Analytical performance of a highly sensitive system to detect gene variants using next-generation sequencing for lung cancer companion diagnostics

Kikuya Kato\textsuperscript{a}, Jiro Okami\textsuperscript{b}, Harumi Nakamura\textsuperscript{c}, Keiichiro Honma\textsuperscript{d}, Yoshiharu Sato\textsuperscript{e}, Seiji Nakamura\textsuperscript{a}, Yoji Kukita\textsuperscript{a,c}, Shuichi Nakatsuka\textsuperscript{d}, Masahiko Higashiyama\textsuperscript{b,f}

\textsuperscript{a} Laboratory of Medical Genomics, Nara Institute of Science and Technology, Nara, Japan.
\textsuperscript{b} Department of Thoracic Surgery, Osaka International Cancer Institute, Osaka, Japan.
\textsuperscript{c} Laboratory of Genomic Pathology, Osaka International Cancer Institute, Osaka, Japan.
\textsuperscript{d} Department of Diagnostic Pathology and Cytology, Osaka International Cancer Institute, Osaka, Japan.
\textsuperscript{e} DNA Chip Research Inc., Tokyo, Japan.
\textsuperscript{f} Department of Thoracic Surgery, Higashiosaka City Medical Center, Osaka, Japan.

Corresponding author: Kikuya Kato, kikuya@kato.nifty.jp

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Companion diagnostics, which predict the efficacy of molecular targeted agents based on genetic information, are indispensable for the treatment of advanced non-small cell lung carcinoma. Recent increase in the number of molecular targeted agents and the corresponding target genes have led to the demand for the simultaneous testing of multiple genes. Although gene panels using next-generation sequencing (NGS panels) are ideal for this purpose, conventional panels require high tumor content, and biopsy samples often do not meet this requirement. We developed a new NGS panel called a compact panel, to accommodate biopsy samples without the restriction of tumor content. The compact panel is characterized by high sensitivity, with detection limits for mutations of 0.14%, 0.20%, 0.48%, 0.24%, and 0.20% for EGFR exon 19 deletion, L858R, T790M, BRAF V600E, and KRAS G12C, respectively. Mutation detection also has a high quantitative ability, with correlation coefficients ranging from 0.966 to 0.992. The panel detected fusions in samples whose tumor cell content was as low as 1%. The compact panel exhibited good concordance with approved tests as follows: EGFR positive, 100.0 (95% confidence interval 95.5-100); EGFR negative, 90.9 (82.2-96.3); ALK positive, 96.7 (83.8-99.9); ALK negative, 98.4 (97.2-99.2); ROS1 positive, 100 (66.4-100); ROS1 negative, 99.0 (97.1-99.2); MET positive, 98.0 (89.0-100); MET negative 100 (92.8-100). The analytical performance demonstrated that the compact panel can handle various
types of biopsy samples obtained by routine clinical practice, without requiring strict pathological monitoring like in case of conventional NGS panels.

Keywords: non-small cell lung carcinoma, companion diagnostics, next-generation sequencing.
Introduction

One of the major paradigms in modern cancer therapies is the prediction of the efficacy of anti-cancer agents, based on genomic information of the patient. The foremost example of this paradigm is the therapy for advanced non-small cell lung carcinoma (NSCLC). Molecular targeted agents are selected for patients based on mutations in the target genes, a process called companion diagnostics. The era of precision medicine began with the discovery of a correlation between EGFR mutations and gefitinib efficacy [1]. Since its discovery, the number of molecular targeted agents for NSCLC has continuously increased. Current molecular targeted agents with companion diagnostics include osimertinib [2], alectinib [3], crizotinib [4], dabrafenib plus trametinib [5], tepotinib [6], and selpertinib [7] for EGFR, ALK, ROS1, BRAF, MET, and RET, respectively. Other agents targeting KRAS [8] and EGFR with exon 20 insertion [9] will be available in the near future.

In the early days, only EGFR tyrosine kinase inhibitors were available, and diagnostic tests for one gene were sufficient for practical use. With the increase in the number of agents, however, there is now a strong demand for testing multiple genes simultaneously. In this context, gene panels, or next-generation sequencing (NGS) panels, analytic/diagnostic systems to detect variants in multiple cancer-related genes using NGS, such as Foundation One CDx (Foundation Medicine, Inc.) and Oncomine™ Dx Target Test (Thermo
Fisher Scientific), should be useful.

However, these panels are not necessarily optimized for diagnostics. The main limitation is the restriction of tumor samples: high tumor content (> 20%) is required to perform the panel tests. This restriction prevents a considerable proportion of samples from being tested, and such samples must be subjected to single-gene tests based on real-time polymerase chain reaction (PCR). These NGS panels were originally developed for samples collected for research purposes, and are not optimized for testing various types of clinical samples in the real world. This shortcoming is primarily due to the low sensitivity of variant detection. We developed a NGS panel, named the compact panel for companion diagnostics of NSCLC. The compact panel achieved higher sensitivity than that of conventional NGS panels and single-gene tests.

Results

Design of the compact panel

The compact panel limits the number of target genes to reduce costs and improve the flexibility of panel design. The panel consists of four modules: two DNA modules and two RNA modules, using DNA and RNA as templates, respectively. The target mutations were as follows: DNA module I for EGFR exon 19 deletion, EGFR L858R, BRAF V600E, and KRAS G12C; DNA module II for other EGFR mutations, Her2 exon 20, and MET exon 14
skipping/amplification; RNA module I for ALK, fusion and MET exon 14 skipping; and RNA module II for ROS1 and RET fusions. The current design of the compact panel is presented in Table 1. Detailed lists of the variants of ALK, ROS1 and RET fusions are shown in Supplementary File 1. The modular structure of the panel simplifies the process of obtaining official authorization, since the analytical performance can be evaluated for each module, without the need to re-evaluate the entire panel. In addition, replacement or addition of modules can help deal with new diagnostic genes in the future.

Table 1. Modules of the compact panel. RNA module I includes HGPRT1 for the control of amplification.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Target mutations / fusion variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA module I</td>
<td>EGFR Exon 19 deletion, L858R, T790M, L861Q, L861R</td>
</tr>
<tr>
<td></td>
<td>BRAF V600E</td>
</tr>
<tr>
<td></td>
<td>KRAS G12C</td>
</tr>
<tr>
<td>DNA module II</td>
<td>EGFR G719X, S768I, exon 20 insertion, L861Q, L861R</td>
</tr>
<tr>
<td></td>
<td>HER2 Exon 20 insertion</td>
</tr>
<tr>
<td></td>
<td>MET Exon 14 skipping, amplification</td>
</tr>
<tr>
<td>RNA module I</td>
<td>ALK EML4, 22 variants; KIF5, 3; TFG, 1; HIP, 3; KLC1, 1</td>
</tr>
<tr>
<td></td>
<td>MET Exon 14 skipping</td>
</tr>
<tr>
<td>DNA module II</td>
<td>ROS1 CD74, 2 variants; SLC34A2, 7; EZR, 1; GOPC, 2; SDC4, 4; LRIG, 1; TPM3, 1; CCDC6, 1; KDEL2, 1; CD74, 1</td>
</tr>
<tr>
<td></td>
<td>RET KIF5B, 7 variants; CCDC6, 1; NCOA4, 1</td>
</tr>
</tbody>
</table>

To increase the sensitivity of detection, we applied different strategies for mutations and fusions. To detect mutations, templates were amplified from genomic DNA, and deep sequencing, that is, repeated sequencing of target regions, was performed. Because sensitivity depends on the number of sequence reads [10], we set the number to 5,000 instead of 700 with
Oncomine™ Dx Target Test. To detect fusions, multiplex PCR was used to amplify junctions of known variants using RNAs as templates. Multiplex PCR is designed to optimize the amplification of fusions, and not normal genes, thereby reducing the concentration of normal genes. Consequently, high sensitivity was achieved.

PCR primers were designed such that the sizes of the amplified products were less than 100 base pairs, which enabled PCR amplification with deteriorated DNA/RNA templates.

**Sensitivity of mutation detection using DNA as template**

The thresholds for mutation detection were set using an anomaly detection algorithm [10]. In anomaly detection, the probability of false positives was estimated from the measured values of normal DNA, assuming a Poisson distribution. The threshold values were set such that the probability of false positives was $10^{-10}$. The allele frequency values are listed in Table 2.

To assess the sensitivity of DNA module I, we measured 24 samples of 1% and 0% mutant allele DNA for each test. There were no false positives, that is, mutation positives in 0% mutated DNA, and no false negatives, that is, mutation negatives in 1% mutated DNA (Table 2).

**Sensitivity of fusion detection using RNA as template**

The thresholds for fusion detection were also set using an anomaly detection
algorithm. The threshold values were set such that the probability based on the Poisson distribution of false positives was $10^{-5}$. The values were defined by the TM score, the number of positive reads in 100,000 reads, and are shown in Table 2.

Fusion-positive samples for sensitivity tests were prepared so that 1% of the RNA was derived from fusion-positive cells. In the sensitivity tests, we measured 24 or 48 samples of fusion-positive and fusion-negative RNA for each test. The false positive/negative rates were less than 0.5% (Table 2).

Table 2. Sensitivity test of the compact panel. The number of analyzed samples was 24 except ALK false negative (48 samples). Positive samples for the DNA module includes 1% mutated DNA; those for the RNA module were prepared so that 1% of the RNA was derived from fusion-positive cells.

<table>
<thead>
<tr>
<th></th>
<th>Threshold of detection DNA, % allele frequency; RNA, TM score</th>
<th>Sensitivity test</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA module</td>
<td></td>
<td>False negative</td>
</tr>
<tr>
<td>EGFR exon 19 del</td>
<td>0.14</td>
<td>0</td>
</tr>
<tr>
<td>EGFR L858R</td>
<td>0.20</td>
<td>0</td>
</tr>
<tr>
<td>EGFR T790M</td>
<td>0.48</td>
<td>0</td>
</tr>
<tr>
<td>BRAF V600E</td>
<td>0.24</td>
<td>0</td>
</tr>
<tr>
<td>KRAS G12C</td>
<td>0.20</td>
<td>0</td>
</tr>
<tr>
<td>RNA module</td>
<td></td>
<td>False negative</td>
</tr>
<tr>
<td>ALK fusion</td>
<td>188</td>
<td>0</td>
</tr>
<tr>
<td>ROS1 fusion</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>RET fusion</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>MET exon 14 skipping</td>
<td>28</td>
<td>0</td>
</tr>
</tbody>
</table>

Quantification of mutation detection

Deep sequencing enables excellent quantification of mutation alleles. The
quantitative ability of DNA module I was examined using artificial DNA samples prepared such that 1-8% of the total DNA consisted of mutant alleles. The results are shown in Figure 1.

![Figure 1](https://example.com/figure1.png)

Figure 1. Quantitative ability of DNA module I. Horizontal axis, mutation allele frequency of artificial DNA (%); vertical axis, mutation allele frequency measured with the compact panel (%). Each data point is average of eight samples.

There was very good linearity between the inoculated amounts of the mutant alleles, and the observed mutant allele frequency deduced from deep sequencing. The correlation coefficients were as follows: *EGFR* exon 19 deletion, 0.973; *EGFR* L858R, 0.972; *EGFR* T790M, 0.966; *BRAF* V600E, 0.992; *KRAS* G12C, 0.991. The difference of the slopes was likely to be due to that in amplification efficiency.
Concordance with conventional diagnostic tests

The performance of the compact panel was then compared with that of approved diagnostic tests. The reference diagnostic tests were as follows: EGFR, Cobas® EGFR Mutation Test v2 (Roche Diagnostics K.K.); ALK, Histofine ALK iAEP® kit (Nichirei Bioscience Inc.); Vysis® ALK Break Apart FISH probe kit (Abbott Laboratories); MET, Archer®MET (Archer DX, Inc.). Formalin-fixed paraffin-embedded (FFPE) NSCLC samples were analyzed using a compact panel and a reference test simultaneously, and the concordance of both tests was examined. The results are shown in Table 3. The good concordance proved the feasibility of the compact panel for practical use.

Table 3. Concordance of the compact panel with approved diagnostic tests.

<table>
<thead>
<tr>
<th></th>
<th>Compact panel positive</th>
<th>Compact panel negative</th>
<th>Reference test positive identity rate</th>
<th>Reference test negative identity rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFR</td>
<td>73/0</td>
<td>7/70</td>
<td>100.0% (95.5-100)</td>
<td>90.9% (82.2-96.3)</td>
</tr>
<tr>
<td>ALK</td>
<td>29/1</td>
<td>11/692</td>
<td>98.7% (83.8-99.9)</td>
<td>98.4% (97.2-99.2)</td>
</tr>
<tr>
<td>ROS1</td>
<td>9/0</td>
<td>1/99</td>
<td>100.0% (99.4-100.0)</td>
<td>99.0% (97.1-99.2)</td>
</tr>
<tr>
<td>MET</td>
<td>48/1</td>
<td>0/50</td>
<td>98.0% (89.0-100)</td>
<td>100.0% (92.8-100)</td>
</tr>
</tbody>
</table>

Seven discordant EGFR samples (compact panel positive and reference test negative) were analyzed using digital PCR, and all the samples were found to be mutation-positive, agreeing with the results of the compact panel. One discordant ALK sample was analyzed with the non-overlapping
integrated read sequencing system (NOIR-SS) [11], and was found to be an
ALK fusion with C2ORF71, a rare fusion type not covered in the compact panel
and in Oncomine™ Dx Target Test.

Incidence of mutations and fusions
NSCLC samples stored at the Osaka International Cancer Institute were
screened with the compact panel, and the incidence of mutations and fusions at
the population level was determined (Table 4). The figures were compared with
those obtained from Oncomine™ comprehensive assay in a nationwide project,
named Scrum-Japan [12].

Table 4. Frequencies of mutations and fusions (%). Samples of the Osaka International Cancer
Institute were obtained during the period from October 26th 2015 to August 24th 2020. Because
samples from Scrum-Japan were restricted to EGFR mutation-negative category, the frequency
of EGFR mutations is assumed to be 50%. The frequency of BRAF V600E is assumed to be
50% of all BRAF mutations.

<table>
<thead>
<tr>
<th></th>
<th>Osaka International Cancer Institute</th>
<th>Scrum-Japan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assay system</td>
<td>Compact panel</td>
<td>Oncomine Comprehensive Assay</td>
</tr>
<tr>
<td>Number of samples</td>
<td>1000</td>
<td>3919</td>
</tr>
<tr>
<td>Frequencies (%)</td>
<td>EGFR 43.0</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>BRAF V600E 0.6</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>ALK fusion 2.8</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>ROS1 fusion 9</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>MET exon 14 skipping 4.4</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>RET fusion 1.5</td>
<td>1.3</td>
</tr>
</tbody>
</table>
Both results are generally consistent, except for a higher incidence of ALK fusion and MET skipping and low incidence of ROS1 fusion in the compact panel cohort.

Discussion

Companion diagnostics have long been performed using real-time PCR-based tests for individual genes. However, owing to the increasing number of target genes, there is now strong demand for testing multiple genes simultaneously. In this context, the simultaneous analysis of multiple genes by NGS is ideal. In the USA, NGS panels were performed with 44% of NSCLC patients requiring genetic tests [13], and are becoming indispensable for NSCLC treatment.

However, the introduction of NGS panels into Japanese medical practice has not been straightforward. In the USA, NSCLC samples subjected to genetic tests are usually obtained by core needle biopsy or transthoracic fine needle aspiration biopsy, guided by computed tomography or ultrasound [14]. In contrast, biopsy is mainly performed using bronchoscopy in Japan. This difference in biopsy practice causes a unique problem that is not apparent in the medical environment in the USA. In Japan, the feasibility of NGS panels is generally lower than that of other diagnostics in clinical practice, especially with regard to nonsurgical biopsy [15]. The success rate of the Oncomine™ Dx Target Test is influenced by tissue size and tumor cell count [16, 17]. Thus, strict
pathological monitoring is necessary to select samples subjected to the test, and consequently, a considerable fraction of the samples is excluded from the test.

This issue can be addressed by introducing a new NGS panel with high sensitivity. We developed a compact panel, which is an NGS panel that accommodates various types of NSCLC samples without requiring strict pathological monitoring. The sensitivity of the compact panel suggests detection of mutation/fusion in samples with 1% tumor content. The confirmed detection limit of the DNA module I of the compact panel, which is 1% mutation allele frequency, is lower than that of Oncomine™ Dx Target Test, which is 6-13%, and Cobas® EGFR Mutation Test v2, which is 1.26 to 6.81% (FDA Summary of Safety and Effectiveness Data). Because the amount of DNA/RNA required is only 5 ng, lesser amount of tissue would need be obtained by routine biopsy practice. Although the compact panel was originally designed for the Japanese medical environment, such high-sensitivity NGS panels would be beneficial in other medical environments as well.

The compact panel can handle various sample types, including cytological samples. Formalin fixation, which may deteriorate the quality of nucleic acids, is not necessary for genetic testing. Ammonium sulfate solution enables efficient preservation of RNA/DNA for several days at room temperature [18], and is expected to replace formalin fixation. It would be particularly useful as a washing solution for samples obtained by cytodiagnostic brushing or curette washing.
Materials and methods

DNA and RNA samples

*Sensitivity tests.* DNA samples, including various fractions of mutant alleles, were prepared using the following solid tumor analysis reference standards (Horizon Discovery Ltd.): HDx FFPE *EGFR* e19 del. 50%; HDx FFPE *EGFR* L858R 50%; FFPE *EGFR* T790M 50%; HDx FFPE *BRAF* V600E 50%. To adjust allele frequencies, wild-type FFPE genomic DNA from the lung, including bronchioles (Cureline, Inc.), was used. RNA samples, including fusions, were constructed from *ALK-RET-ROS1* targeted FFPE RNA Fusion Reference Standards (Horizon Discovery Ltd.). For *MET* exon 14 skipping, synthesized DNA was used, skipping the need for reverse transcription. To adjust RNA concentration, wild-type FFPE RNA from the lungs, including bronchioles (Cureline, Inc.), was used.

*Concordance tests.* FFPE NSCLC samples were obtained by surgical resection, and stored at the Osaka International Cancer Institute. DNA and RNA were purified from the samples using a Maxwell® RSC instrument (Promega Corporation). Studies using clinical samples were approved by the ethics committees of the Osaka International Cancer Institute, and the Nara Institute of Science and Technology.
Library preparation and sequencing for DNA modules

PCR amplification was performed in 50 µL of reaction mixture containing 1 x buffer (Toyobo, Inc.), 0.2 mM dNTPs, 1.5 mM MgSO₄, 5 ng of DNA purified from FFPE NSCLC, 0.3 µM each of primer mixture and 0.02 U of KOD -Plus-Neo (Toyobo, Inc.). Forty cycles of 98 °C for 10 s and 62 °C for 30 s were performed. After purification with AMPure XP (Beckman Coulter Life Sciences), the amplified products were subjected to library construction using the GenNext® NGS Library Prep Kit (TOYOBO, Inc.). To discriminate individual samples, index sequences were introduced using TruSeq DNA Single Indexes Set A or TruSeq DNA CD Indexes Set (Illumina, Inc.). Sequencing was performed using the MiSeq System (Illumina, Inc.). The minimum number of reads per fragment was 5,000.

Library preparation and sequencing for RNA modules

Reverse transcription was performed in 20 µL of reaction mixture containing 1 x ReverTra buffer (Toyobo, Inc.), 1 mM dNTPs, 10 ng of RNA purified from FFPE NSCLC, and 100 U of ReverTra Ace (Toyobo, Inc.). After denaturation with RNA and 9-base random primer (Toyobo, Inc.) at 65 °C for 5 min, the reaction mixture was incubated at 30 °C for 10 min, and then at 42 °C for 60 min. PCR amplification was performed in 50 µL of reaction mixture containing 1 x buffer (Toyobo, Inc.), 0.4 mM dNTPs, 1.5 mM MgSO₄, the 20 µl reaction mixture mentioned above, 0.25 µM each of the primer mixtures, and 1 U of KOD -Fx-
Neo (Toyobo, Inc.). Forty cycles of 98 °C for 15 s, 60 °C for 30 s, and 68 °C for 10 s, followed by extension at 68 °C for 1 min, were performed. After purification with AMPure XP (Beckman Coulter Life Sciences), the amplified products were subjected to library construction and sequencing as described in the previous section. The minimum number of reads per sample was 300. All primer sequences will be supplied upon request.

Analysis for discordant samples

Digital PCR was performed using the QX200 Droplet Digital PCR System (Bio-Rad Laboratories, Inc.), following the supplier’s protocol. NOIR-SS assay was performed as previously described [11].

Acknowledgements

The authors thank Miho Ishii, Yuki Mori, Yui Nose and Yumi Ueda for laboratory work. The authors also thank Hiroyuki Sato and Motohiko Tanino for data analysis. The authors are profoundly grateful for Yuji Horikawa for administration. Laboratory of Medical Genomics in Nara Institute of Science and Technology is an endowed chair for Kikuya Kato and Yoji Kukita provided by Gene Metrics LLC.
References


