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

Takashi Sato1,2, Eri Arai1,*, Takashi Kohno3, Koji Tsuta4, Shun-ichi Watanabe5, Kenzo Soejima2, Tomoko Betsuyaku2, Yae Kanai1

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 to 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 \( \beta \)-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 (\( \Delta \beta_{\text{N-T}} > 0.1 \)). 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 ADCYS, EVK1, GFR1, 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 pre-cancerous stages determine tumor aggressiveness and outcome through silencing of specific genes.


Introduction
Lung adenocarcinoma (LADC) is increasingly recognized as a clinicopathologically and molecularly heterogeneous disease: frequent mutations of the EGFR, KRAS, BRAF, TP53, ERBB2, PKR3CA and MET genes and EMLA-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, RARB, 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...
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 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).
 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, rhabdomyosarcoma, 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

### Table 1. The 28 probes for which the average β-value in N samples (βN) was higher in recurrence-positive patients than in recurrence-negative patients, for which the average β-value in C samples (βc) was less than 0.2, and for which the average β-value in T samples minus that in corresponding N samples (∆βT−N) was more than 0.1.

<table>
<thead>
<tr>
<th>Target ID</th>
<th>Chromosome</th>
<th>Position</th>
<th>Gene symbol</th>
<th>$P$</th>
<th>Adjusted $P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>cg00516481 21</td>
<td>44,073,202</td>
<td>PDE9A</td>
<td>$9.96 \times 10^{-4}$</td>
<td>$1.19 \times 10^{-2}$</td>
<td></td>
</tr>
<tr>
<td>cg01295203 8</td>
<td>70,984,199</td>
<td>PRDM14</td>
<td>$8.15 \times 10^{-5}$</td>
<td>$3.23 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg02008154 7</td>
<td>35,293,537</td>
<td>TBX20</td>
<td>$7.18 \times 10^{-4}$</td>
<td>$9.91 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg02909790 6</td>
<td>26,271,587</td>
<td>HIST1H3G</td>
<td>$2.89 \times 10^{-6}$</td>
<td>$5.78 \times 10^{-4}$</td>
<td></td>
</tr>
<tr>
<td>cg03538436 12</td>
<td>117,799,370</td>
<td>NOS1</td>
<td>$7.22 \times 10^{-4}$</td>
<td>$9.64 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg03963198 5</td>
<td>1,882,871</td>
<td>IRX4</td>
<td>$5.13 \times 10^{-3}$</td>
<td>$8.16 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg06005396 19</td>
<td>590,541</td>
<td>HCN2</td>
<td>$9.48 \times 10^{-4}$</td>
<td>$1.15 \times 10^{-2}$</td>
<td></td>
</tr>
<tr>
<td>cg06269753 8</td>
<td>72,557,851</td>
<td>MSC</td>
<td>$8.62 \times 10^{-7}$</td>
<td>$3.42 \times 10^{-4}$</td>
<td></td>
</tr>
<tr>
<td>cg07651242 7</td>
<td>45,614,720</td>
<td>ADCY1</td>
<td>$7.40 \times 10^{-4}$</td>
<td>$1.01 \times 10^{-2}$</td>
<td></td>
</tr>
<tr>
<td>cg11612345 6</td>
<td>168,842,491</td>
<td>SMO2</td>
<td>$4.26 \times 10^{-5}$</td>
<td>$6.73 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg12087643 10</td>
<td>118,033,370</td>
<td>GFR1A</td>
<td>$3.14 \times 10^{-9}$</td>
<td>$2.48 \times 10^{-5}$</td>
<td></td>
</tr>
<tr>
<td>cg12265829 14</td>
<td>24,804,022</td>
<td>ADCY4</td>
<td>$7.26 \times 10^{-4}$</td>
<td>$9.84 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg13262687 4</td>
<td>147,559,579</td>
<td>POU4F2</td>
<td>$6.24 \times 10^{-5}$</td>
<td>$2.79 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg13449779 1</td>
<td>179,712,298</td>
<td>FAM163A</td>
<td>$8.19 \times 10^{-7}$</td>
<td>$3.42 \times 10^{-4}$</td>
<td></td>
</tr>
<tr>
<td>cg13787010 3</td>
<td>123,167,276</td>
<td>ADYS</td>
<td>$2.33 \times 10^{-5}$</td>
<td>$5.22 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg16254309 7</td>
<td>145,814,152</td>
<td>CTNAP2</td>
<td>$2.91 \times 10^{-4}$</td>
<td>$6.24 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg16387606 1</td>
<td>149,804,293</td>
<td>HIST2HA</td>
<td>$7.44 \times 10^{-4}$</td>
<td>$1.01 \times 10^{-2}$</td>
<td></td>
</tr>
<tr>
<td>cg16604516 3</td>
<td>13,590,419</td>
<td>FBLN2</td>
<td>$7.18 \times 10^{-4}$</td>
<td>$9.91 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg16652259 2</td>
<td>172,949,501</td>
<td>DLX1</td>
<td>$4.17 \times 10^{-4}$</td>
<td>$7.36 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg17191178 3</td>
<td>157,824,217</td>
<td>SHOX2</td>
<td>$8.87 \times 10^{-5}$</td>
<td>$3.33 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg18454685 17</td>
<td>48,639,239</td>
<td>CACNA1G</td>
<td>$4.32 \times 10^{-4}$</td>
<td>$7.48 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg20286200 6</td>
<td>133,562,267</td>
<td>EYA4</td>
<td>$1.54 \times 10^{-9}$</td>
<td>$2.03 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg21087137 12</td>
<td>75,728,469</td>
<td>GLPR1L1</td>
<td>$1.96 \times 10^{-5}$</td>
<td>$4.82 \times 10^{-4}$</td>
<td></td>
</tr>
<tr>
<td>cg22461835 8</td>
<td>26,723,365</td>
<td>ADRA1A</td>
<td>$7.29 \times 10^{-4}$</td>
<td>$9.98 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg23418591 20</td>
<td>57,090,317</td>
<td>LOC149773</td>
<td>$4.07 \times 10^{-4}$</td>
<td>$7.28 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg25302419 5</td>
<td>11,904,015</td>
<td>CTNNCD2</td>
<td>$5.14 \times 10^{-4}$</td>
<td>$8.19 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg25764191 10</td>
<td>105,037,215</td>
<td>INA</td>
<td>$3.83 \times 10^{-4}$</td>
<td>$7.02 \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>cg27626299 7</td>
<td>27,282,431</td>
<td>EVX1</td>
<td>$2.16 \times 10^{-4}$</td>
<td>$5.42 \times 10^{-3}$</td>
<td></td>
</tr>
</tbody>
</table>

*Probe ID for the Infinium HumanMethylation27 Bead Array (Illumina).

1National Center for Biotechnology Information database (Genome Build 37).

2Non-adjusted P-values and.

3Benjamini-Hochberg-adjusted P-values for the Cox regression model used for evaluation of correlation with recurrence.

doi:10.1371/journal.pone.0059444.t001
DNA Methylation Profiles at Precancerous Stages

**a**

**ADCY5**

DNA methylation level (I-value)

- N
- T

P = 5.69 x 10^{-31}

**b**

**CNTNAP2**

DNA methylation level (I-value)

- N
- T

P = 1.40 x 10^{-31}

**EVX1**

DNA methylation level (I-value)

- N
- T

P = 1.22 x 10^{-22}

**GFRA1**

DNA methylation level (I-value)

- N
- T

P = 2.46 x 10^{-30}

**PDE9A**

DNA methylation level (I-value)

- N
- T

P = 1.49 x 10^{-30}

**TBX20**

DNA methylation level (I-value)

- N
- T

P = 5.65 x 10^{-22}

**mRNA expression level (GAPDH)**

- N
- T

P = 1.26 x 10^{-27}

P = 7.70 x 10^{-1}

P = 1.38 x 10^{-8}

P = 1.11 x 10^{-30}

P = 5.92 x 10^{-31}

P = 3.02 x 10^{-2}

P = 5.69 x 10^{-31}

P = 1.40 x 10^{-31}

P = 1.22 x 10^{-22}

P = 2.46 x 10^{-30}

P = 1.49 x 10^{-30}

P = 5.65 x 10^{-22}
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

A349, PC9, VMRC-LCD and EBC-1 cells were seeded at a density of 9×10⁴ 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 26,453 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<1×10⁻¹⁴). 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, GFR4A1, 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<1×10⁻¹⁴) revealed ordered progression of DNA methylation alterations from C to T, 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 (ΔβT-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 ΔβT-C value of >0.1 or <−0.1, the false positive 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 425 probes (red dots in Figure 1B), higher β-values in T samples minus average β-values in C samples (ΔβT-C) 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 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-negative 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-negative 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 (βC) were less than 0.2 and the average β-values in T samples minus that in the corresponding N samples (ΔβT-N) were more than 0.1 (Table 1).
DNA Methylation Profiles at Precancerous Stages

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 the ADCY5, EVX1, GFRA1 and TBX20 genes (Figure 3).

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.

<table>
<thead>
<tr>
<th>Clinicopathological parameters</th>
<th>Number of tumors</th>
<th>ADCYS Expressiona</th>
<th>EVX1 Expressiona</th>
<th>GFRα1 Expressiona</th>
<th>PDE9A Expressiona</th>
<th>TBX20 Expressiona</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor diameter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;2.5 cm</td>
<td>43</td>
<td>-4.95 ± 2.10</td>
<td>1.63 × 10⁻¹⁰</td>
<td>-2.80 ± 1.82</td>
<td>4.24 × 10⁻¹⁰</td>
<td>-3.27 ± 2.67</td>
</tr>
<tr>
<td>≥2.5 cm, &lt;4 cm</td>
<td>61</td>
<td>-5.25 ± 2.39</td>
<td>-2.36 ± 2.16</td>
<td>-4.02 ± 2.86</td>
<td>0.28 ± 1.76</td>
<td>-0.57 ± 2.78</td>
</tr>
<tr>
<td>≥4 cm</td>
<td>47</td>
<td>-6.14 ± 2.76</td>
<td>-3.04 ± 2.02</td>
<td>-5.74 ± 3.28</td>
<td>-0.50 ± 1.59</td>
<td>-6.10 ± 2.84</td>
</tr>
<tr>
<td>Histological subtype</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lepidic</td>
<td>12</td>
<td>-4.19 ± 1.61</td>
<td>2.10 × 10⁻¹⁰</td>
<td>-2.73 ± 2.14</td>
<td>1.79 × 10⁻¹⁰</td>
<td>-1.64 ± 2.35</td>
</tr>
<tr>
<td>Acinar</td>
<td>24</td>
<td>-6.03 ± 2.01</td>
<td>-2.31 ± 2.09</td>
<td>5.41 ± 2.72</td>
<td>-0.20 ± 1.36</td>
<td>-0.57 ± 2.57</td>
</tr>
<tr>
<td>Papillary</td>
<td>69</td>
<td>-5.30 ± 2.35</td>
<td>-2.29 ± 2.19</td>
<td>3.90 ± 2.65</td>
<td>0.23 ± 1.63</td>
<td>-0.52 ± 2.80</td>
</tr>
<tr>
<td>Micropapillary</td>
<td>15</td>
<td>-4.69 ± 2.19</td>
<td>-2.35 ± 2.10</td>
<td>3.47 ± 2.87</td>
<td>1.20 ± 1.27</td>
<td>-0.68 ± 1.79</td>
</tr>
<tr>
<td>Solid</td>
<td>28</td>
<td>-6.34 ± 3.24</td>
<td>-3.62 ± 1.88</td>
<td>6.27 ± 3.62</td>
<td>-0.91 ± 1.81</td>
<td>-0.63 ± 3.11</td>
</tr>
<tr>
<td>Invasive mucinous</td>
<td>3</td>
<td>-4.64 ± 1.46</td>
<td>-3.19 ± 1.24</td>
<td>2.95 ± 1.43</td>
<td>0.16 ± 1.68</td>
<td>-0.82 ± 0.14</td>
</tr>
<tr>
<td>Histological grades</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td>56</td>
<td>-4.60 ± 1.80</td>
<td>1.28 × 10⁻¹⁰</td>
<td>-2.68 ± 1.90</td>
<td>4.57 × 10⁻¹⁰</td>
<td>-2.62 ± 1.93</td>
</tr>
<tr>
<td>G2</td>
<td>65</td>
<td>-5.58 ± 2.33</td>
<td>-2.32 ± 2.10</td>
<td>-4.59 ± 2.71</td>
<td>0.25 ± 1.57</td>
<td>-0.57 ± 2.67</td>
</tr>
<tr>
<td>G3</td>
<td>30</td>
<td>-6.73 ± 3.20</td>
<td>-3.54 ± 1.92</td>
<td>-7.01 ± 3.60</td>
<td>-1.06 ± 1.77</td>
<td>-0.68 ± 2.94</td>
</tr>
<tr>
<td>Lymphatic invasion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>52</td>
<td>-5.31 ± 2.43</td>
<td>1.00e</td>
<td>-2.76 ± 1.71</td>
<td>1.00e</td>
<td>0.06 ± 1.64</td>
</tr>
<tr>
<td>Positive</td>
<td>99</td>
<td>-5.52 ± 2.50</td>
<td>-2.66 ± 2.19</td>
<td>-4.55 ± 2.92</td>
<td>0.12 ± 1.67</td>
<td>-0.58 ± 2.93</td>
</tr>
<tr>
<td>Blood vessel invasion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>43</td>
<td>-4.89 ± 2.40</td>
<td>2.38 × 10⁻¹⁰</td>
<td>-2.77 ± 1.84</td>
<td>7.68 × 10⁻¹⁰</td>
<td>-3.13 ± 3.45</td>
</tr>
<tr>
<td>Positive</td>
<td>108</td>
<td>-5.66 ± 2.47</td>
<td>-2.67 ± 2.11</td>
<td>-4.82 ± 2.81</td>
<td>-0.11 ± 1.60</td>
<td>-0.59 ± 2.83</td>
</tr>
<tr>
<td>Pleural invasion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>78</td>
<td>-5.16 ± 2.42</td>
<td>4.45 × 10⁻¹⁰</td>
<td>-2.77 ± 2.15</td>
<td>4.60 × 10⁻¹⁰</td>
<td>-3.65 ± 2.83</td>
</tr>
<tr>
<td>Invasion to the visceral pleura beyond the elastic fiber</td>
<td>33</td>
<td>-5.88 ± 2.68</td>
<td>-2.62 ± 2.07</td>
<td>-4.99 ± 3.54</td>
<td>0.03 ± 1.93</td>
<td>-0.64 ± 2.96</td>
</tr>
<tr>
<td>Invasion to the surface of the visceral pleura</td>
<td>22</td>
<td>-4.96 ± 2.07</td>
<td>-2.16 ± 1.62</td>
<td>-4.01 ± 2.66</td>
<td>0.18 ± 1.29</td>
<td>-0.49 ± 2.82</td>
</tr>
<tr>
<td>Invasion to the parietal pleura</td>
<td>18</td>
<td>-6.47 ± 2.49</td>
<td>-3.15 ± 1.90</td>
<td>-6.56 ± 2.71</td>
<td>-0.65 ± 1.58</td>
<td>-0.55 ± 3.22</td>
</tr>
<tr>
<td>Tumor anthracosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>68</td>
<td>-4.97 ± 2.57</td>
<td>8.27 × 10⁻¹⁰</td>
<td>-2.37 ± 2.22</td>
<td>2.04 × 10⁻¹⁰</td>
<td>-3.32 ± 2.80</td>
</tr>
<tr>
<td>Positive</td>
<td>82</td>
<td>-5.87 ± 2.30</td>
<td>-2.91 ± 1.79</td>
<td>-5.10 ± 3.02</td>
<td>-0.39 ± 1.66</td>
<td>-0.69 ± 2.69</td>
</tr>
</tbody>
</table>

*aAverage of log2-transformed mRNA expression levels/GAPDH ± standard deviation. bAdjusted P-values using adjusted Bonferroni correction. cAnalysis of variance between groups. dPredominant 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]. eWelch's T-test. fCoal 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 chemotherapysensitive 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 the regulation of TBX20 and of TBX20 dysfunctions 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], P16 [38], SHOX2 [39], 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 (βs 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.
Author Contributions
Conceived and designed the experiments: EA YK. Performed the experiments: TS EA YK. Analyzed the data: TS EA KS TB YK.

References


Contributed reagents/materials/analysis tools: EA KT SW YK. Wrote the paper: TS EA YK.