Comparison of molecular testing methods for the detection of EGFR mutations in formalin-fixed paraffin-embedded tissue specimens of non-small cell lung cancer

Fernando Lopez-Rios,1 Barbara Angulo,1 Belen Gomez,2 Debbie Mair,2 Rebeca Martinez,1 Esther Conde,1 Felice Shieh,3 Julie Tsai,3 Jeffrey Vaks,3 Robert Current,3 H Jeffrey Lawrence,3 David Gonzalez de Castro2

ABSTRACT

AIM To conduct a methods correlation study of three different assays for the detection of mutations at EGFR gene in human formalin-fixed paraffin-embedded tumour (FFPET) specimens of non-small cell lung carcinomas (NSCLC).

METHODS We conducted a 2-site method comparison study of two European conformity (CE) in vitro diagnostic (IVD)-marked assays, the cobas EGFR Mutation Test and the Therascreen EGFR29 Mutation Kit, and 2x bidirectional Sanger sequencing. We blind-tested 124 NSCLC FFPET specimens with all three methods; the cobas test was performed at both sites. Positive (PPA) and negative percent agreements (NPA) were determined for the cobas test versus each of the other two methods. Specimens yielding discordant test results between methods were further tested using quantitative massively parallel pyrosequencing (MPP).

RESULTS PPA between cobas and Sanger was 98.8%; NPA was 79.3%. Overall there were seven discordant results. MPP confirmed an exon 19 deletion in two cases and L858R mutation in four cases. PPA between cobas and Therascreen was 98.9% and NPA was 100%. There was one discordant result. Reproducibility of the cobas test between the two sites was 99.2%.

CONCLUSIONS The invalid rates for the cobas test and Therascreen were lower than Sanger sequencing. The cobas and Therascreen assays showed a high degree of concordance, and both were more sensitive for the detection of exon 19 deletion and L858R mutations than Sanger. The cobas test was highly reproducible between the two testing sites, used the least amount of DNA input and was the only test with automated results reporting.

INTRODUCTION

Anti-EGFR tyrosine kinase inhibitors (TKIs) are indicated for first-line therapy in patients with EGFR mutation positive non-small cell lung carcinomas (NSCLC).1–6 In Europe, the European Medicines Agency labelling states that gefitinib and erlotinib are indicated for patients with locally advanced or metastatic NSCLC with activating mutations of EGFR. In the USA, the Food and Drug Administration has not yet approved the use of erlotinib as first-line therapy, though EGFR testing is recommended in the NSCLC guidelines from major US oncology organisations, such as the National Comprehensive Cancer Network and the American Society of Clinical Oncology to guide treatment regime. No guidance recommending specific testing methods or assay attributes currently exists. Approximately 90% of all activating mutations are exon 19 deletions and a point mutation (L858R) in exon 21.7 At present, although it is recommended that the presence of activating mutations in the EGFR gene should be identified before treatment with gefitinib or erlotinib, there is no consensus regarding the specific mutations that should be tested nor methods that should be used.8 A number of sequencing and PCR-based methods for detecting EGFR mutations are currently in clinical use; however, it is not clear which technique offers the best performance in terms of sensitivity, specificity, reproducibility and accuracy.9,10 Additional attributes that are desirable for EGFR tests include short turnaround time; low input DNA requirement, as large specimens are often difficult to obtain; test robustness, so that results are not influenced by pathological factors, for instance, varying tumour content, lymphocyte infiltration, tissue necrosis, and mucin content; and automated reporting, which can reduce variability in the interpretation of results.

We performed a two-centre study to compare the analytic performance and workflow characteristics of the cobas EGFR Mutation Test against two other methods commonly used in the clinical setting: 2x bidirectional Sanger sequencing and the Therascreen EGFR29 Mutation Kit, using a blinded panel of formalin-fixed paraffin-embedded tumour (FFPET) NSCLC specimens.

MATERIALS AND METHODS

Mutation testing methods

The cobas EGFR Mutation Test kit (AS-PCR test), Allele-Specific PCR test, Roche Molecular Systems, Inc., Branchburg, New Jersey, USA) is a CE-IVD-marked allele-specific PCR test designed to detect the presence of 41 mutations in exons 18, 19, 20 and 21 of the EGFR gene in NSCLC FFPET specimens. The test requires 150 ng total DNA input, which can typically be obtained by using one 5 μm FFPET section. All analysis and results reporting is fully automated.

The Therascreen EGFR29 Mutation Kit (‘ARMS test’, Amplification Refractory Mutation System,
Qiagen, Manchester, UK) is CE-IVD-marked real-time PCR assay that combines the ARMS and Scorpion fluorescent primer/probe systems to detect 29 mutations in exons 18, 19, 20 and 21 of the EGFR gene. Reactions were run in the ABI 7500 Real-Time PCR System (Life Technologies, Warrington, UK) and analysed using ABI 7500 software (V2.0.5). Per the package insert, the ARMS test requires ≥20 ng of amplifiable genomic DNA from FFPET specimens. According to the laboratory-validated clinical protocol, ~100 ng of total DNA, as measured by spectrophotometry, is used per PCR reaction to account for the partial degradation of FFPET DNA without resulting in oversaturation of the reaction. The test thus requires a total DNA input of 800 ng.

Sanger sequencing (‘Sanger’): Mutation screening for exons 18, 19, 20 and 21 of the EGFR gene was carried out using PCR conditions and 2× bidirectional direct sequencing following previously described protocols.11,12 According to the laboratory-validated protocol, 150 ng of DNA is used per PCR performed, for a total of 600 ng total DNA input.

454 sequencing (454 Life Sciences, Branford, Connecticut, USA) is a quantitative method that involves clonal amplification by emulsion PCR of target sequences followed by massively parallel pyrosequencing (MPP).13

**Materials:** Of the total 133 specimens, 127 of the tested FFPET specimens were residual NSCLC tumour specimens from Biochain Institute (Hayward, California, USA). Six FFPET specimens were purchased from INDIVUMED Inc. (Kensington, Maryland, USA), Careline, Inc. (South San Francisco, California, USA), Bioserve (Beltsville, Maryland, USA), and ConversantBio (Huntsville, Alabama, USA).

**Study design**
From a panel of 300 vendor-purchased FFPET NSCLC tumour specimens, 133 specimens were selected for this study based on histologic and clinical characteristics. All selected specimens were from a cohort of Asian non-smoking women with adenocarcinoma, characteristics of which have been shown to have higher frequency of EGFR mutation positive lung cancer.14,15 The study design is depicted in figure 1. Of the 133 specimens, 114 specimens were selected at random; 13 specimens were previously characterised using real-time PCR, and six specimens that were known to contain rare mutations (L861Q, S768I and exon 20 insertions). Testing at both sites was performed blinded.

The 133 tumour specimens were each partitioned to yield five 5 μm sections per panel member. One section was mounted on a slide and stained with haematoxylin and eosin, and reviewed by two pathologists (FL-R and EC) who assessed the tumour content, lymphocyte infiltration, extent of necrosis and mucin content. None of the specimens was subjected to macrodissection in order to avoid bias and intersite variability of DNA sample preparation.

**Sample Preparation Kit (Roche Molecular Systems).** The DNA eluate was subsequently tested using the AS-PCR test according to the package insert.17 DNA for Sanger sequencing and the ARMS test was isolated from separate single 5 μm sections per panel member using the QIAamp DNA FFPET tissue kit and automated QIAcube robot (Qiagen, Hilden, Germany). The DNA eluate was then tested with Sanger according to a standard laboratory protocol at Site 1 and the ARMS test according to the vendor-provided protocol at Site 2.

Specimens with invalid test results or with discordant results between any of the methods were retested according to manufacturer/procedure instructions. Criteria for retesting were:

- **AS-PCR test:** <10% tumour content by area; insufficient DNA concentration
- **Sanger:** no PCR amplification or difficult sequence interpretation
- **ARMS test:** positive controls have not amplified specific product; mixed standard delta Ct is not within specified range; no template control has Ct<38; any sample with control gene Ct>37–38.

Quantitative MPP was performed on all discordant and invalid specimens.

**Study objectives**
Specific objectives of the study were to:

1. Compare analytical performance based on methods correlation parameters of PPA and NPA between the AS-PCR test with the ARMS test and Sanger sequencing (respectively, at each site)
2. Assess the reproducibility of the AS-PCR test at two independent laboratories
3. Assess the frequency of invalid test results for all methods
4. Assess the effects of pathological factors on the analytical performance of the AS-PCR test
5. Compare turnaround times between all methods.

**Methods correlation**
The PPA and NPA of the AS-PCR test were compared with the two other testing methods (Sanger and ARMS) for 124 evaluable FFPET NSCLC specimens. Agreement analysis for all methods is based on reportable results (table 1). Quantitative MPP was performed for all specimens for which the AS-PCR test and the comparison method gave discordant results and/or for which one of the two testing methods gave an invalid result. False positive (FP) and false negative (FN) rates were calculated for both methods using the AS-PCR test as the reference.

**Reproducibility**
The reproducibility of the AS-PCR test was evaluated by comparing the results at the two independent clinical laboratory sites. Discrepant analysis was performed using MPP on all specimens for which the AS-PCR test gave discordant results and/or when an invalid result was obtained.

**Invalid test rate**
The number of invalid test results from the evaluable tumour panel were recorded and compared across the three testing methods.

**Impact of specimen attributes on analytic performance**
The following pathological characteristics were assessed and graded according to the criteria:

- **Tumour content:** high (≥50%) vs low (<50%)
- **Lymphocyte infiltration:** high (≥10%) vs low (<10%)
Tumour necrosis: high (≥ 50%) vs low (<50%)

- Mucin content: high (≥ 50%) vs low (<50%).

**Workflow Assessment**

Assay turnaround time from DNA isolation to results reporting was compared for all methods, assuming one 8 h shift/day.

**RESULTS**

**Invalid test rate**

Of the 248 specimens that were evaluated by the AS-PCR test (124 at each site), two specimens gave initially invalid test results at both sites, upon retest one specimen remained invalid, resulting in an invalid rate of 0.4% (1/248) for both sites. Of the 124 specimens evaluated by Sanger, 23 specimens (18.5%) were initially invalid; upon retesting, 4.0% (5/124) remained invalid. Interestingly, the five samples which remained invalid by Sanger sequencing were characterised by poor quality of the extracted DNA (ratio A260/A230 ≤ 1). Of the 124 specimens evaluated by the ARMS test, one specimen was originally invalid and was resolved upon retest. The proportions of invalids for AS-PCR and ARMS were equivalent.

**Methods correlation with Sanger**

Of the 124 specimens tested at Clinical Site 1 using the AS-PCR test and Sanger, six specimens were invalid by either the AS-PCR test or Sanger, and five contained mutations that were not reportable by AS-PCR test (L861Q mutations), leaving 113 evaluable specimens for comparison. The initial agreement analysis showed a PPA of 98.8%, NPA of 79.3%, and an overall percent agreement (OPA) of 89.8% (table 2).

Seven specimens with discordant test results were subjected to MPP. One specimen reported as ‘mutation not detected’ (MND) by the AS-PCR test and ‘mutation detected’ (MD) (exon 20 insertion) by Sanger sequencing was confirmed to contain an L858R mutation. Following discrepant resolution with MPP, the PPA was 100%, NPA was 100%, and OPA was 100% (table 2).

**Methods correlation with the ARMS test**

Of the 124 specimens tested at Clinical Site 2 using the AS-PCR test and the ARMS test, three contained mutations that were not reportable by AS-PCR test (L861Q mutations), leaving 121 evaluable specimens for comparison. ARMS test did not call two specimens reported by Sanger as L861Q. The initial agreement analysis showed a PPA of 98.9%, NPA of 100%, and an OPA of 99.2% (table 3).

One specimen with a discordant test result, MD L858R mutation with the ARMS test and MND with the AS-PCR test was subjected to MPP. This specimen was confirmed by MPP to contain an L858R mutation. Following discrepant resolution with MPP, the PPA was 98.9%, NPA was 100%, and OPA was 99.2% (table 3). The FP and FN rates for ARMS test were 1.1% and 0%, respectively, using AS-PCR test as the reference.

**Reproducibility**

On the 124 evaluable specimens, one specimen remained invalid upon testing with the AS-PCR test, leaving 123 specimens for the

---

**Table 1**

<table>
<thead>
<tr>
<th>Exon</th>
<th>Mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>G719x</td>
</tr>
<tr>
<td>19</td>
<td>Exon 19 deletions</td>
</tr>
<tr>
<td>20</td>
<td>Exon 20 insertions; T790M</td>
</tr>
<tr>
<td>21</td>
<td>L858R</td>
</tr>
</tbody>
</table>

**Table 2**

<table>
<thead>
<tr>
<th></th>
<th>Mutation detected</th>
<th>Mutation not detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sanger sequencing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AS-PCR test</td>
<td>Mutation detected</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>Mutation not detected</td>
<td>6*</td>
</tr>
<tr>
<td>Sanger sequencing with MPP resolution</td>
<td>Mutation detected</td>
<td>89</td>
</tr>
<tr>
<td>AS-PCR test</td>
<td>Mutation not detected</td>
<td>0</td>
</tr>
</tbody>
</table>

*Two exon 19 deletions; four L858R mutations.
†Exon 20 insertion.

Sanger Sequencing
Positive percent agreement=98.8% (95% CI 94.8 to 99.7).
Negative percent agreement=79.3% (95% CI 64.7 to 88.9).
Overall percent agreement=89.8% (95% CI 89.0 to 96.6).
Sanger sequencing with MPP resolution
Positive percent agreement=100% (95% CI 97.0 to 100).
Negative percent agreement=100% (95% CI 98.9 to 100).
Overall percent agreement=100% (95% CI 97.7 to 100).
AS-PCR test, allele-specific PCR; MPP, massively parallel pyrosequencing.
analysis of AS-PCR test reproducibility between the two sites. Of these, 122/123 (99.2%) produced concordant results. The one discordant sample contained an L858R mutation according to MPP.

Impact of specimen attributes on analytic performance

Pathologic assessment of the 124 FFPET specimens revealed varying degrees of tumour content, lymphocyte infiltration, necrosis and mucin content (table 4). Regarding tumour content, it is important to note that mean, median and range were as follows: 59.9%, 60% and 5–95%, respectively. None of these pathologic characteristics had an observable effect on agreement analysis, reproducibility, or the invalid rate of the AS-PCR test (data not shown).

**Tumour content**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumour content*</td>
<td>33</td>
<td>91</td>
</tr>
<tr>
<td>Lymphocyte infiltration†</td>
<td>79</td>
<td>45</td>
</tr>
<tr>
<td>Necrosis*</td>
<td>115</td>
<td>9</td>
</tr>
<tr>
<td>Mucin content*</td>
<td>122</td>
<td>2</td>
</tr>
</tbody>
</table>

*Low<50%; High≥50%.
†Low<10%; High≥10%.

**NEW** Mutation not detected 1 29

**AS-PCR test, allele-specific PCR; MPP, massively parallel pyrosequencing.**

Workflow

Turnaround time per run for the AS-PCR test was approximately 1 day for 10 samples; for Sanger it was approximately 7 days for 24 samples; and for the ARMS test it was approximately 1 day for 24 samples.

**DISCUSSION**

The AS-PCR test was highly reproducible (>99%) between sites, a notable attribute given the high variability of other methods between laboratories. The high reproducibility may be attributed to the extensive validation studies and automated reporting. Given the current concerns around intratumour heterogeneity, minimising variability in results interpretation and achieving high lab-to-lab reproducibility is critical in ensuring accurate patient diagnosis and treatment. The AS-PCR test requires 150 ng of total DNA input and detects 41 mutations in exons 18–21. This amount of DNA can be isolated typically from a single 5 μm section of NSCLC FFPET specimen, a significant assay attribute given the current specimen tissue constraints and need for effective testing prioritisation. Although the current study only involved surgical specimens, the AS-PCR test has a similar performance when using small biopsies despite the lower amount of DNA obtained in this setting (D González de Castro, R Angulo, E Conde and F Lopez-Rios, unpublished data). In patients with advanced lung NSCLCs there is often not enough sample left for all the predictive biomarkers (eg, EGFR and Anaplastic Lymphoma Kinase (ALK) testing) after routine pathology workflow.

The ARMS test requires >160 ng of amplifiable DNA (equivalent to 800 ng of total DNA input in our experience) and detects 29 mutations in exons 18–21. With the Sanger sequencing method, four PCR reactions were performed (one reaction for each of the four exons) using approximately 150 ng DNA per PCR reaction (600 ng total) to theoretically detect any mutation in these four exons.

The results of the analysis of discordant results show that the AS-PCR test and the ARMS test are more sensitive than Sanger for detecting EGFR mutations. The AS-PCR test detects >90% of the activating mutations in exons 19 and 21 according to frequencies based on COSMIC database v54, but the tested version of the kit does not detect the L861Q mutation. Based on published data, it is not possible to conclude whether this mutation is sensitising for erlotinib and gefitinib.

In summary, the low invalid rates for the AS-PCR test and ARMS test allowed both tests to provide results in the majority of cases. The AS-PCR test and the ARMS test results were highly concordant with one another and highly sensitive. Of the three methods, the AS-PCR test had the lowest DNA input requirements and was the only test with automated results reporting.

**What the paper adds**

- New targeted therapies are being approved with companion diagnostic tests. Unfortunately, this has not been the case for EGFR inhibitors. Therefore, it is not clear which technique offers the best performance in terms of sensitivity, specificity, reproducibility and accuracy. We performed a two-centre study to compare the analytic performance and workflow characteristics of the three most widely used methodologies for EGFR mutation testing.

**Take-home messages**

- The low invalid rates for the AS-PCR test and ARMS test allowed both tests to provide results in the majority of cases.
- The AS-PCR test and the ARMS test results were highly concordant with one another and highly sensitive.
- Of the three methods, the AS-PCR test had the lowest DNA input requirements and was the only test with automated results reporting.
Acknowledgements  We would like to thank the Sequencing Service at the Spanish National Cancer Centre (CNIO) for performing the direct sequencing assays. Miller Medical was contracted by Roche Molecular Systems and contributed in the drafting of this manuscript. This study was sponsored by Roche Molecular Systems, Pleasanton, CA.


Funding  This study was sponsored by Roche Molecular Systems, Pleasanton, CA.

Competing interests  DGC declares the receipt of honoraria from Roche Molecular Systems (Pleasanton, CA) and Roche Products Ltd (UK). FLR declares the receipt of honoraria from Roche Molecular Systems (Pleasanton, CA) and Roche Diagnostics (Spain). FS, JV, RC, JT are employees of Roche Molecular Systems. HJL is a former employee of Roche Molecular Systems. BA, RM, EC, BG and DM declare that they have no conflicts of interest.

Provenance and peer review  Not commissioned; externally peer reviewed.

Open Access  This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/3.0/

REFERENCES


4 Agency EM. Kirensk (gefitinib)—EPAP summary for the public 2009.

5 Genentech IAPU, Inc. Tarceva Prescribing Information 2012.


16 Cobas DNA Sample Preparation Kit CE-IVD package insert 2011.

17 Cobas EGFR Mutation Kit CE-IVD package insert 2011.


