Genetic Classification of Lung Adenocarcinoma Based on Array-Based Comparative Genomic Hybridization Analysis: Its Association with Clinicopathologic Features

Issei Imoto,7,8 Johji Inazawa,7,8 and Setsuo Hirohashi1

Lung cancer is one of the most lethal and increasing cancers in Western countries as well as in Japan (1). Lung cancer is histopathologically divided in two subgroups, small cell lung carcinoma and non–small cell lung carcinoma, and lung adenocarcinoma comprises >40% of the latter (1).

Previous genetic analyses using allelotyping, comparative genomic hybridization (CGH), or the candidate gene approach revealed many genomic (genetic and epigenetic) alterations of tumor suppressor genes (such as p53, p16INK4a, FHIT, LKB1, and PTEN) and oncogenes [such as K-ras, B-RAF, MYC, epidermal growth factor receptor (EGFR), and ERBB2] as well as many chromosomal imbalances (such as on 3p, 8p, 9p, 17p, 18q, and 19p) in lung adenocarcinomas (2–10). However, overall understanding of genomic alterations in lung adenocarcinomas is far from complete and analysis of the relationship between the overall profile and combinations of genetic alterations with clinicopathologic parameters is still lacking. Recently, genome-wide gene expression analyses have uncovered a novel dimension of cancer profiling and helped define the nature of the heterogeneous subgroups of lung adenocarcinoma, each of which shows distinct tumor histology and patient prognosis (11–14). However, it is unclear whether there exist multiple genomic pathways in lung adenocarcinoma because of the lack of a genome-wide view of genetic alterations. It is clinically important to examine the correlations of certain molecular-genetic pathways with cancer cell traits relating to patient prognosis or chemotherapy sensitivity because it is possible that genetic alteration profiling may predict tumor recurrence/metastasis or sensitivity to molecular-target therapies as well as mRNA or protein expression profiles (15–17).

The recently developed array-based CGH method using microarrayed bacterial artificial chromosome clones allows high-resolution analysis of genome-wide copy number changes in various tumors (18, 19). To define and analyze the genetic alterations of lung adenocarcinoma in a more detailed way,

Abstract
The array-based comparative genomic hybridization using microarrayed bacterial artificial chromosome clones allows high-resolution analysis of genome-wide copy number changes in tumors. To analyze the genetic alterations of primary lung adenocarcinoma in a high-throughput way, we used laser-capture microdissection of cancer cells and array comparative genomic hybridization focusing on 800 chromosomal loci containing cancer-related genes. We identified a large number of chromosomal numerical alterations, including frequent amplifications on 7p12, 7q13, 12q44-15, and 17q21, and two homozygous deletions on 9p21 and one on 8p23. Unsupervised hierarchical clustering analysis of multiple alterations revealed three subgroups of lung adenocarcinoma that were characterized by the accumulation of distinct genetic alterations and associated with smoking history and gender. The mutation status of the epidermal growth factor receptor (EGFR) gene was significantly associated with specific genetic alterations and supervised clustering analysis based on EGFR gene mutations elucidated a subgroup including all EGFR gene mutated tumors, which showed significantly shorter disease-free survival. Our results suggest that there exist multiple molecular carcinogenesis pathways in lung adenocarcinoma that may associate with smoking habits and gender, and that genetic cancer profiling will reveal previously uncharacterized genetic heterogeneity of cancer and be beneficial in estimating patient prognosis and discovering novel cancer-related genes including therapeutic targets.

Lung cancer is one of the most lethal and increasing cancers in Western countries as well as in Japan (1). Lung cancer is histopathologically divided in two subgroups, small cell lung carcinoma and non–small cell lung carcinoma, and lung adenocarcinoma comprises >40% of the latter (1).
we used the array CGH technique and laser-capture microdissection of cancer cells, a combination that we have successfully used in other tumor types (20). Using these two powerful methodologies enabled us to detect a considerable number of novel chromosomal numerical alterations in primary lung adenocarcinoma. Unsupervised and supervised hierarchical clustering analyses of genome-wide genetic alterations revealed the presence of heterogeneous groups of lung adenocarcinoma, which are characterized by specific combinations of genetic alterations, varying EGFR gene mutation status, and tumor recurrence rates.

**Materials and Methods**

**Patient materials.** Surgical specimens of 55 lung adenocarcinoma patients who had been diagnosed and undergone operation between June 2001 and May 2002 at the National Cancer Center Hospital were examined. Fragments of tumor and corresponding normal lung tissue were taken immediately after surgery, fixed with 100% methanol, and embedded in paraffin. This study was approved by the institutional review boards of the National Cancer Center. The clinicopathologic data of the patients are shown in Table 1.

**Laser-capture microdissection and whole-genome amplification.** Laser-capture microdissection was done using LM200 (Arcturus, Mount View, CA) as described (21). Only cancer cells were microdissected and lymphocytes, fibroblasts, and endothelial cells were carefully excluded. Corresponding normal lung epithelial cells were similarly microdissected and used as reference. To amplify the genomic DNA fragments, we used an adaptor-ligated whole-genome PCR as previously reported (22).

**Array-based comparative genomic hybridization.** This study used a custom-made CGH array called “MC CancerArray-800 ver.2,” which consists of 800 duplicated bacterial artificial chromosome clones corresponding to various chromosomal loci that have been reported or considered to be altered in various human cancers (20, 23). Details of hybridization procedures have been previously reported (20). Sixteen-bit fluorescence intensity TIF images were obtained using a scanner (FLA8000, Fuji Film, Tokyo, Japan) and analyzed using GenePix Pro 5.0 (Axon Instruments, Inc., Foster City, CA). Thresholds for chromosomal gain (ratio >1.25) and loss (ratio <0.75) were determined by “normal versus normal experiments” (23, 24). We also validated our array CGH data by other methods. Loss of the 17p13 locus was confirmed by loss of heterozygosity of the p53 gene, which is located within that bacterial artificial chromosome using a microsatellite marker (TP53CA). Gene amplification of a representative gene, cyclin D1, was validated by fluorescence in situ hybridization analysis (24). We applied multiplex ligation-dependent probe amplification (MLPA) to validate our array data. Copy number alterations of multiple loci were analyzed using MLPA-SALSA kit (MRC-Holland, Amsterdam, the Netherlands) as per the recommendation of the manufacturer (25).

**Mutational analysis.** We amplified exons 18, 19, 21, and 23 of the EGFR gene; exons 2 and 3 (covering codons 12, 13, and 61) of the K-ras gene; exons 20 and 21 of the ERBB2 gene; and exons 10, 14, 16, 17, 18, 19, and 20 of the MET gene from microdissected tumor and corresponding normal DNA samples with PCR using High Fidelity Taq polymerase (Roche, Mannheim, Germany) and appropriate primers (primer sequences are available on request). All PCR products were purified and analyzed by sequencing. PCR products showing deletions were subcloned in TA-vector (Invitrogen, Carlsbad, CA) and sequenced.

---

**Immunohistochemical analysis.** Four-micrometer sections of formalin-fixed, paraffin-embedded specimens of lung adenocarcinoma were stained with an anti-MET mouse monoclonal antibody (×100 dilution, Zymed, San Francisco, CA) as the suppliers recommended.

**Statistical analysis.** Two-dimensional hierarchical clustering analysis of the samples and signal ratios was done using the ImaGene software (Gene Data, Basel, Switzerland) and GeneMaths (Applied Maths, Sint-Martens-Latem, Belgium) software programs as described (26, 27). Data were standardized by dividing by the root means and dendirgrams were produced using the Pearson Correlation algorithm. For supervised clustering, we first selected loci that were significantly different between EGFR wild-type and mutated tumors based on the average ratio by Student’s t test. We then used a machine-learning method, in which the leave-one-out cross-validation was done with all combinations of loci and multiple independent classifier algorithms, and selected 46 loci that could discriminate EGFR mutation status most accurately to classify the tumors. The Kaplan-Meier method was used to estimate the probability of disease-free survival. Cox proportional hazards regression model and multivariate analysis were done to detect the association between the presence of chromosomal alterations and disease-free survival. Log-rank analysis was used to assess the significance of the difference between subgroups.

**Results**

**Array-based comparative genomic hybridization analysis of primary lung adenocarcinoma.** We analyzed 55 cases of lung...
adenocarcinoma by array-based CGH and the chromosomal alteration profiles of 800 loci are shown in Fig. 1. We identified 32 loci that were lost in >40% of cases (Table 2). Among them, the 9p21 locus containing the p16INK4a gene and the 17p13.1 locus containing the p53 gene were lost in 54% and 40% of analyzed cases, respectively. We found homozygous deletions of three loci, including two on 9p21 and one on 8p23.3. We also identified 19 loci that were gained in >50% of cases (Table 3) and recurrent (>4 cases) amplifications (>4 copies) on 12q14-15 (9 of 55, 16.3%) followed by 7p12.3 (5 of 55), 11q13 (5 of 55), 17q12 (5 of 55), 1p36.1 (4 of 55), 1q21 (4 of 55), 5p15 (4 of 55), 7q31 (4 of 55), 8q24 (4 of 55), 14q12 (4 of 55), and 17q21.2 (4 of 55). These included genes previously reported to be amplified in lung cancer, such as the cyclin D1 (1q13), EGFR (7p12.3), and ERBB2 (17q21.2) genes (28, 29). We further validated copy number alterations on 8q24.3, 17q21.2, 3p21, and 17p13.1 by MLPA method. Chromosomal copy number changes (both gains and losses) detected by array CGH corresponded to those by MLPA (Fig. 1C).

Unsupervised hierarchical clustering of array comparative genomic hybridization data. To examine whether there exist multiple carcinogenesis pathways in lung adenocarcinoma, we attempted two-dimensional hierarchical profiling of the chromosomal alterations detected. We first plotted the number of loci showing various incidences of alterations and found that there exist two peaks (loci altered in 10-15% and 20-25% of cases; Fig. 1D). We assumed that alterations appearing in <20% of cases reflect mostly random alterations as observed in genome-wide allelotyping analyses (30), whereas alterations affecting >20% of cases probably represent nonrandom (cancer-specific) alterations. Therefore, to exclude random changes that may be caused by the intrinsic genetic instability of cancer, we selected the loci that were affected in >25% of analyzed cases (397 loci in total) and did the unsupervised hierarchical clustering analysis. When analyzed by using loci affected in >5% and 15% of cases, we obtained almost the same classification as described below (data not shown).

Our hierarchical clustering yielded three distinct subclasses of primary lung adenocarcinoma (clusters A, B, and C shown in Fig. 1E). Cluster B exhibited significantly fewer genetic alterations (losses and gains) in all examined clones than the other two clusters; the average number of alterations (losses, gains, total) in cluster A was 102, 141, and 244, respectively; in B 43, 81, and 125; and in C 85, 131, and 216 (versus B: P < 0.0001; versus C: P < 0.0001). Frequencies of the various lost or gained loci were significantly different among these three cluster groups (P < 0.01). Cluster A was characterized by gains on 1p32-26, 4p16.3, 11p15, 12q13-14, 16p11.2-13.3, 17q11.1-25, 19q13.2, 20q11, 20q11.2, and 22q12.2 and losses on 1p22, 6q26, 10q24-26, 13q22.1-34, 15q21-25, and 18p11.2. Cluster C significantly showed gains on 5p13-14.3, 7p13.2-21, 7q22, 7q31, 8q12-21, and 14q11-24, and losses on 1q23.3-41, 10q22.1, and Xq. Some loci were similarly altered in both clusters A and C, including losses on 3p21-24, 6q26, 8q24.3, 9p21, 10p15, 10q11, 10q26, 15q21.1, 15q26.1, and 19p13.3 (containing LKB1), and gains on 5p15, 6p21, 7p21-22, 7q21, 8q21, 8q22-24 (containing MYC), 9q21-22, 11q13 (containing cyclin D1), and 20q13.1. Losses on 3p14 (containing FHIT), 8p22-23.3, 9p21 (containing p16INK4a), 13q11-34, 17p13.1 (containing p53), 18q21, and gains on 1q21-23, 1q42, 7p15, 17q12, 17q21.2, and 17q25 were observed in all subgroups with similar frequency. Two alterations (a gain on 19q13.1 and a loss on 22q12.2) were more frequently observed in cluster B than in clusters A and C. The above classification into cluster groups showed significant correlation with the patients’ smoking history (P < 0.01) and gender (P < 0.001); cluster A frequently contained female patients without any smoking history (female: 17 of 20 cases and never smoker: 14 of 20 cases), whereas cluster B included male patients with current or former smoking history (male: 11 of 15 cases and smoker: 11 of 15 cases). Cluster C included more male patients (male: 14 of 20 cases), but showed no significant association with smoking history (smoker: 11 of 20 cases). No significant differences were observed between the groups with regard to other clinical features (histologic differentiation, clinical stage, and disease-free survival). Multivariate analysis revealed that two chromosomal alterations showed significant association with disease-free survival: a loss on 13q14.1 (P = 0.01, hazard ratio, 3.21; 95% confidence interval, 1.30-7.91) and a gain on 8q24.2 (P = 0.02, hazard ratio, 2.92; 95% confidence interval, 1.16-7.37).

It has been reported that somatic mutations of the K-ras, EGFR, and ERBB2 genes are frequent in lung adenocarcinoma (2, 8-10, 31-33). We attempted to determine the correlation of these oncogenic mutations with the above classification. We sequenced exons covering the kinase domain of the EGFR gene and found somatic mutations in 26 cases (47%; Table 1). EGFR mutations were more frequently observed in never-smoker patients (P < 0.001) and in the A and C cluster groups (P = 0.01 and 0.02). We detected K-ras activating mutations in six cases (11% Table 1) and EGFR and K-ras mutations were mutually exclusive in our cases as reported by others (31, 32). No mutation in the reported exons of the ERBB2 gene was detected.

Supervised clustering analysis revealed correlation of EGFR gene mutation with specific genetic alterations. Tumors with EGFR mutations showed significantly more genetic alterations (losses and gains) than those without EGFR mutations; the average numbers of alterations (losses, gains, total) were 68, 110, and 178 in EGFR wild-type tumors and 93, 134, and 227 in EGFR mutated tumors, respectively (wild-type versus mutated P = 0.01, 0.03, and 0.01). We found 58 loci that showed significant differences in the frequency of copy number alterations between EGFR wild-type and mutated tumors. To further examine the genetic profile of EGFR mutated tumors, we classified lung adenocarcinomas based on their EGFR mutation status with the use of supervised hierarchical clustering. We preformed a machine-learning method with leave-one-out cross-validation and selected 46 loci that could discriminate EGFR mutation status. EGFR wild-type and mutated tumors were clustered in distinct groups using the ratios of 46 selected loci (Fig. 2A). Tumors carrying K-ras mutations were segregated from the EGFR mutant branch (Fig. 2A). Interestingly, some cases without EGFR mutation were clustered with the mutant branch, and tumors carrying EGFR mutations were separated in two subbranches (EGFR-MUT-A and EGFR-MUT-B; Fig. 2A). One branch (EGFR-MUT-A) was characterized by amplification of 12q14 or 1p36.1, whereas the other (EGFR-MUT-B) contained frequent amplification of 7q12.3 (containing the EGFR gene), 1q44-23, 5p12, 14q31, and 16p13.3. Poorly differentiated tumors were significantly (P = 0.01) segregated in the EGFR-MUT-B subgroup (EGFR-MUT-A: 1 of 15 cases and EGFR-MUT-B: 5 of 21 cases).
Fig. 1. Copy number alteration profiles of primary lung adenocarcinoma. Frequencies (%) of chromosomal losses (A) and gains (B) detected by array CGH analyses are plotted from chromosome 1p (left) to Y (right). Each chromosome is represented by underlining boxes (short and long arms are indicated by closed and open boxes, respectively). Chromosomal arms that contain frequent losses or gains are also indicated.

C. Validation of array CGH analysis. Distributions of copy number on 8q24.3, 17q21.2, 3p21, and 17p13.1 loci detected by MLPA method in five normal samples (Normal), tumors with alterations detected by array CGH (Amp/Gain or Loss), and tumors without alteration by array CGH method (No change) were shown.
Fig. 1. Continued. D and E, genetic classification of lung adenocarcinoma. D, number of loci as a function of the percentage of altered loci in lung adenocarcinoma. Two modes, which can be separated by the vertical dotted line, can be observed. Loci within the first mode (altered in <20% of cases) were considered random, whereas those in the second mode (altered in >20% of cases) were considered pathogenic. E, unsupervised genetic profiling of lung adenocarcinoma. Fifty-five lung adenocarcinomas were clustered hierarchically on the basis of copy number changes in 397 loci. Multiple lost (green) or gained (red) loci are observed in individual tumors. Overall pattern of standardized gene copy numbers and the cluster tree of individual tumors are shown. Tumors were clustered in three subgroups (clusters A, B, and C). Loci that are specific for (black letters) or shared with (blue letters) each subgroup are indicated on the right side.
The EGFR wild-type tumors that were clustered with the EGFR-MUT branch led us to hypothesize that aberrant activation of tyrosine kinases other than EGFR have an effect equivalent to that of EGFR mutations in these tumors. We examined the copy number changes of loci containing oncogenic receptor–type tyrosine kinases (FGFR1, FGFR2, FGFR3, PDGFR, KIT, MET, ERBB2, FLT3, NTRK1, and NTRK3) detected by our arrays. We found that amplification of a locus containing the MET gene (7q31) was observed in EGFR wild-type tumors (three of four amplified cases; a representative case was shown in Fig. 3A) and that these tumors were clustered with the EGFR-MUT branch (Fig. 2A). Overexpression of MET protein was immunohistochemically detected in 24% (13 of 55) of cases including all cases with amplification of the MET gene (Fig. 3B), although there was no significant association between MET overexpression and the clustering. To determine whether somatic mutations of this kinase also occur in lung adenocarcinoma, we sequenced the exons of the MET gene, which have been reported to exhibit activating mutations in various tumors (34–36); however, no mutation was found in any of examined 55 cases.

We further examined whether this classification is of clinical significance. Kaplan-Meier plots showed a statistically significant difference in disease-free survival between the two groups (EGFR-WT and EGFR-MUT; log-rank analysis, P = 0.01; Fig. 2B) although EGFR mutation status alone did not (log-rank analysis, P = 0.06; data not shown).

### Table 3. Loci frequently gained in primary lung adenocarcinoma

<table>
<thead>
<tr>
<th>Chromosomal location</th>
<th>Covered candidate gene</th>
<th>Chromosomal gain in lung adenocarcinoma (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>17q25</td>
<td>MAFG</td>
<td>67.3</td>
</tr>
<tr>
<td>1q21</td>
<td>MUC1*</td>
<td>63.6</td>
</tr>
<tr>
<td>1q21</td>
<td>MCL1*</td>
<td>61.8</td>
</tr>
<tr>
<td>7p21</td>
<td>IL6</td>
<td>58.2</td>
</tr>
<tr>
<td>1q21</td>
<td>ARHGEF2</td>
<td>58.2</td>
</tr>
<tr>
<td>16p13.3</td>
<td>ABCA3</td>
<td>56.4</td>
</tr>
<tr>
<td>17q11</td>
<td>ITGB4</td>
<td>56.4</td>
</tr>
<tr>
<td>20q13</td>
<td>Livin-2</td>
<td>56.4</td>
</tr>
<tr>
<td>5p15</td>
<td>TERT</td>
<td>54.5</td>
</tr>
<tr>
<td>8q24</td>
<td>GLI4</td>
<td>54.5</td>
</tr>
<tr>
<td>16p13.3</td>
<td>IGFALS</td>
<td>54.5</td>
</tr>
<tr>
<td>17q24-25</td>
<td>GRB2</td>
<td>54.5</td>
</tr>
<tr>
<td>1q21</td>
<td>AFIQ</td>
<td>54.5</td>
</tr>
<tr>
<td>1q23.1</td>
<td>PMF1</td>
<td>54.5</td>
</tr>
<tr>
<td>12q24</td>
<td>stSGB935</td>
<td>52.7</td>
</tr>
<tr>
<td>17q12</td>
<td>PPARB</td>
<td>52.7</td>
</tr>
<tr>
<td>17q25</td>
<td>Survivin*</td>
<td>50.9</td>
</tr>
<tr>
<td>8q24</td>
<td>RECOL4</td>
<td>50.9</td>
</tr>
<tr>
<td>11q2-13</td>
<td>RELA</td>
<td>50.9</td>
</tr>
</tbody>
</table>

*Aberrant expression was previously reported in lung cancer (3–5).

### Discussion

This study is the first high-resolution copy number analyses of primary lung adenocarcinoma by array CGH method. To extract common and specific genetic alterations, we collected and analyzed 55 cases of primary lung adenocarcinoma by combining the array-based CGH analysis with laser-capture microdissection of tumor cells. Importantly, our results were validated by the elucidation of frequent alterations of previously reported cancer-related genes in lung adenocarcinoma, including losses on the 10q telomere, p53, and FHIT loci, and amplifications on the cyclin D1 and EGFR loci. Moreover, we elucidated novel frequent alterations in small chromosomal regions such as losses on 13q11-14 and 15q12-25 and gains on 17q25, 1q21, and 16p13.3, which have not been detected by previous studies. Novel recurrent amplification, which may be a landmark for the existence of oncogenes, was also detected on loci, including 1p36.1, 1q21, 5p15, 12q14-15, and 14q12. We also found a homozygous deletion on 8p23.3 accompanied with frequent chromosomal loss and identification of candidate tumor suppressor genes in this locus is in progress.
It has been argued that there are distinct subclasses of lung adenocarcinoma by histopathologic observations and recent gene expression profilings (2, 13). Girard et al. (30) reported the possibility of classification of lung cancer by genome-wide allelotyping although their study only examined lung cancer cell lines and could not discriminate between copy number gain and loss. In our study, we analyzed primary lung adenocarcinoma and used unsupervised hierarchical cluster analysis to identify three groups of lung adenocarcinoma based on their distinct genetic changes. Among them, two subclasses (clusters A and C, 20 of

Fig. 2. Supervised genetic profiling of EGFR mutation – related loci. A, Hierarchical clustering determined copy number change patterns against 46 loci that were identified using the training testing, cross-validation analysis. Lost (green) or gained (red) loci were indicated in individual tumors. Tumors are classified into two branches (EGFR-WT and EGFR-MUT). All EGFR mutated tumors (red spot) are clustered in the EGFR-MUT branch. Most tumors without EGFR mutations (blue spot) are clustered with the EGFR-WT branch, although some are clustered with the EGFR-MUT branch. Six tumors carrying K-ras mutations (yellow spot) are clustered with the EGFR-WT branch and four MET-amplified tumors (green spot) are clustered with the EGFR-MUT branch. The EGFR-MUT branch is subdivided into two subgroups (EGFR-MUT-A and EGFR-MUT-B) with distinctive genetic changes. B, genetic profiles and patient disease-free survival. Relationship between patients’ disease-free survival and genetic classification based on the EGFR mutation – related loci. EGFR-WT and EGFR-MUT groups were significantly different ($P = 0.01$).
55 cases and 36.3% of total cases, respectively) shared many genetic alterations but also had changes unique to each other. This implied that they may be derived from a common precursor and diverge via the acquisition of specific genetic alterations during tumor development. In contrast, the third subclass (cluster B, 15 of 55 cases, 27.3%) showed characteristically fewer genetic alterations than the other two. Although one would expect this group to consist of tumors of an earlier stage, it contained tumors that varied clinically (from stage I to stage III) and there was no significant correlation of histopathologic features with the above classification. Interestingly, this clustering classification is significantly associated with smoking habits, suggesting that the specific carcinogen exposure may affect overall genetic profile of lung cancer. We propose three possibilities for this group, focusing on the above mechanisms, will provide a more complete view of lung carcinogenesis.

Because the EGFR gene is frequently altered in lung adenocarcinoma and its mutation status is correlated to the sensitivity to the specific inhibitor, Gefitinib (8–10), we assumed that the EGFR pathway plays important roles in lung cancer and examined whether EGFR mutated tumors have any genetic characteristics in nature. We detected the EGFR gene mutations at similar frequency as reported (31, 32, 40) and the presence of somatic mutations was significantly associated with never-smoking history as previous studies reported (8–10, 31, 32, 40). We detected K-ras mutations relatively less frequent than previously reported (41) but comparatively to other study (42) probably because our analyzed cases contained more female and nonsmokers. We found that EGFR mutation and K-ras mutation were mutually exclusive as reported (31, 32, 40) and this finding is consistent with the notion that activation of both EGFR and K-ras stimulates the same downstream pathway (43).

We identified 58 loci whose alterations significantly correlated with the presence of EGFR mutations. It is interesting to note that amplification of the EGFR gene itself is significantly observed in EGFR mutated tumors, indicating that both somatic mutation and amplification of the EGFR gene simultaneously occur in part of lung adenocarcinoma. Using these selected loci, we classified the tumors by supervised hierarchical clustering. This classification revealed two groups: one containing only EGFR wild-type tumors (EGFR-WT) and the other (EGFR-MUT) containing all EGFR mutated and some EGFR wild-type tumors. Because the EGFR wild-type tumors that were grouped with the EGFR-MUT group shared similar genetic alterations with the EGFR mutated tumors, we hypothesized that they may have unknown genetic alterations complementary to EGFR activation and subsequently examined loci containing oncogenic receptor–type tyrosine kinases in our arrays. We found that a locus (7q21) containing the MET gene was amplified in part of these EGFR wild-type tumors and immunohistochemically validated overexpression of MET protein in these tumors. MET was shown to be implicated in ras-mediated tumorigenicity (44, 45) and activated in many tumors (34–36). Although the number of cases with MET amplification is small in this study, it is tempting to speculate that amplification of the MET gene may play a role similar to EGFR mutation in lung adenocarcinoma. Recently, somatic alterations of the MET gene were detected in lung cancer and pharmacologic inhibitors specific to the MET kinase have been...
reported (46–48). Our results also support the idea that the MET oncoprotein is a potent new candidate for therapeutic target in lung adenocarcinoma although there was no somatic mutation in the analyzed exons of our cases. In our cases, there are seven tumors without either EGFR mutations or MET amplification in the EGFR-MUT group. Somatic mutations in the kinase domain of ERBB2 were reported in EGFR wild-type lung adenocarcinomas (7, 33). Therefore, we searched for ERBB2 mutations in all 55 cases and found no somatic mutations, suggesting that other oncogenic kinases might be involved in these tumors.

EGFR mutation status could not predict tumor recurrence, which is consistent with a previous report on the insignificant relationship between EGFR mutation and patient outcome (40). However, we found that EGFR-MUT group, which is revealed by genetic classification, showed significantly shorter disease-free survival than EGFR-WT group. Our results imply the possibility that specific combinations of genetic alterations (genetic code) selected by genome-wide analysis could evaluate tumor characteristics and estimation of such codes would be applicable for diagnostic purposes. Our classification also revealed that there are two genetically distinctive subgroups in the EGFR mutated lung adenocarcinoma, which were associated with tumor histologic differentiation. Because Gefitinib is one of the most promising molecular target drugs against lung cancer and molecular mechanisms determining its efficacy are still unclear (49), further analysis of a larger cohort is warranted to determine any possible relationship of genetic profiling with sensitivity to chemotherapeutic agents, including tyrosine kinase inhibitors.

References


www.aacrjournals.org

6185

Clin Cancer Res 2005;11(17) September 1, 2005

Downloaded from clinicanas.cancers.org on July 20, 2017. © 2005 American Association for Cancer Research.
Genetic Classification of Lung Adenocarcinoma Based on Array-Based Comparative Genomic Hybridization Analysis: Its Association with Clinicopathologic Features

Tatsuhiro Shibata, Satomi Uryu, Akiko Kokubu, et al.


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/11/17/6177

Cited articles
This article cites 45 articles, 25 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/11/17/6177.full#ref-list-1

Citing articles
This article has been cited by 10 HighWire-hosted articles. Access the articles at:
http://clincancerres.aacrjournals.org/content/11/17/6177.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.