Table 2. Practical Application of the New Assay

	Clinical response		
EGFR status	PD	NC	PR
Wild type	5	11	1
Mutated	0	2	10
Deletion in exon 19	0	2	5
Point mutation of codon 858	0	0	5

PD, progressive disease; NC, no change; PR, partial response.

mutation independently of gefitinib treatment,[16,17] was also positive in this assay. In one of the other three recurrent tumors, this assay clearly demonstrated the mutation (Figure 3), although it was often difficult to detect the mutated signal with direct sequencing of the PCR product.

Discussion

Paez et al[5] and Lynch et al[4] simultaneously published the result that somatic mutation of EGFR in lung adenocarcinoma predicts a clinical response to gefitinib. Erlotinib is another targeted small-molecule inhibitor of EGFR, and lung adenocarcinoma sensitive to erlotinib also harbored EGFR mutations. In addition, *in vitro* studies support the observation that EGFR mutations make tumor cells significantly sensitive to gefitinib[18] and erlotinib. This increased sensitivity may be explained by the "addiction to oncogene" hypothesis proposed by Weinstein.[19] Tumor cells with EGFR mutation are highly dependent on the activated EGFR pathway and are thus very susceptible to inhibition of this dependence. We have reported that patients with EGFR mutations survived longer than those without mutations after the initiation of gefitinib treatment.[7] Recently, failure to show a survival benefit in the IRESSA Survival Evaluation in Lung Cancer was announced. Gefitinib may not be effective enough to kill tumor cells that are not under a state of "addiction to EGFR mutation." Conversely, these findings suggest that selection of patients with EGFR-mutated tumors has the advantage of increasing the response rate of EGFR-targeted therapy. Furthermore, selection may also be efficient at preventing serious interstitial pneumonia occurring as a side effect.[20]

Although an assay using paraffin sections is very practical, immunohistochemical analysis of the tumors failed to predict the response. Currently, the microdissection of tumor cells and direct sequencing of PCR products is commonly used as a standard method. Regarding practical applications, the new assay reported here provides two benefits compared with the conventional method. First, microdissection is not necessary for the assay because a positive mutated signal makes this assay very sensitive. Second, this assay is rapid, does not require a purification step, and is usually completed within 4 hours: digestion with proteinase K for 1 hour, real-time PCR or regular PCR for 3 hours, and electrophoresis for 1 hour. In addition to paraffin sections, pleural effusion and specimens for fine needle aspiration cytology can be used. All three specimens of pleural effusion for the T790M cycleave assay were successfully analyzed, whereas direct sequencing occasionally resulted in an ambiguous result (Figure 3). The main targets for gefitinib or erlotinib therapy are recurrent and refractory tumors, and an assay using such specimens is therefore quite useful. However, the examination of limited regions of the EGFR gene appears to be a disadvantage of this study. Recent studies suggested that an insertion of exon 20 was shown to be resistant to EGFR inhibitors[25], whereas the gefitinib sensitivity of cells expressing the G719S mutant was significantly less than that of cells expressing the L858R mutant form[26]. Therefore, these results suggest that examination of the L858R mutation and deletion in exon 19 is reasonable, because these two mutations are likely to be a major target of the EGFR inhibitors.

A few approaches for the detection of EGFR mutation have been reported recently.[21–23] Comparing the assays, the advantage of the method presented here is its practical clinical use. Biopsy specimens frequently result in small, fragmented tissues containing only a few cancer cells. Using such biopsy specimens, the assay successfully demonstrated the EGFR mutations that correlate with gefitinib response, in contrast to failure of the direct sequencing of some biopsy specimens. Furthermore, the cycleave technique can be simultaneously applied for the detection of the K-ras mutation, which has been proposed to be an adverse prognostic marker for chemotherapy with erlotinib.[24]

In summary, we have introduced a new practical approach for the detection of EGFR mutations. This assay is very sensitive and useful for predicting gefitinib response. This rapid screening assay uses paraffin sections from biopsy without the need for a microdissection

Table 3. Detection of T790M Mutation Associated with Acquired Resistance to Gefitinib

Patient	Gefitinib treatment	Tissue examined	T790M mutation	Comments
Case 1*	No	Primary tumor	Yes	A rare case, harboring T790M mutation independent of gefitinib treatment
Case 2	Yes	Pleural effusion	Yes	15-bp deletion of exon 19 in the primary and recurrent cancers (Figure 3)
Case 3	Yes	Pleural effusion	No	9-bp deletion of exon 19 in the primary and recurrent cancers
Case 4	Yes	Pleural effusion	No	15-bp deletion of exon 19 in the primary and recurrent cancers

*The mutation of codon 790 in a primary cancer, which was demonstrated with RT-PCR direct sequencing, has been reported previously.

procedure and has significant advantages over other methods.

Acknowledgments

We thank Kaori Hayashi and Noriko Shibata for excellent technical assistance with the molecular genetic experiments on EGFR, respectively, and Hiroji Ishida for assistance in constructing the tissue array.

References

1. Baselga J, Arteaga CL: Critical update and emerging trends in epidermal growth factor receptor targeting in cancer. J Clin Oncol 2005, 23:2445–2459
2. Arteaga C: Targeting HER1/EGFR: a molecular approach to cancer therapy. Semin Oncol 2003, 30:3–14
3. Arteaga CL: Overview of epidermal growth factor receptor biology and its role as a therapeutic target in human neoplasia. Semin Oncol 2002, 29:3–9
4. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG, Louis DN, Christiani DC, Settleman J, Haber DA: Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004, 350:2129–2139
5. Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ, Naoki K, Sasaki H, Fujii Y, Eck MJ, Sellers WR, Johnson BE, Meyerson M: EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004, 304:1497–1500
6. Han SW, Kim TY, Hwang PG, Jeong S, Kim J, Choi IS, Oh DY, Kim JH, Kim DW, Chung DH, Im SA, Kim YT, Lee JS, Heo DS, Bang YJ, Kim NK: Predictive and prognostic impact of epidermal growth factor receptor mutation in non-small-cell lung cancer patients treated with gefitinib. J Clin Oncol 2005, 23:2493–2501
7. Mitsudomi T, Kosaka T, Endoh H, Horio Y, Hida T, Mori S, Hatooka S, Shinoda M, Takahashi T, Yatabe Y: Mutations of the epidermal growth factor receptor gene predict prolonged survival after gefitinib treatment in patients with non-small-cell lung cancer with postoperative recurrence. J Clin Oncol 2005, 23:2513–2520
8. Marchetti A, Martella C, Felicioni L, Barassi F, Salvatore S, Chella A, Camplese PP, Iarussi T, Mucilli F, Mezzetti A, Cuccurullo F, Sacco R, Buttitta F: EGFR mutations in non-small-cell lung cancer: analysis of a large series of cases and development of a rapid and sensitive method for diagnostic screening with potential implications on pharmacologic treatment. J Clin Oncol 2005, 23:857–865
9. Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, Singh B, Heelan R, Rusch V, Fulton L, Mardis E, Kupfer D, Wilson R, Kris M, Varmus H: EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci USA 2004, 101: 13306–13311
10. Santoro A, Cavina R, Latteri F, Zucali PA, Ginanni V, Campagnoli E, Ferrari B, Morenghi E, Pedicini V, Roncalli M, Alloisio M, Ravasi G, Soto Parra HJ: Activity of a specific inhibitor, gefitinib (Iressa, ZD1839), of epidermal growth factor receptor in refractory non-small-cell lung cancer. Ann Oncol 2004, 15:33–37
11. Parra HS, Cavina R, Latteri F, Zucali PA, Campagnoli E, Morenghi E, Grimaldi GC, Roncalli M, Santoro A: Analysis of epidermal growth factor receptor expression as a predictive factor for response to gefitinib ('Iressa', ZD1839) in non-small-cell lung cancer. Br J Cancer 2004, 91:208–212
12. Bailey LR, Kris M, Wolf MK, Kay AC, Averbuch S, Askaa J, Janas M, Schmidt K, Fukuoka M: Tumor EGFR membrane staining is not clinically relevant for predicting response in patients receiving gefitinib ("Iressa", ZD1839) monotherapy for pretreated advanced non-small-cell lung cancer: IDEAL 1 and 2. Presented at the American Association for Cancer Research annual meeting, July 11–14, 2003, Washington, DC, 2003
13. Kosaka T, Yatabe Y, Endoh H, Kuwano H, Takahashi T, Mitsudomi T: Mutations of the epidermal growth factor receptor gene in lung cancer: biological and clinical implications. Cancer Res 2004, 64:8919–8923
14. Toyooka S, Kiura K, Mitsudomi T: EGFR mutation and response of lung cancer to gefitinib [letter]. N Engl J Med 2005, 352:2136; author reply, 2136
15. Huang SF, Liu HP, Li LH, Ku YC, Fu YN, Tsai HY, Chen YT, Lin YF, Chang WC, Kuo HP, Wu YC, Chen YR, Tsai SF: High frequency of epidermal growth factor receptor mutations with complex patterns in non-small cell lung cancers related to gefitinib responsiveness in Taiwan. Clin Cancer Res 2004, 10:8195–8203
16. Kobayashi S, Boggon TJ, Dayaram T, Janne PA, Kocher O, Meyerson M, Johnson BE, Eck MJ, Tenen DG, Halmos B: EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. N Engl J Med 2005, 352:786–792
17. Pao W, Miller VA, Politi KA, Riely GJ, Somwar R, Zakowski MF, Kris MG, Varmus H: Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. PLoS Med 2005, 2:e73
18. Tracy S, Mukohara T, Hansen M, Meyerson M, Johnson BE, Janne PA: Gefitinib induces apoptosis in the EGFRL858R non-small-cell lung cancer cell line H3255. Cancer Res 2004, 64:7241–7244
19. Weinstein IB: Cancer. Addiction to oncogenes: the Achilles heal of cancer. Science 2002, 297:63–64
20. Inoue A, Saijo Y, Maemondo M, Gomi K, Tokue Y, Kimura Y, Ebina M, Kikuchi T, Moriya T, Nukiwa T: Severe acute interstitial pneumonia and gefitinib. Lancet 2003, 361:137–139
21. Pan Q, Pao W, Ladanyi M: Rapid polymerase chain reaction-based detection of epidermal growth factor receptor gene mutations in lung adenocarcinomas. J Mol Diagn 2005, 7:396–403
22. Sasaki H, Endo K, Konishi A, Takada M, Kawahara M, Iuchi K, Matsumura A, Okumura M, Tanaka H, Kawaguchi T, Shimizu T, Takeuchi H, Yano M, Fukai I, Fujii Y: EGFR mutation status in Japanese lung cancer patients: genotyping analysis using LightCycler. Clin Cancer Res 2005, 11:2924–2929
23. Nagai Y, Miyazawa H, Huqun, Tanaka T, Udagawa K, Kato M, Fukuyama S, Yokote A, Kobayashi K, Kanazawa M, Hagiwara K: Genetic heterogeneity of the epidermal growth factor receptor in non-small cell lung cancer cell lines revealed by a rapid and sensitive detection system, the peptide nucleic acid-locked nucleic acid PCR clamp. Cancer Res 2005, 65:7276–7282
24. Eberhard DA, Johnson BE, Amler LC, Goddard AD, Heldens SL, Herbst RS, Ince WL, Janne PA, Januario T, Johnson DH, Klein P, Miller VA, Ostland MA, Ramies DA, Sebisanovic D, Stinson JA, Zhang YR, Seshagiri S, Hillan KJ: Mutations in the epidermal growth factor receptor and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib. J Clin Oncol 2005, 23:5900–5909
25. Greulich H, Chen TH, Feng W, Janne PA, Alvarez JV, Zappaterra M, Bulmer SE, Frank DA, Hahn WC, Sellers WR, Meyerson M: Oncogenic transformation by inhibitor-sensitive and -resistant EGFR mutants. PLoS Med 2005, 2:e313
26. Jiang J, Greulich H, Janne PA, Sellers WR, Meyerson M, Griffin JD: Epidermal growth factor-independent transformation of Ba/F3 cells with cancer-derived epidermal growth factor receptor mutants induces gefitinib-sensitive cell cycle progression. Cancer Res 2005, 65:8968–8974

Int. J. Cancer: **118,** 1588–1590 (2006)

Letter to the Editor

Mutations of epidermal growth factor receptor and *K-ras* genes in adenosquamous carcinoma of the lung

Shinichi Toyooka[1]*, Yasushi Yatabe[2], Masaki Tokumo[1], Kouichi Ichimura[3], Hiroaki Asano[1], Kunitoshi Tomii[1], Motoi Aoe[1], Hiroyuki Yanai[3], Hiroshi Date[1], Tetsuya Mitsudomi[4] and Nobuyoshi Shimizu[1]

[1]*Department of Cancer and Thoracic Surgery, Graduate School of Medicine and Dentistry, Okayama University, Okayama 700-8558, Japan*
[2]*Department of Pathology and Molecular Diagnosis, Aichi Cancer Center Hospital, Nagoya 464-8681, Japan*
[3]*Department of Pathology, Graduate School of Medicine and Dentistry, Okayama University, Okayama 700-8558, Japan*
[4]*Department of Thoracic Surgery, Aichi Cancer Center Hospital, Nagoya 464-8681, Japan*

Dear Sir,

Lung cancer is one of the most common malignant diseases in developed countries and is classified into nonsmall cell lung cancer (NSCLC) and small cell lung cancer.[1] Adenocarcinoma and squamous cell carcinoma are the 2 major subtypes of NSCLC. However, significant differences in etiology and genetic and epigenetic alterations exist between these 2 subtypes.[2–4] For example, squamous cell carcinoma generally arises in smokers and rarely occurs in never-smokers; mutations in the *K-ras* and *EGFR* genes are also infrequent. In contrast, adenocarcinoma is the dominant histological subtype in female never-smokers, and *K-ras* or *EGFR* mutations are frequently present in this subtype.[5,6] Adenosquamous carcinoma of the lung is a rather rare subtype of NSCLC, comprising 0.4–4% of pulmonary carcinomas.[3] According to the World Health Organization's classification, adenosquamous carcinoma is defined as a carcinoma showing components of both adenocarcinoma and squamous cell carcinoma, with each component comprising at least 10% of the tumor.[3] The etiology of adenosquamous carcinoma, including age, smoking status and race, is similar to that of other types of lung cancers.[7] A clinicopathological analysis has demonstrated that adenosquamous carcinoma is more aggressive and results in a poorer prognosis than does adenocarcinoma or squamous cell carcinoma,[7] indicating that its biological features are different from these major types of NSCLCs. Regarding the histogenesis of adenosquamous carcinoma, monoclonal or polyclonal pathways have been proposed. Monoclonality is considered to be a fundamental feature of neoplasms and consists of the transformation of 1 component to the other, whereas the polyclonal pathway may result from a collision of 2 types of independent tumors.[8,9] However, little is known about the progenitor cells and the process of tumorigenesis in adenosquamous carcinoma of the lung.[3] Information on genetic alterations is also limited; only *TP53*, *K-ras* mutation and loss of heterozygosity at several loci have been reported in a limited number of adenosquamous carcinomas.[10–12] In our study, we investigated the molecular features of adenosquamous carcinoma, a typical heterogeneous tumor of the lung.

In our previous analysis for *EGFR* mutation in 397 cases of NSCLCs, we found somatic mutations in 2 cases for *EGFR* and 1 for *K-ras* out of 6 adenosquamous carcinomas by direct sequence of exon 18–21 for *EGFR* and codon 12 and 13 for *K-ras* genes.[13,14] The rates of *EGFR* mutation in each histology of lung cancer are shown in Table I. In our study, we added 5 new cases of adenosquamous carcinomas of the lung for *EGFR* and *K-ras* analysis, and *EGFR* mutation was present in 1 case and *K-ras* mutation was absent in 5 cases. Thus, 4 of 11 cases showed either *EGFR* or *K-ras* mutations (27% for *EGFR* and 9% for *K-ras*). The characteristics of 11 cases are exhibited in Table II. These results provoked considerable interest, because mutations in these genes were assumed to be usually present in adenocarcinoma and rarely present in squamous cell carcinoma. We examined whether the same mutations were present in both the components of these 4 adenosquamous carcinomas by microdissecting each component and sequencing the DNA obtained from the separate components. The microdissection was performed on paraffin-embedded sec-

TABLE I – *EGFR* MUTATION IN LUNG CANCER

Histology	*EGFR* mutation
Adenocarcinoma (n = 306)	147 (48%)
Squamous cell carcinoma (n = 70)	0
Adenosquamous carcinoma (n = 11)	3 (27%)
Others[1] (n = 15)	0

[1]Others include 11 cases of large cell carcinomas, 3 small cell carcinomas, and 1 of carcinoid.

TABLE II – PATIENT CHARACTERISTICS

Case	Sex	Age	Smoking status	p-Stage	*EGFR*	*K-ras*
1	F	77	Never	IIIA	Mut	Wt
2	M	61	Ever	IIIB	Mut	Wt
3	M	65	Ever	IIIA	Wt	Mut
4	F	56	Never	IIB	Mut	Wt
5	F	45	Never	IA	Wt	Wt
6	M	74	Ever	IIIA	Wt	Wt
7	M	76	Ever	IB	Wt	Wt
8	M	70	Ever	IIA	Wt	Wt
9	M	76	Ever	IIB	Wt	Wt
10	M	62	Ever	IIB	Wt	Wt
11	M	67	Ever	IA	Wt	Wt

Never, never–smoker; Ever, ever–smoker; Wt, wild type; Mut, mutation.

***Correspondence to:** Department of Cancer and Thoracic Surgery, Okayama University Graduate School of Medicine and Dentistry, 2-5-1 Shikata-cho, Okayama 700-8558, Japan. Fax: +81-86-235-7269.
E-mail: s_toyooka@nigeka2.hospital.okayama-u.ac.jp
Received 15 June 2005; Accepted after revision 27 July 2005
DOI 10.1002/ijc.21500
Published online 26 September 2005 in Wiley InterScience (www.interscience.wiley.com).

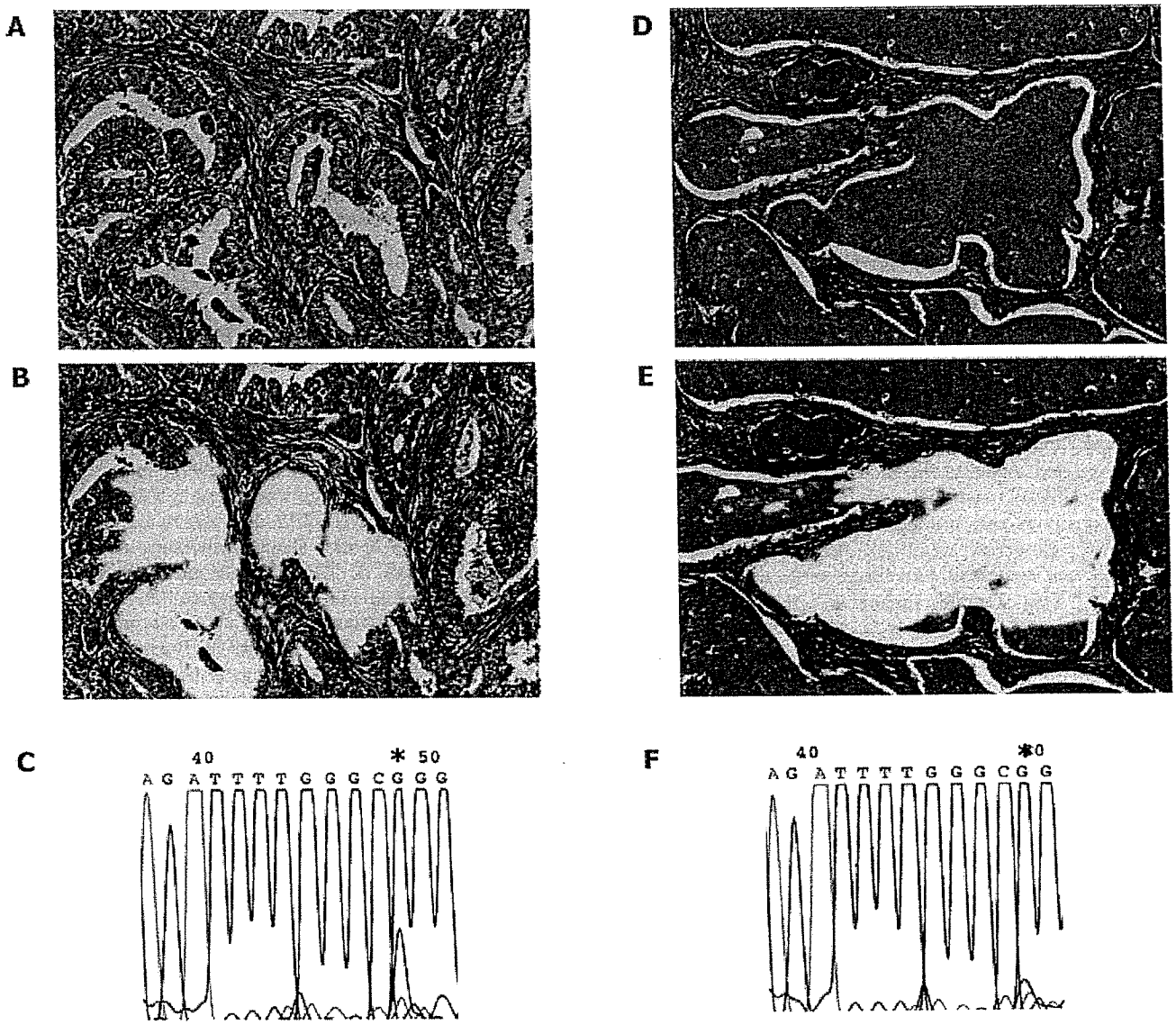

FIGURE 1 – Microdissection and direct sequence of case 1.(*a*) Microscopic finding of adenocarcinomatous component. (*b*). Microdissection of adenocarcinomatous component. (*c*). Sequence of microdissected adenocarcinomatous component. (*d*). Microscopic finding of squamous cell carcinomatous component. (*e*). Microdissection of squamous cell carcinomatous component. (*f*). Sequence of microdissected squamous cell carcinomatous component. * indicates T to G point mutation at second letter of EGFR codon 858.

tions, using a laser capture microdissection (PixCell II; Arcturus Engineering, Mountain View, CA) or by manual microdissection, and DNA was extracted from microdissected cells by incubation with proteinase K (200 μg/ml) for 36 hr at 37°C. The precise information of mutant cases is as follows: Case 1 was a 77-year-old female never-smoker who had been diagnosed as having NSCLC and had surgical resection. A mutation analysis of her resected specimen showed a mutation at codon 858 (CTG–CGG) of *EGFR*. Microdissection for each component was performed showing the same type of mutation (Fig. 1). Case 2 was a 61-year-old male current smoker and showed exon 19 deletion in surgically resected specimen. Mutational pattern of the 2 microdissected components harbored the same type of deletion (delE746–A750). Case 3 was a 65-year-old male former smoker who underwent surgery showing *K-ras* codon 12-point mutation (GGT–GTT) in both the components of microdissected cells. Case 4 was a 56-year-old female with no smoking history and underwent surgery, and the sequence analysis showed exon 19 mutation (delE746–A750) in both components.

On the basis of the mutational status of *EGFR* and *K-ras* genes, our findings suggested 2 interesting notions. First, identical mutation patterns in the 2 tumor components in each of the 4 cases suggested the monoclonality of the adenocarcinomatous and squamous cell carcinomatous components. To our knowledge, 2 reports have also described the monoclonality of 2 components in adenosquamous carcinoma by genetic analysis.[10,11] Immunohistochemical analyses have indicated that the

protein expression characteristics of the adenocarcinomatous and squamous cell carcinomatous components differ.[3,9,11] Taking together these findings, the tumorigenesis in some population of adenosquamous carcinoma is hypothesized; the critical event causing oncogenesis, such as a mutation in *K-ras* or *EGFR*, occurs in the progenitor cells of the adenosquamous carcinoma, then, in the process of multistep tumorigenesis, subsequent differences in the protein expression profiles may cause the cells to differentiate into 2 different phenotypes. Second, although adenosquamous carcinoma is believed to arise from pluripotential bronchial reserve cells, little is known about the progenitor cells of adenosquamous carcinoma. Because *EGFR* and *K-ras* mutations are frequently observed in adenocarcinoma and are rarely present in squamous cell carcinoma, the ancestor of some adenosquamous carcinomas may have features that are similar to those of adenocarcinoma, rather than squamous cell carcinoma, from the viewpoint of genetic alterations. Of interest, the incidence of adenosquamous carcinoma seems to be rising in accordance with the increase in adenocarcinoma.[3,7,10]

Our findings may be critical to understand the carcinogenesis of adenosquamous carcinoma. Further accumulation of specimens for genetic analyses should clarify the mechanism of tumorigenesis in the adenosquamous carcinoma. Such information may suggest novel therapeutic strategies for the adenosquamous carcinoma, of which clinical behavior is quite different from adenocarcinoma or squamous cell carcinoma of the lung.

Yours sincerely,

Shinichi TOYOOKA, Yasushi YATABE, Masaki TOKUMO, Kouichi ICHIMURA, Hiroaki ASANO, Kunitoshi TOMII, Motoi AOE, Hiroyuki YANAI, Hiroshi DATE, Tetsuya MITSUDOMI, and Nobuyoshi SHIMIZU

Acknowledgements

We thank Mayumi Okada, Cancer and Thoracic Surgery, and Yukinari Isomoto, Central Research Laboratory, Graduate School of Medicine and Dentistry, Okayama University, for excellent technical support.

References

1. Parkin DM, Bray FI, Devesa SS. Cancer burden in the year 2000. The global picture. Eur J Cancer 2001;37(Suppl 8):4–66.
2. Zochbauer-Muller S, Gazdar AF, Minna JD. Molecular pathogenesis of lung cancer. Annu Rev Physiol 2002;64:681–708.
3. Travis WD, Branbilla E, Muller-Hermelink HK, Harris HH. World Health Organization classification of tumours. Pathology and genetics of tumours of the lung, pleura, thymus and heart. Lyon: IARC, 2004. 10–52.
4. Toyooka S, Toyooka KO, Maruyama R, Virmani AK, Girard L, Miyajima K, Harada K, Ariyoshi Y, Takahashi T, Sugio K, Brambilla M, Gilcrease M, et al. DNA methylation profiles of lung tumors. Mol Cancer Ther 2001;1:61–7.
5. Slebos RJ, Kibbelaar RE, Dalesio O, Kooistra A, Stam J, Meijer CJ, Wagenaar SS, Vanderschueren RG, van Zandwijk N, Mooi WJ, Bos JL, Rodenhuis S, et al. *K-ras* oncogene activation as a prognostic marker in adenocarcinoma of the lung. N Engl J Med 1990;323: 561–5.
6. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG, Louis DN, Christiani DC, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004;350:2129–39.
7. Sridhar KS, Raub WA, Jr, Duncan RC, Hilsenbeck S. The increasing recognition of adenosquamous lung carcinoma (1977–1986). Am J Clin Oncol 1992;15:356–62.
8. Nowell PC. The clonal evolution of tumor cell populations. Science 1976;194:23–8.
9. Ishida T, Kaneko S, Yokoyama H, Inoue T, Sugio K, Sugimachi K. Adenosquamous carcinoma of the lung. Clinicopathologic and immunohistochemical features. Am J Clin Pathol 1992;97:678–85.
10. Niho S, Yokose T, Kodama T, Nishiwaki Y, Mukai K. Clonal analysis of adenosquamous carcinoma of the lung. Jpn J Cancer Res 1999;90: 1244–7.
11. Kanazawa H, Ebina M, Ino-Oka N, Shimizukawa M, Takahashi T, Fujimura S, Imai T, Nukiwa T. Transition from squamous cell carcinoma to adenocarcinoma in adenosquamous carcinoma of the lung. Am J Pathol 2000;156:1289–98.
12. Graziano SL, Gamble GP, Newman NB, Abbott LZ, Rooney M, Mookherjee S, Lamb ML, Kohman LJ, Poiesz BJ. Prognostic significance of *K-ras* codon 12 mutations in patients with resected stage I and II non-small-cell lung cancer. J Clin Oncol 1999;17:668–75.
13. Kosaka T, Yatabe Y, Endoh H, Kuwano H, Takahashi T, Mitsudomi T. Mutations of the epidermal growth factor receptor gene in lung cancer: biological and clinical implications. Cancer Res 2004;64: 8919–23.
14. Tokumo M, Toyooka S, Kiura K, Shigematsu H, Tomii K, Aoe M, Ishimura K, Tsuda T, Yano M, Tsukuda K, Tabata M, Ueoka H, et al. The relationship between epidermal growth factor receptor mutations and clinicopathologic features in non-small cell lung cancers. Clin Cancer Res 2005;11:1167–73.

VOLUME 24 · NUMBER 11 · APRIL 10 2006

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Expression Profile–Defined Classification of Lung Adenocarcinoma Shows Close Relationship With Underlying Major Genetic Changes and Clinicopathologic Behaviors

Toshiyuki Takeuchi, Shuta Tomida, Yasushi Yatabe, Takayuki Kosaka, Hirotaka Osada, Kiyoshi Yanagisawa, Tetsuya Mitsudomi, and Takashi Takahashi

From the Division of Molecular Carcinogenesis, Center for Neurological Diseases and Cancer, Nagoya University Graduate School of Medicine; Department of Pathology and Molecular Diagnostics, Aichi Cancer Center Hospital; Department of Thoracic Surgery, Aichi Cancer Center Hospital; Division of Molecular Oncology, Aichi Cancer Center Research Institute, Nagoya, Japan; and the Japan Biological Informatics Consortium, Tokyo, Japan.

Submitted August 12, 2005; accepted October 27, 2005.

Supported by in part by a grant-in-aid for scientific research on priority areas and a grant-in-aid for scientific research from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and a grant from the Japan New Energy and Industrial Technology Development Organization.

T. Takeuchi, S.T., and Y.Y. contributed equally to this work.

Terms in blue are defined in the glossary, found at the end of this article and online at www.jco.org.

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

Address reprint requests to Takashi Takahashi, MD, PhD, Division of Molecular Carcinogenesis, Center for Neurological Diseases and Cancer, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan; e-mail: tak@med.nagoya-u.ac.jp.

0732-183X/06/2411-1679/$20.00

DOI: 10.1200/JCO.2005.03.8224

ABSTRACT

Purpose
This study was conducted to gain insight into the relationship between expression profiles and underlying genetic changes, which are known to be important for the pathogenesis of lung cancers.

Methods
Expression profiles of 18,175 unique genes and three major targets for genetic changes, *p53*, epidermal growth factor receptor (*EGFR*), and *K-ras*, were investigated in 149 patients with non–small-cell lung cancer, including 90 patients with adenocarcinoma to determine their relationships with various clinicopathologic features and Gene Ontology (GO) terms.

Results
This study successfully established a basis for expression profile-defined classification, which can classify adenocarcinomas into two major types, terminal respiratory unit (TRU) type and non–TRU type. Our GO term–based identifier of particular biologic processes, molecular functions, and cellular compartments clearly showed characteristic retention of normal peripheral lung features in TRU type, in sharp contrast to the significant association of non–TRU type with cell cycling and proliferation-related features. While significantly higher frequency of EGFR mutation was observed in TRU type, we found that the presence of EGFR mutations was a significant predictor of shorter postoperative survival for TRU type, independent of disease stage. We were also able to identify a set of genes in vivo with significant upregulation in the presence of EGFR mutations.

Conclusion
This study has shed light on heterogeneity in lung cancers, especially in adenocarcinomas, by establishing a molecularly, genetically, and clinically relevant, expression profile-defined classification. Future studies using independent patient cohorts are warranted to confirm the prognostic significance of EGFR mutations in TRU-type adenocarcinoma.

J Clin Oncol 24:1679-1688. © 2006 by American Society of Clinical Oncology

INTRODUCTION

Lung cancer is the leading cause of cancer-related death in developed countries.[1,2] While the current classification of neoplasia is largely based on the histologic features observed under the microscope, marked variations in clinical behavior are sometimes evident even within a particular histologic type, and adenocarcinomas are known to exhibit the highest degree of morphologic and clinical diversities.[3] Thus, more detailed, accurate, and objective means to classify non–small-cell lung cancer (NSCLC) tumors, especially adenocarcinomas, are greatly anticipated not only for a better understanding of the pathogenesis, but also for improving our currently inadequate diagnostic capabilities to help develop more effective treatment modalities.

The recent development of microarray technologies has made it possible to correlate gene expression profiles in individual cases with various clinical parameters.[4-9] To date, various groups including our own have reported that expression profiling can recapitulate morphologic classification of NSCLCs, and some studies also showed that adenocarcinomas can be subclassified additionally.[10-13] However, these previously reported subclassifications vary considerably from study to study, making it difficult to reconcile their findings or reach any definite conclusions. It should also be noted that previous expression profiling studies provided little

information with regard to the relationship of expression profiles with underlying genetic changes, which are known to be important for the pathogenesis of lung cancers.[14]

The identification of activating mutations of epidermal growth factor receptor (*EGFR*) is one of the most intriguing recent discoveries in the field of lung cancer research.[15,16] *EGFR* mutations are present in a subset of pulmonary adenocarcinomas,[15-20] and tumors with this mutation have been shown to be highly sensitive to gefitinib.[15-17,19] Good clinical response to gefitinib has been observed most frequently in female, Japanese and other Asian ethnicities, nonsmoking patients with adenocarcinomas.[21,22] We have also reported that *EGFR* mutations are more prevalent in the terminal respiratory unit (TRU) -type adenocarcinomas,[23] which we have proposed as a characteristic subset, representing adenocarcinomas with the peripheral airway epithelium as their putative origin.[23,24]

In this study, we concurrently analyzed global expression profiles and *EGFR*, *p53*, and *K-ras* mutation status in 149 patients with NSCLC, including 90 patients with adenocarcinoma, in order to establish a both genetically and clinicopathologically relevant expression profile-defined classification. We used this classification to identify the significantly higher prevalence and clear prognostic impact of *EGFR* mutations in one of the two major subtypes identified in adenocarcinoma.

METHODS

Patients

A series of 149 patients with NSCLC, comprising 90 patients with adenocarcinomas, 35 patients with squamous cell carcinomas, 18 patients with large-cell carcinomas, four patients with adenosquamous carcinomas, and two patients with large cell neuroendocrine carcinomas, who successfully underwent potential curative resections between December 1995 and December 1999, were obtained from a file at the Department of Pathology and Molecular Diagnostics, Aichi Cancer Center, Nagoya, Japan. For the postoperative survival analysis, 82 of the patients with adenocarcinoma met the criteria for inclusion, while the remaining eight patients were excluded because of postoperative treatment with gefitinib. The median follow-up period was 77 months (range, 6 to 108 months) when all the eligible patients were included and 92 months (range, 64 to 108 months) when deceased patients were excluded. All the tumor specimens were embedded in OCT compound (Sakura Finetechnical Co Ltd, Tokyo, Japan) and stored at −80°C after the requisite approval from the institutional review board and patients' written informed consent had been obtained.

Acquisition of Expression Profiles

Frozen tissues of the tumor specimens were subjected to gross microdissection under the guidance of a pathologist (Y.Y.) by using every tenth section stained with Giemsa. Total RNA was extracted using the RNeasy kit (Qiagen, Valencia, CA), followed by treatment with DNase I. A large batch of common reference RNA was prepared by using 20 lung cell lines representing all major histologic types of lung cancers. Double-stranded cDNA was synthesized from 250 ng of total RNA using Moloney murine leukemia virus-reverse transcriptase (Agilent Technologies, Palo Alto, CA) and poly dT primer incorporating the T7 promoter. cRNA was generated and labeled with Cy3 or Cy5 (CyDye, Amersham Pharmacia Biotech, Piscataway, NJ) using the Low RNA Fluorescent Linear Amplification kit (Agilent Technologies). Cy5-sample cRNA and Cy3-common reference cRNA were hybridized to a custom Agilent oligonucleotide microarray, containing a total of 21,619 spots corresponding to 18,175 unique genes, followed by confocal laser scanning (Agilent Technologies). Fluorescence intensities on scanned images were quantified, and the values were corrected for background level and normalized.

Bioinformatic Analysis

Details of the bioinformatics analysis are provided in the Appendix (online only). In brief, genes that were flagged in more than 10 samples were excluded from additional analyses. In addition, genes whose expression levels did varied by a factor of less than three across the sample set of interest were eliminated because they were unlikely to be informative. We used the Cluster program (http://rana.lbl.gov/EisenSoftware.htm) to perform average linkage hierarchical clustering of both genes and cases, using median centering and normalization, and displayed the results with the aid of TreeView software (http://rana.lbl.gov/EisenSoftware.htm).[25] We used significance analysis of microarrays (SAM; www-stat.stanford.edu/~tibs/SAM/index.html) to perform gene ranking specific for each of the patient subtypes.[26]

Mutation Analysis of EGFR, p53, and K-ras Genes

The *p53* (exons 4 to 10), *EGFR* (exons 15 to 24), and *K-ras* (exons 1 and 2) genes were amplified from the same RNA used for the microarray analysis, and the resulting polymerase chain reaction products were directly sequenced essentially as described previously.[18]

Gene Ontology Term–Based Identifier and Other Statistical Analysis

Gene Ontology (GO; http://www.geneontology.org/) analysis was employed to highlight functionally distinct biologic features of gene sets specific for each of the patient subtypes (Appendix).[27] Database files used for this GO analysis were downloaded from the UniGene FTP site (Appendix). Eventually, 12,745 known genes among the 18,175 unique genes on the microarray chip were linked to about 67,000 GO terms by parsing the database files including Hs.seq.all, Hs.data, and LL_tmpl with the aid of our newly developed program written with Perl (practical extraction and report language; www.activestate.com/Products/ActivePerl/). These terms were subjected to Fisher's exact test in order to identify which GO terms were over- or under-represented in a gene set of interest.

The χ^2 test or Fisher's exact test were used for comparisons of proportions, while multivariate logistic regression analysis was performed to examine associations of the expression of expression profile-defined subtypes with various clinical parameters. The Kaplan-Meier method was used to estimate survival as a function of time, and survival differences were analyzed with the log-rank test. Cox proportional hazards modeling was performed to identify which independent factors might jointly have a significant effect on survival. All the analyses were performed with Stata software (version 7; Stata Corp, College Station, TX), and the two-sided significance level was set at $P < .05$.

RESULTS

Expression Profiled-Defined Classification of NSCLCs

We first used unsupervised hierarchical clustering to classify all 149 samples using the 4,834 most variably expressed transcripts in order to attain a molecular classification based on the similarity of genome-wide expression patterns in individual tumors. The resultant clusters accurately recapitulated the well-established, widely used histologic classification of NSCLC, while the large cell cluster was considerably mixed with other histologic types (Fig 1). SAM analysis disclosed the presence of distinct sets of genes with up- or downregulation for each of the clusters, and the gene sets were found to be similar to those previously reported by us[13] and others[10,11] (data not shown). Expression profiling data were confirmed by both real-time reverse transcriptase polymerase chain reaction and immunohistochemical analyses.

In this study, we concurrently performed extensive searches for mutations in the *EGFR*, *p53*, and *K-ras* genes, which are strongly believed to contribute to the development of lung cancers, in order to investigate whether the expression profile-defined subtypes have

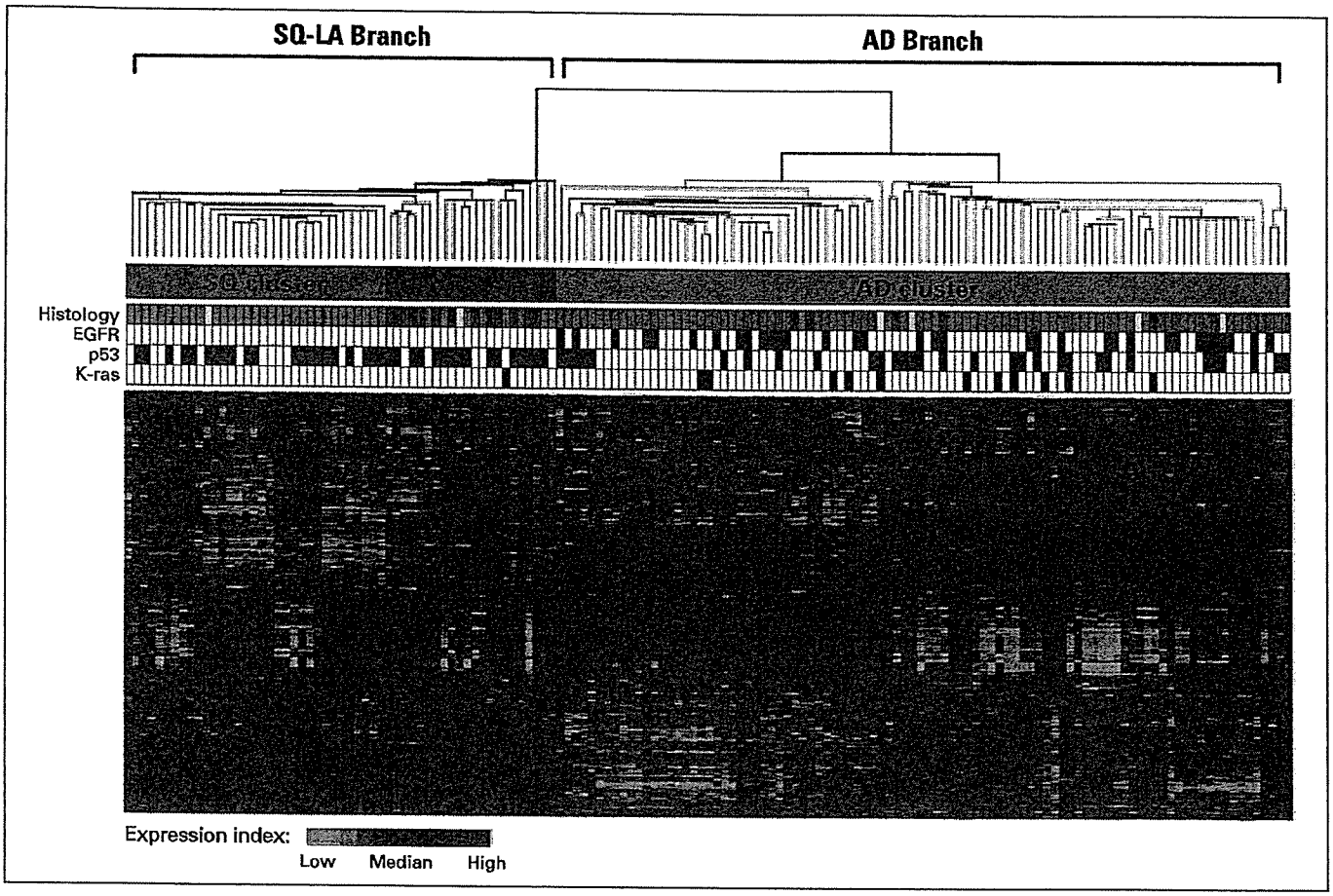

Fig 1. Unsupervised hierarchical clustering and analysis of three major genetic changes in 149 patients with non–small-cell lung cancer. Boxes in the histology row represent squamous cell carcinoma (SQ; blue), large cell carcinoma (LA; red), adenocarcinoma (AD; orange), adenosquamous cell carcinoma (gray), and large cell neuroendocrine carcinoma (yellow). Black boxes indicate the presence of *EGFR*, *p53*, and *K-ras* mutations in their respective row.

any relationship to the presence or absence of these genetic changes. As was expected from the clearly accurate recapitulation of the conventional histologic classification of NSCLC and previous reports on these three gene alterations, highly significant associations were observed between *EGFR* mutations and the corresponding adenocarcinoma branches (36% *v* 0%; $P < .001$) as well as between *p53* mutations and the branch consisting mostly of squamous cell carcinomas and large cell carcinomas (67% *vs* 33%; $P < .001$).

Expression Profile-Defined Two Major Types of Adenocarcinomas

We noted during that, although adenocarcinoma cases clustered together as a large single branch, there were two major subclusters. This finding led us to perform a separate analysis of adenocarcinomas based on the reasoning that strong signatures in other histologic types of NSCLC may obscure subtle differences within adenocarcinomas. To this end, hierarchical clustering, performed with the 4,138 transcripts most variably expressed within adenocarcinomas, clearly showed the presence of two major branches as well as of two additional subclusters in the right branch (Fig 2A). Morphologic analysis showed that, although well-differentiated tumors and adenocarcinomas with bronchioloalveolar carcinomas (BAC) features tended to be more prevalent in the right branch and BAC residing within the extreme right hand subcluster, clustering of other adenocarcinoma subtypes, or variants according to the WHO classification were not noticeable in any particular branches (Fig 2A). We noticed, however, that the morphologic characteristics of cases in the right branch resembled those of TRU-type adenocarcinoma, a distinctive adenocarcinoma subset that we previously proposed based on its distinct cellular morphology and expression of TITF1 (TTF-1) and surfactant proteins[23,24,28] (for the sake of convenience, tumors in the right and left branches will be referred to as, respectively, TRU- and non–TRU type adenocarcinomas).

To gain additional insight into the molecular and biologic nature of these two major expression profile-defined adenocarcinoma subtypes (ie, non-TRU and TRU types), SAM analysis was performed first to select differentially expressed genes. A total of 1,657 genes passed prefiltering at a significance level of < 0.1% false-discovery rates in the SAM analysis, and 293 of these genes showed differences by a factor of more than 2 between their expression levels in TRU- and non-TRU types. These 293 genes consisted of 201 with higher expression in the TRU type and 92 with higher expression in the non-TRU type. In order to better understand the underlying functional distinctions between TRU and non–TRU types, the SAM-identified gene sets were subjected to our GO term identifier of differentially utilized functions

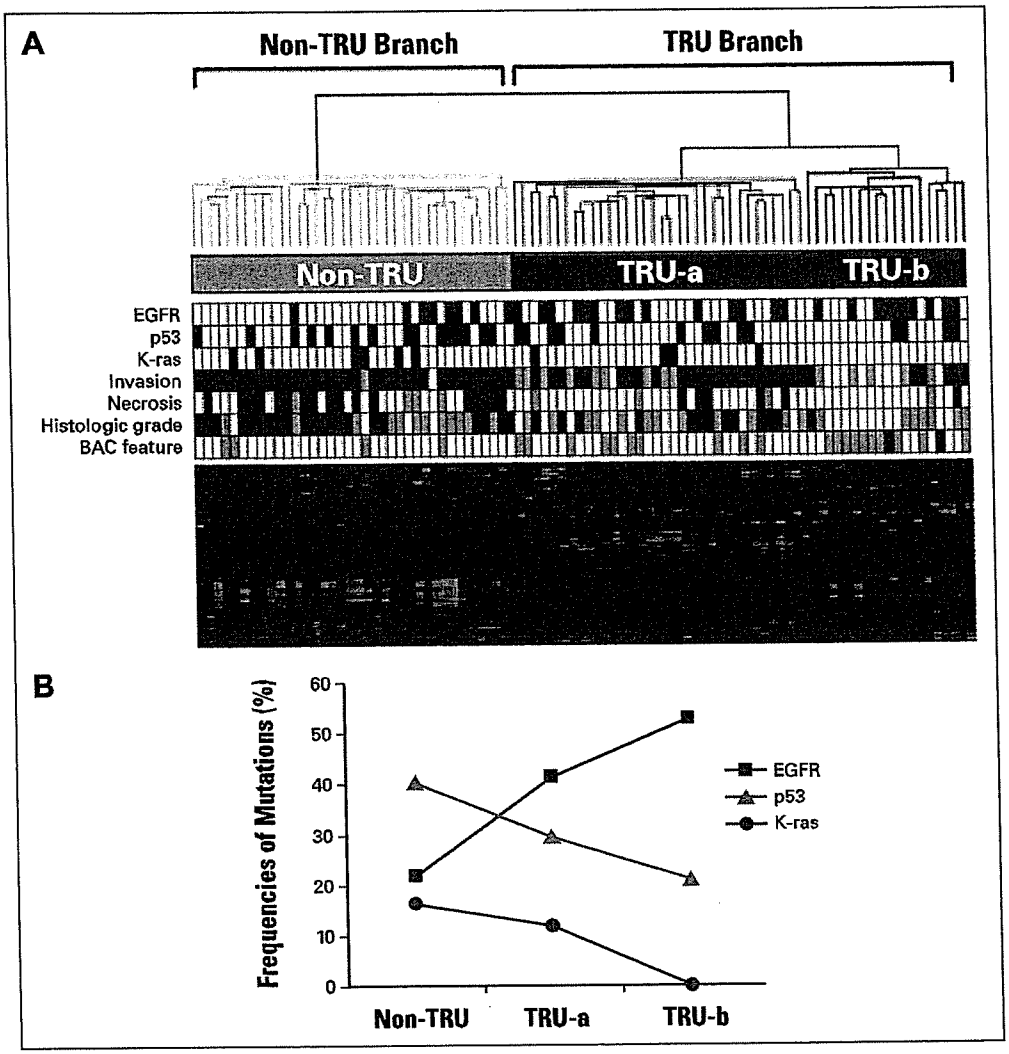

Fig 2. Unsupervised hierarchical clustering and analysis of three major genetic changes in 90 patients with adenocarcinoma. Prominent invasion and necrosis, poorly differentiated tumor, and pure BAC, black boxes; focal invasion, moderate necrosis, moderately differentiated tumor, and adenocarcinoma with BAC features, gray boxes; and lack or negligible invasion and necrosis, well-differentiated tumor, and those without lepidic growth, open boxes. TRU, terminal respiratory unit; BAC, bronchioloalveolar carcinomas.

and other characteristics, which was developed in our laboratory for this study. With this identifier, GO terms related to nine biologic processes, six molecular functions, and two cellular components were extracted as occurring significantly more frequently in TRU-type adenocarcinomas, exhibiting clear relation to normal lung functions (Table 1). In contrast, GO terms identified as those specific to non-TRU type distinctively included those related to cell cycling and proliferation, suggesting an inherently aggressive nature of non–TRU type adenocarcinomas.

The identification of these highly robust expression profile-defined subtypes of adenocarcinomas led us to conduct a similar unsupervised hierarchical clustering analysis using an independent data set. The Stanford data set consisting of 34 patients with lung adenocarcinoma[10] was analyzed by means of unsupervised hierarchical clustering based on the expression profiles of the 30 top-ranked genes in terms of distinctive expression in the non-TRU, TRU-a, or TRU-b types. This resulted in the clear visualization of two major branches with a subcluster, which appeared to correspond well to the three expression profile-defined adenocarcinoma subtypes identified in this study (Fig 3).

Significant Association of Expression Profile-Defined Adenocarcinoma Subtypes With Clinicopathologic Characteristics

We examined the relationship between various clinicopathologic features and the two expression profile-defined adenocarcinoma subtypes (ie, TRU and non-TRU adenocarcinomas; Table 2). TRU-type adenocarcinomas were seen significantly more frequently than non-TRU types in females ($P = .005$) and never-smokers ($P < .001$). In contrast, a detailed microscopic examination showed that invasive growth and the presence of necrosis, indicative of higher malignant potential/appearance and a rapidly proliferating nature, were characteristically more prevalent in non-TRU types ($P < .001$ for both invasive growth and necrosis), in accordance with the predominance of proliferation and cell cycling-related GO terms distinctive for non–TRU type tumors. Multivariate logistic regression analysis with age, sex, smoking status, and pathologic stage as variables identified never-smoking status as the only significantly associated variable ($P = .001$). As for postoperative prognosis, we noted that patients belonging to the two apparent clusters under the TRU-type adenocarcinoma branch seen in Figure 2A, namely TRU-a and TRU-b, were

Table 1. GO Terms Significantly Related to TRU or Non-TRU Types

Aspect	Term	Acession No.	*P*
Related to TRU type			
Biological processes			
	Regulation of lipid surface tension	0050828	.002
	Sex differentiation	0007548	.002
	Lipoprotein metabolism	0042157	.013
	Innate immune response	0045087	.014
	Mesoderm development	0007498	.018
	Phosphate transport	0006817	.019
	Respiratory gaseous exchange	0007585	.022
	Lipid metabolism	0006629	.027
	Fertilization (sensu Metazoa)	0007338	.030
Molecular functions			
	Antigen binding	0003823	.001
	Oxidoreductase activity	0016712	.002
	Oxygen binding	0019825	.003
	Phospholipids-translocating ATPase activity	0004012	.013
	Unspecific mono-oxygenase activity	0050381	.032
	Steroid binding	0005496	.037
Cellular components			
	Extra cellular region	0005576	.011
	Microsome	0005792	.032
Related to non-TRU type			
Biological processes			
	Nucleotide biosynthesis	0009165	.002
	Cell cycle	0007049	.003
	Circulation	0008015	.003
	Establishment and/or maintenance of chromatin	0006325	.005
	Mitosis	0007067	.014
	Pregnancy	0007565	.021
	Cell surface receptor linked signal transduction	0007166	.036
	Epidermis development	0008544	.046
Molecular functions			
	Polypeptide *N*-acetyl-galactosaminyl transferase	0004653	.006
	ATP binding	0005524	.038
Cellular components			
	Soluble fraction	0005625	< .001
	Chromosome, pericentric region	0000775	.004
	Kinetochore	0000776	.005
	Nucleolus	0005730	.032

Abbreviation: TRU, terminal respiratory unit; GO, Gene Ontology (http://www.geneontology.org/); ATP, adenosine triphosphate.

distinct in terms of their postoperative prognoses. While TRU-a type had a prognosis similar to that of non–TRU type, the prognosis for TRU-b type adenocarcinomas was significantly better than for non–TRU type ($P = .021$; Fig 4). This seemed to be consistent with the fact that microscopic examination showed that apparent invasive growth occurred much less frequently in TRU-b type adenocarcinomas than in the other types of adenocarcinomas ($P < .001$).

Significant Association of Expression Profile-Defined Adenocarcinoma Subtypes With EGFR *Mutation Status*

We examined whether any of the three major genetic alterations in NSCLCs were significantly associated with the expression profile-defined adenocarcinoma subtypes (Fig 2B). The presence of *EGFR* mutations was found to be significantly more prevalent in TRU-type adenocarcinomas than in non–TRU type adenocarcinomas (45.3% *v* 21.6%; $P = .026$). We also noted that the two apparent clusters (ie, TRU-a and TRU-b) under the TRU-type adenocarcinomas branch modestly differed in the prevalence of *EGFR* mutations, with a higher *EGFR* mutation frequency for TRU-b type (52.6%) than for TRU-a type (41.2%) adenocarcinomas. In contrast, *K-ras* and *p53* did not correlate with the subtypes of adenocarcinomas ($P = .33$ for *p53*; $P = .17$ for *K-ras*), although interesting inverse correlations of mutation frequencies were observed between these genetic alterations and *EGFR* mutations. Non–TRU type adenocarcinomas contained the highest percentage of tumors carrying *p53* and/or *K-ras* mutations (41% for *p53*; 16% for *K-ras*), followed by TRU-a type (29% and 12%, respectively) and TRU-b type (21% and 0%, respectively).

Prognostic Significance of EGFR *Mutations in TRU-Type Adenocarcinoma*

We, as well as others, have reported that the presence of *EGFR* mutations does not affect postoperative prognosis for NSCLC

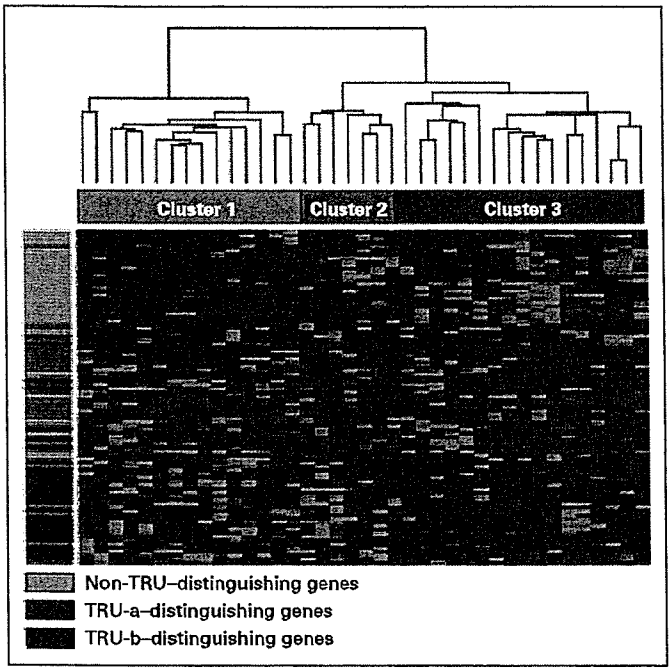

Fig 3. Unsupervised hierarchical clustering based on the expression profiles of the 30 top-ranked genes in terms of distinctive expression in the non–TRU, TRU-a, or TRU-b types, and profiling-defined adenocarcinoma subsets identified with the Stanford data set.[10] The expression index is indicated as in Fig. 1. TRU, terminal respiratory unit.

Table 2. Relationships Between Expression Profile–Defined Subtypes of Adenocarcinomas and Clinicopathologic Characteristics

Clinical Feature	EP-Defined Subtypes		
	Non-TRU	TRU	*P*
No. of patients	37	53	
Age, years			
≤ 62	22	28	.67
> 62	15	25	
Sex			
Male	26	21	.005
Female	11	32	
Smoking history			
Never-smoker	8	37	< .001
Current and former smoker	29	16	
pT			
T1	14	26	.72
T2	17	20	
T3	4	4	
T4	2	3	
pN			
N0	25	35	.99
N1	3	5	
N2	9	13	
pStage			
I	22	30	.83
II	6	7	
III	9	16	
Histologic grade			
Well differentiated	4	21	.002
Moderately differentiated	12	19	
Poorly differentiated	21	13	
Invasive growth			
Positive	35	30	< .001
Positive, but focal	1	14	
Negative or negligible	1	9	
Necrosis			
Positive	17	6	< .001
Focal	4	4	
Negative	16	43	

Abbreviations: EP, expression-profiled; TRU, terminal respiratory unit.

patients,[18,20] and this was confirmed in this independent data set (Fig 5; $P = .42$). On the basis of the reasoning that the present findings may suggest more important roles for *EGFR* mutations, especially in the development of TRU-type adenocarcinomas rather than in that of other types of lung cancers, we performed a separate analysis of TRU-type adenocarcinoma cases to examine the potential association between *EGFR* mutations and postoperative prognosis. A significant association was detected between poor postoperative prognosis and the presence of *EGFR* mutations in TRU-type adenocarcinomas ($P = .024$; Fig 5B). This association was confirmed additionally by the results of multivariate Cox regression analysis (Table 3). The presence of *EGFR* mutations in TRU-type adenocarcinoma was shown to be an independent prognostic factor (hazard ratio [HR], 7.87; $P = .003$) in addition to disease stage (HR, 7.85; $P = .001$) and expression-profile-defined histologic subtypes (HR, 8.81; $P = .005$), whereas sex, age, smoking status, and histologic grade did not show any significant associations.

Search for Genes With Significant Differential Expression in the Presence or Absence of EGFR *Mutations*

We used a significance level of 5% false-discovery rate for selecting genes associated with the presence of *EGFR* mutations in all adenocarcinoma cases, since none were selected when a 0.1% false discovery rate was used (Table 4). Five genes were identified with more than two fold upregulation in tumors with *EGFR* mutations, while additional 11 genes showed upregulation of 1.5 to two times in association with the presence of *EGFR* mutations. We also searched for the genes differentially expressed specifically within TRU-type adenocarcinomas with significantly higher prevalence of *EGFR* mutations. Although it was necessary to use a quite high false-positive discovery rate (10%) for selecting genes by SAM, a single gene, *GGTLA4*, was identified with more than two fold upregulation in the presence of *EGFR* mutations. In addition, five genes showed upregulation of 1.5 to two times at the same level of significance (Table 4).

DISCUSSION

Expression profiles in a given tumor can be regarded as the outcome of complex influences resulting from the accumulated genetic changes important for the pathogenesis as well as from differentiation commitment of the progenitor cells. In this comprehensive study, we were able to establish an expression profile-defined, highly robust classification of adenocarcinomas, the relevance of which is clearly supported by the inclusion of various molecular genetic and clinicopathologic distinctions. The two major subtypes (ie, TRU and non-TRU types) feature differential expressions of a large number of genes, thus

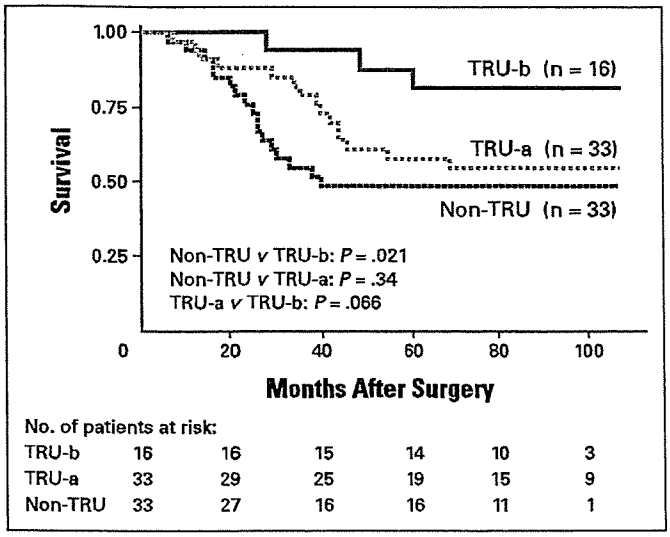

Fig 4. Kaplan-Meier survival curves for the expression profile-defined adenocarcinoma subtypes. TRU-b-type adenocarcinomas had significantly better prognosis than non–TRU type, while prognosis for TRU-a type was similar to that for non–TRU type. TRU, terminal respiratory unit.

Table 3. Multivariate Cox Regression Analysis of Potential Prognostic Factors for TRU-Type Adenocarcinoma

Unfavorable/Favorable Variable	Hazard Ratio	95% CI	*P*
Age, years			
> 62/≤ 62	1.14	0.39 to 3.32	.82
Sex			
Male/female	1.11	0.32 to 3.89	.87
Smoking			
Current and former/never	1.34	0.36 to 4.99	.66
Stage			
II or III/I	7.85	2.38 to 25.9	.001
Histologic grade			
Poor/moderate and well	0.82	0.20 to 3.45	.79
Subcluster			
TRU-a/TRU-b	8.81	1.96 to 39.6	.005
EGFR status			
Mutant/wild-type	7.87	2.02 to 30.7	.003

Abbreviations: TRU, terminal respiratory unit; EGFR, epidermal growth factor receptor.

indicating their significant difference in terms of gene usage.[23] The results obtained by using our GO term–based identifier support their different nature additionally. TRU-type tumors are characterized by biologic processes, which are important for the maintenance of peripheral lung functions.[29] Similarly, molecular functions related to TRU type also appear to reflect retention of their progenitor cells' characteristics. In contrast, most non–TRU type associated GO terms are related to cell cycling and cellular proliferation, which appear to be consistent with their microscopic appearance of high-grade characteristics.

The marked distinctions in clinical features between TRU and non-TRU subtypes also support the robustness of the present expression profile-defined classification. Notable clinical features of TRU type are significantly higher proportions of females and nonsmokers, whereas the result of multivariate analysis indicates nonsmoker status, but not sex as an independently associated factor. Thus, this tumor type appears to arise from peripheral lung airway cells under much less influence of smoking and to retain its progenitor's characteristics as indicated by the GO term-based analysis. TRU-b type had the most favorable prognosis with less frequent invasive growth and expressed various differentiation markers at higher levels even when compared with TRU-a type, suggesting that TRU-b type retains features of normal peripheral lung airway cells better than the other types, but may progress to the TRU-a type.

Our study clearly shows that the presence of *EGFR* mutations is significantly associated with TRU type adenocarcinoma. In fact, 45.3% of TRU-type adenocarcinomas carried *EGFR* mutations in contrast to a 21.6% occurrence in non-TRU type. In contrast, *p53* and *K-ras* did not show significant differences in terms of their mutation frequencies, but there was an interesting inverse propensity with the highest occurrence in non-TRU type. It should also be noted that the presence of *EGFR* mutations was significantly associated with shortened postoperative survival, specifically for TRU-type adenocarcinoma, independently of disease stage. Indeed, 5-year survival rates of patients with stage II/III disease of TRU type were found to be 25% in the presence of *EGFR* mutations and 78% in their absence. These findings suggest the potential clinical usefulness of concurrent analysis of expression profile-defined subtypes and *EGFR* mutation status to select candidates for intensive adjuvant therapy, possibly with gefitinib.

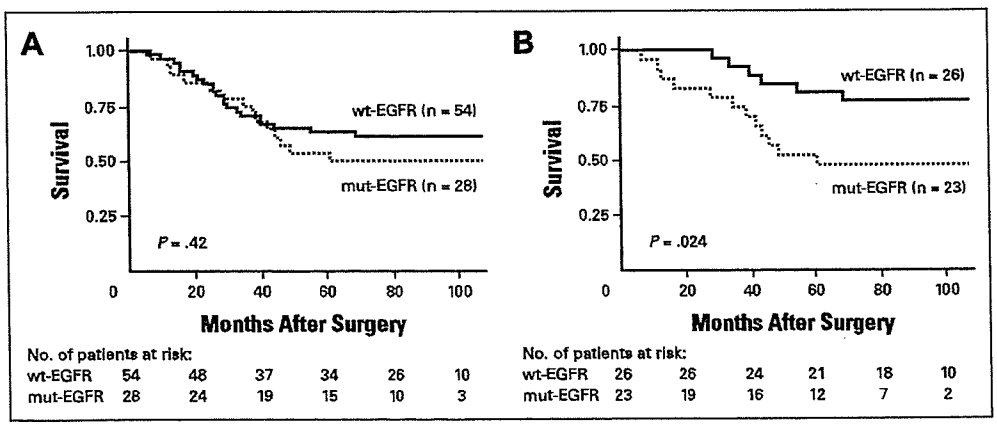

Fig 5. Kaplan-Meier survival curves for the presence or absence of *EGFR* mutations (mut) in all cases of adenocarcinomas as well as in TRU-type adenocarcinoma. (A) Postoperative survival curves of adenocarcinoma cases with and without *EGFR* mutations; (B) postoperative survival curves for the presence or absence of *EGFR* mutations in TRU-type adenocarcinomas. wt, wild type.

Table 4. Genes Identified As Those Upregulated in Association With the Presence of EGFR Mutations in All Adenocarcinomas or in TRU-Type Adenocarcinoma

Gene Symbol	Gene Name	UniGene* ID	Difference
All adenocarcinomas			
ZDHHC11	Zinc finger, DHHC domain containing 11	Hs.368851	3.0
GGTLA4†	Gamma-glutamyltransferase-like activity 4	Hs.355394	2.3
LOC401022†	Hypothetical LOC401022	Hs.98661	2.1
CPAMD8	C3 and PZP-like, alpha-2-macroglobulin domain containing 8	Hs.529075	2.1
EST	Transcribed locus	Hs.449965	2.0
UNQ541†	GSGL541	Hs.211267	1.9
RGMA	RGM domain family, member A	Hs.271277	1.8
DNALI1	Dynein, axonemal, light intermediate polypeptide 1	Hs.406050	1.8
MESP1	Mesoderm posterior 1	Hs.447531	1.7
CDKL2†	Cyclin-dependent kinase-like 2	Hs.310540	1.7
EST	Transcribed locus	Hs.553240	1.7
EST	Transcribed locus	Hs.98587	1.6
APOH	Apolipoprotein H (beta-2-glycoprotein I)	Hs.445358	1.6
LGALS3BP	Lectin, galactoside-binding, soluble, 3 binding protein	Hs.514535	1.5
PEX3	Peroxisomal biogenesis factor 3	Hs.7277	1.5
C6orf60	Chromosome 6 open reading frame 60	Hs.443789	1.5
TRU-type adenocarcinoma			
GGTLA4†	Gamma-glutamyltransferase-like activity 4	Hs.355394	2.4
RAMP1	Receptor (calcitonin) activity modifying protein 1	Hs.471783	1.9
APOH	Apolipoprotein H (beta-2-glycoprotein I)	Hs.445358	1.7
PEX3	Peroxisomal biogenesis factor 3	Hs.7277	1.7
EST	Transcribed locus	Hs.553240	1.6
DHRS7	Dehydrogenase/reductase (SDR family) member 7	Hs.59719	1.5

Abbreviations: EGFR, epidermal growth factor receptor; TRU, terminal respiratory unit; DHHC, Asp-His-His-Cys; RGM, repulsive guidance molecule; SDR, short-chain dehydrogenase/reductase.
*http://www.ncbi.nlm.nih.gov/UniGene.
†Genes also differentially expressed in TRU and non-TRU type adenocarcinomas.

In conclusion, this comprehensive study has shed light on the existence of heterogeneity in lung cancers, especially adenocarcinomas, by establishing a genetically and clinicopathologically relevant, expression profile-defined molecular classification. Additional studies using independent patient cohorts will be necessary before any definitive conclusion can be reached with regard to the prognostic significance of the presence of *EGFR* mutations in TRU-type adenocarcinoma.

REFERENCES

1. Minna JD: Neoplasms of the lung, in Kasper DL, Braunwald E, Fauci AS, et al (eds): Harrison's Principals of Internal Medicine (Ed 16). New York, NY, McGraw-Hill, 2001, pp 506-516
2. Satistics and Information Department, Minister's Secretariat: Vital Statistics of Japan 2001. Tokyo, Japan, Ministry of Health, Labor and Welfare, 2003; pp 384-411,
3. Travis WD, Brambilla E, Mueller-Hermelink HK, et al: Pathology and genetics of tumours of the lung, pleura, thymus and heart, World Health Organization classification of tumors. Lyon, France, IARC Press, 2004
4. Perou CM, Sorlie T, Eisen MB, et al: Molecular portraits of human breast tumours. Nature 406: 747-752, 2000
5. Alizadeh AA, Eisen MB, Davis RE, et al: Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403: 503-511, 2000
6. Shipp MA, Ross KN, Tamayo P, et al: Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. Nat Med 8:68-74, 2002
7. Singh D, Febbo PG, Ross K, et al: Gene expression correlates of clinical prostate cancer behavior. Cancer Cell 1:203-209, 2002
8. van 't Veer LJ, Dai H, van de Vijver MJ, et al: Gene expression profiling predicts clinical outcome of breast cancer. Nature 415:530-536, 2002
9. Hedenfalk I, Ringner M, Ben-Dor A, et al: Molecular classification of familial non-BRCA1/BRCA2 breast cancer. Proc Natl Acad Sci U S A 100:2532-2537, 2003
10. Garber ME, Troyanskaya OG, Schluens K, et al: Diversity of gene expression in adenocarcinoma of the lung. Proc Natl Acad Sci U S A 98:13784-13789, 2001
11. Bhattacharjee A, Richards WG, Staunton J, et al: Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. Proc Natl Acad Sci U S A 98:13790-13795, 2001
12. Beer DG, Kardia SL, Huang CC, et al: Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat Med 8:816-824, 2002
13. Tomida S, Koshikawa K, Yatabe Y, et al: Gene expression-based, individualized outcome prediction for surgically treated lung cancer patients. Oncogene 23:5360-5370, 2004
14. Meyerson M, Carbone D: Genomic and proteomic profiling of lung cancers: Lung cancer classification in the age of targeted therapy. J Clin Oncol 23:3219-3226, 2005
15. Paez JG, Janne PA, Lee JC, et al: EGFR mutations in lung cancer: Correlation with clinical response to gefitinib therapy. Science 304:1497-1500, 2004
16. Lynch TJ, Bell DW, Sordella R, et al: Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 350:2129-2139, 2004
17. Pao W, Miller V, Zakowski M, et al: EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci U S A 101:13306-13311, 2004
18. Kosaka T, Yatabe Y, Endoh H, et al: Mutations of the epidermal growth factor receptor gene in lung cancer: Biological and clinical implications. Cancer Res 64:8919-8923, 2004
19. Mitsudomi T, Kosaka T, Endoh H, et al: Mutations of the epidermal growth factor receptor gene predict prolonged survival after gefitinib treatment in patients with non-small-cell lung cancer with postoperative recurrence. J Clin Oncol 23:2513-2520, 2005
20. Shigematsu H, Lin L, Takahashi T, et al: Clinical and biological features associated with epider-

mal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst 97:339-346, 2005

21. Fukuoka M, Yano S, Giaccone G, et al: Multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small-cell lung cancer. J Clin Oncol 21:2237-2246, 2003

22. Kris MG, Natale RB, Herbst RS, et al: Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: A randomized trial. JAMA 290:2149-2158, 2003

23. Yatabe Y, Kosaka T, Takahashi T, et al: EGFR mutation is specific for terminal respiratory unit type adenocarcinoma. Am J Surg Pathol 29:633-639, 2005

24. Yatabe Y, Mitsudomi T, Takahashi T: TTF-1 expression in pulmonary adenocarcinomas. Am J Surg Pathol 26:767-773, 2002

25. Eisen MB, Spellman PT, Brown PO, et al: Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95:14863-14868, 1998

26. Tusher VG, Tibshirani R, Chu G: Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 98:5116-5121, 2001

27. Ashburner M, Ball CA, Blake JA, et al: Gene Ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25:25-29, 2000

28. Yatabe Y: Role of expression of thyroid transcription factor-1 in pulmonary adenocarcinoma, in Hayat MA (ed): Immunohistochemistry and In Situ Hybridization of Human Carcinomas. New York, NY, Elsevier Science/Academic Press, 2004, pp 169-179

29. Veldhuizen EJ, Haagsman HP: Role of pulmonary surfactant components in surface film formation and dynamics. Biochim Biophys Acta 1467:255-270, 2000

Acknowledgment

We thank Tetsuo Noda, MD, PhD, and Yoshio Miki, MD, PhD, (Cancer Institute, Tokyo, Japan) for their encouragement throughout the study, and Kaori Hayashi (Aichi Cancer Center, Nagoya, Japan) for her excellent technical assistance with the molecular genetic experiments.

Appendix

The Appendix is included in the full-text version of this article, available online at www.jco.org. It is not included in the PDF version (via Adobe® Reader®).

Authors' Disclosures of Potential Conflicts of Interest

Although all authors completed the disclosure declaration, the following author or immediate family members indicated a financial interest. No conflict exists for drugs or devices used in a study if they are not being evaluated as part of the investigation. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Authors	Employment	Leadership	Consultant	Stock	Honoraria	Research Funds	Testimony	Other
Tetsuya Mitsudomi					AstraZeneca Japan (A)			

Dollar Amount Codes (A) < $10,000 (B) $10,000-99,999 (C) ≥ $100,000 (N/R) Not Required

Author Contributions

Conception and design: Takashi Takahashi
Financial support: Takashi Takahashi
Administrative support: Takashi Takahashi
Provision of study materials or patients: Yasushi Yatabe, Takayuki Kosaka, Tetsuya Mitsudomi
Collection and assembly of data: Toshiyuki Takeuchi, Yasushi Yatabe, Takayuki Kosaka
Data analysis and interpretation: Toshiyuki Takeuchi, Shuta Tomida, Yasushi Yatabe, Hirotaka Osada, Kiyoshi Yanagisawa, Takashi Takahashi
Manuscript writing: Toshiyuki Takeuchi, Shuta Tomida, Yasushi Yatabe, Hirotaka Osada, Kiyoshi Yanagisawa, Tetsuya Mitsudomi, Takashi Takahashi
Final approval of manuscript: Toshiyuki Takeuchi, Shuta Tomida, Yasushi Yatabe, Tetsuya Mitsudomi, Takashi Takahashi

GLOSSARY

EGFR (epidermal growth factor receptor): Also known as HER-1, EGFR belongs to a family of receptors (HER-2, HER-3, HER-4 are other members of the family) and binds to the EGF, TGF-α, and other related proteins, leading to the generation of proliferative and survival signals within the cell. It also belongs to the larger family of tyrosine kinase receptors and is generally overexpressed in several solid tumors of epithelial origin.

TRU (terminal respiratory unit): A physioanatomical unit of the terminal airway in the lung that is conducted by single respiratory bronchioles to a few alveoli. This unit is a primary site of gas exchange. The surface is covered by a characteristic epithelium (ie, pneumocytes).

Expression profile: The expression pattern of genes, selected from a particular cell or tissue type, generally obtained by a variety of high-throughput methods, such as microarray and serial analysis gene expression (SAGE).

Hierarchical clustering: An analytical tool used to find the closest associations among gene profiles and specimens under evaluation.

Gene ontology: Allows for annotating genes and their products with a limited set of attributes, with the three organizing principles being molecular function, biological process, and cellular component. The development of structured, controlled vocabularies (ontologies) that describe gene products in terms of these organizing principles in a species-independent manner is a constantly evolving process.

Hs.seq.all, Hs.data, and LL_tmpl: UniGene (www.ncbi.nlm.nih.gov/UniGene) is an experimental system maintained at the National Cen-

ter for Biotechnology Information for automatically partitioning GenBank sequences into a nonredundant set of gene-oriented clusters. Each UniGene cluster contains sequences that represent a unique gene, as well as related information such as the tissue types in which the gene has been expressed and map location. Hs.seq.all contains all Homo Sapience sequences and the information included in clusters, Hs.data contains various information related to each UniGene ID, and LL_tmpl contains various information related to each Locus ID.

Perl (practical extraction and report language): A programming language especially designed for processing text. Because of its strong text processing abilities, Perl is used extensively in areas such as bioinformatics and Web programming.

SAM (significance analysis of microarrays): A statistical technique using established software that determines the significance in changes of gene expression seen in microarray analysis (eg, cDNA and oligonucleotide microarrays), which measures the expression of thousands of genes and identifies changes in expression between different biologic states. On the basis of changes in gene expression relative to the standard deviation of repeated measurements, SAM assigns a score to each gene. When scores are greater than an adjustable threshold, permutations of repeated measurements are used by SAM to estimate the percentage of such genes identified by chance, the false discovery rate (FDR). In addition, SAM correlates gene expression data to a wide range of clinical parameters, including treatment, diagnosis categories, and survival time.

Int J Clin Oncol (2006) 11:190–198
DOI 10.1007/s10147-006-0583-4

REVIEW ARTICLE

Tetsuya Mitsudomi · Takayuki Kosaka · Yasushi Yatabe

Biological and clinical implications of EGFR mutations in lung cancer

Received: April 17, 2006

Abstract

Background. Patients with non-small-cell lung cancer sometimes show a dramatic clinical response to gefitinib or erlotinib, small-molecule tyrosine kinase inhibitors (TKI) specific for the epidermal growth factor receptor (EGFR). However, until April 2004, it was unclear how to identify patients who would benefit from these drugs. Then, two groups from Boston reported that *EGFR* gene mutations in the kinase domain are strongly associated with gefitinib sensitivity. *EGFR* mutations are more frequent in Asians, females, nonsmokers, and adenocarcinomas than in their counterparts. These populations precisely coincide with those populations with higher response rates to TKIs. We and others subsequently confirmed and extended these findings.

Methods. We reviewed recent literatures on EGFR mutations and EGFR-TKIs. We discuss topics including the molecular epidemiology and biology of *EGFR* mutations in relation to EGFR-TKIs, the controversy about whether *EGFR* mutations account for all the clinical activity of EGFR-TKIs, and the mechanisms of acquired resistance to gefitinib or erlotinib.

Results. The discovery of EGFR mutations has great biologic and clinical implications in lung cancer. However, all but one phase III trials have so far failed to show a survival advantage of the treatment arm involving EGFR-TKIs.

Conclusion. It would be possible to individualize EGFR-TKI treatment of lung cancer by selecting patients according to EGFR mutational status and other biomarkers.

Key words Molecular targeted therapy · Tyrosine kinase inhibitor · Gefitinib · Individualized therapy · Predictive factor

T. Mitsudomi (✉) · T. Kosaka · Y. Yatabe
Department of Thoracic Surgery and Department of Pathology and Molecular Diagnostics, Aichi Cancer Center Hospital, Aichi Cancer Center, 1-1 Kanokoden, Chikusa-ku, Nagoya 464-8681, Japan
Tel. +81-52-762-6111; Fax +81-52-764-2963
e-mail: mitsudom@aichi-cc.jp

Introduction

Lung cancer is the leading cause of cancer-related death in Japan, as in Western countries, claiming nearly 60000 lives annually. Although various chemotherapeutic agents were developed in the 1990s, platinum doublet therapy reached a therapeutic plateau with an objective response rate of 30%–40% and a median survival time (MST) of 8–10 months for patients with stage IIIB or IV disease.[1]

To circumvent this situation, a new class of drugs that specifically targets certain molecular pathways leading to cancer phenotypes is being actively developed. Epidermal growth factor receptor (EGFR) is one such target for the treatment of non-small-cell lung cancer (NSCLC), because EGFR is frequently overexpressed and aberrantly activated in NSCLC. When EGFR binds to several specific ligands, multiple signaling pathways are activated including the RAS/RAF/ERK/MAPK pathway, resulting in cell proliferation, and the PI3K/AKT pathway, STAT (signal transducers and activators of transcription) 3 and 5 signal transduction pathways, resulting in the evasion of apoptosis.[2] Antibodies directed against the extracellular domain of EGFR (such as cetuximab, matuzumab, and panitumab) and small-molecule tyrosine kinase inhibitors (TKIs) that target the kinase domain (such as gefitinib and erlotinib) are in clinical use or in a late developmental stage.[3]

In the phase II trials of gefitinib, IDEAL 1 and 2, certain patient subgroups appeared to have a higher response rate: female and Japanese patients, and adenocarcinomas.[4,5] Miller et al. reported that smoking history and bronchioloalveolar pathological subtype predict sensitivity to gefitinib.[6] Overall, a partial radiographic response was observed in 21 (15%) of 139 patients with advanced NSCLC. Never-smokers had a significantly higher response rate than former/current smokers (36% vs 8%, respectively; $P < 0.001$) and multivariate analysis confirmed this association ($P = 0.006$).[6] Following the report of these findings, various groups confirmed that a response to gefitinib or erlotinib is consistently seen in a certain patient subgroup. An analysis

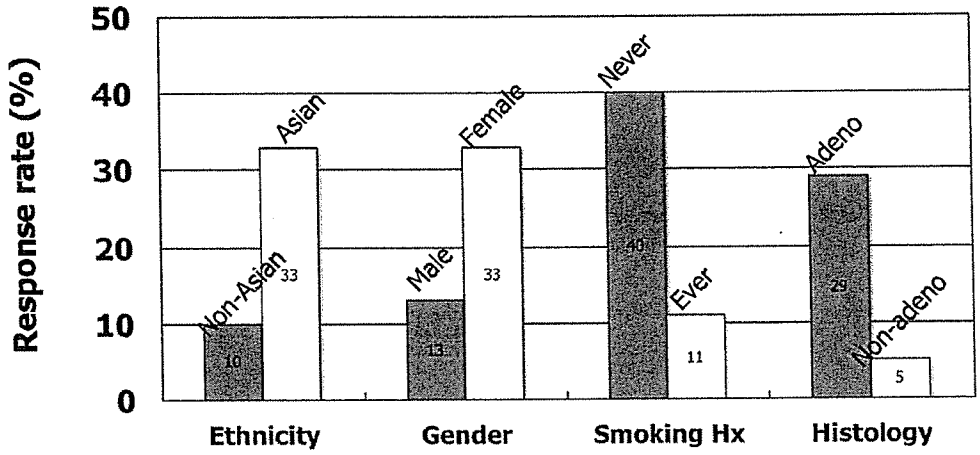

Fig. 1. Relationship between response rate and various clinical backgrounds. Data on 1974 patients compiled from the literature[4–6,23–25,28,31,44,46,52–59]

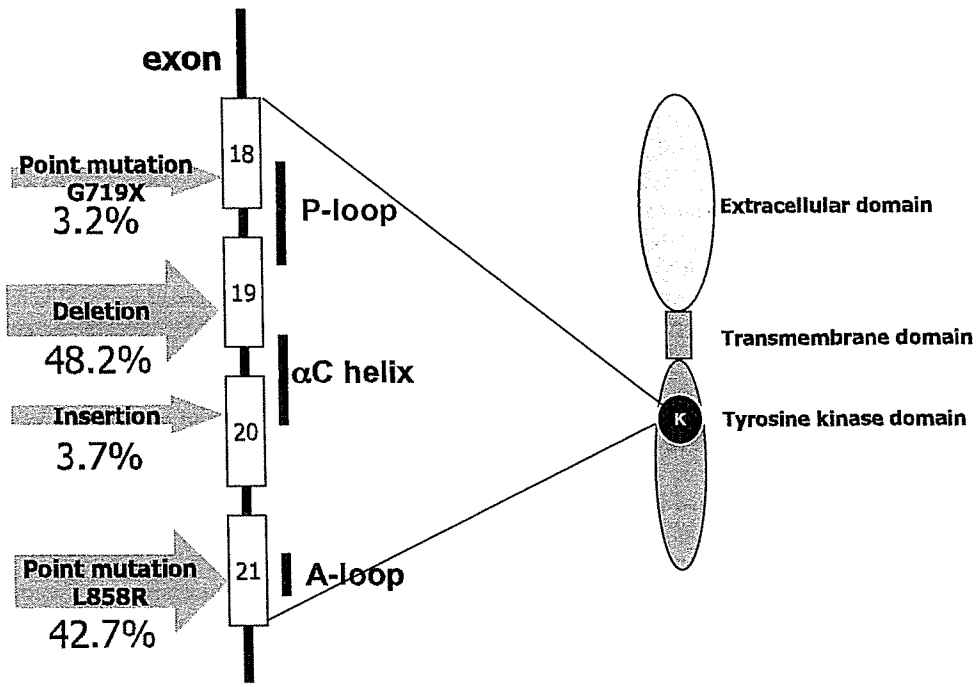

Fig. 2. Distribution of *EGFR* mutations ($n = 569$)[9,10,15,20,21,27,54,59–65]

of 1974 patients taken from previously published analyses (Fig. 1) indicated that the TKI response is significantly dependent on ethnicity, sex, smoking history, and histological type. However, it was not possible to predict gefitinib sensitivity by levels of EGFR overexpression, determined by immunohistochemistry or immunoblotting.[7,8] The factors that determine gefitinib sensitivity have long been an enigma.

In April 2004, two groups of researchers in Boston reported that activating mutations of the *EGFR* gene are present in a subset of NSCLCs and that tumors with *EGFR* mutations are highly sensitive to EGFR-TKI.[9,10] The populations with a higher response to gefitinib described above correspond to those with a higher incidence of *EGFR* mutations. Interestingly, *EGFR* mutations are the first molecular aberration identified as more frequent in nonsmoking patients. In this review, we discuss the current status of EGFR research in relation to kinase inhibitor therapy for lung cancer.

EGFR mutations

Figure 2 shows the distribution of the *EGFR* mutations reported so far ($n = 569$ from 14 studies). Mutations are present in the tyrosine kinase domain of *EGFR*. There are four main types of mutations: point mutations at codon 719 (G719X), deletions in exon 19, insertion mutations in exon 20, and a point mutation at codon 858 in exon 21. In exon

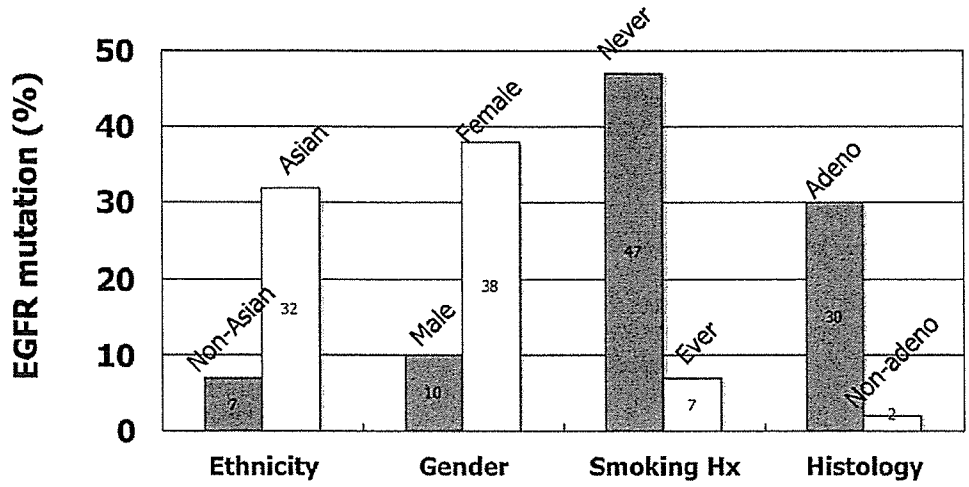

Fig. 3. Incidence of *EGFR* mutations according to particular clinical characteristics (compilation of the data from the literature used for Fig. 2; $n = 2880$). Overall incidence of *EGFR* mutations was 569/2880 = 19.8%

18, mutations are frequent at codon 719 (3.2%) and the patterns of amino acid substitutions are not uniform at this codon, resulting in changes from glycine to cysteine, serine, or alanine. Mutations resulting in the deletion of typically five amino acids at codons 746–750 (ELREA) in exon 19 and a leucine-to-arginine mutation at codon 858 (L858R) are two major types of mutations, which account for 90% of all mutations. These two types of *EGFR* mutations cause increased and sustained phosphorylation of EGFR itself and the phosphorylation of downstream molecules involved in antiapoptotic pathways (PI3K/AKT and STAT).[11] However, *EGFR* mutations have less effect on proliferation through the RAS/RAF/ERK/MAPK pathway.[11] There are several variant types of deletion mutations in exon 19, e.g., a larger deletion, deletion plus point mutation, deletion plus insertion, etc. There are also rarer point mutations and some patients have double mutations, but these usually accompany L858R. Interestingly, it is very rare for double mutations to occur among the four predominant types of mutations.

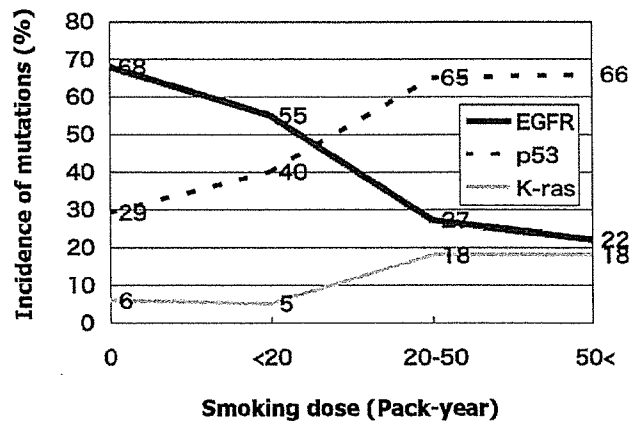

Fig. 4. Incidence of *EGFR*, *KRAS*, and *TP53* gene mutations in pulmonary adenocarcinomas according to smoking dose (data from Kosaka et al.[15])

EGFR mutations and clinical features

Originally, *EGFR* mutations were predominantly found in females, nonsmokers, adenocarcinomas, and Japanese patients.[9,10] Subsequently, many different research groups have confirmed and extended these findings and their results, based on the 2880 mutations reported so far, are summarized in Fig. 3. The strong similarity of the graphs in Figs. 1 and 3 indicates that *EGFR* mutations are frequent in patient subsets that have a high response rate to TKIs.

Previously described genetic changes in lung cancer are almost always more frequent in smokers than nonsmokers. For example, mutations of the *TP53* gene,[12] or *KRAS* genes,[13] or deletion of the short arm of chromosome 3[14] are known to be more frequent in smokers. Indeed, we first showed that the frequency of *EGFR* mutations is inversely associated with smoking dose.[15] When we divided smokers into three categories according to smoke exposure, there was a trend such that the higher the exposure, the lower the incidence of *EGFR* mutations (Fig. 4). This is in distinct contrast to the *KRAS* and *TP53* mutations.

However, it cannot be construed that smoking has a preventive effect on *EGFR* mutations. Instead, it is reasonable to assume that *EGFR* mutations are caused by carcinogen(s) other than those contained in tobacco smoke, and that this apparent negative correlation with smoking dose results from the dilution of EGFR-positive tumors with the increased incidence of tumors with wild-type *EGFR* that occurs as smoking dose increases. This was the case in our recent case–control study of 152 patients with lung cancer and *EGFR* mutations, 283 patients with lung cancer and wild-type *EGFR*, and 2175 age- and sex-matched controls. For example, when the cumulative smoking exposure was divided into three groups, the odds ratio for lung cancer with wild-type *EGFR* increased from 1.00 to 2.72 (1–40 pack years) and further to 10.0 (>40 pack years; $P < 0.001$ for trend). In contrast, the odds ratios for patients with *EGFR* mutations were 1.00, 0.68, and 0.79 ($P = 0.303$ for trend) (K. Matsuo et al., unpublished).

It appears that the marked difference in the incidence of *EGFR* mutations with ethnicity might be at least partially attributable to differences in the incidence of nonsmoking patients among Japanese and American females. In our Japanese cohort, 83% of female patients and 10% of male patients were never-smokers.[15] In contrast, only 15% of 706 U.S. female patients and 6% of 1347 male patients with lung cancer were never-smokers.[16] However, smoking may not be the sole factor explaining these ethnic differences. It is known that the *EGFR* intron 1 polymorphic CA repeat (CA-SSR1) is longer in Asians than in Caucasians[17] and that a longer CA repeat leads to less gene transcription.[18] One can infer that lower transcription of *EGFR* may require mutational activation to obtain growth advantage in Asian patients.

EGFR mutations and pathology

In terms of morphological and pathological features, we found that *EGFR* mutations predominantly occur in adenocarcinomas of the terminal respiratory unit type.[19] We have proposed that these form a characteristic subset of adenocarcinomas putatively originating in the peripheral airway epithelium.[19] They have a papillolepidic growth pattern and frequently express thyroid transcription factor 1 or surfactant apoproteins. Interestingly, some atypical adenomatous hyperplasias, precursor lesions of this type of adenocarcinoma, occasionally harbor *EGFR* mutations, suggesting that *EGFR* mutations occur relatively early in pathogenesis.[19] This observation is closely associated with reports that gefitinib sensitivity is high in patients with adenocarcinomas with bronchioloalveolar cell carcinoma features.[6]

Relationship between *EGFR* mutations and mutations of other cancer genes

We and others have found that *EGFR* mutations never occur in tumors with *KRAS* mutations, thus exhibiting a mutually exclusive relationship.[15,20,21] Furthermore, mutations of *BRAF* and *ERBB2* are present in a very small fraction of adenocarcinomas of the lung (1% and 4%, respectively[22]) and these mutations also have a mutually exclusive relationship with *EGFR* and *KRAS* mutations. It is noteworthy that one of the major downstream pathways from EGFR is the RAS/RAF/ERK/MAPK pathway.

In contrast, *EGFR* mutations and *TP53* mutations appear to occur independently.[15] However, 23 *TP53* mutations relating to tobacco carcinogens (G-to-T transversions, mutations occurring at codons 157, 248, and 273) also have a mutually exclusive relationship with *EGFR* mutations, with two exceptions.[15] Again, this suggests that *EGFR* mutations occur independently of tobacco carcinogens.

EGFR mutations and the TKI response

When *EGFR* mutations were first reported, the most interesting and exciting finding was that patients with this genetic alteration showed a striking response to EGFR-TKIs.[9,10] Figure 5 summarizes the relationship between

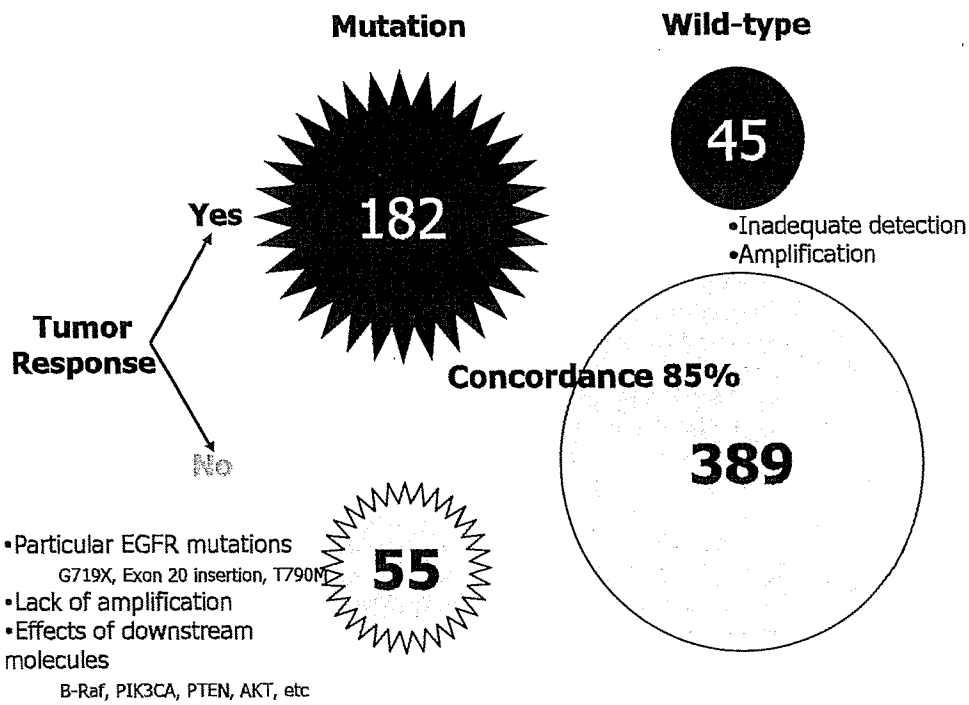

Fig. 5. Relationship between *EGFR* mutations and clinical responses to EGFR-TKI. Numbers of patients in each category are from a compilation of published data (n = 671).[9,10,23–28,31,54,58–60,64,66,67] The positive and negative predictive values of *EGFR* mutations (response rates in patients with or without *EGFR* mutations) were 77% and 10%, respectively. Possible reasons for the discrepancies are also listed

EGFR mutations and the response to EGFR-TKIs in 671 patients, compiled from the literature. In general, about 80% of NSCLCs with *EGFR* mutations respond to EGFR-TKIs, whereas 10% of tumors without *EGFR* mutations do so. Furthermore, several investigators have reported that patients with *EGFR* mutations have a significantly longer survival than those with wild-type *EGFR* when treated with gefitinib or erlotinib.[23–28] These results indicate that *EGFR* mutations are important in determining EGFR-TKI sensitivity. At the same time, they suggest that *EGFR* mutations are not the sole factor determining TKI sensitivity.

We first reported that gefitinib is more effective in patients with deletional *EGFR* mutations than in patients with other types of mutations, predominantly L858R.[23] Figure 6 shows the differences in response rates by classes of *EGFR* mutations from a compilation of 224 patients. The response rates of patients with an exon 19 deletion and L858R were 84% and 71%, respectively. In contrast, only about half the patients with G719X responded to gefitinib. Furthermore, patients with *EGFR* exon 19 deletions had significantly longer MSTs after treatment with erlotinib or gefitinib than those of patients with EGFR L858R (34 vs 8 months, respectively; log-rank $P = 0.01$).[29] Greulich et al. measured erlotinib sensitivity by the inhibition of cell-line transformation in vitro, using various *EGFR* mutant constructs and varying concentrations of erlotinib. They found that the order of sensitivity was: exon 19 deletion = L858R > G719X ≫ exon 20 insertion = wild type, which accords well with the clinical observations described above.[30]

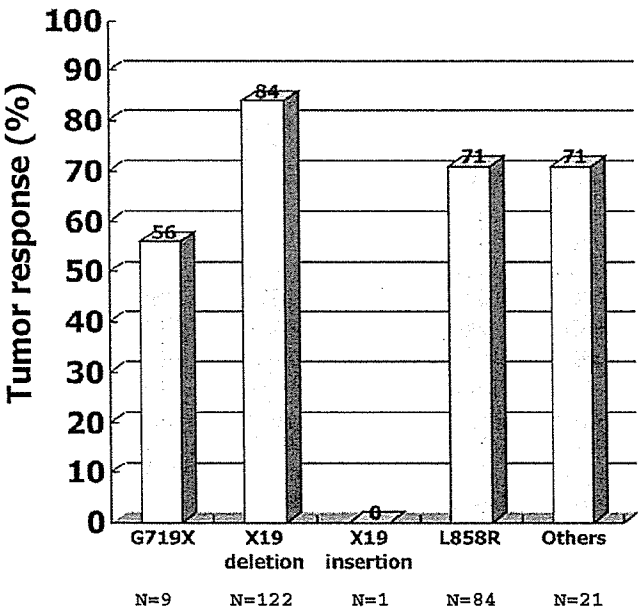

Fig. 6. Response rates to TKIs by classes of *EGFR* mutations (compilation of data from 16 papers used for Fig. 5; $n = 224$)

EGFR gene copy number and TKI sensitivity

In May 2005, Cappuzzo et al. reported that *EGFR* gene amplification, as measured by fluorescence in situ hybridization (FISH), is more predictive of patient survival after gefitinib treatment than *EGFR* mutations.[31] However, this report does not necessarily refute the role of *EGFR* mutations as a predictive factor, because *EGFR* mutations only failed to significantly affect overall survival ($P = 0.09$); *EGFR* mutations were predictive of response rate and time to progression.[31] However, it should be noted that FISH positivity is defined as tumors in which more than 40% of tumor cells have more than four copies (high polysomy) in addition to those with *EGFR* gene amplification. It is biologically unclear whether high polysomy indicates the activation of the *EGFR* gene, resulting in effects similar to those caused by gene amplification. As shown in Table 1, whether mutation or copy number is more predictive of response and useful in patient selection remains controversial. Tsao et al. reported that *EGFR* gene amplification was most predictive of a stronger response and a longer survival in patients who received erlotinib in a phase III clinical trial (BR.21) that compared erlotinib with best supportive care.[32] They concluded that the detection of *EGFR* mutations is not necessary in selecting patients who will benefit from erlotinib therapy.[32] However, many investigators, particularly those in Japan, refute this point. In general, tumors with *EGFR* mutations tend also to have gene amplification.

Table 1. Effectiveness of EGFR-TKIs, and mutation or copy number of the *EGFR* gene in predicting the effectiveness of EGFR-TKIs, reported in selected recently published studies

First author[Ref.]	TKI	*n*	Mutation			Copy number		
			Response	TTP	OS	Response	TTP	OS
Han[24]	G	90	65% vs 14%	Yes	Yes	–	–	–
Mitsudomi[23]	G	59	83% vs 10%	–	Yes	–	–	–
Cappuzzo[31a]	G	89	53% vs 5%	Yes	No	36% vs 3%	Yes	Yes
Bell[47b]	G	79/90	46% vs 10%	Yes	No	29% vs 15%	Yes	No
Tsao[32a]	E	177/125	16% vs 7%	–	No	20% vs 2.4%	–	Yes
Hirsch[48a]	G	100	–	–	–	26% vs 11%	No	Yes
Takano[28b]	G	66	82% vs 11%	Yes	Yes	72% vs 38%	Yes	No

G, gefitinib; E, erlotinib; TTP, time to progression; OS, overall survival
[a] Copy number was examined by fluorescence in situ hybridization
[b] Copy number was examined by quantitative polymerase chain reaction