Screening of Epidermal Growth Factor Receptor Gene Mutation in Non-Small Cell Lung Cancer Using a PCR-Based Enzymatic Digestion Method

Purpose: We applied a simplified method using polymerase chain reaction (PCR)-based enzymatic digestion for the detection of epidermal growth factor receptor (EGFR) mutation. Materials and Methods: We selected 74 samples of adenocarcinoma of the lung with EGFR exons 19 and 21 that had been previously sequenced. We designed PCR primers and chose a DNA restriction enzyme. Seventy four additional lung cancer samples were tested as a test set. For test sets, the PCR-based method was performed first, followed by validation of the result by DNA sequencing. Results: In the first sample group, we found 15 (20.3%) mutations in exon 19, and 9 (12.2%) mutations in exon 21 using the sequencing method. By using the PCR-based method, we were able to identify all of the mutated samples detected by the sequencing method. The PCR-based method also detected mutations in exon 19 in three additional samples and in exon 21 in one additional sample. In the second sample group, by performing the PCR-based method, we found 10 (13.5%) and 7 (9.5%) mutations in exons 19 and 21, respectively. Additional mutations in exon 19 were identified in 2 samples by the sequencing method. However, the sequencing method failed to identify a mutation in exon 21 in one sample. Conclusion: The sensitivity of the PCR-based enzymatic digestion method seems to be comparable to that of the traditional sequencing method for detecting EGFR mutations. Our method can be widely used as a screening test to select patients who may benefit from EGFR targeted therapy. (J Lung Cancer 2012;11(2):77–83)

Key Words: Lung neoplasms, EGFR genes, Mutation, Polymerase chain reaction, Restriction mapping

INTRODUCTION

Among the various molecular targeted therapeutic drugs, epidermal growth factor receptor tyrosine kinase inhibitor (EGFR TKI) is one of the most important agents for the treatment of non-small cell lung cancer (NSCLC) (1). Not all patients, however, respond to EGFR TKIs. Since only a
subgroup of patients with NSCLC experience a clinical benefit from treatment with EGFR inhibitors, the identification of the patients who are most likely to respond to these drugs is very important. Clinical features have been known to be the most important parameters for such prediction. Clinical studies of gefitinib or erlotinib in NSCLC suggest that patients that are most likely to have a positive clinical response to anti-EGFR therapy (1-3) are non-smoking Asian, females with tumors that have histologic features of adenocarcinoma. The subsequent discovery that somatic mutations within the tyrosine kinase of the EGFR gene that are associated with a response to EGFR TKIs in NSCLC suggest that the selection of patients through molecular screening might be feasible (4,5). Publications have documented various mutations of EGFR genes. Approximately 90% of EGFR mutations affect small regions of exons 18 to 24 that code for the EGFR tyrosine kinase domain. Among them, the most common mutations are in-frame deletion in exon 19 around the codons 746 to 750, and a missense mutation leading to a substitution of arginine for leucine at codon 858 (L858R) in exon 21 (6). Therefore, mutations of the EGFR gene in exons 19 and 21 may be useful molecular markers for predicting a good response to EGFR TKIs in NSCLC.

At the present time, the most common method to detect such mutations is the direct sequencing of DNA that has been isolated from tumor cells. However, this method is complicated and time-consuming, limiting its wide use for screening purposes. Also, there are regions in the world where the sequencing method is not available. We have designed a simplified method using polymerase chain reaction (PCR)-based enzymatic digestion for the detection of mutations in exons 19 and 21, and have validated its usefulness as an initial screening tool.

**MATERIALS AND METHODS**

1) Sample selection

Among the patients who underwent surgical resection for primary lung cancer between December 2000 and May 2006 at Seoul National University Hospital, we selected 148 patients whose frozen specimens were available from our Lung Cancer Tissue Bank.

Informed consent for tissue collection and gene analyses for research purposes was obtained from individual patients pre-operatively, according to the policy of the Lung Cancer Tissue Bank at The Cancer Research Institute of Seoul National University. We also followed the recommendations of the Declaration of Helsinki for biomedical research involving human subjects. The study protocol, as well as ethical issues, were reviewed and approved by Seoul National University Hospital IRB (H-0504/147-016).

2) Tissue sample collection and DNA extraction

Tissue sampling was performed immediately after lung resection, and a gross dissection of tumor tissue was performed. The samples were quickly frozen in liquid nitrogen and stored at −80°C until the analysis. Genomic DNA samples from the frozen lung cancer tissue were isolated by using the QIAmp tissue kit (Qiagen, Hilden, Germany), according to manufacturer’s protocol.

3) PCR amplification and direct sequencing

For EGFR mutations, 10 ng of DNA were amplified in a 20 μL reaction solution, containing 2 μL of 10× buffer (Roche, Mannheim, Germany), 1.7 to 2.5 mM/L of MgCl₂, 0.3 μM of each primer pairs (exon 19, F: 5’-ccagtgtcctcaccttc, R: 5’-gcagggtctagagcagagca; exon 21, F: 5’-gctcagagcctggcatggaa, R: 5’-catcctcccctgcatgtgt), 250 μM of deoxynucleoside triphosphate, and 2.5 units of DNA polymerase (Roche). Amplifications for exon 19 were performed using a 15-minute initial denaturation at 95°C, followed by 30 cycles of 30 seconds at 94°C, 30 seconds at 55°C, and 45 seconds at 72°C, and a 10-minute final extension at 72°C.

Amplifications for exon 21 were performed using a 15-minute initial denaturation at 95°C, followed by 32 cycles of 45 seconds at 94°C, 45 seconds at 60°C, and 45 seconds at 72°C, and a 7-minute final extension at 72°C.

PCR products were then 2% gel-purified with a QIAgen gel extraction kit (Qiagen). DNA templates were processed for the DNA sequencing reaction using the ABI-PRISM BigDye Terminator version 3.1 (Applied Biosystems, Foster, CA, USA) with both forward and reverse sequence-specific primers. Twenty nanograms of purified PCR products were used in a 20 μL sequencing reaction solution containing 8 μL of BigDye Terminator version 3.1 and 0.1 μM of the same PCR primer. Sequencing reactions were performed using a 2-minute initial denaturation at 96°C, followed by 25 cycles of 10 seconds at
Table 1. *EGFR* Mutations in Exons 18∼21 Tested by DNA Sequencing in the Training Set

<table>
<thead>
<tr>
<th>Exon</th>
<th>Type of mutation</th>
<th>Nucleotide number and sequence (5’-3’)</th>
<th>Amino acid</th>
<th>Number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>Single-nucleotide substitution</td>
<td>2126A&gt;C and 2155G&gt;A</td>
<td>E709A and G719S</td>
<td>26 (35.1)</td>
</tr>
<tr>
<td>19</td>
<td>In-frame deletion</td>
<td>2233∼2247 del AAGGAATTAAGAGA</td>
<td>K745∼E749 del</td>
<td>1 (1.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2235∼2249 del GGAATTAAGAGAAGGC</td>
<td>E746∼A750 del</td>
<td>3 (4.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2236∼2250 del GAATTAAGAGAAGCA</td>
<td>E746∼A750 del</td>
<td>2 (2.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2237∼2251 del ATTAAGAGAAGCAG</td>
<td>E746∼T751 del</td>
<td>5 (6.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2240∼2257 del TAAGAGAAGCAACATCTC</td>
<td>L747∼P753 del</td>
<td>3 (4.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2248∼2256 del GCAACATCTCC</td>
<td>A750∼S752 del</td>
<td>1 (1.4)</td>
</tr>
<tr>
<td>20</td>
<td>In-frame insertion</td>
<td>2321∼2326 ins CCCACG</td>
<td>A772∼H773 ins</td>
<td>1 (1.4)</td>
</tr>
<tr>
<td>21</td>
<td>Single-nucleotide substitution</td>
<td>2573T&gt;G</td>
<td>L858R</td>
<td>9 (12.2)</td>
</tr>
</tbody>
</table>

*EGFR*: epidermal growth factor receptor, del: deletion, ins: insertion.

94°C, 15 seconds at 50°C, and 3 minutes at 60°C. Sequence data were generated with the ABI PRISM 3100 DNA analyzer (Applied Biosystems). Sequences were analyzed by Sequencer 3.1.1. software (Applied Biosystems) to compare variations.

4) PCR-based and enzymatic digestion method

Based on the initial *EGFR* sequencing result (Table 1) for 74 patients, we designed PCR primers with the capability to detect mutations without sequencing. For exon 19, we selected primers encompassing the commonly deleted sequences and the resulting size of PCR products for the wild type to be 157 bp. The correct size of the amplified PCR products was checked by electrophoresis on 3.5% metaphor agarose gel. If there was a deletion in exon 19, two bands, one band 157 bp in size and the other smaller, were observed. If there was no mutation, only a single band of 157 bp was observed (Fig. 1).

For exon 21, the point mutation of L858R resulted in the abolishment of the restriction site for Msc I. As a consequence, we designed PCR primers to include the area of the mutation for PCR amplification. If there was a mutation of the L858R locus, the restriction of the endonuclease Msc I resulted in three DNA fragments of 125, 245, and 370 bp. In the wild-type profile, only two fragments of 125 and 245 bp (Fig. 1) were evident.

For PCR reaction, 20 ng of DNA was amplified in a 25 μL reaction solution containing 10× buffer (Qiagen), 1.5 mM/L of MgCl2, 0.3 μM of each primer pairs (exon 19, F: 5’t-acaaattgccccgacctgcgt, R: 5’-agcaagcagaaactcatac; exon 21, F: 5’-agtactacatgcggtcg, R: 5’-caatacagctagtgggaagg), 200 μM of deoxynucleoside triphosphate, and 1.25 units of HotStart Taq DNA polymerase (Qiagen). The PCR conditions were an initial denaturation at 95°C for 15 minutes, followed by 35 cycles of denaturation at 94°C for 30 seconds, annealing at 54°C for 45 seconds and extension at 72°C for 1 minute, and a final incubation at 72°C for 5 minutes.

5) Comparison of two methods

First, we selected 74 samples of adenocarcinoma whose *EGFR* exons 19 and 21 had been previously sequenced. Based on the sequencing result, we designed PCR primers and selected a restriction enzyme. PCR products were tested on the 3.5% metaphor agarose gel directly for exon 19 and on the 2% agarose gel after enzymatic digestion for exon 21. To validate its accuracy, we set up test sets of 74 additional lung cancer samples. For those samples, the PCR-based method was
performed first, followed by sequencing. The sensitivities of the two methods were then compared by using the McNemar test. The gold standard was presumed by the final result obtained from re-testing of the sequencing method.

RESULTS

1) The training sample group

*EGFR* mutations were detected in 26 patients in the first set of 74 samples. Among these samples, using the sequencing method, we found 15 (20.3%) mutations in exon 19, and 9 (12.2%) mutations in exon 21. All mutations of exon 19 were deletion mutations, and the mutations in exon 21 were L858R substitution mutations (Table 1).

By using the PCR-based method, we were able to identify all the mutated samples detected by the sequencing method. At the same time, we were able to detect mutations in exon 19 in three samples, and in exon 21 in one additional sample which were not detected in initial Sanger sequence method. Those additional samples were confirmed by performing a repeated sequencing analysis (Table 2).

2) The test sample group

In the second set of samples, the PCR-based method was performed first. It detected 10 (13.5%) and 7 (9.5%) mutations in exons 19 and 21, respectively. Additional mutations in exon 19 were identified in 2 samples by the sequencing method. However, the sequencing method failed to identify a mutation in exon 21 in one sample (Table 3).

3) Sensitivity of two methods

When we compared the two methods, the sensitivity of the sequencing method for exon 19 mutation was 90%, whereas that of the PCR-base method was 93%. The two methods were not statistically different (McNemar test). For exon 21, the sequencing method had a sensitivity of 88% and the PCR-based method showed 100% sensitivity.

DISCUSSION AND CONCLUSION

Since the initial identification of a close correlation between *EGFR* mutations and radiographic regression in NSCLC pa-
tients treated with gefitinib, many retrospective studies have suggested that the EGFR mutation status is one of the most important predictors in identifying patients who may benefit from EGFR TKIs (4,5). At the present time, mutation detection requires isolation of the tumor cells, preparation of DNA, and sequencing of the EGFR tyrosine kinase domain. Because specialized expertise is required to perform this complex analysis, testing is limited in many institutions. Many laboratories around the world are now offering EGFR testing. However, as long as EGFR mutation testing is based on sequencing, this test will not be widely available. Our method, however, is relatively simple and can be performed in any laboratory where the conventional PCR experiments are available.

To validate our method, we designed a test set and examined it by the PCR-based method first, then confirmed it by direct sequencing. For exon 21, direct sequencing failed to detect a mutation in one case, compared to the PCR-based method in the training set. However, for exon 19, two additional mutations, which had not been detected in the PCR-based method, were detected with direct sequencing. These two samples were re-tested with both methods and the result was the same. There are several possible explanations. As both methods were performed using the same DNA extracted at one time, the difference in concentration of DNA cannot explain the result. We believe that the two false-negative results can be attributed to an error during the gel running. Both samples possessed 9 bp deletion, which was the smallest deletion in our series. In the literature, 5 base pair deletion mutation has been reported as the smallest mutation in exon 19. We re-tested two false negative samples using a higher density agarose gel and were able to identify a faint band underneath the wild type DNA band (Fig. 2). As a consequence, we presumed that the reason of false negative was due to the difference in size between the mutated and wild-type DNA was too small to be differentiated on the 3.5% agarose gel. The detection of such small size differences may require either a different electrophoresis running protocol on the higher concentration of agarose/polymer gel or the use of capillary electrophoresis (7).

Another possible explanation is an intrinsic low sensitivity of our method. As shown in Fig. 2, bands of two false-negative samples were not clear, even in the re-testing result. Because the sequencing method failed to detect exon 19 in the training set and exon 21 in both the training and test sets, the sensitivity issue is also a problem of the direct sequencing method. The direct sequencing method requires a sufficiently large specimen, at least a core biopsy, to provide enough DNA for sequencing. However, due to the trend towards minimally invasive diagnostic procedures in patients with NSCLC, endobronchial biopsies, needle core biopsies, or cytology specimens, such as bronchial brushings or fine-needle aspirates are often the only tumor samples available. Although more tissue could facilitate molecular testing, obtaining larger samples is often not realistic or practical. A low copy number DNA template can result in PCR errors, leading to sequencing artifacts. This may have implications for patients and physicians, who want to determine the presence of an EGFR mutation before making a therapeutic decision. To overcome such disadvantages of direct sequencing, a variety of alternative methods of EGFR mutation detection in NSCLC have been devised that have advantages and disadvantages, in terms of complexity, performance, and sensitivity. Some technologies, such as single-stranded conformational polymorphism, denaturing high-performance liquid chromatography, and high-resolution melting analysis, have the advantages of allowing for rapid mutation screening of large numbers of samples and having high sensitivity. However, these methods require direct sequencing to confirm the identity of the detected mutations (8-11).

To improve sensitivity, many researchers have developed various methods. Sasaki et al. (12) examined 118 freshly frozen tumor specimens using an allele-specific real-time reverse transcription-PCR assay system. Similarly, Pan et al. (7) reported a PCR-based method for exons 19 and 21. They found 29 cases with mutations in exons 19 and 21, out of 39 lung cancer samples using the PCR-based methods, whereas the sequencing
method detected only 25 cases (7). Our method is similar to other PCR-based methods in that we use PCR reactions for the amplification and restriction of endonuclease for the detection of mutations. The difference between these methods and ours is an application of real-time PCR, which may improve sensitivity in the detection in a small amount of DNA. Unlike the other methods, however, our method does not require complex analysis techniques, such as real-time PCR. We believe that our simplified method can serve as a fast and easy initial screening tool for EGFR mutation in NSCLC patients who have undergone surgical resection.

Another consideration is the content in a sample of the tumor cells as compared with that of the non-tumor cells, including the stromal, vascular, and inflammatory cells. Since the mutations are only found in the tumor cells, a mutation can be missed if the tumor cells are not enriched by a micro- or gross dissection. In clinical samples, such as pleural fluids or serum, where the tumor cell amount is small, the challenge is to overcome background interference by large amounts of non-malignant cells. Various techniques have been developed for the simple but highly sensitive detection of specific EGFR mutations, such as the amplification refractory mutation system and the peptide nucleic acid-locked PCR clamping. Others selectively digest wild-type DNA templates with restriction endonucleases to enrich mutant alleles by PCR (13-15). For example, Asano et al. (13) described additional mutations that were identified by applying mutant-enriched PCR assay for surgical specimens, as well as cytology or biopsy specimens. We had applied mutant enrichment methods for these two false negative samples. However, the band was too strong to discriminate a small 9 bp deletion mutation from the wild DNA in exon 19. Based on our result, we believe it may not be necessary to apply the time-consuming mutant-enrichment method for the freshly frozen, well prepared surgical specimen for the screening purpose of EGFR mutation. For the most part, it is not yet known whether or not these low-abundance mutations will ultimately have predictive or prognostic significance. Although these methods are of great value, in terms of improved sensitivity and accuracy, because of limited resources and expertise, most institutions will find routine screening for mutations with these methods to be a major challenge and costly to implement.

Infrequent mutations or new mutations are not detectable using our method. Although many EGFR mutations have been reported, not all have been responsive to EGFR TKIs. The two most common EGFR mutations that have been identified, representing 85% to 90% of EGFR mutations, consist of either a deletion in exon 19 of a leucine-arginine-glutamic acid-alanine or a leucine-to-arginine substitution at position 858 (L858R) (16). These mutations occur near the ATP cleft of the tyrosine kinase domain, where 4-anilinoquinazoline compounds, such as gefitinib, compete with ATP for binding (4,5,17-19). Consequently, EGFR mutations in exons 19 and 21 may be useful molecular markers for treatment of NSCLC. In our own series of the Korean population, EGFR mutations on these two loci were found in 32.4% of adenocarcinoma patients and were 92.3% of the entire EGFR mutations found (20). This means that we can identify a majority of patients likely to respond to EGFR TKI by detecting mutations in exons 19 and 21. In our first set, the PCR-based method found 4 additional mutations, compared to that of direct sequencing. Thus, the false-negative rate for direct sequencing was 14.3% (4/28).

Compared to direct sequencing, our method can be performed more easily with less cost, and can still achieve superior sensitivity. If our tests are negative, direct sequencing of EGFR exons 18 and 20 can be performed to detect the remaining 10% of EGFR mutations. EGFR mutation screening is currently not routinely recommended in patient management. However, it is possible that testing of tumor tissue for EGFR mutations may soon be indicated in all patients with adenocarcinoma of the lung to aid in selecting therapy in various neoadjuvant, adjuvant, or palliative settings. In western countries, at least 10% of lung adenocarcinomas will harbor EGFR mutations. In Korea and other Asian countries, 30~35% of adenocarcinomas are expected to have EGFR mutations, and 20~25% of these will be detectable using our methods. Our method is simple and easy to set up, requiring only an ordinary laboratory where PCR reaction can be performed. Hence, our PCR-based enzymatic digestion method can be used as an initial screening test in virtually every institution.

In conclusion, for well-prepared freshly frozen tumor samples, our conventional PCR-based method could detect EGFR mutations in exons 19 and 21 without more expensive and time-consuming techniques, and can be used as an initial screening test for the selection of potential candidates of EGFR TKI.
ACKNOWLEDGMENTS

We appreciate Mrs. Yung Min Kim for English proof-reading. The statistical analysis was performed by Medical Research Collaborating Center, Seoul National University Hospital.

REFERENCES