wide DNA methylation profiles in liver tissue at the precancerous stage and in hepatocellular carcinoma. *Int J Cancer* 2009;125:2854–62.
18. Ushijima T, Hattori N. Molecular pathways: involvement of Helicobacter pylori-triggered inflammation in the formation of an epigenetic field defect, and its usefulness as cancer risk and exposure markers. *Clin Cancer Res* 2012;18:923–9.
19. Kanai Y, Hirohashi S. Alterations of DNA methylation associated with abnormalities of DNA methyltransferases in human cancers during transition from a precancerous to a malignant state. *Carcinogenesis* 2007;28:2434–42.
20. Garibaldi BT, Illei P, Danoff SK. Bronchiolitis. *Immunol Allergy Clin North Am* 2012;32:601–19.
21. Travis WD, Colby TV, Koss MN, et al. Obstructive pulmonary disease. In: King DW, ed. Atlas of nontumor pathology. Non-neoplastic disorders of the lower respiratory tract. Washington, DC: American Registry of Pathology, 2002:435–71.
22. Kerr KM. Clinical relevance of the new IASLC/ERS/ATS adenocarcinoma classification. *J Clin Pathol* 2013;66:832–8.
23. Noguchi M, Shimosato Y. The development and progression of adenocarcinoma of the lung. *Cancer Treat Res* 1995;72:131–42.
24. Colby TV, Noguchi M, Henschke C, et al. Adenocarcinoma. In: Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC, eds. World health classification of tumours pathology and genetics of tumours of the lung, pleura, thymus and heart. Lyon: IARC Press, 2004:35–44.
25. Wang D, Minami Y, Shu Y, et al. The implication of background anthracosis in the development and progression of pulmonary adenocarcinoma. *Cancer Sci* 2003;94:707–11.
26. Noguchi M, Morikawa A, Kawasaki M, et al. Small adenocarcinoma of the lung. Histologic characteristics and prognosis. *Cancer* 1995;75:2844–52.
27. UICC International Union Against Cancer. Lung and pleural tumours. In: Sobin LH, Gospodarowicz MK, Wittekind C, eds. TNM classification of malignant tumours, 7th edn. Oxford: Wiley-Blackwell, 2009:138–46.
28. Hu L, Sekine M, Gaina A, et al. Association of smoking behavior and socio-demographic factors, work, lifestyle and mental health of Japanese civil servants. *J Occup Health* 2007;49:443–52.
29. Kadara H, Kabbout M, Wistuba II. Pulmonary adenocarcinoma: a renewed entity in 2011. *Respirology* 2012;17:50–65.
30. Decramer M, Janssens W, Miravitlles M. Chronic obstructive pulmonary disease. *Lancet* 2012;379:1341–51.
31. Peng DF, Kanai Y, Sawada M, et al. DNA methylation of multiple tumor-related genes in association with overexpression of DNA methyltransferase 1 (DNMT1) during multistage carcinogenesis of the pancreas. *Carcinogenesis* 2006;27:1160–8.
32. Peng DF, Kanai Y, Sawada M, et al. Increased DNA methyltransferase 1 (DNMT1) protein expression in precancerous conditions and ductal carcinomas of the pancreas. *Cancer Sci* 2005;96:403–8.
33. Suzuki M, Wada H, Yoshino M, et al. Molecular characterization of chronic obstructive pulmonary disease-related non-small cell lung cancer through aberrant methylation and alterations of EGFR signaling. *Ann Surg Oncol* 2010;17:878–88.
34. Sharma SV, Bell DW, Settleman J, et al. Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer* 2007;7:169–81.
35. Mummenhoff J, Houweling AC, Peters T, et al. Expression of Irx6 during mouse morphogenesis. *Mech Dev* 2001;103:193–5.
36. Leshchenko VV, Kuo PY, Shaknovich R, et al. Genomewide DNA methylation analysis reveals novel targets for drug development in mantle cell lymphoma. *Blood* 2010;116:1025–34.
37. Kanai M, Hamada J, Takada M, et al. Aberrant expressions of HOX genes in colorectal and hepatocellular carcinomas. *Oncol Rep* 2010;23:843–51.
38. Hurley PJ, Marchionni L, Simons BW, et al. Secreted protein, acidic and rich in cysteine-like 1 (SPARCL1) is down regulated in aggressive prostate cancers and is prognostic for poor clinical outcome. *Proc Natl Acad Sci USA* 2012;109:14977–82.
39. Turtoi A, Musmeci D, Naccarato AG, et al. Sparc-like protein 1 is a new marker of human glioma progression. *J Proteome Res* 2012;11:5011–21.
40. Jin Y, An X, Ye Z, et al. RGS5, a hypoxia-inducible apoptotic stimulator in endothelial cells. *J Biol Chem* 2009;284:23436–43.
41. Huang G, Song H, Wang R, et al. The relationship between RGS5 expression and cancer differentiation and metastasis in non-small cell lung cancer. *J Surg Oncol* 2012;105:420–4.
42. Wang JH, Huang WS, Hu CR, et al. Relationship between RGS5 expression and differentiation and angiogenesis of gastric carcinoma. *World J Gastroenterol* 2010;16:5642–6.
43. Zhao X, Ayer RE, Davis SL, et al. Apoptosis factor EI24/PIG8 is a novel endoplasmic reticulum-localized Bcl-2-binding protein which is associated with suppression of breast cancer invasiveness. *Cancer Res* 2005;65:2125–9.

frontiers in
GENETICS

MINI REVIEW ARTICLE
published: 14 February 2014
doi: 10.3389/fgene.2014.00024

Multilayer-omics analyses of human cancers: exploration of biomarkers and drug targets based on the activities of the International Human Epigenome Consortium

Yae Kanai[1,2] * *and Eri Arai*[1,2]

[1] *Division of Molecular Pathology, National Cancer Center Research Institute, Tokyo, Japan*
[2] *Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Tokyo, Japan*

Edited by:
Yoshimasa Saito, Keio University, Japan

Reviewed by:
Makoto Chuma, Hokkaido University, Japan
Masaaki Takamura, Niigata University Graduate School of Medical and Dental Sciences, Japan

***Correspondence:**
Yae Kanai, Division of Molecular Pathology, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan
e-mail: ykanai@ncc.go.jp

Epigenetic alterations consisting mainly of DNA methylation alterations and histone modification alterations are frequently observed in cancers associated with chronic inflammation and/or persistent infection with viruses or other pathogenic microorganisms, or with cigarette smoking. Accumulating evidence suggests that alterations of DNA methylation are involved even in the early and precancerous stages. On the other hand, in patients with cancers, aberrant DNA methylation is frequently associated with tumor aggressiveness and poor patient outcome. Recently, epigenome alterations have been attracting a great deal of attention from researchers who are focusing on not only cancers but also neuronal, immune and metabolic disorders. In order to accurately identify disease-specific epigenome profiles that could be potentially applicable for disease prevention, diagnosis and therapy, strict comparison with standard epigenome profiles of normal tissues is indispensable. However, epigenome mechanisms show heterogeneity among tissues and cell lineages. Therefore, it is not easy to obtain a comprehensive picture of standard epigenome profiles of normal tissues. In 2010, the International Human Epigenome Consortium (IHEC) was established to coordinate the production of reference maps of human epigenomes for key cellular states. In order to gain substantial coverage of the human epigenome, the IHEC has set an ambitious goal to decipher at least 1000 epigenomes within the next 7–10 years. We consider that pathway analysis using genes showing multilayer-omics abnormalities, including genome, epigenome, transcriptome, proteome and metabolome abnormalities, may be useful for elucidating the molecular background of pathogenesis and for exploring possible therapeutic targets for each disease.

Keywords: epigenetics, epigenome, DNA methylation, International Human Epigenome Consortium (IHEC), multilayer/integrated disease omics analyses

MICRO RNAs AND HUMAN DISEASES

The Encyclopedia of DNA Elements (ENCODE) Consortium[1] data have revealed in more detail the high degree of complexity of the mammalian transcriptome: 75% of the genome is transcribed into different types of RNA molecules, e.g., protein-coding, long non-coding, pseudogenes, and small RNA genes (Djebali et al., 2012). RNA molecules show much greater variety than previously suspected. Among such RNA molecules, microRNAs (miRNAs) are non-coding RNAs comprising about 22 nucleotides initially transcribed by RNA polymerase II as primary miRNA (pri-miRNA) molecule precursors that possess a stem loop structure (Jinek and Doudna, 2009). RNase III Drosha acts over pri-mRNAs generating a pre-miRNA containing the hairpin (Jinek and Doudna, 2009). The pre-miRNAs are then exported to the cytoplasm and processed by Dicer into mature miRNAs, which are subsequently translocated into the RNA-induced silencing complex (RISC; Gomes et al., 2013). Each miRNA has multiple tasks, such as transcriptional repression via binding to partially complementary sequences in the 3′-untranslated regions of the target mRNAs and direct mRNA degradation via binding to perfectly complementary sequences (He and Hannon, 2004). Therefore, deregulation of miRNA levels may disturb the expression profiles in cells, thereby playing a key role in induction of diseases, such as cancers, neurodegenerative diseases, and autoimmune diseases.

Abbreviations: CIMP, CpG island methylator phenotype; COPD, chronic obstructive pulmonary disease; IHEC, International Human Epigenome Consortium; LC, normal lung tissue; LN, non-cancerous lung tissue obtained from patients with lung adenocarcinoma; LT, lung adenocarcinoma tissue; PBAT, post-bisulfite adaptor-tagging; RC, normal renal cortex tissue; RN, non-cancerous renal cortex tissue obtained from patients with clear cell renal cell carcinoma; RT, clear cell renal cell carcinoma tissue.

[1]https://genome.ucsc.edu/ENCODE/

EPIGENETICS AND miRNAs

Saito et al. revealed that treatment with the DNA demethylating agent 5-aza-2′-deoxycytidine and the histone deacetylase inhibitor 4-phenylbutyric acid induced marked changes in the expression profiles of miRNAs in human cancer cell lines. In particular, DNA hypermethylation and induction of active histone marks in the

promoter region of miR-127 resulted in decreased and increased expression of miR-127, respectively (Saito et al., 2006). Activation of miR-512-5p by epigenetic treatment induced apoptosis of human gastric cancer cell lines via suppression of the *MCL1* gene (Saito et al., 2009). In the human colon cancer cell line HCT116, disturbance of miRNA expression patterns has been reported after disruption of both *DNA methyltransferase* (*DNMT*) *1* and *DNMT3B* (Lujambio et al., 2007). Findings accumulated to date clearly indicate that expression levels of multiple miRNAs, such as let-7a-3, miR-1, miR-9-1, miR-9-3, miR-34a, mir34a*, mir34b/c, miR-124a, miR-126. miR127, miR-342, and miR-512-5p, are regulated epigenetically (Saito et al., 2013).

On the other hand, the expression of many proteins involved in epigenetics is regulated by miRNAs. For example, miR-152 acts as a tumor suppressor via suppression of *DNMT1* (Huang et al., 2010). The miR-29 family targets *DNMT3A* and *DNMT3B*, whereas miR-101 targets *EZH2* and may alter global chromatin structure (Fabbri et al., 2007). In addition, it has been shown that miRNA has the capacity to recognize chromatin by increasing the methylation of histone, for example through histone H3 lysine 27 tri-methylation (Kim et al., 2008). Thus the close connection between epigenetic alterations and miRNA dysregulation may have a great impact on human diseases.

PARTICIPATION OF EPIGENETIC ALTERATIONS IN MULTISTAGE HUMAN CARCINOGENESIS

Epigenetic alterations, consisting mainly of DNA methylation alterations and histone modification alterations, are often observed in cancers that are associated with chronic inflammation and/or persistent infection with viruses, such as hepatitis B or C viruses, Epstein–Barr virus, and human papillomavirus, or with cigarette smoking (Kanai and Hirohashi, 2007). Accumulating evidence suggests that alterations of DNA methylation are involved even in the early and precancerous stages (Arai and Kanai, 2010). On the other hand, in patients with cancers, aberrant DNA methylation is frequently associated with tumor aggressiveness and poor patient outcome (Kanai, 2008). Precancerous conditions showing alterations of DNA methylation may progress rapidly and generate more malignant cancers (Kanai, 2010).

As we described in the webpage of our laboratory[2], even though genetic alterations, such as activation of oncogenes and inactivation of tumor suppressor genes, have been considered to provide the molecular framework of multistage human carcinogenesis, genetic events alone may not explain the histological heterogeneity underlying the complex biological characteristics of tumors. Therefore, in the 1990s, we began to focus on epigenetic events that can be reversible, in an attempt to explain why cancers show such histopathological heterogeneity. At a time when only two genes, *RB* and *VHL*, were known as tumor suppressor genes silenced by DNA methylation, we showed for the first time that the *CDH1* gene, which encodes the E-cadherin cell adhesion molecule and acts as tumor suppressor, is silenced by DNA methylation around the promoter region in human cancers (Yoshiura et al., 1995). In 1996, we demonstrated that DNA methylation alterations frequently occurred at multiple loci on chromosome 16, one of the hot spots for loss of heterozygosity in liver cancers. This preceded loss of heterozygosity even at the chronic hepatitis or liver cirrhosis stages, which are widely considered to be precancerous conditions. This was one of the earliest reports of aberrant DNA methylation at the precancerous stage (Kanai et al., 1996).

Since then, we have reported DNA methylation alterations in tissue specimens at precancerous stages and in cancers using a candidate-gene approach. As an example of inflammation-associated carcinogenesis, ductal adenocarcinomas of the pancreas frequently develop in a background of chronic pancreatitis. Under these conditions, at least a proportion of peripheral pancreatic duct epithelia may be at the precancerous stage. It has been reported that the average number of methylated tumor-related genes, the incidence of DNA methylation of at least one of such genes, and the expression level of *DNMT1* protein are increased in pancreatic duct epithelia with an inflammatory background, and in another precancerous lesion, pancreatic intraductal neoplasia (PanIN), in comparison with normal pancreatic duct epithelia (Peng et al., 2006).

Urothelial carcinomas of the urinary bladder, renal pelvis, and ureter are clinically remarkable because of their multicentricity and tendency to recur. Such multiplicity may be attributable to the "field effect." Even non-cancerous urothelia showing no marked histological findings from patients with urothelial carcinomas can be considered precancerous, because they may have been exposed to carcinogens in the urine. It has been reported that the average number of methylated tumor-related genes and the expression level of *DNMT1* protein are increased in non-cancerous urothelia showing no marked histological findings from patients with urothelial carcinomas, in comparison with normal urothelia from patients without urothelial carcinomas (Nakagawa et al., 2005). Thus, overexpression of the major DNMT, *DNMT1*, may result in accumulated hypermethylation of DNA for tumor-related genes (Etoh et al., 2004). On the other hand, splicing alteration of *DNMT3B* may induce chromosomal instability through DNA hypomethylation of pericentromeric satellite regions (Saito et al., 2002).

As we described in the webpage of our laboratory[2], after genome-wide epigenetic (epigenome) analysis had become practical, we employed the bacterial artificial chromosome array-based methylated CpG island amplification (BAMCA) method for overviewing the DNA methylation tendency of large individual chromosomal regions. Although precancerous conditions in the kidney have rarely been described, despite the lack of any marked histological findings or association with chronic inflammation or persistent infection with pathogens, it can be considered that non-cancerous renal cortex tissue obtained from patients with renal cancers is already at the precancerous stage showing genome-wide DNA methylation alterations (Arai et al., 2006). We showed that DNA methylation profiles at the precancerous stage are inherited by renal cancers developing in individual patients, and that DNA methylation alterations at the precancerous stage determine both the aggressiveness of subsequently developing cancers and patient outcome through inducing further epigenetic and genetic alterations (Arai et al., 2009a). In addition, we have developed indicators for carcinogenetic risk estimation in patients with chronic hepatitis and liver cirrhosis (Arai et al., 2009b), indicators

[2]http://www.ncc.go.jp/en/nccri/divisions/01path/01path01.html

for estimating the risk of development of urothelial carcinomas that can be determined from urine samples (Nishiyama et al., 2010), diagnostic markers of pancreatic cancer that can be assessed from specimens of pancreatic juice (Gotoh et al., 2011), and indicators for prognostication of kidney, liver, pancreas, and urinary bladder cancers based on DNA methylation profiling. Based on these findings, we have filed patent applications for epigenome diagnosis techniques, and are now attempting to apply them practically.

ACTIVITIES OF THE INTERNATIONAL HUMAN EPIGENOME CONSORTIUM (IHEC)

Recently, epigenome alterations have been attracting a great deal of attention from researchers who are focusing on not only cancers but also neuronal, immune, and metabolic disorders. On the basis of epigenome profiling, attempts are now being made to elucidate the molecular pathogenesis of such diseases and to explore possible biomarkers and drug targets. In order to accurately identify such disease-specific epigenome profiles that could be potentially applicable for disease prevention, diagnosis, and therapy, strict comparison with standard epigenome profiles of normal tissues is indispensable. However, epigenome mechanisms show heterogeneity among tissues and cell lineages. Therefore, it is not easy to obtain a comprehensive picture of standard epigenome profiles of normal tissues. Based on improvements in next-generation sequencing technology, international collaboration will likely help to reveal standard epigenome profiles.

In 2010, the IHEC was established by researchers and founding agencies from Canada, South Korea, the EU, Italy, Germany, Japan, and the USA (Bae, 2013). As described in the webpage of IHEC[3], the primary goal of the IHEC is "to coordinate the production of reference maps of human epigenomes for key cellular states that are relevant to health and diseases." In order to achieve substantial coverage of the human epigenome, the IHEC has set an ambitious goal to decipher at least 1000 epigenomes[3]. To attain this goal, IHEC will use robust techniques to generate (1) high-resolution maps of histone modifications, H3K4me3, H3K9me3, H3K27me3, H3K27ac, H3K4me1, and H3K36me3, (2) high-resolution DNA methylation maps, (3) landmark maps of transcription start sites for all protein-encoding genes, and (4) a comprehensive catalog of non-coding and small RNAs and their patterns of expression[3]. The target cell types being studied by each team in the participating countries are shown on the IHEC website[4].

In Japan, three Japanese IHEC teams[5] including our team are supported by the Core Research for Evolutional Science and Technology division of the Japan Science and Technology Agency. To strengthen the research bases for cancers of digestive organs, including hepatocellular carcinomas and gastric carcinomas, which show high incidences in Japan, we are now performing standard epigenome analyses of normal epithelial cell lineages in digestive organs (**Figure 1**). Target cells of sufficient quality and quantity are being obtained from materials surgically resected from a range of Japanese patients. For example, for liver, we have obtained samples of normal liver tissue distant from sites of liver metastases from primary colon cancers in partial hepatectomy specimens from patients without viral hepatitis, chronic hepatitis, or liver cirrhosis. To isolate hepatocytes, we have performed collagenase perfusion of cannulated branches of the hepatic vein, followed by low-velocity centrifugation. On average, more than 10^7 dispersed cells can be obtained from each case, and immunocytochemistry has confirmed that the hepatocytes are more than 95% pure. In the stomach and colorectum, we initially employ the crypt isolation technique and collagenase digestion. Thereafter, each normal cell lineage is purified by fluorescence activated cell sorting using appropriate antibodies.

Members of our IHEC team have originally developed the post-bisulfite adaptor-tagging method (PBAT), which is an efficient library preparation method for whole-genome bisulfite sequencing (Miura et al., 2012). For the PBAT method, we first perform bisulfite modification followed by adaptor ligation using random priming. The PBAT method minimally requires sub-microgram DNA for mammalian whole-genome bisulfite sequencing without global PCR amplification. A good correlation of the DNA methylation pattern was observed among PBAT, the standard Methyl C-seq methodology developed by Lister et al. (2008), and the Illumina beads chip Infinium assay. The PBAT method is advantageous in that it requires only a small amount of genomic DNA but has good coverage of GC-rich regions, especially in CpG islands and gene-rich chromosomes. We now propose to make the PBAT method one of the standard protocols for IHEC. Under the supervision of the IHEC, we intend to disclose the data we obtain through the National Bioscience Database Center supported by the Japan Science and Technology Agency. Accurate standard epigenome profiles of digestive organ epithelial cells obtained through IHEC activities will be used to explore more useful biomarkers and drug targets of digestive organ cancers.

MULTILAYER/INTEGRATIVE DISEASE OMICS ANALYSES FOR EXPLORATION OF BIOMARKERS AND DRUG TARGETS

Recently big data analysis has impacted various fields of bioscience, especially disease research. It may not be appropriate to perform epigenome analysis including miRNA analysis using clinical samples. Instead, simultaneous multilayer/integrative disease omics analyses would seem more appropriate, including genome, epigenome, transcriptome, proteome, and metabolome analyses for exploration of drug targets. Since 2010, researchers at six National Centers in Japan, i.e., the National Cancer Center, National Cerebral and Cardiovascular Center, National Center for Neurology and Psychiatry, National Center for Global Health and Medicine, National Center for Child Health and Development and National Center for Geriatrics and Gerontology, have been engaged in a research project "Comprehensive exploration of drug targets based on multilayer/integrative disease omics analyses" supported by the Program for Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation (NiBio) (**Figure 2**). This project has been divided among a number of centers specializing in genome, epigenome, transcriptome, proteome, and metabolome analyses of tissue specimens from patients with various diseases that show a high incidence in the Japanese population. Tissue and body fluid

[3]http://ihec-epigenomes.org/about/objectives/

[4]http://ihec-epigenomes.org/research/cell-types/

[5]http://crest-ihec.jp/

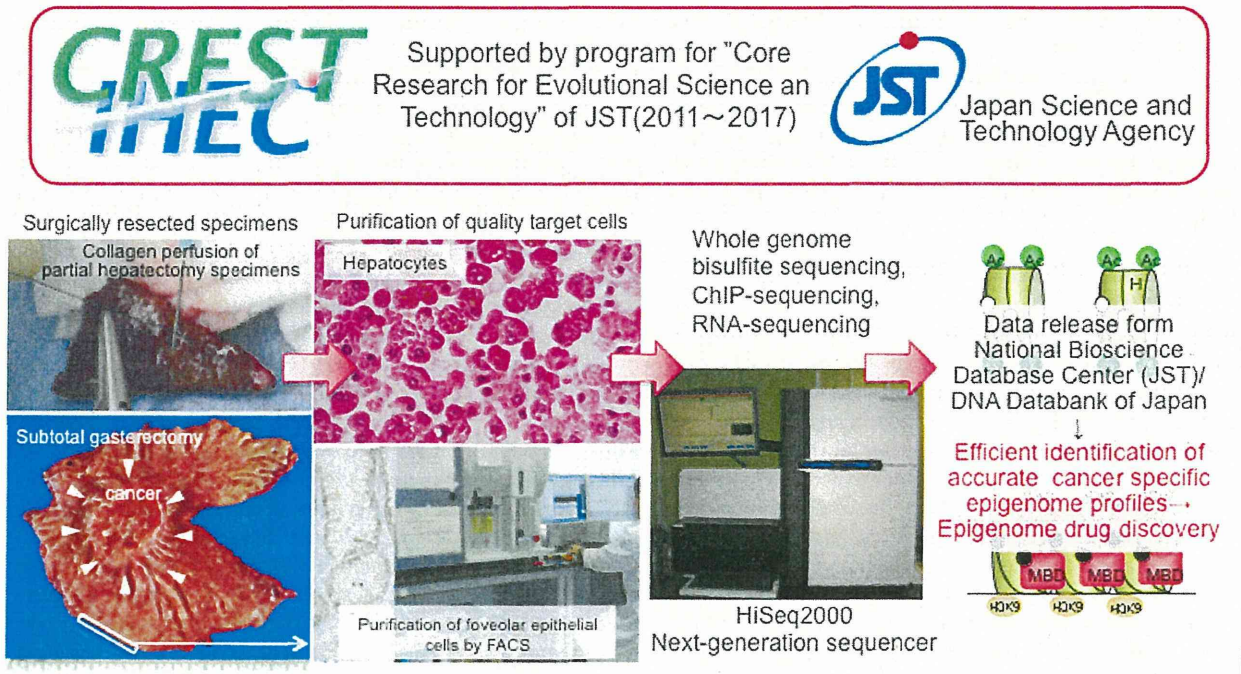

FIGURE 1 | Activity of our team, one of the Japan Teams of the International Human Epigenome Consortium supported by the Core Research for Evolutional Science and Technology division of the Japan Science and Technology Agency (modified from the webpage of our laboratory [http://www.ncc.go.jp/en/nccri/divisions/01path/01path01.html]). We are now performing whole-genome bisulfite sequencing using the post-bisulfite adaptor-tagging method, chromatin immunoprecipitation-sequencing, and RNA-sequencing of purified target cells, i.e., hepatocytes and other live cell lineages, foveolar epithelial cells, and other gastric epithelial cell lineages, and absorptive epithelial cells from the ascending and descending colon and rectum. Accurate epigenome profiling of normal cells will allow the identification of disease-specific epigenome profiles, thus facilitating a potential breakthrough in the prevention, diagnosis, and therapy of diseases.

specimens, cultured cells and animal models of adult cancers, infant leukemia, allergic disease, dilated cardiomyopathy, aortic aneurysm, epilepsy, obesity, non-alcoholic steatohepatitis, spinal canal stenosis, and Alzheimer's disease have been subjected to multilayer-omics analyses. As we described in the webpage of our laboratory[2], we are especially focusing on molecules or molecular pathways which are impaired as a result of multiple mechanisms, such as events in all five omics layers, which may participate in the molecular pathogeneses of diseases and might become potential biomarkers and/or druggable targets (**Figure 2**).

With regard to epigenome analysis of adult cancers in this research project, 414 lung tissue specimens including normal lung tissue (LC) obtained from patients without any primary lung tumor, non-cancerous lung tissue (LN) obtained from patients with lung adenocarcinomas, and lung adenocarcinoma tissue (LT) itself have been subjected to single-CpG resolution Infinium assay. DNA methylation alterations on many probes were evident in LN samples relative to LC samples, and were inherited by, or strengthened in, LT samples. Unsupervised hierarchical clustering using DNA methylation levels in LN samples subclustered patients into clusters I, II, and III. Lung adenocarcinomas in cluster I developed from an inflammatory background in chronic obstructive pulmonary disease (COPD) in heavy smokers, and were locally invasive. Most patients in cluster II were non-smokers and had a favorable outcome. Lung adenocarcinomas in cluster III were most aggressive cancers in light smokers that developed before accumulation of the long-term effects of cigarette smoking, and were probably due to the direct actions of carcinogens, rather than the effects of inflammation. DNA methylation profiles reflecting carcinogenetic factors such as smoking and COPD appear to be established in LNs and may determine the aggressiveness of tumors developing in individual patients, and thus patient outcome (Sato et al., 2014). Among the genes for which DNA methylation status in LN samples was significantly correlated with recurrence of lung adenocarcinomas in individual patients, we focused on *ADCY5*, *EVX1*, and other genes that were involved in apoptosis and cell adhesion. The mRNA expression levels of these genes were directly regulated by DNA methylation, and a decrease of their mRNA expression in LT samples was significantly correlated with tumor aggressiveness (Sato et al., 2013). When these genes were ectopically expressed in lung cancer cell lines, growth suppression, and apoptosis were induced, indicating that these genes could become therapeutic targets of lung adenocarcinomas.

With regard to epigenome analysis during renal carcinogenesis, 245 renal tissue specimens including normal renal cortex tissue (RC) obtained from patients without any primary renal cancer, non-cancerous renal cortex tissue (RN) obtained from patients with clear cell renal cell carcinomas, and clear cell

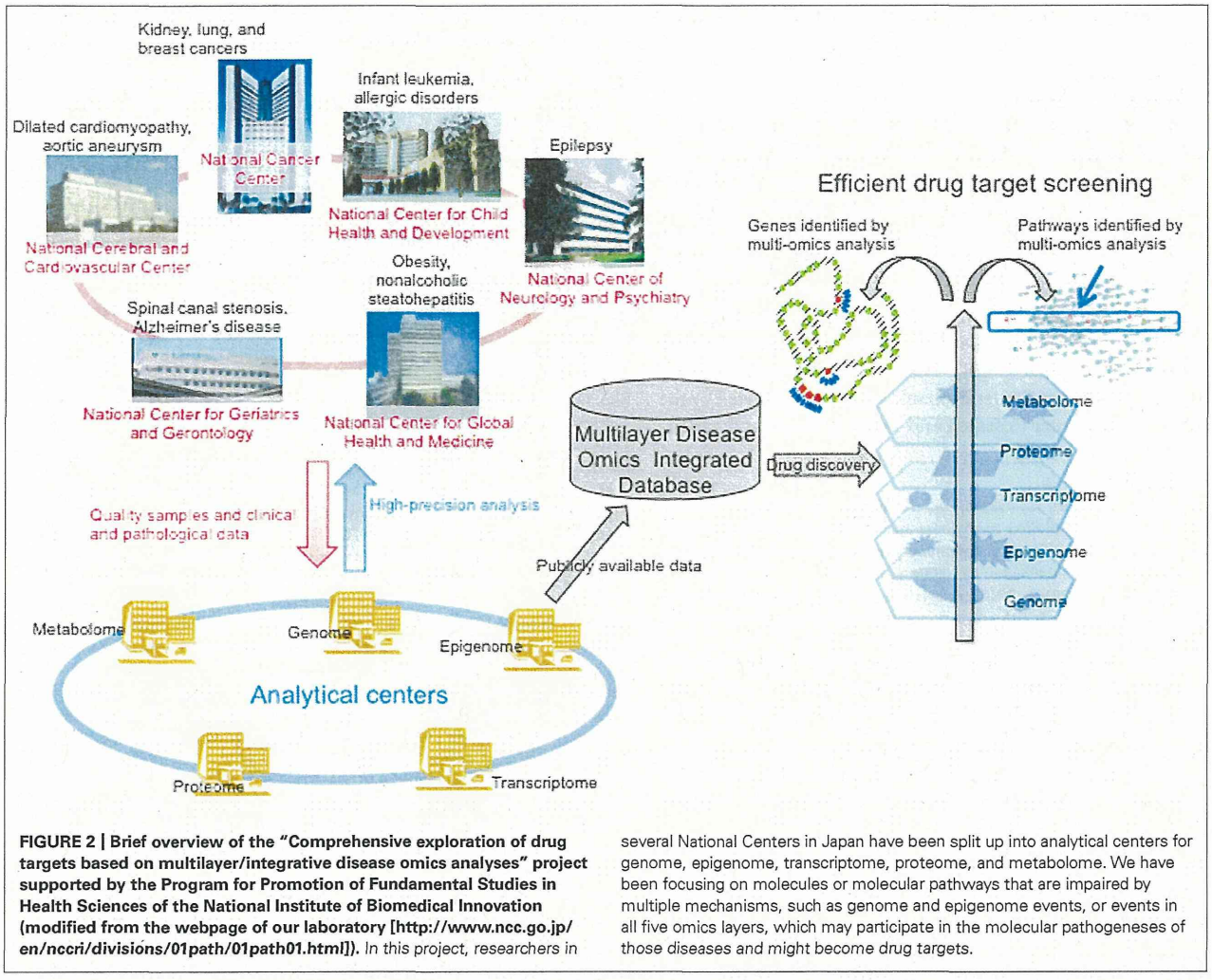

FIGURE 2 | Brief overview of the "Comprehensive exploration of drug targets based on multilayer/integrative disease omics analyses" project supported by the Program for Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation (modified from the webpage of our laboratory [http://www.ncc.go.jp/en/nccri/divisions/01path/01path01.html]). *In this project, researchers in several National Centers in Japan have been split up into analytical centers for genome, epigenome, transcriptome, proteome, and metabolome. We have been focusing on molecules or molecular pathways that are impaired by multiple mechanisms, such as genome and epigenome events, or events in all five omics layers, which may participate in the molecular pathogeneses of those diseases and might become drug targets.*

renal cell carcinoma tissue (RT) itself were subjected to the Infinium assay. DNA methylation levels at multiple Infinium probe sites were already altered in RN samples relative to RC samples. Unsupervised hierarchical clustering analysis based on DNA methylation levels at the CpG sites where DNA methylation alterations had occurred even in RN samples and were inherited by, and strengthened in, RT samples divided the clear cell renal cell carcinomas into CpG island methylator phenotype (CIMP)-positive and -negative clusters (Arai et al., 2012). Clinicopathologically aggressive cancers were accumulated in the CIMP-positive cluster, where the cancer-free and overall survival rates of the patients were significantly lower than in the CIMP-negative cluster. *FAM150A, GRM6, ZNF540, ZFP42, ZNF154, RIMS4, PCDHAC1, KHDRBS2, ASCL2, KCNQ1, PRAC, WNT3A, TRH, FAM78A, ZNF671, SLC13A5*, and *NKX6-2* have been identified as renal cell carcinoma-specific CIMP marker genes (Arai et al., 2012). Since CIMP-positive renal cell carcinomas show tumor aggressiveness and poorer patient outcome, we established criteria for prognostication of patients with clear cell renal cell carcinomas using renal cell carcinoma-specific CIMP marker genes. We are now performing pathway analysis based on a Bayesian estimation model using multiple genes showing frequent mutations and alterations of expression at the mRNA, miRNA, and protein levels based on multilayer-omics analyses in each of the CIMP-negative and CIMP-positive renal cell carcinomas for exploration of possible drug targets.

PERSPECTIVES

Once DNA methylation alterations occur during multistage carcinogenesis, such alterations are stably preserved on DNA double strands through maintenance methylation mechanisms by *DNMT1*. Therefore, stable stratification of cancers reflecting clinicopathological diversity may be possible based on epigenome profiling. Genes showing epigenome alterations, such as CIMP-marker genes, may become excellent biomarkers discriminating each tumor type stratified on the basis of epigenome profiling. We consider that pathway analysis using genes showing multilayer-omics abnormalities after stratification based on epigenome profiling may be useful for elucidating the molecular background of carcinogenetic pathways

and for exploring possible therapeutic targets for each tumor type.

ACKNOWLEDGMENTS

This study was supported by the Core Research for Evolutional Science and Technology (CREST) division "Development of Fundamental Technologies for Diagnosis and Therapy Based upon Epigenome Analysis" of the Japan Science and Technology Agency (JST), the Program for Promotion of Fundamental Studies in Health Sciences from the NiBio, a Grant in Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health, Labor and Welfare of Japan, and Grants in Aid for Scientific Research (B) and (C) from the Japan Society for the Promotion of Science (JSPS). Tissue specimens were provided by the National Cancer Center Biobank, which is supported by the National Cancer Center Research and Development Fund, Japan.

REFERENCES

Arai, E., Chiku, S., Mori, T., Gotoh, M., Nakagawa, T., Fujimoto, H., et al. (2012). Single-CpG-resolution methylome analysis identifies clinicopathologically aggressive CpG island methylator phenotype clear cell renal cell carcinomas. *Carcinogenesis* 33, 1487–1493. doi: 10.1093/carcin/bgs177

Arai, E., and Kanai, Y. (2010). DNA methylation profiles in precancerous tissue and cancers: carcinogenetic risk estimation and prognostication based on DNA methylation status. *Epigenomics* 2, 467–481. doi: 10.2217/epi.10.16

Arai, E., Kanai, Y., Ushijima, S., Fujimoto, H., Mukai, K., and Hirohashi, S. (2006). Regional DNA hypermethylation and DNA methyltransferase (DNMT) 1 protein overexpression in both renal tumors and corresponding nontumorous renal tissues. *Int. J. Cancer* 119, 288–296. doi:10.1002/ijc.21807

Arai, E., Ushijima, S., Fujimoto, H., Hosoda, F., Shibata, T., Kondo, T., et al. (2009a). Genome-wide DNA methylation profiles in both precancerous conditions and clear cell renal cell carcinomas are correlated with malignant potential and patient outcome. *Carcinogenesis* 30, 214–221. doi: 10.1093/carcin/bgn268

Arai, E., Ushijima, S., Gotoh, M., Ojima, H., Kosuge, T., Hosoda, F., et al. (2009b). Genome-wide DNA methylation profiles in liver tissue at the precancerous stage and in hepatocellular carcinoma. *Int. J. Cancer* 125, 2854–2862. doi: 10.1002/ijc.24708

Bae, J. B. (2013). Perspectives of international human epigenome consortium. *Genomics Inform.* 11, 7–14. doi: 10.5808/GI.2013.11.1.7

Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., et al. (2012). Landscape of transcription in human cells. *Nature* 489, 101–108. doi: 10.1038/nature11233

Etoh, T., Kanai, Y., Ushijima, S., Nakagawa, T., Nakanishi, Y., Sasako, M., et al. (2004). Increased DNA methyltransferase 1 (DNMT1) protein expression correlates significantly with poorer tumor differentiation and frequent DNA hypermethylation of multiple CpG islands in gastric cancers. *Am. J. Pathol.* 164, 689–699. doi: 10.1016/S0002-9440(10)63156-2

Fabbri, M., Garzon, R., Cimmino, A., Liu, Z., Zanesi, N., Callegari, E., et al. (2007). MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15805–15810. doi:10.1073/pnas.0707628104

Gomes, A. Q., Nolasco, S., and Soares, H. (2013). Non-coding RNAs: multitasking molecules in the cell. *Int. J. Mol. Sci.* 14, 16010–16039. doi: 10.3390/ijms140816010

Gotoh, M., Arai, E., Wakai-Ushijima, S., Hiraoka, N., Kosuge, T., Hosoda, F., et al. (2011). Diagnosis and prognostication of ductal adenocarcinomas of the pancreas based on genome-wide DNA methylation profiling by bacterial artificial chromosome array-based methylated CpG island amplification. *J. Biomed. Biotechnol.* 2011, 780836. doi: 10.1155/2011/780836

He, L., and Hannon, G. J. (2004). MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* 5, 522–531. doi: 10.1038/nrg1379

Huang, J., Wang, Y., Guo, Y., and Sun, S. (2010). Down-regulated microRNA-152 induces aberrant DNA methylation in hepatitis B virus-related hepatocellular carcinoma by targeting DNA methyltransferase 1. *Hepatology* 52, 60–70. doi: 10.1002/hep.23660

Jinek, M., and Doudna, J. A. (2009). A three-dimensional view of the molecular machinery of RNA interference. *Nature* 457, 405–412. doi: 10.1038/nature07755

Kanai, Y. (2008). Alterations of DNA methylation and clinicopathological diversity of human cancers. *Pathol. Int.* 58, 544–558. doi: 10.1111/j.1440-1827.2008.02270.x

Kanai, Y. (2010). Genome-wide DNA methylation profiles in precancerous conditions and cancers. *Cancer Sci.* 101, 36–45. doi: 10.1111/j.1349-7006.2009.01383.x

Kanai, Y., and Hirohashi, S. (2007). Alterations of DNA methylation associated with abnormalities of DNA methyltransferases in human cancers during transition from a precancerous to a malignant state. *Carcinogenesis* 28, 2434–2442. doi: 10.1093/carcin/bgm206

Kanai, Y., Ushijima, S., Tsuda, H., Sakamoto, M., Sugimura, T., and Hirohashi, S. (1996). Aberrant DNA methylation on chromosome 16 is an early event in hepatocarcinogenesis. *Jpn. J. Cancer Res.* 87, 1210–1217. doi: 10.1111/j.1349-7006.1996.tb03135.x

Kim, D. H., Saetrom, P., Snøve, O. Jr., and Rossi, J. J. (2008). MicroRNA-directed transcriptional gene silencing in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 105, 16230–16235. doi: 10.1073/pnas.0808830105

Lister, R., O'Malley, R. C., Tonti-Filippini, J., Gregory, B. D., Berry, C. C., Millar, A. H., et al. (2008). Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* 133, 523–536. doi: 10.1016/j.cell.2008.03.029

Lujambio, A., Ropero, S., Ballestar, E., Fraga, M. F., Cerrato, C., Setién, F., et al. (2007). Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res.* 67, 1424–1429. doi: 10.1158/0008-5472.CAN-06-4218

Miura, F., Enomoto, Y., Dairiki, R., and Ito, T. (2012). Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. *Nucleic Acids Res.* 40, e136. doi: 10.1093/nar/gks454

Nakagawa, T., Kanai, Y., Ushijima, S., Kitamura, T., Kakizoe, T., and Hirohashi, S. (2005). DNA hypermethylation on multiple CpG islands associated with increased DNA methyltransferase DNMT1 protein expression during multistage urothelial carcinogenesis. *J. Urol.* 173, 1767–1771. doi: 10.1097/01.ju.0000154632.11824.4d

Nishiyama, N., Arai, E., Chihara, Y., Fujimoto, H., Hosoda, F., Shibata, T., et al. (2010). Genome-wide DNA methylation profiles in urothelial carcinomas and urothelia at the precancerous stage. *Cancer Sci.* 101, 231–240. doi: 10.1111/j.1349-7006.2009.01330.x

Peng, D. F., Kanai, Y., Sawada, M., Ushijima, S., Hiraoka, N., Kitazawa, S., et al. (2006). DNA methylation of multiple tumor-related genes in association with overexpression of DNA methyltransferase 1 (DNMT1) during multistage carcinogenesis of the pancreas. *Carcinogenesis* 27, 1160–1168. doi: 10.1093/carcin/bgi361

Saito, Y., Kanai, Y., Sakamoto, M., Saito, H., Ishii, H., and Hirohashi, S. (2002). Overexpression of a splice variant of DNA methyltransferase 3b, DNMT3b4, associated with DNA hypomethylation on pericentromeric satellite regions during human hepatocarcinogenesis. *Proc. Natl. Acad. Sci. U.S.A.* 99, 10060–10065. doi: 10.1073/pnas.152121799

Saito, Y., Liang, G., Egger, G., Friedman, J. M., Chuang, J. C., and Coetzee. G. A., et al. (2006). Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell* 9, 435–443. doi: 10.1016/j.ccr.2006.04.020

Saito, Y., Saito, H., Liang, G., and Friedman, J. M. (2013). Epigenetic alterations and microRNA misexpression in cancer and autoimmune diseases: a critical review. *Clin. Rev. Allergy Immunol.* doi: 10.1007/s12016-013-8401-z [Epub ahead of print].

Saito, Y., Suzuki, H., Tsugawa, H., Nakagawa, I., Matsuzaki, J., Kanai, Y., et al. (2009). Chromatin remodeling at Alu repeats by epigenetic treatment activates silenced microRNA-512-5p with downregulation of Mcl-1 in human gastric cancer cells. *Oncogene* 28, 2738–2744. doi: 10.1038/onc.2009.140

Sato, T., Arai, E., Kohno T., Takahashi, Y., Miyata, S., Tsuta, K., et al. (2014). Epigenetic clustering of lung adenocarcinomas based on DNA methylation profiles in adjacent lung tissue: its correlation with smoking history and chronic obstructive pulmonary disease. *Int. J. Cancer* doi: 10.1002/ijc.28684

Sato, T., Arai, E., Kohno, T., Tsuta, K., Watanabe, S., Soejima, K., et al. (2013). DNA methylation profiles at precancerous stages associated with recurrence of lung adenocarcinoma. *PLoS ONE* 8:e59444. doi: 10.1371/journal.pone.0059444

Yoshiura, K., Kanai, Y., Ochiai, A., Shimoyama, Y., Sugimura, T., and Hirohashi, S. (1995). Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. *Proc. Natl. Acad. Sci. U.S.A.* 92, 7416–7419. doi: 10.1073/pnas.92.16.7416

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 January 2014; accepted: 24 January 2014; published online: 14 February 2014.

Citation: Kanai Y and Arai E (2014) Multilayer-omics analyses of human cancers: exploration of biomarkers and drug targets based on the activities of the International Human Epigenome Consortium. Front. Genet. 5:24. doi: 10.3389/fgene.2014.00024
This article was submitted to Epigenomics and Epigenetics, a section of the journal Frontiers in Genetics.

OPEN ACCESS Freely available online

PLOS | ONE

DNA Methylation Profiles at Precancerous Stages Associated with Recurrence of Lung Adenocarcinoma

Takashi Sato[1,2], Eri Arai[1]*, Takashi Kohno[3], Koji Tsuta[4], Shun-ichi Watanabe[5], Kenzo Soejima[2], Tomoko Betsuyaku[2], Yae Kanai[1]

1 Division of Molecular Pathology, National Cancer Center Research Institute, Tokyo, Japan, **2** Division of Pulmonary Medicine, Department of Medicine, Keio University School of Medicine, Tokyo, Japan, **3** Division of Genome Biology, National Cancer Center Research Institute, Tokyo, Japan, **4** Department of Pathology and Clinical Laboratories, Pathology Division, National Cancer Center Hospital, Tokyo, Japan, **5** Department of Thoracic Oncology, Thoracic Surgery Division, National Cancer Center Hospital, Tokyo, Japan

Abstract

The aim of this study was to clarify the significance of DNA methylation alterations at precancerous stages of lung adenocarcinoma. Using single-CpG resolution Infinium array, genome-wide DNA methylation analysis was performed in 36 samples of normal lung tissue obtained from patients without any primary lung tumor, 145 samples of non-cancerous lung tissue (N) obtained from patients with lung adenocarcinomas, and 145 samples of tumorous tissue (T). Stepwise progression of DNA methylation alterations from normal lung tissue to non-cancerous lung tissue obtained from patients with lung adenocarcinomas, and then tumorous tissue samples, was observed at 3,270 CpG sites, suggesting that non-cancerous lung tissue obtained from patients with lung adenocarcinomas was at precancerous stages with DNA methylation alterations. At CpG sites of 2,083 genes, DNA methylation status in samples of non-cancerous lung tissue obtained from patients with lung adenocarcinomas was significantly correlated with recurrence after establishment of lung adenocarcinomas. Among such recurrence-related genes, 28 genes are normally unmethylated (average β-values based on Infinium assay in normal lung tissue samples was less than 0.2) and their DNA hypermethylation at precancerous stages was strengthened during progression to lung adenocarcinomas ($\Delta\beta_{T-N}>0.1$). Among these 28 genes, we focused on 6 for which implications in transcription regulation, apoptosis or cell adhesion had been reported. DNA hypermethylation of the *ADCY5, EVX1, GFRA1, PDE9A,* and *TBX20* genes resulted in reduced mRNA expression in tumorous tissue samples. 5-Aza-2′-deoxycytidine treatment of lung cancer cell lines restored the mRNA expression levels of these 5 genes. Reduced mRNA expression in tumorous tissue samples was significantly correlated with tumor aggressiveness. These data suggest that DNA methylation alterations at precancerous stages determine tumor aggressiveness and outcome through silencing of specific genes.

Citation: Sato T, Arai E, Kohno T, Tsuta K, Watanabe S-i, et al. (2013) DNA Methylation Profiles at Precancerous Stages Associated with Recurrence of Lung Adenocarcinoma. PLoS ONE 8(3): e59444. doi:10.1371/journal.pone.0059444

Editor: Bernard W Futscher, The University of Arizona, United States of America

Received September 25, 2012; **Accepted** February 14, 2013; **Published** March 27, 2013

Funding: This study was supported by the Program for Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation (NiBio, 10–42, http://www.nibio.go.jp/part/promote/fundamental/doc/index.html) and partially supported by a Grant-in-Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health, Labor and Welfare of Japan (2, http://www.mhlw.go.jp/bunya/kenkyuujigyou/hojokin-koubo14/09.html), National Cancer Center Research and Development Fund (23-A-11, http://www.ncc.go.jp/jp/about/rinri/kaihatsu/files/h24_ncc_research_list.pdf), and Grants-in-Aid for Scientific Research (B, 233900 90, http://www.jsps.go.jp/j-grantsinaid/) and for Young Scientists (B, 23790456, http://www.jsps.go.jp/j-grantsinaid/) from the Japan Society for the Promotion of Science (JSPS). National Cancer Center Biobank is supported by the National Cancer Center Research and Development Fund (23-A-1, http://www.ncc.go.jp/jp/about/rinri/kaihatsu/files/h24_ncc_research_list.pdf), Japan. T. Sato is an awardee of a research resident fellowship from the Foundation for Promotion of Cancer Research in Japan (http://www.fpcr.or.jp/index.html). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: earai@ncc.go.jp

Introduction

Lung adenocarcinoma (LADC) is increasingly recognized as a clinicopathologically and molecularly heterogeneous disease: frequent mutations of the *EGFR, KRAS, BRAF, TP53, ERBB2, PIK3CA* and *MET* genes and *EML4-ALK* fusions have been reported in LADCs [1,2]. In addition, recent whole-exome sequencing has revealed frequent mutation of the *CSMD3* gene [3]. However, the molecular background responsible for the clinicopathological diversity of LADCs is not yet fully understood.

As well as genetic abnormalities, epigenetic changes in human cancers have also been described [4–7]. In LADCs, silencing of the *RASSF1A, CDKN2A, RARβ, MGMT, APC, DAPK, FHIT* and *CDH13* genes due to DNA hypermethylation around their promoter regions has been frequently observed [8]. In addition, DNA methylation alterations are known to occur even at the early and precancerous stages of carcinogenesis in various organs [5–7,9]. For example, we have reported that DNA hypermethylation at the D17S5 locus, where the *HIC-1* tumor suppressor gene has been identified, is evident even in non-cancerous lung tissue obtained from patients with non-small cell lung cancers, and is correlated with smoking history [10]. Other researchers have also reported DNA hypermethylation of specific tumor-related genes at precancerous stages associated with cigarette smoking [8,11]. However, it has been unclear whether DNA methylation status is simply altered at precancerous stages or whether DNA methylation alterations at these stages actually result in gene expression alterations in established LADCs. Moreover, in organs other than

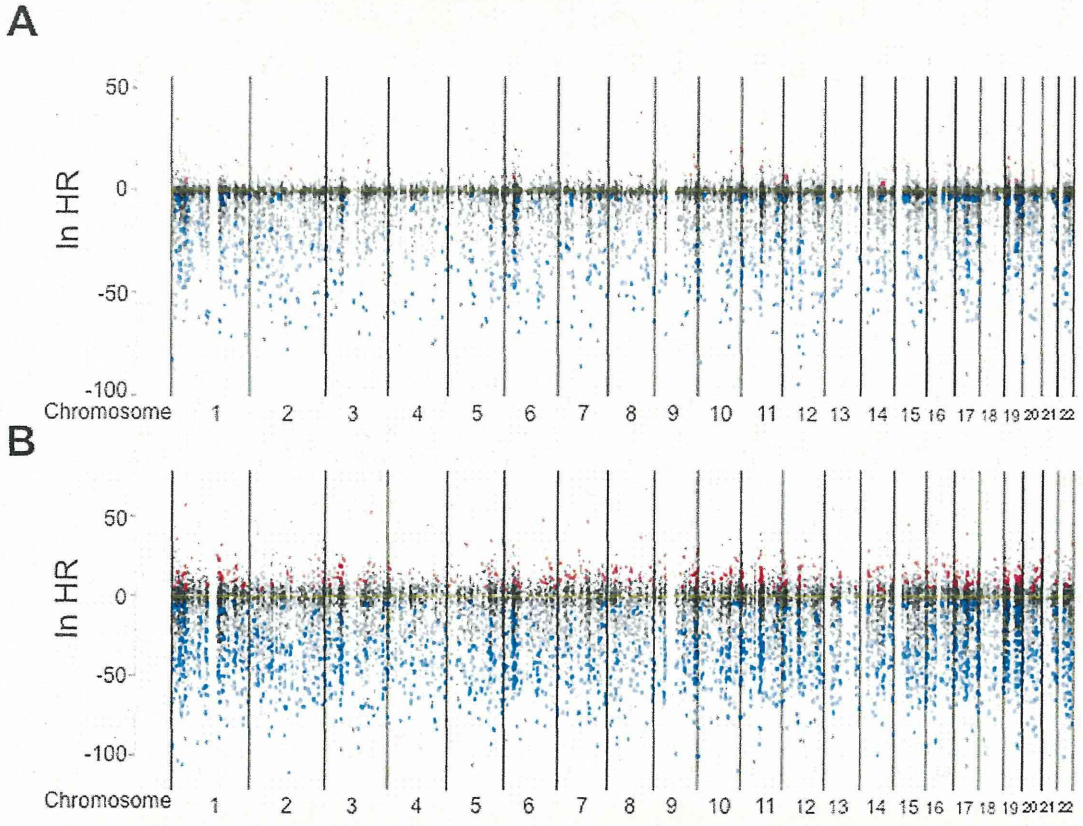

Figure 1. Hazard ratio (HR) obtained from the Cox regression model. Correlation between DNA methylation status (average β-values) and recurrence was examined in 145 samples of tumorous tissue (**A**) and 145 samples of the corresponding non-cancerous lung tissue (**B**) obtained from patients with lung adenocarcinomas who had undergone complete resection and had not received any adjuvant therapy after surgery. All of the examined 26,455 probes of the Infinium array are shown along the chromosomes. Red color means that higher β-values of the probes were observed in recurrence-positive patients than in recurrence-negative patients ($P<0.001$). Blue color means that lower β-values of the probes were observed in recurrence-positive patients than in recurrence-negative patients ($P<0.001$).
doi:10.1371/journal.pone.0059444.g001

the lung, it has been suggested that DNA methylation profiles at precancerous stages may determine tumor aggressiveness and outcome [12–14]. However, the clinicopathological impact of DNA methylation alterations at precancerous stages during lung carcinogenesis has not been clarified.

Recently, genome-wide DNA methylation analysis using the single-CpG resolution Infinium array has made it possible to interrogate 27,000 highly informative CpG sites, i.e. an average of two CpG sites within the proximal promoter regions of the transcription start sites of each of 14,475 consensus coding sequences in the National Center for Biotechnology Information Database, especially 3 to 20 CpG sites for more than 200 cancer-related and imprinted genes [15]. Although a few studies of lung cancers employing the Infinium assay have been reported [16,17], they did not focus on precancerous stages. In order to clarify the significance of DNA methylation alterations at precancerous stages of lung carcinogenesis, we performed the Infinium assay in association with mRNA expression and clinicopathological analyses of 36 samples of normal lung tissue (C) obtained from patients without any primary lung tumors, 145 samples of non-cancerous lung tissue (N) from patients with LADCs, and 145 corresponding samples of tissue from the tumors (T) themselves. Although the molecular classification of LADCs based on the results of the Infinium assay will be published elsewhere, we focused on specific genes methylated at precancerous stages in the present study.

Materials and Methods

Patients and Tissue Samples

The 145 paired samples of N and the corresponding T were obtained from patients with primary LADCs who underwent lung resection at the National Cancer Center Hospital, Japan, between December 1997 and March 2008. These patients had undergone complete resection and had not received any preoperative treatment or adjuvant therapy after surgery. Eighty-one patients were males and 64 were females with a median age of 61 years (range, 30–81 yr). Histological diagnosis and grading were based on the 2004 World Health Organization classification [18]. Recurrence was diagnosed by clinicians on the basis of physical examination and imaging modalities such as computed tomography, magnetic resonance imaging, scintigraphy or positron-emission tomography, and sometimes confirmed histopathologically by biopsy.

For comparison, 36 C samples were obtained from materials that had been surgically resected from patients without any primary lung tumor. Twenty-two of these patients were males and 14 were females, with a median age of 63 years (range, 27–83 yr).

Table 1. The 28 probes for which the average β-value in N samples (β_N) was higher in recurrence-positive patients than in recurrence-negative patients, for which the average β-value in C samples (β_C) was less than 0.2, and for which the average β-value in T samples minus that in corresponding N samples ($\Delta\beta_{T-N}$) was more than 0.1.

Target ID[a]	Chromosome	Position[b]	Gene symbol	P[c]	Adjusted P[d]
cg00516481	21	44,073,202	*PDE9A*	9.996×10^{-4}	1.193×10^{-2}
cg01295203	8	70,984,199	*PRDM14*	8.156×10^{-5}	3.236×10^{-3}
cg02008154	7	35,293,537	*TBX20*	7.188×10^{-4}	9.917×10^{-3}
cg02909790	6	26,271,587	*HIST1H3G*	2.896×10^{-6}	5.787×10^{-4}
cg03538436	12	117,799,370	*NOS1*	7.227×10^{-4}	9.944×10^{-3}
cg03963198	5	1,882,871	*IRX4*	5.133×10^{-4}	8.186×10^{-3}
cg06005396	19	590,541	*HCN2*	9.485×10^{-4}	1.158×10^{-2}
cg06269753	8	72,755,871	*MSC*	8.627×10^{-7}	3.426×10^{-4}
cg07651242	7	45,614,720	*ADCY1*	7.403×10^{-4}	1.010×10^{-2}
cg11612345	6	168,842,491	*SMOC2*	4.264×10^{-6}	6.753×10^{-4}
cg12087643	10	118,033,370	*GFRA1*	3.144×10^{-9}	2.489×10^{-5}
cg12265829	14	24,804,022	*ADCY4*	7.265×10^{-4}	9.984×10^{-3}
cg13262687	4	147,559,579	*POU4F2*	6.246×10^{-5}	2.795×10^{-3}
cg13449778	1	179,712,298	*FAM163A*	8.195×10^{-7}	3.426×10^{-4}
cg13878010	3	123,167,276	*ADCY5*	2.339×10^{-6}	5.220×10^{-4}
cg16254309	7	145,814,152	*CNTNAP2*	2.917×10^{-4}	6.240×10^{-3}
cg16387606	1	149,804,293	*HIST2H4A*	7.441×10^{-4}	1.011×10^{-2}
cg16604516	3	13,590,419	*FBLN2*	7.181×10^{-4}	9.916×10^{-3}
cg16652259	2	172,949,501	*DLX1*	4.175×10^{-4}	7.386×10^{-3}
cg17191178	3	157,824,217	*SHOX2*	8.872×10^{-5}	3.353×10^{-3}
cg18454685	17	48,639,239	*CACNA1G*	4.326×10^{-4}	7.487×10^{-3}
cg20286200	6	133,562,267	*EYA4*	1.541×10^{-9}	2.038×10^{-5}
cg21087137	12	75,728,469	*GLIPR1L1*	1.965×10^{-6}	4.821×10^{-4}
cg22461835	8	26,723,365	*ADRA1A*	7.297×10^{-4}	9.998×10^{-3}
cg23418591	20	57,090,317	*LOC149773*	4.075×10^{-4}	7.287×10^{-3}
cg25302419	5	11,904,015	*CTNND2*	5.149×10^{-4}	8.194×10^{-3}
cg25764191	10	105,037,215	*INA*	3.834×10^{-4}	7.026×10^{-3}
cg27626299	7	27,282,431	*EVX1*	2.164×10^{-4}	5.423×10^{-3}

[a]Probe ID for the Infinium HumanMethylation27 Bead Array (Illumina).
[b]National Center for Biotechnology Information database (Genome Build 37).
[c]Non-adjusted *P*-values and.
[d]Benjamini-Hochberg-adjusted *P*-values for the Cox regression model used for evaluation of correlation with recurrence.
doi:10.1371/journal.pone.0059444.t001

Thirty-five had undergone lung resection for metastatic lesions of primary cancers of the colon, rectum, kidney, urinary bladder, thyroid, breast, pancreas, ampulla of Vater and salivary gland, osteosarcoma, synovial sarcoma, leiomyosarcoma, rhabdmyosarcoma, liposarcoma, dermatofibrosarcoma, and myxofibrosarcoma. The remaining one patient had undergone chest wall resection for lipoma with removal of adjacent lung tissue. Histological observation confirmed that all of the C samples showed no remarkable histological abnormality and did not contain any contaminating tumor cells that had metastasized from organs other than the lung.

Tissue specimens were provided by the National Cancer Center Biobank, Japan. This study was approved by the Ethics Committee of the National Cancer Center, Tokyo, Japan, and was performed in accordance with the Declaration of Helsinki 1975. All patients included in this study provided written informed consent.

Cell Lines

The characteristics of the four lung cancer cell lines used in this study are summarized in Table S1.

Infinium Assay

Genomic DNA was extracted using a QIAamp DNA Mini kit (Qiagen, Valencia, CA, USA) and phenol-chloroform extraction followed by dialysis [19] from all tissue samples and cell lines, respectively. Five-hundred-nanogram aliquots of DNA were subjected to bisulfite conversion using an EZ DNA Methylation-Gold Kit (Zymo Research, Irvine, CA, USA). DNA methylation status at 27,578 CpG loci was examined at single-CpG resolution using the Infinium HumanMethylation27 Bead Array (Illumina, San Diego, CA, USA). After hybridization, the specifically hybridized DNA was fluorescence-labeled by a single-base extension reaction and detected using a BeadScan reader (Illumina) in accordance with the manufacturer's protocols. The data were then assembled using GenomeStudio methylation

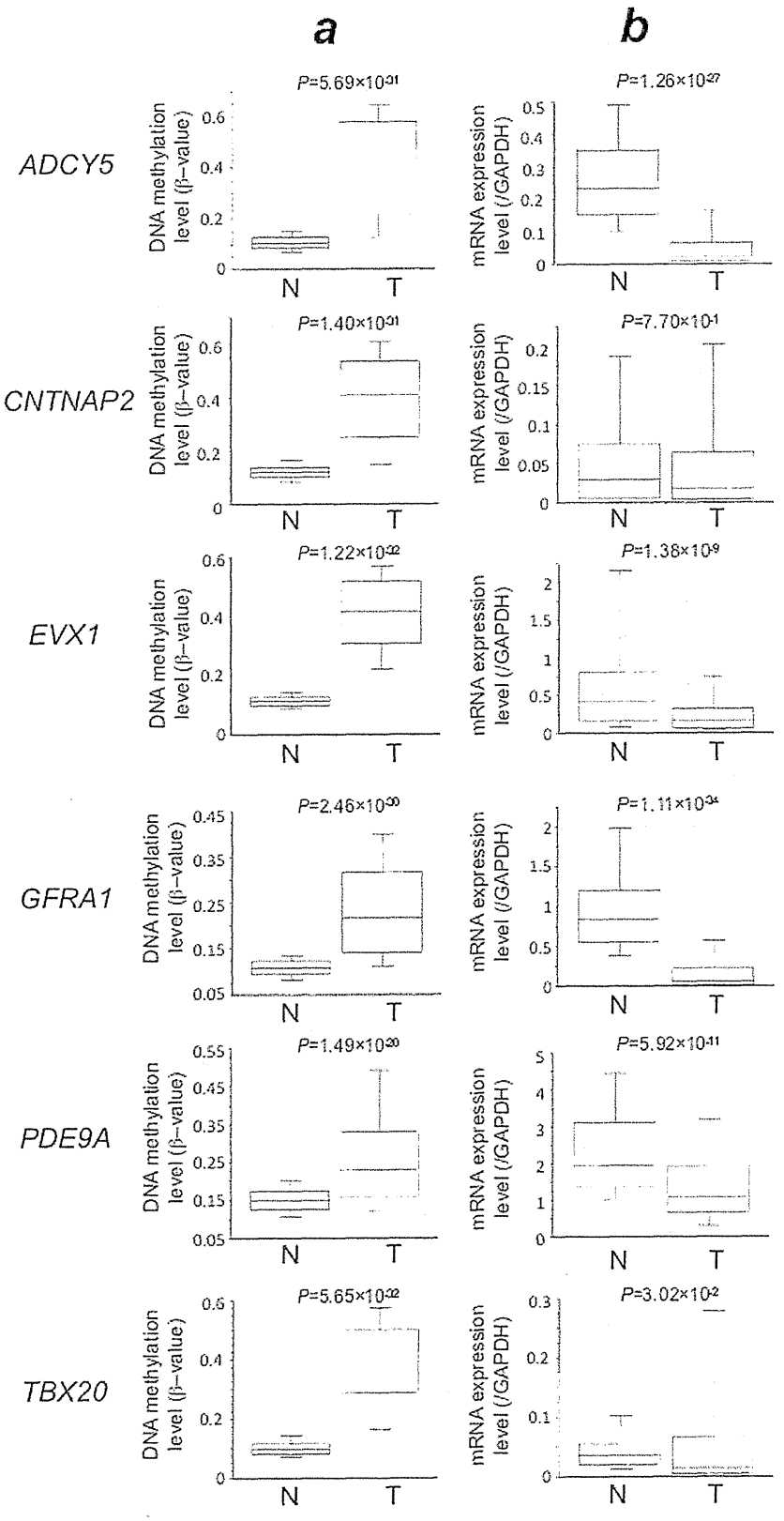
a
b
ADCY5
CNTNAP2
EVX1
GFRA1
PDE9A
TBX20
DNA methylation level (β-value)
mRNA expression level (/GAPDH)
N
T
P=5.69×10-31
P=1.26×10-27
P=1.40×10-31
P=7.70×10-1
P=1.22×10-32
P=1.38×10-9
P=2.46×10-30
P=1.11×10-34
P=1.49×10-30
P=5.92×10-11
P=5.65×10-32
P=3.02×10-2

Figure 2. Correlation between DNA methylation levels and mRNA expression levels. DNA methylation levels (average β-values) (**A**) and mRNA expression levels (**B**) for the *ADCY5, CNTNAP, EVX1, GFRA1, PDE9A* and *TBX20* genes in samples of non-cancerous lung tissue (N) from patients with lung adenocarcinomas and samples of the corresponding tumorous tissue (T) were determined by Infinium assay and quantitative real-time reverse transcription-PCR analysis, respectively. DNA methylation levels for all six genes were significantly higher in T samples than in N samples, and levels of expression of mRNAs for the *ADCY5, EVX1, GFRA1, PDE9A* and *TBX20* genes were significantly lower in T samples than in N samples, although the reduction in the expression of the *CNTNAP2* gene did not reach statistical significance. These results suggested that DNA hypermethylation of the *ADCY5, EVX1, GFRA1, PDE9A* and *TBX20* genes may result in reduced mRNA expression in tissue samples from the same cohort.
doi:10.1371/journal.pone.0059444.g002

software (Illumina). At each CpG site, the ratio of the fluorescence signal was measured using a methylated probe relative to the sum of the methylated and unmethylated probes, i.e. the so-called β-value, which ranges from 0.00 to 1.00, reflecting the methylation level of an individual CpG site.

Quantitative Real-time Reverse Transcription (RT)-PCR Analysis

Total RNA was extracted from 132 N and 151 T samples for which additional tissue specimens were available and cell lines using TRIzol reagent (Life Technologies, Carlsbad, CA, USA) in accordance with the manufacturer's instructions. cDNA was synthesized from total RNA with random primers using SuperScript III Reverse Transcriptase (Life Technologies) and pre-amplified using TaqMan PreAmp Master Mix (Life Technologies).

To evaluate mRNA expression levels, fluorescence-labeled locked nucleic acid hydrolysis probes were selected from the Universal Probe Library collection (Roche Applied Science, Mannheim, Germany) and specific PCR primers yielding intron-spanning amplicons were designed using ProbeFinder assay design software (https://www.roche-applied-science.com/sis/rtpcr/upl/index.jsp?id=UP030000). The probe ID and primer sequences are summarized in Table S2. Quantitative real-time PCR was performed using TaqMan Universal Master Mix II (Life Technologies) and the relative standard curve method in the BioMark HD System (Fluidigm, South San Francisco, CA, USA). Ct values were normalized to that of *GAPDH* in the same sample. All assays were performed in triplicate.

5-aza-2′-deoxycytidine (5-aza-dC) Treatment

A549, PC9, VMRC-LCD and EBC-1 cells were seeded at a density of 9×10^5 cells per 15-cm dish on day 0 and then allowed to attach for a 24-h period. Then, 5-aza-dC (Sigma-Aldrich, St. Louis, MO, USA) was added to a final concentration of 1 μM. Cells were passaged at a subculture ratio of 1:2 on day 3. At 24 h after replating, 5-aza-dC was added again to the same final concentration. Since toxicity had been obvious during preliminary experiments, the final concentration of 5-aza-dC was reduced to 0.1 μM for EBC-1 cells. Genomic DNA and total RNA were extracted from all cells on days 3 and 6.

Statistics

In the Infinium assay, all CpG sites on chromosomes X and Y were excluded, to avoid any gender-specific methylation bias. The call proportions (*P*-values for detection of signals above the background <0.01) for 31 probes (shown in Table S3) in all of the tissue samples examined were less than 90%. Since such a low proportion may be attributable to polymorphism at the probe CpG sites, these 31 probes were excluded from the present assay, leaving a final total of 26,455 autosomal CpG sites.

Infinium probes showing ordered differences from 36 C to 145 N, and then to the 145 T samples themselves, were examined by the cumulative logit model adjusted by sex, age and experimental batch ($P<1\times10^{-14}$). Correlations between β-values in N and T samples and recurrence were assessed by the Cox regression model adjusted by sex, age and experimental batch ($P<0.001$). Benjamini-Hochberg correction was performed to adjust for multiple testing. Differences of β-values and mRNA expression levels between N and T samples were examined by Mann-Whitney U test. Correlations between mRNA expression levels and clinicopathological parameters were assayed by analysis of variance between groups (ANOVA) and Welch's T-test: after adjusted Bonferroni correction to adjust for multiple testing, corrected *P* values of <0.05 were considered to be significant. All statistical analyses were performed using programming language R.

Results

DNA Methylation Profiles Associated with Recurrence are Established at Precancerous Stages

DNA methylation levels of CpG sites of the *CNTNAP2*, *EVX1*, *GFRA1*, *PDE9A* and *TBX20* genes based on the Infinium assay were clearly verified using the quantitative pyrosequencing method (Figure S1), indicating the reliability of the Infinium assay. The cumulative logit model ($P<1\times10^{-14}$) revealed ordered progression of DNA methylation alterations from C to N, and then to T samples, on 3,270 probes; DNA methylation alterations occurred even in N samples compared to C samples, and such DNA methylation alterations were inherited by, or strengthened in, T samples, indicating that Ns were at precancerous stages with DNA methylation alterations. Among the 3,270 probes, the number showing average β-values in T samples minus average β-values in C samples ($\Delta\beta_{T\text{-}C}$) of >0.1 and <−0.1 were 1,209 and 1,056, respectively. Thus, when we defined differentially methylated probes as probes showing a $\Delta\beta_{T\text{-}C}$ value of >0.1 or <−0.1, the false positivity rate by the cumulative logit model was 4.3%.

Correlations between DNA methylation status and recurrence were examined using the Cox regression model ($P<0.001$) for 145 patients. In T samples, DNA methylation status on 944 probes for the 916 genes was significantly correlated with recurrence: on 87 probes (red dots in Figure 1A), higher β-values were observed in recurrence-positive patients than in recurrence-negative patients, whereas lower β-values on 857 probes (blue dots in Figure 1A) were observed in recurrence-positive patients. Surprisingly, even in N samples, the DNA methylation status on 2,215 probes for the 2,083 genes was significantly correlated with recurrence: on 425 probes (red dots in Figure 1B), higher β-values were observed in recurrence-positive patients than in recurrence-negative patients, whereas lower β-values on 1,790 probes (blue dots in Figure 1B) were observed in recurrence-positive patients.

In order to identify recurrence-related genes that are normally unmethylated and for which DNA hypermethylation at precancerous stages is strengthened in the established LADCs, among the 425 probes (red dots in Figure 1B), we initially focused on 28 probes for which the average β-values in C samples (β_C) were less than 0.2 and the average β-values in T samples minus that in the corresponding N samples ($\Delta\beta_{T\text{-}N}$) were more than 0.1 (Table 1).

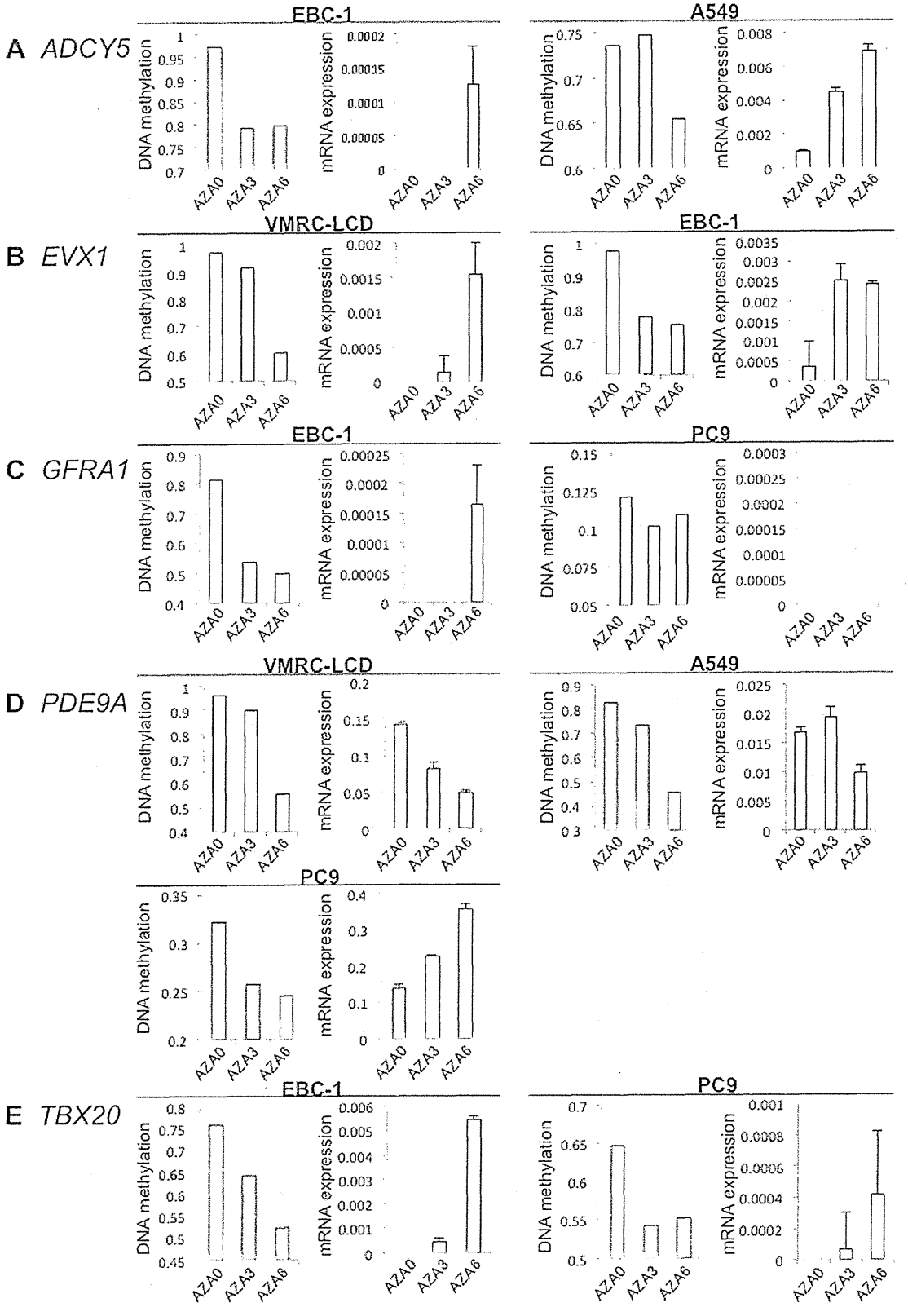
A ADCY5
EBC-1
A549
B EVX1
VMRC-LCD
EBC-1
C GFRA1
EBC-1
PC9
D PDE9A
VMRC-LCD
A549
PC9
E TBX20
EBC-1
PC9
DNA methylation
mRNA expression
AZA0
AZA3
AZA6

Figure 3. DNA methylation levels and mRNA expression levels after 5-aza-2′-deoxycytidine (5-aza-dC) treatment. DNA methylation levels (β-values) and mRNA expression levels for the *ADCY5* (**A**), *EVX1* (**B**), *GFRA1* (**C**), *PDE9A* (**D**) and *TBX20* (**E**) genes were determined by Infinium assay and quantitative real-time reverse transcription-PCR analysis, respectively. The error bars represent the standard deviation for triplicate quantitative real-time RT-PCR analyses. DNA methylation levels and mRNA expression levels on days 3 (AZA3) and 6 (AZA6) were compared with those of untreated cells (AZA0). After 5-aza-dC treatment, reduction of DNA methylation levels and restoration of the mRNA expression levels of *ADCY5* (**A**), *EVX1* (**B**) and *TBX20* (**E**) were observed in both of the cell lines used. In panel **C**, since reduction of the DNA methylation level was not induced by 5-aza-dC in PC9 cells, restoration of mRNA expression did not occur in these cells. Panel **D** shows reduction of the DNA methylation level and restoration of the mRNA expression level in PC9 cells.
doi:10.1371/journal.pone.0059444.g003

Silencing of Recurrence-related Genes due to DNA Hypermethylation

Among the 28 genes listed in Table 1, we further focused on 6 genes (*ADCY5* [20,21], *CNTNAP2* [22,23], *EVX1* [24], *GFRA1* [25], *PDE9A* [26,27] and *TBX20* [28]) for which implications in transcription regulation, apoptosis or cell adhesion had been reported. Quantitative real-time RT-PCR analysis of these 6 genes was performed in 132 N and 151 T samples for which total RNA was available. mRNA expression levels for the *ADCY5, EVX1, GFRA1, PDE9A* and *TBX20* genes in T samples were significantly lower than those in N samples, although the reduced expression of the *CNTNAP2* gene did not reach statistical significance (Figure 2B). The DNA methylation statuses (β-values in N and T samples) of the 6 genes are also shown in Figure 2A; the data suggested that DNA hypermethylation of these genes might result in reduction of mRNA expression in tissue samples from the same cohort.

DNA methylation levels of the *ADCY5, EVX1, GFRA1, PDE9A* and *TBX20* genes in lung cancer cell lines A549, PC9, VMRC-LCD and EBC-1 are shown in Figure S2. To examine the effects of the DNA methylation inhibitor, the top two cell lines showing the highest DNA methylation levels (β-values) were selected for each gene. In fact, mRNA expression levels determined by quantitative real-time RT-PCR analysis of the genes (with the exception of PDE9A) were extremely low in the cell lines selected. 5-aza-dC treatment induced marked reduction of DNA methylation levels and restored the mRNA expression levels of *ADCY5, EVX1, GFRA1* and *TBX20* (Figure 3). With regard to *GFRA1*, since reduction of the DNA methylation level was not induced by 5-aza-dC, restoration of mRNA expression did not occur in PC9 cells. Taken together with Figures 2 and 3, the data suggested that the examined genes were silenced due to DNA hypermethylation in the lung cancers. With regard to *PDE9A*, for which the mRNA expression levels were high even in the top two cell lines showing the highest levels of DNA methylation, 5-aza-dC treatment did not further increase the mRNA expression level in the two cell lines. The PC9 cells were then additionally treated with 5-aza-dC, and restoration of *PDE9A* mRNA expression due to DNA demethylation was confirmed in the cells (Figure 3).

Clinicopathological Impact of Reduced Expression of mRNA for Recurrence-related Genes

Reduced expression of mRNA for *ADCY5, EVX1, GFRA1* and *PDE9A* in T samples was correlated with clinicopathological parameters reflecting tumor aggressiveness, such as a larger tumor diameter, higher histological grade, blood vessel invasion, pleural invasion and tumor anthracosis (Table 2), although their mRNA expression levels were not predictors of recurrence that were independent of known parameters such as pathological-TNM stage and lymph node metastasis (Table S4). With regard to the correlation with histological subtype, levels of mRNA expression for *GFRA1* and *PDE9A* were significantly higher in lepidic-type LADCs showing a less invasive growth pattern than in other histological subtypes.

Discussion

The 'field cancerization' phenomenon in the lung has become evident, being especially associated with cigarette smoking [31]. We and other groups have reported DNA methylation of specific genes or chromosomal loci in non-cancerous lung tissue obtained from lung cancer patients, or in lung tissue from cancer-free smokers [8–10]. These previous data drew our attention to DNA methylation alterations at precancerous stages of LADC. However, the impact of DNA methylation alterations at precancerous stages on the expression of specific gene and clinicopathological parameters of established cancers has remained unclear. Moreover, previous examinations focusing on precancerous stages in the lung have not involved a genome-wide approach. Although Selamat et al. and Lockwood et al. have reported Infinium assay results for 59 and 43 lung cancer samples, respectively [16,17], they did not focus on precancerous stages.

Here we have reported the results of the Infinium assay for 326 lung tissue samples including 145 N samples. Our cumulative logit model analysis revealed stepwise progression of DNA methylation alterations from C to N, and then T samples on 3,270 probes. Genome-wide analysis at single-CpG resolution confirmed that DNA methylation alterations actually occurred even at precancerous stages, although the possibility that such alterations observed in N samples had been influenced by differences in tissue composition between C and N samples cannot be completely excluded. Moreover, it was clearly shown that such DNA methylation alterations had clinicopathological impact, since many probes in N samples were significantly correlated with recurrence after establishment of LADCs (Figure 1B). DNA methylation profiles determining outcome are already established at precancerous stages. The finding that the number of probes showing DNA methylation alterations significantly associated with recurrence in N samples was larger than that in T samples may have been due to the fact that passenger DNA methylation alterations occurring during progression from the precancerous stages to established cancers may have masked any clinicopathologically significant DNA methylation profiles in T samples.

Next, we focused on 28 recurrence-related genes that are normally unmethylated and for which DNA hypermethylation in N samples was strengthened in T samples (Table 1). Among these 28 genes, we further focused on *ADCY5, CNTNAP2, EVX1, GFRA1, PDE9A* and *TBX20*, based on their previously reported implications in transcription regulation, apoptosis or cell adhesion. (a) The data in *adcy5*-knockout mice indicated that ADCY5 promotes apoptosis in cardiomyocytes [20,21]. Although large-scale screening studies of leukemias have identified *ADCY5* as one of the genes that are methylated in leukemic cells [32,33], the clinicopathological impact of DNA methylation of the *ADCY5* gene has not yet been clarified in human malignancies. (b) CNTNAP2, a glial adhesion molecule, binds extracellularly to contactin 2, an immunoglobulin superfamily neural recognition protein [22]. *CNTNAP2* is known to be a tumor-suppressor gene for gliomas, and is disrupted by chromosomal translocations and gene mutations [23]. Although a large-scale screening study of

Table 2. Correlation between mRNA expression levels of recurrence-related genes and clinicopathological factors.

Clinicopathological parameters	Number of tumors	*ADCY5*		*EVX1*		*GFRA1*		*PDE9A*		*TBX20*	
		Expression[a]	P^b	Expression[a]	P^b	Expression[a]	P^b	Expression[a]	P^b	Expression[a]	P^b
Tumor diameter											
<2.5 cm	43	−4.95±2.10	1.63×10^{-1c}	−2.80±1.82	4.24×10^{-1c}	−3.27±2.67	1.60×10^{-3c}	0.50±1.39	3.40×10^{-2c}	−5.73±2.68	7.65×10^{-1c}
≥2.5 cm, <4 cm	61	−5.25±2.39		−2.36±2.16		−4.02±2.86		0.28±1.76		−5.76±2.78	
≥4 cm	47	−6.14±2.76		−3.04±2.02		−5.74±3.28		−0.50±1.59		−6.10±2.84	
Histological subtype[d]											
Lepidic	12	−4.19±1.61	2.10×10^{-1c}	−2.73±2.14	1.79×10^{-1c}	−1.64±2.35	1.71×10^{-4c}	0.93±1.07	1.73×10^{-3c}	−6.24±1.79	1.28×10^{-1c}
Acinar	24	−6.03±2.01		−2.91±2.09		−5.41±2.72		−0.20±1.36		−5.77±2.57	
Papillary	69	−5.30±2.35		−2.29±1.99		−3.90±2.65		0.23±1.63		−5.28±2.80	
Micropapillary	15	−4.69±2.19		−2.35±2.10		−3.47±2.87		1.20±1.27		−6.89±1.79	
Solid	28	−6.34±3.24		−3.62±1.88		−6.27±3.62		−0.91±1.81		−6.37±3.11	
Invasive mucinous	3	−4.64±1.46		−3.19±1.24		−2.95±1.43		0.16±1.68		−8.21±0.14	
Histological grades											
G1	56	−4.60±1.80	1.28×10^{-3c}	−2.68±1.90	4.57×10^{-2c}	−2.62±1.93	4.80×10^{-10}	0.55±1.41	1.32×10^{-4}	−5.48±2.69	9.50×10^{-2}
G2	65	−5.58±2.33		−2.32±2.10		−4.59±2.71		0.25±1.57		−5.74±2.67	
G3	30	−6.73±3.20		−3.54±1.92		−7.01±3.60		−1.06±1.77		−6.81±2.94	
Lymphatic invasion											
Negative	52	−5.31±2.43	1.00[e]	−2.76±1.71	1.00[e]	−3.95±3.38	1.00[e]	0.06±1.64	1.00[e]	−5.86±2.42	9.82×10^{-1e}
Positive	99	−5.52±2.50		−2.66±2.19		−4.55±2.92		0.12±1.67		−5.85±2.93	
Blood vessel invasion											
Negative	43	−4.89±2.40	2.38×10^{-1e}	−2.77±1.84	7.68×10^{-1e}	−3.13±3.45	2.81×10^{-2e}	0.63±1.68	6.39×10^{-2e}	−5.71±2.60	1.00[e]
Positive	108	−5.66±2.47		−2.67±2.11		−4.82±2.81		−0.11±1.60		−5.92±2.83	
Pleural invasion											
Negative	78	−5.16±2.42	4.45×10^{-1c}	−2.77±2.15	4.60×10^{-1c}	−3.65±2.83	7.28×10^{-3c}	0.28±1.61	5.82×10^{-1c}	−5.95±2.51	5.38×10^{-1c}
Invasion to the visceral pleura beyond the elastic fiber	33	−5.88±2.68		−2.62±2.07		−4.99±3.54		0.03±1.93		−6.41±2.96	
Invasion to the surface of the visceral pleura	22	−4.96±2.07		−2.16±1.62		−4.01±2.66		0.18±1.29		−4.98±2.82	
Invasion to the parietal pleura	18	−6.47±2.49		−3.15±1.90		−6.56±2.71		−0.65±1.58		−5.52±3.22	
Tumor anthracosis [f]											
Negative	68	−4.97±2.57	8.27×10^{-2e}	−2.37±2.22	2.04×10^{-1e}	−3.32±2.80	1.08×10^{-3e}	0.70±1.46	1.66×10^{-4e}	−5.51±2.78	1.99×10^{-1e}
Positive	82	−5.87±2.30		−2.91±1.79		−5.10±3.02		−0.39±1.66		−6.09±2.69	

[a]Average of log2-transformed mRNA expression levels/GAPDH ± standard deviation. [b]Adjusted *P*-values using adjusted Bonferroni correction. [c]Analysis of variance between groups. [d]Predominant histological subtypes of LADCs were diagnosed according to the classification devised by the International Association for the Study of Lung Cancer, the American Thoracic Society and the European Respiratory Society [29]. [e]Welch's T-test. [f]Coal dust is accumulated in active fibroblast proliferation foci, which is associated with poorer prognosis of lung adenocarcinoma patients and reflects an active cancer-stromal interaction [30]. *P* values of <0.05 are underlined.

doi:10.1371/journal.pone.0059444.t002

pancreatic cancers has identified *CNTNAP2* as one of the genes that are methylated in cancer cells [34], the clinicopathological impact of DNA methylation of the *CNTNAP2* gene has not yet been elucidated in human malignancies. (c) *EVX1*, encoding a homeobox protein, functions as a potent repressor of gene transcription and plays an important role during mouse embryogenesis [24]. Although Truong et al. have recently suggested that DNA hypermethylation of the *EVX1* gene may be a predictor of recurrence of prostatic cancers [35], the implications of *EVX1* in human cancers other than prostatic cancer have been unclear. (d) GFRA1 is a receptor for glial cell-derived neurotrophic factor (GDNF) and enriched in the pre- and post-synaptic compartments. GDNF triggers *trans*-homophilic binding between GFRA1 molecules, resulting in adhesion between *GFRA1*-expressing cells [25]. Overexpression of *GFRA1* has been reported in chemotherapy-sensitive oligodendrogliomas [36]. Although Salamat et al. described *GFRA1* as one of the genes differentially methylated between cancerous and non-cancerous tissue obtained from lung cancer patients subclustered on the basis of DNA methylation profiles [16], the clinicopathological impact of DNA methylation of the *GFRA1* gene was not examined in LADCs. (e) *PDE9A* encodes cGMP-specific phosphodiesterase [26]. PDE9A inhibitor has been reported to induce apoptosis of breast cancer cell lines through caspase 3 activation [27]. On the other hand, breakpoints within the *PDE9A* gene have been frequently observed in B-cell precursor acute lymphoblastic leukemia [36]. However, the implications of DNA methylation in the regulation of *PDE9A* have never been reported in human malignancies or other diseases. (f) TBX20, a member of the T-box transcription factors, is a critical regulator of heart development, and mutations of the human *TBX20* gene result in cardiac malformations [28]. However, the implications of DNA methylation in *TBX20* regulation and of *TBX20* dysfunction in human malignancies have never been reported.

Among these 6 genes, DNA hypermethylation of the *ADCY5, EVX1, GFRA1, PDE9A* and *TBX20* genes was associated with reduced mRNA expression in tissue samples of the same cohort. 5-aza-dC treatment of human lung cancer cell lines restored the expression of the 5 genes, indicating that genes showing DNA hypermethylation at precancerous stages are actually silenced due to DNA hypermethylation during lung carcinogenesis. With regard to *PDE9A*, the level of mRNA expression was not necessarily low and was not always restored by 5-aza-dC treatment in any of the cell lines examined, indicating that DNA methylation is not the only mechanism responsible for *PDE9A* regulation in lung cancers. In addition, there are gaps between the timing of the reduction of DNA methylation and recovery of mRNA expression in *ADCY5* in EBC-1 cells and A549 cells and *GFRA1* in EBC-1 cells, again indicating the possibility that there are alternative mechanisms regulating the mRNA expression levels of these genes other than DNA methylation. Reduced expression of the mRNAs for the *ADCY5, EVX1, GFRA1* and *PDE9A* genes was significantly correlated with clinicopathological parameters (Table 2), indicating that DNA methylation alterations, even from the precancerous stage, ultimately determine the tumor phenotype through gene silencing.

Unlike the present study, which was conducted to clarify the significance of DNA methylation alterations at precancerous stages, many previous studies attempted to establish prognostic factors based on DNA methylation status using candidate gene approaches [37–41]. Microarray studies are generally considered to be useful for establishing prognostic biomarkers. In lung cancers, array-based DNA methylation screening [42] has been performed for prognostication. Next-generation sequencing associated with bioinformatics analysis [43] has also been reported for lung cancers. Such previous studies identified the *RASSF1A* [37], *PITX2* [38], *SHOX2* [38], *TFPI-2* [39], *FHIT* [40], *p16* [41], *CDH13* [41], and *APC* [41] genes as predictors of recurrence of lung cancers. The genes for which DNA methylation status at precancerous stages (β_N values) was associated with recurrence in the present study are different from the prognostic biomarkers selected in previous studies based on the DNA methylation status of tumorous tissues themselves.

In summary, DNA methylation status is not simply altered at precancerous stages, and is significantly correlated with recurrence after establishment of LADCs. DNA methylation alterations at precancerous stages are strengthened in comparison to normal lung tissues during progression to established LADC. DNA methylation profiles at precancerous stages may determine tumor aggressiveness through alterations in the expression of mRNAs for specific genes.

Supporting Information

Figure S1 Verification of the results of the Infinium assay using pyrosequencing methods. To overcome the PCR bias in pyrosequencing, the PCR conditions were optimized for each primer set, as described previously (Nagashio R, et al. Int J Cancer 129:1170, 2011). Pyrosequencing was performed for the *CNTNAP2, EVX1, GFRA1, PDE9A* and *TBX20* genes using 9 representative T samples and 9 corresponding N samples. β-values obtained from the Infinium assay were strongly correlated with DNA methylation levels obtained by pyrosequencing in all 5 genes, indicating that the results of the Infinium assay were successfully verified.
(TIF)

Figure S2 DNA methylation levels for the *ADCY5, EVX1, GFRA1, PDE9A* and *TBX20* genes in lung cancer cell lines. DNA methylation levels (β-values) for the *ADCY5, EVX1, GFRA1, PDE9A* and *TBX20* genes in all 4 of the lung cancer cell lines were examined by Infinium assay. To examine the effects of the DNA methylation inhibitor, 5-aza-2′-deoxycytidine, the top two cell lines showing the highest DNA methylation levels were selected for each gene: EBC-1 and A549 cells for the *ADCY5* gene, VMRC-LCD and EBC-1 cells for the *EVX1* gene, EBC-1 and PC9 cells for the *GFRA1* gene, VMRC-LCD and A549 cells for the *PDE9A* gene, and EBC-1 and PC9 cells for the *TBX20* gene.
(TIF)

Table S1 Characteristics of the lung cancer cell lines.
(PDF)

Table S2 Probe ID and primer sequences for quantitative real-time reverse transcription-PCR.
(PDF)

Table S3 The probes for which call proportions in all examined tissue samples were less than 90%.
(PDF)

Table S4 Multivariate analysis of clinicopathological parameters and mRNA expression levels of selected genes associated with recurrence in patients with lung adenocarcinomas.
(PDF)

Acknowledgments

The authors thank Y. Shimada for providing technical assistance.