Hypermethylation of the CpG dinucleotide in epidermal growth factor receptor codon 790: implications for a mutational hotspot leading to the T790M mutation in non−small-cell lung cancer

藤井, 亜希子

https://doi.org/10.15017/1500593

出版情報：九州大学, 2014, 博士（医学）, 課程博士
バージョン：
権利関係：やむを得ない事由により本文ファイル非公開（2）

Hypermethylation of CpG dinucleotide in *epidermal growth factor receptor* codon 790: implications for a mutational hotspot leading to the T790M mutation in non-small-cell lung cancer

Akiko Fujii [a], Taishi Harada [a, *], Eiji Iwama [a, b], Keiichi Ota [a], Kazuto Furuyama [a], Kayo Ijichi [a, c], Tatsuro Okamoto [d], Isamu Okamoto [a, e], Koichi Takayama [a], and Yoichi Nakanishi [a]

[a] Research Institute for Diseases of the Chest, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka, 812-8582, Japan

[b] Faculty of Medical Sciences Department of Comprehensive Clinical Oncology, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka, 812-8582, Japan

[c] Division of Pathophysiological and Experimental Pathology, Department of Pathology, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka, 812-8582, Japan

[d] Department of Surgery and Science, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka, 812-8582, Japan

[e] Center for Clinical and Traditional Research, Kyushu University Hospital, 3-1-1 Maidashi, Higashi-ku, Fukuoka, 812-8582, Japan

***Correspondence:**

Dr. Taishi Harada, Research Institute for Diseases of the Chest, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka, 812-8582, Japan.

Phone: +81-92-642-5378; FAX: +81-92-642-5377

E-mail address: harada-t@kokyu.med.kyushu-u.ac.jp

Abstract

Nearly half of all cases of acquired resistance to epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKIs) for non–small-cell lung cancer (NSCLC) are due to the T790M mutation in *EGFR* exon 20. The T790M mutation is a C → T transition mutation at a CpG dinucleotide. DNA methylation of cytosine (5-methylcytosine; 5-mC) in CpG dinucleotides is a common DNA modification; CpG dinucleotides are considered to be mutational hotspots that cause genetic diseases and cancers through spontaneous deamination of 5-mC, resulting in C → T transition mutations. This study aimed to examine the methylation level of cytosine of *EGFR* codon 790 and investigate whether DNA methylation was involved in acquiring the T790M mutation. We examined 18 NSCLC tumor tissues, 7 normal lymph node tissues and 4 NSCLC cell lines (PC9, HCC827, 11-18 and A549). 5-mC was checked by bisulfite sequencing and quantified by pyrosequencing. We found that all tissue samples and cell lines had 5-mC in *EGFR* codon 790. The 5-mC range was 58.4–90.8%. Our results imply that hypermethylation of the CpG dinucleotide in *EGFR* codon 790 leads to the C → T transition mutation, causing resistance to EGFR-TKI treatment.

Keywords EGFR-TKI, non-small-cell lung cancer, T790M, methylation, resistance

Introduction

Epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKIs) have dramatic anti-tumor effects on non–small-cell lung cancer (NSCLC) in patients harboring activating *EGFR* mutations, such as in-frame deletion in exon 19 and L858 point mutation (L858R) in

exon 21 (1, 2). However, despite initial response, these patients tend to relapse after around 10 months from development of acquired resistance (2). Nearly half of all cases of acquired resistance to EGFR-TKIs are due to a secondary mutation, T790M, which substitutes methionine for threonine at position 790 in the kinase domain (2-4). The T790M mutation is a gatekeeper mutation, which alters the binding kinetics of EGFR-TKIs to the target molecule (2, 5). The T790M mutation has a higher incidence than that of gatekeeper mutations for other receptor tyrosine kinases such as BCR-ABL or KIT (6-9). Furthermore, the T790M mutation reportedly exists at low frequency within tumor cells before EGFR-TKI treatment (2-4). These observations imply that some underlying mechanism leads to frequent acquisition of the T790M mutation. Elucidating this mechanism could increase expectations for EGFR-targeted therapies in NSCLC.

DNA methylation at the 5-carbon of cytosine bases (5-methylcytosine; 5-mC) in CpG dinucleotides is a common DNA modification in mammalian genomes (10, 11). 5-mC at CpG dinucleotides is widely known as a mutational hotspot, leading to C → T transition mutations (12). A third of all transition mutations are thought to be because of C → T transition mutations at these sites, instigating various genetic diseases and cancers through gene silencing and mutagenesis (10, 12, 13). In germline and colon tumors, approximately 47% of mutations that inactivate the p53 tumor suppressor gene are C→T transition mutations at CpG dinucleotides; five of the six mutational hotspots between exon 5 and exon 8 are located within CpG dinucleotides (13).

The T790M mutation is a C→T transition mutation at a CpG dinucleotide, causing a missense mutation from ACG (threonine) to ATG (methionine) (2-4). We hypothesized that CpG dinucleotide of *EGFR* codon 790 is methylated and this site is a mutational hotspot, leading to the T790M mutation. To study 5-mC in *EGFR* codon 790, we examined 18

NSCLC tumor tissues, 7 normal lymph node tissues and 4 NSCLC cell lines (PC9, HCC827, 11-18, and A549) using bisulfite sequencing and pyrosequencing. Herein, we present the high methylation level of cytosine in *EGFR* codon 790.

Materials and Methods

Patient samples.

In the present study, we analyzed specimens from 18 patients with NSCLC who had undergone surgical resection for lung cancer at the Department of Surgery and Science, Kyushu University Hospital, Japan, from January 2008 to January 2011. All 18 tumors were confirmed to be lung adenocarcinoma. Formalin-fixed paraffin-embedded (FFPE) tissues were collected from each patient. None of these patients underwent preoperative chemotherapy, preoperative radiotherapy, or preoperative chemoradiotherapy. The study design was approved by the ethics reviews board of our university. Informed consent for the use of tissue specimens was obtained from all patients.

Cell culture.

We selected PC9 (*EGFR* Del E746_A750; Ex19-del), HCC827 (Ex19-del), 11-18 (*EGFR* L858R) and A549 (*EGFR* Wild type; WT) cells. PC9, HCC827 and 11-18 cells were cultured in RPMI 1640 medium (Gibco, Carlsbad, CA) supplemented with 10% fetal bovine serum (FBS), 100 U/mL penicillin (PC) and 1 µg/mL streptomycin (SM) (Gibco). A549 cells were maintained in DMEM medium (Gibco) supplemented with 10% FBS, 100 U/mL PC and 1 µg/mL SM. All cells were grown at 37°C in a humidified atmosphere of 5% CO_2.

DNA extraction and bisulfite modification.

All 18 tumor specimens were obtained by tumor dissection. Normal lymph node tissues were obtained from 7 out of 18 patients by lymphadenectomy. Genomic DNA samples from FFPE blocks of these specimens were prepared by slicing several paraffin sections, adding xylene to dissolve the paraffin and extracting DNA using QIAamp DNA FFPE Tissue Kit (QIAGEN, Hilden, Germany). Genomic DNA from cell lines was extracted using NucleoSpin Tissue (Takara, Ohtsu, Japan). Bisulfite treatment for DNA was performed using MethylEasy Xceed Rapid DNA Bisulphite Modification Kit (Takara). All these procedures were done according to manufacturers' instructions.

Mutant–enriched PCR assay for *EGFR* T790M.

A mutant-enriched polymerase chain reaction (PCR) assay for *EGFR* T790M was a two-step PCR with intermittent restriction enzyme digestion to selectively eliminate wild-type genes, enriching the mutated genes. The assay was performed as described previously (3). Briefly, the first PCR were performed using the primers (5′-ACTGACGTGCCTCTCCCTCC-3′ as forward and 5′-CGAAGGGCATGAGCCGC-3′ as reverse). The first PCR products were digested with *Bst*UI to remove the wild-type products. The second PCR was performed using a forward primer (5′-CCTCCAGGAAGCCTACGTGA-3′) and the same reverse primer used in the first PCR. The products were purified and directly sequenced.

Bisulfite sequencing.

Semi-nested PCR for *EGFR* codon 790 was carried out using bisulfite-treated DNA. The primers for the first PCR were (forward) 5′-TGTTGGGTATTTGTTTTATTTTTAT-3′ and (reverse) 5′-ACAAACTCTTACTATCCCAAAAAC-3′. Initial PCR conditions were 94°C for 1 minute; followed by 35 cycles at 94°C for 30 seconds, at 55°C for 30 seconds, at 72°C for 30 seconds; and a final extension at 72°C for 5 minutes. The second PCR was carried out using the same forward primer used in the initial PCR and a reverse primer (5′-AAAAATCCTAACTCCTTATCTCCC-3′). The PCR conditions were 94°C for 1 minute; followed by 35 cycles at 94°C for 30 seconds, at 57°C for 30 seconds, at 72°C for 30 seconds; and a final extension at 72°C for 5 minutes. The PCR products were electrophoresed on 2.0% agarose gels and purified by QIAquick PCR Purification Kit (QIAGEN). Cycle sequencing reactions were performed using the reverse primer for the second PCR and the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) with ABI 3130×1 DNA Sequencer (Applied Biosystems).

Pyrosequencing analysis.

To quantify 5-mC of *EGFR* codon 790, pyrosequencing analysis was performed. 4 CpG dinucleotides including codon 790 in *EGFR* exon 20 were examined. PCR preparations for pyrosequencing and pyrosequencing reactions were conducted by EpigenDx (Hopkinton, MA). These samples were amplified by PCR using primer pairs designed by EpigenDx, one of which was biotinylated. The PCR products were sequenced using the Pyrosequencing PSQ96 HS System (QIAGEN) following the manufacturer's instructions (QIAGEN). The methylation status of targeted loci was analyzed as T/C SNPs using QCpG software (QIAGEN).

Plasmid construction and site-directed mutagenesis.

Ex19-del and Ex19-del/T790M yellow fluorescent protein (YFP)-tagged fragments of the EGFR intracellular domain (YFP-EGFR-ICD) constructs were generated as described previously (14-16). To construct *EGFR* mutants, QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA) was used according to the manufacturer's protocol with WT YFP-EGFR-ICD as a template. Primers for the R803W mutation are 5′-CTCCTGGACTATGTCTGGGAACACAAAGACAATATTG-3′ and 5′-CAATATTGTCTTTGTGTTCCCAGACATAGTCCAGGAG-3′.

Cell culture, transfection and drug treatment.

Human breast cancer cells MCF-7 were maintained in DMEM medium (Gibco) supplemented with 10% FBS, 100 U/mL PC and 1 µg/mL SM and grown at 37°C in a humidified atmosphere of 5% CO_2. Cells were seeded in twelve-well trays and transfected with 2.5 µg of plasmid DNA using Xfect transfection reagent (Takara), following the manufacturer's protocol. Gefitinib (AG Scientific, San Diego, CA) was added at the indicated concentration 4 hours after transfection, and the cells were incubated 20 hours before being processed for immunofluorescence analysis. Gefitinib treatment was always performed in standard culture medium containing 10% FBS.

Immunofluorescence and microscopy analysis.

To evaluate the sensitivity to gefitinib, MCF-7 cells transfected with plasmids encoding for Ex19-del, Ex19-del/T790M and Ex19-del/R803W were examined with a Keyence BZ-8100 fluorescence microscope (Keyence, Osaka, Japan). BZ Analyzer software (Keyence) was used to collect images. Exposure time was kept constant to allow for comparison of the signal intensities among different samples.

Statistical analysis.

All statistical analysis was performed using JMP 9.0 software (SAS Institute Inc., Cary, NC). *P*-values were two-tailed; $P < 0.05$ was considered statistically significant.

Results

Patient characteristics

A total of 18 NSCLC patients were analyzed. The characteristics of NSCLC patients are shown in Table 1. Mean age (± standard deviation) was 67.9 ± 11.3 years old. Eleven patients (61.1%) were stage I, 4 (22.2%) were stage II, 2 (11.1%) were stage III and 1 (5.5%) was stage VI. Three patients (16.6%) were smokers and 15 (83.3%) had never smoked. *EGFR* mutations in exon 19 (Ex19-del), exon 21 (L858R) and exon 20 (T790M) were analyzed at contracted laboratory (Bio Medical Laboratories, Tokyo, Japan), using PCR invader assay. Thirteen patients (72.2%) had *EGFR* mutations: 4 (22.2%) and 9 (50%) patients had Ex19-del and L858R respectively. None of the 18 patients had T790M mutations.

Hypermethylation of CpG dinucleotide in *EGFR* codon 790

We used bisulfite sequencing to investigate cytosine methylation status of *EGFR* codon 790. Bisulfite sequencing could not discriminate 5-mC (A5-mCG) from thymine (ATG) in *EGFR* codon 790. To find whether the T790M mutation existed in a minor tumor population, we first checked for T790M mutations in tumor tissues using a mutant-enriched PCR assay, which can detect one copy of mutant allele among 10^3 copies of wild-type alleles (2-4). This sensitive assay did not detect the T790M mutation in any tumor tissue (Table 1). We next performed bisulfite sequencing. We found cytosine of *EGFR* codon 790 to be methylated in all tissue samples and 4 cell lines (Table 1 and Figure 1A).

To quantify the precise proportion of 5-mC in *EGFR* codon 790, we performed pyrosequencing. 4 CpG dinucleotides including codon 790 in *EGFR* exon 20 were examined (Figure 1B). The range of 5-mC in *EGFR* codon 790 was 58.4 -90.8% (Table 2 and Figure 1C). Hypermethylation did not significantly differ among NSCLC cell lines (range: 85.0 -86.6%), tumor tissues (range: 58.4 -89.2%) and non-malignant tissues (range: 64.9 -90.8%; $P = 0.434$). Furthermore, the proportion of 5-mC in *EGFR* mutation positive (range: 60.2 -87.8%) and negative tumor tissues (range: 58.4 -89.2%) did not significantly differ ($P = 0.178$). Other three CpG dinucleotides were also hypermethylated (Table 2, Figure 1B and 1C).

Sensitivity to Gefitinib.

Pyrosequencing demonstrated that 3 CpG dinucleotides besides codon 790 in *EGFR* exon 20 were hypermethylated (Table 2). When C → T transition mutations at these CpG dinucleotides (cytosine in codon 785, 795 and 803) occurs, amino acid substitution is caused only in codon 803, changing from arginine to tryptophan (R803W) (Figure 1B). To evaluate

sensitivity to gefitinib, we used an YFP-EGFR-ICD based assay. As reported previously, this assay is a useful method of evaluating the sensitivity of EGFR-TKI of novel EGFR mutants (16). The YFP-tagged EGFR fragments lacked the extracellular and juxtamembrane domains of the receptor. Consequently, we could reduce interference from the experimental context and introduce mutagenesis more efficiently to shorter EGFR fragments. When treated with EGFR-TKI, these YFP-EGFR-ICD fusion proteins relocate to fibril-like formation. Although the mechanism of this relocation is unclear, it parallels the sensitivity to the EGFR-TKI (14-16). In this study, when 70% of cells changed the fusion protein location, we judged that there was sensitivity to EGFR-TKI (15). We introduced Ex19-del, Ex19-del /T790M and Ex19-del/R803W mutations into the WT YFP-EGFR-ICD vector. MCF-7 human breast cancer cells were transfected with the plasmids encoding these three mutations. At four hours after transfection, gefitinib at a final concentration ranging from 20 nM to 1μM was added to culture medium and incubated for 20 hours. Although gefitinib had no effect on Ex19-del /T790M YFP-EGFR-ICD mutant transfected cells, low concentration of gefitinib (20nM) induced relocation of Ex19-del and Ex19-del /R803W YFP-EGFR-ICD mutant transfected cells (Figure 2). The C → T transition mutation in codon 803 could not confer resistance to EGFR-TKI.

Discussion

Spontaneous deamination of cytosine and 5-mC is a major source of transition mutations. 5-mC spontaneously deaminates and changes to thymine, leading to a T:G mismatch (Figure 3A), whereas uracil yielded by spontaneously deaminated cytosine leads to a U:G mismatch (12). The deamination rates of 5-mC is 2.0- to 3.2- fold more than cytosine (17). Uracil is efficiently repaired by uracil DNA glycosylase (Figure 3B) (12, 18, 19). T:G mismatches are

also recognized by thymine DNA glycosylase, correcting back to C:G (12). However, because thymine is a normal component of DNA, T:G mismatches are often corrected to T:A (12). Since the repair of T:G mismatches is not efficient, the transition rate of C→T transition mutation at CpG dinucleotides is 10- to 50- fold higher than other transitional changes (11, 17) . Accumulating evidence has shown correlation between CpG dinucleotides and somatic mutations in cancer. Moreover, a recent analysis of more than 7000 cancer genomes has demonstrated that CpG dinucleotides are closely associated with somatic mutations found in various cancer types, including lung adenocarcinoma (20).

In the present study, we found that cytosine of *EGFR* codon 790 is highly methylated. This result supports our hypothesis that a CpG dinucleotide of *EGFR* codon 790 can be a mutational hotspot because of DNA methylation. We also found that the other 3 CpG dinucleotides in *EGFR* exon 20 were all methylated and showed similar methylation level. However, C to T changes in codon 785 and 795 were silent mutations. Furthermore, despite substitution from arginine to tryptophan in codon 803, the sensitivity of Ex 19-del/R803W to EGFR-TKI was almost the same as that of Ex 19-del. Resistance to EGFR-TKI treatment owing to the T790M mutation is thought to occur step by step: diverse point mutations may randomly occur throughout *EGFR* gene by chance in minor population, and then a clone harboring mutations, which can confer resistance to EGFR-TKI, is selected and increases in population during EGFR-TKI treatment. The initial step is not truly random, but may be biased because of hypermethylation of CpG dinucleotides in *EGFR* gene. Although C→T transition mutations at CpG dinucleotides spontaneously occur in *EGFR* gene, most of these mutations may not be involved in acquiring resistance to EGFR-TKI. Therefore, a clone harboring the T790M mutation may favorably increase in population under EGFR-TKI treatment.

The T315I mutation in *BCR-ABL* in chronic myeloid leukemia (CML) and Philadelphia-positive acute lymphoblastic leukemia (Ph+ALL), the T670I mutation in *KIT* in gastrointestinal stromal tumors (GIST) and the L1196M mutation in *EML4-ALK* -positive NSCLC are gatekeeper mutations, homologous to the T790M mutation (5, 21). T315I and T670I mutations are associated with acquired resistance to imatinib, and the L1196M mutation with crizotinib (21, 22). Although secondary mutations leading to EGFR-TKI resistance are sparsely reported except for the T790M mutation in NSCLC patients, many secondary mutations are associated with acquired resistance to imatinib and crizotinib (Table 3). The secondary mutation, *BCR-ABL* T315I is reported to be found in 12-20% of imatinib-resistant CML and Ph+ ALL patients who acquired secondary mutations (7, 8), whereas *KIT* T670I is reportedly found in 14-29% of imatinib-resistant GIST patients who acquired secondary mutations (6, 9), and *ALK* L1196M is found in 5-18% of crizotinib-resistant *EML4-ALK*-positive NSCLC patients (5, 23). The frequency of *EGFR* T790M is remarkably higher compared with other gatekeeper mutations.

As shown in Table 3, *BCR-ABL* codon 315, *KIT* codon 670 and *EML-ALK* codon 1196 do not each contain a CpG dinucleotide. Only *EGFR* codon 790 has a CpG dinucleotide. The different frequency of these gatekeeper mutations may be explained in part by the existence of CpG dinucleotide. We observed cytosine hypermethylation in *EGFR* codon 790 in all tissue samples. These results suggest that the cytosine of the CpG dinucleotide in *EGFR* codon 790 is hypermethylated ubiquitously. The T790M mutation reportedly exists not only as a somatic mutation in tumor cells but also as a germline mutation (24). Our results also imply that the hypermethylation state of *EGFR* codon 790 may affect the incidence of occurrence of the T790M mutation in germline cells.

We also found that the proportion of 5-mC in *EGFR* codon 790 in some tissues was relatively low compared with other tissues. The difference in proportion of 5-mC in *EGFR* codon 790 may affect incidence of acquiring resistance to EGFR-TKIs. However, because DNA was extracted from FFPE specimens, low DNA quality might affect the results of analysis for methylation status. In an experimental model, EGFR-TKI-sensitive PC9 cells acquired resistance to a TKI mainly through the T790M mutation (2, 25). However, HCC827 cells acquire resistance to EGFR-TKIs mainly due to c-MET amplification (26). The present study showed that both cell lines had similar methylation profiles for cytosine in *EGFR* codon 790. Our hypothesis cannot explain the different nature of both cell lines. There may be other factors affecting the T790M mutation in NSCLC, such as a lack of T:G mismatch repair system.

This study demonstrated hypermethylation state of cytosine in *EGFR* codon 790. Our findings imply that *EGFR* codon 790 is a mutational hotspot, where ACG tends to be substituted for ATG, resulting in theT790M mutation. These results provide clues to overcoming acquired resistance to EGFR-TKIs in NSCLC. Further investigation is warranted into the underlying mechanism of the T790M mutation.

Acknowledgments

This study was performed as a research program of the Project for Development of Innovative Research on Cancer Therapeutics (P-DIRECT), Ministry of Education, Culture, Sports, Science and Technology of Japan and partially supported by Clinical Research Foundation (2013). We appreciate the technical support from the Research Support Center, Graduate School of Medical Sciences, Kyushu University, Japan.

Conflict of Interest

The authors have no potential conflicts of interest.

References

1. Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med. 2009;361:947-57.

2. Ogino A, Kitao H, Hirano S, Uchida A, Ishiai M, Kozuki T, et al. Emergence of epidermal growth factor receptor T790M mutation during chronic exposure to gefitinib in a non small cell lung cancer cell line. Cancer Res. 2007;67:7807-14.

3. Inukai M, Toyooka S, Ito S, Asano H, Ichihara S, Soh J, et al. Presence of epidermal growth factor receptor gene T790M mutation as a minor clone in non-small cell lung cancer. Cancer Res. 2006;66:7854-8.

4. Ma C, Wei S, Song Y. T790M and acquired resistance of EGFR TKI: a literature review of clinical reports. J Thorac Dis. 2011;3:10-8.

5. Doebele RC, Pilling AB, Aisner DL, Kutateladze TG, Le AT, Weickhardt AJ, et al. Mechanisms of resistance to crizotinib in patients with ALK gene rearranged non-small cell lung cancer. Clin Cancer Res. 2012;18:1472-82.

6. Antonescu CR, Besmer P, Guo T, Arkun K, Hom G, Koryotowski B, et al. Acquired resistance to imatinib in gastrointestinal stromal tumor occurs through secondary gene

mutation. Clin Cancer Res. 2005;11:4182-90.

7. Nicolini FE, Corm S, Le QH, Sorel N, Hayette S, Bories D, et al. Mutation status and clinical outcome of 89 imatinib mesylate-resistant chronic myelogenous leukemia patients: a retrospective analysis from the French intergroup of CML (Fi(phi)-LMC GROUP). Leukemia. 2006;20:1061-6.

8. Soverini S, Colarossi S, Gnani A, Rosti G, Castagnetti F, Poerio A, et al. Contribution of ABL kinase domain mutations to imatinib resistance in different subsets of Philadelphia-positive patients: by the GIMEMA Working Party on Chronic Myeloid Leukemia. Clin Cancer Res. 2006;12:7374-9.

9. Wardelmann E, Merkelbach-Bruse S, Pauls K, Thomas N, Schildhaus HU, Heinicke T, et al. Polyclonal evolution of multiple secondary KIT mutations in gastrointestinal stromal tumors under treatment with imatinib mesylate. Clin Cancer Res. 2006;12:1743-9.

10. Jiang M, Zhang Y, Fei J, Chang X, Fan W, Qian X, et al. Rapid quantification of DNA methylation by measuring relative peak heights in direct bisulfite-PCR sequencing traces. Lab Invest. 2010;90:282-90.

11. Zhao Z, Jiang C. Methylation-dependent transition rates are dependent on local sequence lengths and genomic regions. Mol Biol Evol. 2007;24:23-5.

12. Schmutte C, Yang AS, Beart RW, Jones PA. Base excision repair of U:G mismatches at a

mutational hotspot in the p53 gene is more efficient than base excision repair of T:G mismatches in extracts of human colon tumors. Cancer Res. 1995;55:3742-6.

13. Tornaletti S, Pfeifer GP. Complete and tissue-independent methylation of CpG sites in the p53 gene: implications for mutations in human cancers. Oncogene. 1995;10:1493-9.

14. de Gunst MM, Gallegos-Ruiz MI, Giaccone G, Rodriguez JA. Functional analysis of cancer-associated EGFR mutants using a cellular assay with YFP-tagged EGFR intracellular domain. Molecular cancer. 2007;6:56.

15. Furuyama K, Harada T, Iwama E, Shiraishi Y, Okamura K, Ijichi K, et al. Sensitivity and kinase activity of epidermal growth factor receptor (EGFR) exon 19 and others to EGFR-tyrosine kinase inhibitors. Cancer Sci. 2013;104:584-9.

16. Harada T, Lopez-Chavez A, Xi L, Raffeld M, Wang Y, Giaccone G. Characterization of epidermal growth factor receptor mutations in non-small-cell lung cancer patients of African-American ancestry. Oncogene. 2011;30:1744-52.

17. Xia J, Han L, Zhao Z. Investigating the relationship of DNA methylation with mutation rate and allele frequency in the human genome. BMC Genomics. 2012;13 Suppl 8:S7.

18. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409:860-921.

19. Yang SY, Dobkin C, He XY, Philipp M, Brown WT. A 5-methylcytosine hotspot

responsible for the prevalent HSD17B10 mutation. Gene. 2013;515:380-4.

20. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature. 2013;500:415-21.

21. Pao W, Miller VA, Politi KA, Riely GJ, Somwar R, Zakowski MF, et al. 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.

22. Choi YL, Soda M, Yamashita Y, Ueno T, Takashima J, Nakajima T, et al. EML4-ALK mutations in lung cancer that confer resistance to ALK inhibitors. N Engl J Med. 2010;363:1734-9.

23. Katayama R, Shaw AT, Khan TM, Mino-Kenudson M, Solomon BJ, Halmos B, et al. Mechanisms of acquired crizotinib resistance in ALK-rearranged lung Cancers. Sci Transl Med. 2012;4:120ra17.

24. Oxnard GR, Miller VA, Robson ME, Azzoli CG, Pao W, Ladanyi M, et al. Screening for germline EGFR T790M mutations through lung cancer genotyping. J Thorac Oncol. 2012;7:1049-52.

25. Shien K, Toyooka S, Yamamoto H, Soh J, Jida M, Thu KL, et al. Acquired Resistance to EGFR Inhibitors Is Associated with a Manifestation of Stem Cell-like Properties in Cancer Cells. Cancer Res. 2013;73:3051-61.

26. Engelman JA, Zejnullahu K, Mitsudomi T, Song Y, Hyland C, Park JO, et al. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. Science. 2007;316:1039-43.

27. Elias MH, Baba AA, Husin A, Abdullah AD, Hassan R, Sim GA, et al. Contribution of BCR-ABL kinase domain mutations to imatinib mesylate resistance in Philadelphia chromosome positive Malaysian chronic myeloid leukemia patients. Hematol Rep. 2012;4:e23.

Figure legends

Figure 1: Methylation analysis of *EGFR* exon 20. **(A)** Bisulfite sequencing for *EGFR* codon 790 sequenced with a reverse primer. Underscored sequence: codon 790 (ACG). *: 5-methylcytosine at the second base of *EGFR* codon 790. **(B)** Target CpG dinucleotides analyzed by pyrosequencing. #2355, #2369, #2385 and #2407 are nucleotides from the initiation ATG of *EGFR* gene. **(C)** Representative pyrograms for CpG dinucleotides of *EGFR* exon 20 in tumor tissues. The abscissa shows the dispensation order (“E” and “S” means controls). “C” in the boxes means methylated cytosine. “T” in the boxes means unmethylated cytosine in CpG dinucleotides. “Y” in the sequence to analyze represents cytosine or thymine.

Figure 2: Evaluation of sensitivity of mutations in exon 19 to gefitinib.

YFP-EGFR-ICD-transfected cells were treated with gefitinib at the indicated concentrations for 20 hours. Cells were subjected to image analysis by fluorescence microscopy. YFP signal relocation parallels the sensitivity to gefitinib. Ex19-del YFP-EGFR-ICD-transfected cells showed sensitivity to gefitinib at 20nM-1μM. Ex19-del /T790M YFP-EGFR-ICD-transfected cells did not show any relocation at 1μM gefitinib. Ex19-del /R803W YFP-EGFR-ICD-transfected cells showed YFP signal relocation at gefitinib at 20nM-1μM as low concentrations as Ex19-del YFP-EGFR-ICD-transfected cells. YFP, yellow fluorescent protein; YFP-EGFR-ICD, YFP-tagged fragments of the EGFR intracellular domain; Ex19-del, Del 746_750 in exon 19.

Figure 3: Deamination of 5-methylcytosine and cytosine. **(A)** Spontaneous deamination of 5-methylcytosine to thymine. **(B)** Spontaneous deamination of cytosine to uracil; uracil is repaired to cytosine by uracil-DNA glycosylase.