EGFR mutation analysis in sputum of lung cancer patients: A multitechnique study

A. Jasmijn Hubers, Daniëlle A.M. Heideman, Yasushi Yatabe, Michelle D. Wood, Justyna Tull, Miquel Tarón, Miquel A. Molina, Clara Mayo, Jordi Bertran-Alamillo, Gerarda J.M. Herder, Remco Koning, Daoud Sie, Bauke Ylstra, Gerrit A. Meijer, Peter J.F. Snijders, Birgit I. Witte, Pieter E. Postmus, Egbert F. Smit and Erik Thunnissen
ABSTRACT

Objectives

Epidermal growth factor receptor (EGFR) mutations have been identified in lung adenocarcinomas and are associated with high response chance to EGFR tyrosine kinase inhibitors. EGFR mutations can be detected in tumour tissue, cytology specimens and blood from lung cancer patients. Thus far, EGFR mutation analysis has not been systematically demonstrated for sputum samples. The aim of the present study was to determine whether EGFR mutation analysis is attainable on sputum samples, employing different assays in a multicenter study.

Materials and methods

Sputum DNA from 10 lung cancer patients with confirmed EGFR mutation in their tumour tissue, 10 lung cancer patients without evidence of an EGFR mutation, and 10 patients with chronic obstructive pulmonary disease (COPD) was used for mutation analysis by Cycleave PCR, COLD-PCR, PangaeaBiotech SL Technology (PST), and High Resolution Melting, respectively. Targeted resequencing (TruSeq Amplicon Cancer Panel) and droplet digital PCR were additionally performed on the 10 samples with EGFR mutation.

Results

Dependent on the assay, EGFR mutations could be detected in 30–50% of the sputum samples of patients with EGFR mutations. The different techniques revealed consistent results, with slightly higher sensitivity for PST. Neither the lung cancer patients without EGFR mutation nor the COPD controls tested positive for EGFR mutations in their sputum samples, indicating high clinical specificity of all assays.

Conclusion

EGFR mutations can be detected in sputum samples from patients with EGFR-mutated non-small cell lung cancer, which may replace biopsy procedure for some patients.
INTRODUCTION

Lung cancer has the highest mortality rate of all cancers worldwide in both men and women.\(^1\) From a molecular point of view, lung cancer is a heterogeneous disease. Mutations in the tyrosine kinase domain of the epidermal growth factor receptor \((EGFR)\) have been identified in part of the adenocarcinomas. Deletions in exon 19 and c.2753T>G [p.L858R] point mutation in exon 21 occur most frequently and are associated with a high chance of response to \(EGFR\) tyrosine kinase inhibitors (TKIs).\(^2\)

\(EGFR\) mutation analysis has been performed on tumour cells in biopsy or cytology specimens obtained from bronchoscopy, surgical resection, washings or blood.\(^3\)–\(^5\) In some cases there may be insufficient tumour material available to perform \(EGFR\) mutation analysis, for instance due to inaccessible tumour location. Therefore, development of a novel, non-invasive means to enable detection of \(EGFR\)-mutated lung cancers aids in the potential identification of a higher proportion of lung cancer patients eligible for \(EGFR\)–TKI treatment. Procurement of sputum is non-invasive, simple and inexpensive. In sputum, tumour cells can be detected by sputum cytology with a sensitivity of around 66%.\(^6\) Tumour specific molecular alterations, such as mutations in \(KRAS\) and \(p53\),\(^7\)–\(^9\) and DNA promoter hypermethylation of genes\(^8\),\(^10\)–\(^16\) have already been identified in sputum with different molecular techniques. Although \(EGFR\) mutation analysis has previously been performed in some sputum samples as part of a larger series of other cytological samples, mostly the outcomes were not compared to the original tumour or detailed information was not given.\(^17\)–\(^19\) In total in 3 published studies, 25 sputum samples have been tested, 3 of which were positive in cases with cytologically proven malignant cells.

The aim of this study was to investigate the feasibility of \(EGFR\) mutation analysis on sputum samples of patients with lung cancer. As the analytical sensitivity of the assay is important, we performed a multicenter international study, employing assays with different analytical characteristics, i.e., Cycleave PCR, COLD-PCR, and PangaeaBiotech SL Technology (PST). Additionally, a targeted resequencing assay (TruSeq Amplicon Cancer Panel (TSACP) and droplet digital PCR (ddPCR)) were used to estimate the mutation frequency in the sputum specimens of patients with confirmed \(EGFR\) mutation in their tumour tissue.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Patients with lung cancer</th>
<th>P-values&lt;sup&gt;a&lt;/sup&gt;</th>
<th>COPD (n=10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFR mutation</td>
<td>n=10</td>
<td>n=10</td>
<td>COPD (n=10)</td>
</tr>
<tr>
<td>Age (years)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>61.8 ± 10.9</td>
<td>65.2 ± 8.0</td>
<td>0.53</td>
</tr>
<tr>
<td>Gender (% male)</td>
<td>30</td>
<td>80</td>
<td>0.07</td>
</tr>
<tr>
<td>Smoking status (n)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>0</td>
<td>4</td>
<td>0.007</td>
</tr>
<tr>
<td>Former&lt;sup&gt;c&lt;/sup&gt;</td>
<td>5</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Never</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Smoking history&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median duration (years)</td>
<td>27.0</td>
<td>43.5</td>
<td>0.03</td>
</tr>
<tr>
<td>Median smoking dose (pack-years)</td>
<td>28.0</td>
<td>45.5</td>
<td>0.06</td>
</tr>
<tr>
<td>Tumour histology (n)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>10</td>
<td>3</td>
<td>0.003</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>EGFR mutation (n)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deletion exon 19</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Combined deletion exon 19 &amp; point mutation exon 20 (T790M)</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mutation exon 21 (L858R)</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tumour stage (n)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>4</td>
<td>0.04</td>
</tr>
<tr>
<td>II</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>III</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>IV</td>
<td>7</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>COPD&lt;sup&gt;e&lt;/sup&gt; status (n)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>2</td>
<td>0.01</td>
</tr>
<tr>
<td>II</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>III</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>IV</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Classification not available</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No COPD</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

<sup>a</sup> Comparison between lung cancer cases with and without EGFR mutation. Dichotomous variables: Fisher’s Exact test; ordinal and continuous variables: Mann-Whitney U test.

<sup>b</sup> Mean ± SD.

<sup>c</sup> Defined by quitting smoking at least one year before study enrollment.

<sup>d</sup> Includes only current and former smokers.

<sup>e</sup> COPD = chronic obstructive pulmonary disease, classified according to GOLD criteria."
MATERIALS AND METHODS

Clinical samples

Sputum samples were collected between July 2009 and September 2010, as described before. From this cohort, we selected a three-day pooled sputum of (i) 10 patients with known EGFR mutation status of the primary lung cancer or metastasis, as assessed on tumour tissue-derived DNA by High Resolution Melting followed by Sanger sequencing of aberrantly melting out PCR products (further referred to as HRM-sequencing) as described before, (ii) 10 patients diagnosed with lung cancer without evidence of an EGFR mutation in their tumour specimen, and (iii) 10 COPD-patients without lung cancer (Table 1). The group of lung cancer cases without evidence of an EGFR mutation was composed of adenocarcinomas in which no EGFR mutation was identified by HRM-sequencing, and squamous cell carcinomas, as the chance of an EGFR mutation is very rare in this histological lung cancer subtype. The study protocol was approved by the Institutional Review Boards of participating hospitals and all subjects provided written informed consent. Smoking information was procured from questionnaires completed by participants. Clinicopathological data were retrieved from medical records after a follow-up period of minimal one year after study enrollment. Staging of the lung tumours was performed in accordance with UICC 7th edition.

DNA isolation

Genomic DNA was extracted from the sputum specimens using the QIAamp Blood Mini Kit (Qiagen, Hilden, Germany), according to instructions of the manufacturer. The sputum samples were randomly recoded by the first author (AJH) and the key was only made available after submission of the EGFR mutation data. Aliquots of 500 ng DNA (diluted in a total volume of 5 μl with a concentration of 100 ng/μl) were shipped in a blinded fashion to each of the collaborating laboratories for mutation analysis. After knowledge of overall scores of mutation analysis, two DNA aliquots (40 ng) of each sputum sample were subjected to HRM-sequencing in a similar blinded design.

Analytical sensitivity

In order to allow direct, mutual comparison of analytical sensitivity between the four assays (Cycleave PCR, COLD-PCR, PST, and HRM), EGFR mutation status was assessed on DNA extracted from formalin-fixed, paraffin embedded (FFPE) cell lines with EGFR mutation (H1975, H1650; both heterozygous) and without (SiHa, A549). Samples with 10%, 5%, 1% and 0% (w/v) EGFR mutated cells were composed, representing 5%,
2.5%, 0.5% and 0% mutated DNA per sample, respectively. After randomisation, 500 ng of each sample was distributed to all participating labs. Targeted resequencing (TruSeq Amplicon Cancer Panel) and ddPCR were additionally performed on the 10 sputum samples with *EGFR* mutation in the original tumour for estimation of mutation frequency. For a methodological description of the applied assays, see Supplementary data 1.

**Cytological examination of sputum samples**

From all patients 3 tubes were collected, each tube with 3 days of sputum. Whereas only the tube containing the most DNA is used for this molecular study, cytological examination was performed for all three tubes. Cytospin samples were Papanicolaou stained and cytological examination was performed without knowledge of molecular analysis (RK, ET). Alveolar macrophages, bronchial epithelial cells, squamous metaplasia and presence of tumour cells were scored. Sputum samples were considered representative when alveolar macrophages and/or bronchial epithelial cells were present.

**Statistical analysis**

Fisher’s exact test and Mann–Whitney U test were used for comparison of lung cancer groups. Chi-square tests were used to assess the ability of the assay to discriminate between lung cancer cases with and without *EGFR* mutation and control samples. A two-sided *P*-value of ≤0.05 was considered significant. Statistical analyses were carried out using SPSS version 15.0 (SPSS Inc., Chicago, IL).

**RESULTS**

**Patient characteristics**

The clinicopathological data are shown in Table 1. Age and gender were not different in the two lung cancer groups, while smoking history, COPD, tumour histology and stage differed between the two. In the lung cancer group without *EGFR* mutation, more patients had COPD than in those with *EGFR* mutations. This may be explained by the difference in smoking habits between the two groups.

*EGFR* mutation analysis on sputum samples

Five (50%) sputum samples of lung cancer cases with *EGFR* mutation revealed an
**TABLE 2. Details of EGFR mutation analysis in lung cancer patients with confirmed EGFR mutation on tumour tissue.**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Gender (M/F)</th>
<th>Age (years)</th>
<th>Tumour stage&lt;sup&gt;a&lt;/sup&gt;</th>
<th>EGFR mutation status of tumour tissue&lt;sup&gt;b&lt;/sup&gt;</th>
<th>EGFR mutation analysis on sputum specimens&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Cytology&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cycleave PCR</td>
<td>COLD-PCR</td>
</tr>
<tr>
<td>A</td>
<td>F</td>
<td>72</td>
<td>IV</td>
<td>Del E746-A750</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>M</td>
<td>66</td>
<td>I</td>
<td>Del E746-A750</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>C&lt;sup&gt;g&lt;/sup&gt;</td>
<td>F</td>
<td>78</td>
<td>IV</td>
<td>Del E746-A750</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>46</td>
<td>III</td>
<td>Del E746-A750</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E&lt;sup&gt;g&lt;/sup&gt;</td>
<td>M</td>
<td>54</td>
<td>IV</td>
<td>Del E746-A750</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>F&lt;sup&gt;h&lt;/sup&gt;</td>
<td>F</td>
<td>49</td>
<td>III</td>
<td>Del E746-A750 &amp; c.2369C&gt;T [p.T790M]</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>G&lt;sup&gt;i&lt;/sup&gt;</td>
<td>F</td>
<td>54</td>
<td>IV</td>
<td>Del E746-A750 &amp; c.2369C&gt;T [p.T790M]</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>H</td>
<td>F</td>
<td>73</td>
<td>IV</td>
<td>c.2753T&gt;G [p.L858R]</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>I</td>
<td>F</td>
<td>61</td>
<td>IV</td>
<td>c.2753T&gt;G [p.L858R]</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>J&lt;sup&gt;g&lt;/sup&gt;</td>
<td>M</td>
<td>60</td>
<td>IV</td>
<td>Del E746-A750</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<sup>a</sup> Tumour stage according to UICC 7 ed.<sup>25</sup>

<sup>b</sup> Del E746-A750 = deletion exon 19.

<sup>c</sup> Mutation identified: 0 = no, 1 = yes, 2 = dubious (further called positive).

<sup>d</sup> Tumour cells present: 0 = no, 1 = yes, 2 = not in this sample, but in another sputum sample of the same patient.

<sup>e</sup> In this assay, exclusively deletion exon 19 and L858R were assessed.

<sup>f</sup> Only deletion exon 19 detected.

<sup>g</sup> TSACP and ddPCR both tested EGFR mutation (deletion exon 19) positive.

<sup>h</sup> ddPCR dubious for T790M.

<sup>i</sup> ddPCR positive for T790M.

EGFR mutation with the PST method (Tables 2 and 3; P=0.002). PST did not run T790M, because of insufficient DNA quantity in the sputum samples. Three, three, and four of the EGFR mutated cases were also detected in sputum by means of Cycleave PCR, HRM-sequencing and COLD-PCR, respectively. All sputum positive cases involved 15-bp deletions in exon 19; other mutations (i.e., point mutation in exon 20 and exon 21) were not identified in the sputum samples. The TSACP and ddPCR as performed on the 10
EGFR mutated samples both detected three EGFR mutated cases (relative proportion of positive reads was >15%). None of the sputum samples of lung cancer patients without EGFR mutation in their tumour nor the COPD controls tested positive. Thus for all methods, a high specificity (100%) for EGFR mutation analysis in sputum (Table 3) was achieved.

Cytological examination of sputum samples

Cytological examination showed the presence of a few malignant cells in the sputum of one lung cancer case without EGFR mutation and one lung cancer case with mutation. The latter was also detected (i.e., deletion exon 19) by one of the assays (PST). Cytological evaluation of additional sputum samples collected in subsequent days revealed the presence of tumour cells in two other lung cancer patients with EGFR mutation detected in the tumour tissue. In the original sputum samples of both cases, it was possible to confirm EGFR mutation through each of the assays.

Positive cytology results of both directly and indirectly related sputum samples from EGFR mutated cases correlated significantly with positive outcome of EGFR mutation analysis, irrespective of the assay used (chi-square tests, P-values 0.001–0.008; data not shown). These results indicate that sputum samples are suitable for EGFR mutation analysis, especially when tumour cells are detectable by sputum cytology in at least one of a series of consecutive sputum samples.

Analytical sensitivity

Comparison of the analytical sensitivity rates between the assays revealed that the samples with 5% and 2.5% EGFR mutated DNA in a background of wild-type DNA were detected by all assays. PST was found to be the most sensitive approach, as both
samples with 0.5% mutated DNA were convincingly detected by PST, and HRM the least sensitive, as this method was not able to detect the 0.5% *EGFR* mutated samples. COLD-PCR and Cycleave PCR tested in between.

DISCUSSION

This study demonstrates in a systematic blinded design that *EGFR* mutation analysis in sputum is feasible with consistent results. By means of Cycleave PCR, COLD-PCR, PST, and HRM-sequencing, *EGFR* mutations were detected in sputum of a subset of lung cancer patients with known mutations in their tumour tissue. This offers perspectives for *EGFR* mutation testing in sputum as part of the diagnostic procedure, in particular in patients for whom bronchoscopy is not an option.

The assays used consist of different methods. PST uses PNA probe for preferred amplification of mutated DNA combined with nuclease cleavage of mismatched probes.\(^\text{27,28}\) Cycleave PCR is based upon cleavage of matching probes by RNase H.\(^\text{29–31}\) COLD-PCR makes use of a low denaturation temperature in the first cycles of PCR, aiming at preferred amplification of heteroduplexes.\(^\text{32–35}\) HRM-sequencing is based upon melt curve analysis of PCR products to identify specimens with a high likelihood of harbouring a mutation and subsequent Sanger sequencing is used for genotyping.\(^\text{21}\) The analytical sensitivity of PST is slightly higher than that of the other three methods. This may be an explanation why PST was slightly better able to detect *EGFR* mutations in sputum samples compared to the other techniques. Conceptually, it may be interesting to investigate whether combination of assays, such as COLD-PCR with PNA and Cycleave PCR with PNA, may lead to an even higher sensitivity without loss of specificity.

In our experience, sputum samples usually contain around 1% or less tumour cells. Interestingly, in sputum samples with malignant cells detected by cytological examination, either in the specimen analysed for mutations or in a concomitant specimen of the same patient, *EGFR* mutation was detected by one or more assays. This is in line with Tanaka and coworkers, who conducted PNA-LNA PCR clamp for *EGFR* mutation analysis in samples selected for the presence of tumour cells.\(^\text{36}\) In 3 out of 9 sputum samples, *EGFR* mutations were detected. However, no comparison with a tumour sample was made in this and other studies.\(^\text{18,19}\) The TSACP and ddPCR both confirmed the presence of *EGFR* mutated DNA in the three samples tested positive by the other four methods. Remarkably the remaining 7 sputum samples from *EGFR*
positive tumours tested negative in TSACP and ddPCR, indicative of a very low fraction of mutated DNA, arguing (i) the representativity of these sputum samples and (ii) opening the possibility of chance as explanation for positivity by an individual assay. Previously we demonstrated that representativity is an issue in part of the sputum samples.\textsuperscript{20}

It is promising that all techniques were also capable of detecting tumour-related DNA in a sample that was not judged as containing tumour cells by cytological examination. As such, the presence of tumour cells by cytological examination seems not obligatory for an \textit{EGFR} mutation positive outcome. For detection of other mutations and epigenetic alterations, similar findings in sputum samples have been reported before.\textsuperscript{7–9}

In the \textit{EGFR} mutation negative sputum samples of patients with \textit{EGFR} mutated tumours, the question arises whether tumour DNA was present in the sample and if so, whether the amount of tumour DNA was below the threshold (analytical sensitivity) of the assays. In any case, this has an impact on the clinical sensitivity of \textit{EGFR} mutation analysis in sputum, which was to range from 30\% to 50\% in the current study. This sensitivity range is similar to that of \textit{EGFR} mutation detection in blood.\textsuperscript{27,28,37–40} Further studies are needed to evaluate if assays with a higher analytical sensitivity may lead to a better clinical performance on sputum specimens, without loss of specificity, and to investigate whether tumour related DNA may be a more sensitive biomarker than cytological examination.

Remarkably, only deletions in exon 19 were identified, and none of the other mutations. This is probably not attributable to technical specifications of the assays, as analytical features were highly similar between mutation types per assay. As c.2369C>T [p.T790M] mutation in exon 20 is likely occurring in a later stage of lung carcinogenesis and associated with secondary resistance towards TKI treatment,\textsuperscript{41–43} the non-detection of this mutation type may indicate that only a subset of tumour cells contain this mutation at an amount that was below detection limit. This is in line with a methodological study on c.2369C>T [p.T790M] mutation of Miyazawa \textit{et al.}\textsuperscript{44} Moreover, as only two samples from patients with c.2369C>T [p.T790M] mutation were included, this requires further study.

The fact that none of the assays showed an \textit{EGFR} mutation in sputum samples of patients without \textit{EGFR} mutation in the tumour, nor in the COPD controls, leads to 100\% clinical specificity of these assays. The absence of false-positives is a useful characteristic for clinical diagnostic application, though the number of controls in our study is limited to draw firm conclusions. Nonetheless, our study supports the use of sputum for detection of \textit{EGFR} mutations, particularly in sputum samples with positive
EGFR mutation analysis in sputum of lung cancer patients: A multitechnique study

Cytology as the chance to find an EGFR mutation status seems more successful when tumour cells are detected by means of sputum cytology. Although intuitively sputum samples that are devoid of detectable tumour cells may be indicative of a lower tumour load, previous research has shown that aberrant tumour DNA (e.g., KRAS mutation, DNA hypermethylation) can be detected in sputum without cytological aberrations and is independent of tumour location. To ensure good sputum quality, special attention needs to be paid to provide clear patient instructions before sputum collection. In our study, written and verbal information were given on expectoration technique, moment of sputum collection (preferably in the morning) and the distinction between sputum and saliva. A limitation of the study is the small number of sputum samples. The main intention was to show feasibility of EGFR mutation analysis on sputum and identification of suitable technique(s). Now we have demonstrated that this is possible with more or less similar outcome between different assays, it is interesting to expand the study population to assess the diagnostic value of EGFR mutation analysis in a larger cohort of sputum samples.

Although not frequently used, procurement of sputum is a simple and non-invasive means of obtaining tumour-related material, making it an attractive tool for both the patient and the clinician. In some countries, sputum analysis is considered ‘a poor men’s bronchoscopy’, as it provides a more affordable but still reliable method for lung cancer diagnosis in low-income areas. If sputum EGFR mutation analysis alone is required, sputum analysis may be done in order to avoid in some patients a biopsy procedure. Sputum analysis is most useful for patients from whom a tumour biopsy cannot be taken, in particular patients who are considered unfit for bronchoscopy or transthoracic biopsy. Whether monitoring EGFR-TKI treatment is possible with this approach, remains to be determined.

In conclusion, we demonstrated that EGFR mutation analysis is feasible on sputum samples from patients with lung cancer using different assays. Our findings support the use of sputum as a non-invasive approach to detect EGFR-mutations in patients with lung cancer.

ACKNOWLEDGMENTS

We thank the technicians of the units of Molecular Pathology, Microarray Facility and Cytology of the Department of Pathology at the VU University Medical Center Amsterdam for their excellent technical assistance. We also express our gratitude to Eddy van Collenburg and Bio-Rad Laboratories for the use of ddPCR assay.
REFERENCES


SUPPLEMENTARY DATA 1. DESCRIPTION OF ASSAYS THAT WERE EMPLOYED TO DETECT EGFR MUTATION IN SPUTUM SAMPLES.

Cycleave PCR

Cycleave PCR is actually composed of two assays: Cycleave PCR for point mutations, and fragment analyses for insertion and deletion. The Cycleave PCR for the detection of c.2159C>T [p.G719S/C/A], c.2369C>T [p.T790M], c.2753T>G [p.L858R] and c.2582T>A [p.L861Q], was used with the SmartCycler system (SC-100, Cepheid, Sunnyvale, CA). The method uses two fluorescent probes: wild type and mutant, which are cleaved by RNase H in case of correct matching. To detect the deletion in exon 19 and the insertion in exon 20, common fragment analysis was used. The reported analytical sensitivity was 1% of tumour cells, corresponding to 0.5% of mutant molecules in wild-type background. For this analysis, individual mutation assays were performed in individual test-tubes. DNA was derived from one 3 micron section and diluted with distilled water (no quantification of DNA).

COLD-PCR

Co-amplification at low denaturation temperatures (COLD) PCR is a modified PCR amplification step that improves mutation detection by downstream assays including Sanger sequencing, pyrosequencing, HRM, MALDI-TOF and TaqMan-based real time PCR. COLD-PCR was used in conjunction with pyrosequencing (to detect common point mutations) and fragment length analysis (to detect exon 19 deletions). COLD-PCR is used to enrich detection of specific mutations within the tyrosine kinase domain of the EGFR gene. For the combination COLD-PCR and pyrosequencing, the analytical sensitivity is 5 to 1% tumour nuclei. The following mutations could be detected by the pyrosequencing assays used in this study; c.2155G>T, c.2156G>C, c.2159C>T [p.G719S/C/A], c.2294T>C [p.V765A], c.2303G>T [p.S768I], c.2305G>T [p.V769L], c.2369C>T [p.T790M], c.2753T>G [p.L858R], c.2582T>A [p.L861Q]. The use of COLD-PCR in conjunction with the fragment sizing assay has an analytical sensitivity of approximately 0.3 to 1% mutant to wild type ratio. A minimum of 10 ng of DNA was required per PCR.

PangaeaBiotech SL Technology (PST)

Deletions in exon 19 were established by length analysis after PCR amplification with a FAM-labeled primer in an ABI prism 3130 DNA analyzer (Applied Biosystems,
Foster City, CA, USA); c.2753T>G [p.L858R] mutations in exon 21 were detected with a 5’ nuclease PCR assay (TaqMan assay, Applied Biosystems) with a FAM MGB-labeled probe for the wild-type and a VIC MGB-labeled probe for the mutant sequence. Both techniques, length analysis after PCR amplification for exon 19 deletions and TaqMan assay for c.2753T>G [p.L858R] mutations were done in the presence of a protein nucleic acid (PNA) clamp, which was designed to inhibit the amplification of the wild-type allele (Pangaea Biotech SL patent, covered by European patent EP2046985-B1, US patent application 12/374,307 and corresponding patents and patent applications in other territories).\(^6\)\(^,\)\(^7\) The reported analytical sensitivity for exon 19 and c.2753T>G [p.L858R] mutations is similar: 0.1 to 0.2%. PST detects 20 pg of mutated DNA per reaction for exon 19 deletion and 50 pg of mutated DNA per reaction for c.2753T>G [p.L858R]. Input DNA is not quantified. For sputum analysis, 2-3 μl was used in each PCR reaction.

**HRM-Sequencing**

HRM followed by cycle sequencing of PCR products displaying abnormal melting was performed for \(EGFR\) exon 19, 20 and 21, essentially as described previously.\(^8\)\(^,\)\(^9\) The reported analytical sensitivity for detection of mutations for the HRM-sequencing assays is \(\leq 5\%\), independent of the exon targeted.\(^8\) For sputum analysis, 50 ng of input DNA was used in each PCR reaction.

**Massive Parallel Sequencing (Targeted Resequencing)**

The TruSeq Amplicon Cancer Panel (TSACP, Illumina) is a multiplexed targeted resequencing assay for detection of somatic mutations among 212 amplicons covering 48 genes.

The assay begins with hybridization of probes to the target DNA, followed by an extension-ligation reaction which creates templates for PCR that finally generates the amplicons that are ready for sequencing on a MiSeq Personal Sequencer system (Illumina).\(^10\)\(^–\)\(^12\) The \(EGFR\) exon 19 (i.e., deletion loci), exon 21 (i.e., L858R locus) and exon 20 (i.e., T790M locus) are all included in this panel. The TSACP assay was performed according to the recommendations of the manufacturer (Illumina) on 250 ng of input DNA of the sputum samples of \(EGFR\) mutated lung cancer patients, followed by sequencing 150 cycles paired end on a MiSeq. Each \(EGFR\) locus was inspected with IGV to assess the number of wild-type and mutant sequence reads. If \(EGFR\) mutation was present, the ratio mutated versus wild-type DNA was determined. The detection threshold was set to 0.5% according to the specifications of the manufacturer.
Droplet Digital PCR (ddPCR)

Droplet digital PCR (ddPCR) is an emerging nucleic acid detection method that provides absolute quantification of target sequences without relying on the use of standard curves. To perform ddPCR, the DNA target, fluorescently-labeled probe(s) and the ingredients for a PCR reaction are partitioned into an emulsion of approximately 20,000 droplets (Bio-Rad Laboratories, USA). Following PCR amplification, enumeration of both fluorescing and non-fluorescing droplets is performed, allowing the absolute quantification of target molecules present in the original sample. For rare mutation detection this sample partitioning increases sensitivity by distributing both mutant and normal genes into a large number of isolated reaction compartments. Input DNA for each sample was 200 ng. EGFR deletion exon 19, c.2753T>G [p.L858R] and c.2369C>T [p.T790M] were tested according to manufacturer with obtained analytical sensitivity of 0.1% (data not shown).
REFERENCES


EGFR mutation analysis in sputum of lung cancer patients: A multitechnique study