Application of liquid biopsy-based targeted capture sequencing analysis to improve the precision treatment of non-small cell lung cancer by tyrosine kinase inhibitors

Lei Zhang,1 John Coffin,1 Kim Formenti,1 Quincy Chu,2 Iyare Izevbaye1

ABSTRACT
Background Targeted therapy of patients with non-small cell lung cancer (NSCLC) who harbour sensitising mutations by tyrosine kinase inhibitors (TKIs) has been found more effective than traditional chemotherapies. However, target genes status (eg, epidermal growth factor receptor (EGFR) TKIs sensitising and resistant mutations) need to be tested for choosing appropriate TKIs. This study is to investigate the performance of a liquid biopsy-based targeted capture sequencing assay on the molecular analysis of NSCLC.

Methods Plasma samples from patients with NSCLC who showed resistance to the first/second-generation EGFR TKIs treatment were collected. The AVENIO ctdNA Expanded Kit is a 77 pan-cancer genes detection assay that was used for detecting EGFR TKIs resistance-associated gene mutations. Through comparison of the EGFR gene testing results from the Cobas EGFR Mutation Test v2, and UltraSEEK Lung Panel, the effectiveness of the targeted capture sequencing assay was verified.

Results A total of 24 plasma cell-free DNA (ctDNA) samples were tested by the targeted capture sequencing assay. 33.3% (8/24) ctdNA samples were positive for EGFR exon 20 p.T790M which leads to EGFR independent TKIs resistance. 8.3% (2/24) and 4.2% (1/24) samples were positive for mesenchymal-epithelial transition gene amplification and B-Raf proto-oncogene, serine/threonine kinase exon 15 p.V600E mutations which lead to EGFR independent TKIs resistance. The median value of the p.T790M variant allele fraction and variant copy numbers was 2% and 36.10 copies/mL plasma, respectively. The next-generation sequencing test showed higher than 90% concordance with either MassArray or qPCR-based methods for detecting either EGFR TKIs sensitising or resistance mutations.

Conclusion The targeted capture sequencing test can support comprehensive molecular analysis needed for TKIs treatment, which is promising to be clinically applied for the improved precision treatment of NSCLC.

INTRODUCTION
Lung cancer is the most commonly diagnosed primary malignancy in Canada and is currently the leading cause of cancer-related deaths in the country.1 Non-small cell lung cancer (NSCLC) is the major type of lung cancer, which accounts for more than 80% of all lung cancers.2 As lung cancer patients are often found to have advanced disease at the time of initial diagnosis, palliative therapies with the alleviation of symptoms and extension of survival are the main objectives of clinical treatments.3 Molecular analysis of tumour biomarker genes has become the standard of practice in the treatment of NSCLC. Patients with NSCLC who harbour sensitising mutations in epidermal growth factor receptor (EGFR) genes can be offered small-molecule
tyrosine kinase inhibitors (TKIs), which are significantly more effective than traditional chemotherapies.6 Besides, other genes targeted therapies, specifically against sensitising mutations in ALK, ROS1 and B-Raf proto-oncogene, serine/threonine kinase (BRAF) also have a demonstrable clinical improvement in the outcome of NSCLC.2–5

However, after the median treatment duration (8–10 months), the majority of patients with NSCLC (50%–60%) will acquire a novel EGFR mutation in exon 20 known as p.T790M, which confers resistance to the first-generation and second-generation TKIs’ treatment.6 Osimertinib can be alternatively used to treat patients with the acquired EGFR exon 20 p.T790M mutation.7 Acquired EGFR exon 20 p.C797S mutation which mediates resistance to the third-generation EGFR TKIs treatment may also develop in EGFR exon 20 p.T790M-positive patients with NSCLC.8 Non-EGFR-dependent mechanisms of acquired TKIs resistance, such as gene copy number amplification of mesenchymal-epithelial transition (MET) gene and human epidermal growth factor receptor 2 gene, mutations of phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha and BRAF genes have been reported as well.9–12 It is critical to systematically monitor the TKIs-based targeted therapy-associated gene mutations in patients with NSCLC for the improved outcome of targeted therapy.

Liquid biopsy offers a cost-effective and low-risk alternative to tissue biopsy for testing tumour biomarker genes in patients with NSCLC.13 14 The concordance between the testing of EGFR mutations in tumour DNA and cell-free DNA (cfDNA) of patients with NSCLC has been reported as 79%, though the mutation rates detected in plasma cfDNA reflect approximately 20% of that detected from tumour tissues.15 16 Liquid biopsy has been clinically used as a complementary assay to tissue biopsy for testing EGFR mutations in patients with NSCLC.17 Next-generation sequencing (NGS) was found to be able to detect EGFR exon 20 p.T790M in the liquid biopsy samples of patients with NSCLC with percentages of 0.3%–21.0%.18 More importantly, NGS technology-based assays are superior to currently used PCR-based assays with the capacity to simultaneously test multiple genes in one assay.19 20 Thus, more target genes can be tested by NGS to match patients with appropriate NSCLC treatment drugs for improved precision treatment.21–23

This research project aims to investigate the performance of an optimised hybridisation capture-based NGS assay (AVENIO ctDNA expanded kit, Roche Diagnostics, Canada) on testing somatic mutations in the plasma cfDNA for targeted treatment of patients with NSCLC. Based on the EGFR testing results, the concordance of the NGS test with clinically validated MassArray (UltraSEEK Lung Panel, Agena Bioscience, USA) and qPCR-based (Cobas EGFR Mutation Test v2) assays was evaluated.

MATERIALS AND METHODS

Patient and public involvement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

Patients and samples collection

NSCLC patients in advanced stages who experienced disease progression after the targeted treatment with first or second-generation EGFR TKIs were randomly included in this study. Ten millilitre peripheral blood was collected into Streck tubes. The plasma specimen was separated within 4 hours from peripheral blood collection by two rounds of centrifugation at 20°C, with 10 min at 1600×g and 3000×g respectively, and stored at −80°C before DNA extraction. Twenty-four plasma samples were collected and tested by the NGS test. Due to the availability of sample amount, 23 samples were tested by the MassArray-base assay as a replicate detection and 12 samples were tested by the qPCR-based assay as a second replicate detection.

Testing pan-cancer genes mutations in cfDNA by NGS

The AVENIO ctDNA expanded kit (Roche Diagnostics) is a hybridisation capture sequencing-based 77 genes pan-cancer assay (online supplemental table S1). AVENIO ctDNA Isolation Kit (Roche Diagnostics) was used to extract cfDNA from plasma according to the user’s manual. The extracted cfDNA was analysed by Agilent High Sensitivity DNA Analysis Kit (Agilent Technologies, California, USA) and Qubit dsDNA HS Assay Kit (ThermoFisher Scientific, California, USA) for quality control, and then used for library preparation with 10–50 ng cfDNA input. Prepared libraries were sequenced on the NextSeq 500 500/550 High Output Kit V2 (300 cycles) on Illumina NextSeq sequencing platform (Illumina, California, USA) and analysed by the AVENIO oncology analysis software V.2.0.0 (Roche Diagnostics) according to manufacturer’s instructions.24 25

Testing EGFR variants in cfDNA by qPCR and MassArray

Cobas EGFR Mutation Test v2 (Roche Diagnostics), a real-time qPCR-based in vitro diagnostics assay, was used for testing EGFR variants in 2 mL plasma of patients with NSCLC. A clinically validated UltraSEEK Lung Cancer Panel (Agena Bioscience) was also used as a reference method for testing EGFR mutations in 10 ng input cfDNA. Detailed information about these testing methods is provided in online supplemental file 1.

Statistical analysis

AVENIO ctDNA Analysis Software (Roche Diagnostics) was used for sequencing data analysis and generating diagnostic reports.26 With the data of abundance for different biomarker gene variants, the distribution of gene variants’ abundance is shown in box plots and the
median value was calculated by using an R-based statistical data analysis tool (BoxPlotR). 27

RESULTS

**EGFR gene testing by NGS**

Plasma cfDNA is composed of fragmented double-stranded DNA circulating in the blood, which is mainly generated from the apoptotic and necrotic cells. 28 29 The quality of extracted cfDNA samples was checked on the Bioanalyzer (Agilent Technologies). The major peak of the extracted cfDNA sample was typically located at 160–200 bp (online supplemental figure S1A,B). Several distinguishable cfDNA peaks with a size shift were observed, confirming the successful ligation of adapters to target gene fragments (online supplemental figure S1C,D). The library prepared by the AVENIO ctDNA Expanded Kit (Roche Diagnostics) typically showed the primary cfDNA peak at ~300 bp with additional small cfDNA peaks of larger size present (online supplemental figure S1C,D).

In this study, a total of 24 plasma cfDNA samples were tested by the NGS assay. The metrics of the NGS testing are shown in online supplemental table S2. After alignment analysis, the average ratio of mapped reads, and the on-target rate were 94.31% and 71.62% (online supplemental figure S2A). The typical median value of coverage uniformity was higher than 99% (online supplemental figure S2B). It can be observed that the high-quality sequencing data showed a very low error rate, excellent sequencing depth for target genes and high coverage uniformity, which also lead to a high theoretical sensitivity with a median value of 98.71% (online supplemental figure S2A).

The sequencing detection showed that 66.7% (16/24) ctDNA samples were positive for EGFR sensitising mutations, and 50.0% (8/16) of which were also positive for the first-generation and second-generation TKIs resistance mutation (p.T790M) (online supplemental table S3). The percentage of different EGFR sensitising mutations is shown in figure 1. The range of variant allele fraction for EGFR sensitising mutations was 0.11%–45.16% with the median value of 4% (figure 1A). It can be observed that exon 21 p.L858R and exon 19 dels are the main EGFR sensitising mutations detected in the plasma cfDNA, the percentages of which were 81.25% (13/16). And e.g.T746_A750del was the main exon 19 dels variant detected which compromise 62.5% (5/8) of all detected exon 19 dels variants (online supplemental table S3). And the variant allele fraction for EGFR exon 20 p.T790M was 0.08%–8.28% with the median value of 2% (figure 1A). The range of variant copy numbers for EGFR sensitive mutations was 1.79–4950 with the median value of 135.00 copies/mL plasma (figure 1B). The range of variant copy numbers for EGFR exon 20 p.T790M was 0.68–561 with the median value of 36.10 copies/mL plasma (figure 1B).

**Concordance of NGS, qPCR and MassArray on EGFR gene testing**

The qPCR detection showed 75.0% (9/12) and 33.3% (3/9) cfDNA samples were positive for EGFR sensitising mutations and EGFR exon 20 p.T790M (online supplemental table S3). Except for one inconclusive sample, the MassArray-based assay detected 56.5% (13/23) cfDNA samples that were positive for EGFR sensitising mutations, and 61.5% (8/13) of which were also positive for the EGFR exon 20 p.T790M (online supplemental table S3). The NGS-based assay showed 91.3% (21/23) concordance with the MassArray-based test for detecting either EGFR sensitising or EGFR TKIs resistant mutation (table 1). The qPCR assay showed 83.3% (10/12) and 100% (12/12) concordance with the MassArray-based test for testing EGFR sensitising mutations and exon 20 p.T790M variant, respectively (table 1). The concordance of the three assays on testing EGFR variants in the plasma cfDNA from patients with NSCLC was 83.3% (10/12) (table 1).

---


---
In addition, the median value of variant allele fraction value of 133.00 copies/mL plasma (figures 2A and 2B). And the range of variant copy numbers for TP53 was 66.70 copies/mL plasma (figure 2A) (table 2). The range of variant allele fraction for EGFR was 0.08%–45.16% with the median value of 2% (figures 1A and 2A). And the range of variant allele fraction for TP53 was 0.23%–12.65% with a median value of 4% (figures 2A and 3A). The range of variant copy number for EGFR was 0.679 copies/mL to 4950 copies/mL with the median value of 66.70 copies/mL plasma (figures 1B and 2B). And the range of variant copy numbers for TP53 was 12.2 copies/mL to 1110 copies/mL with the median value of 133.00 copies/mL plasma (figures 2B and 3B). In addition, the median value of variant allele fraction for NTRK1, RB1, PMS2, PTCH1, GNAS was 3%, 8%, 46%, 2% and 1%, respectively. And the median value of variant allele fraction for ROS1, BRAF, ERBB2, CTNNB1, Pten was less than 1% (figure 2A) (table 2). The median value of variant copy number for NTRK1, RB1, PMS2, PTCH1, GNAS, CTNNB1 was 225.00, 598.00, 465.00, 22.50, 21.20, 63.00 copies/mL plasma (figure 2B) (table 2). And the median value of variant copy numbers for ROS1, BRAF, ERBB2, Pten was less than 10 copies/mL plasma (table 2).

According to the sequencing data analysis process for the Avenio ctDNA kit (Roche Diagnostics), the copy number variant (CNV) score is statistically calculated by accounting for the log2 ratio to normal gene copy number, and the higher CNV score indicates the higher confidence value. The NGS detection showed that the copy number amplification variants were detected on EGFR and MET genes. Statistically, the range of the CNV scores for EGFR gene variants was from 4.14 to 29.06 with a median value of 14.45. And the range of the CNV scores

### Table 1

Concordance of EGFR testing results from NGS, MassArray and qPCR

<table>
<thead>
<tr>
<th>Mutations</th>
<th>Variants</th>
<th>NGS vs MassArray</th>
<th>NGS vs qPCR</th>
<th>MassArray vs qPCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFR sensitising mutations</td>
<td>Exon 18 p.G719A</td>
<td>100% (23/23)</td>
<td>100% (12/12)</td>
<td>100% (12/12)</td>
</tr>
<tr>
<td></td>
<td>Exon 20 p.S768I</td>
<td>100% (23/23)</td>
<td>100% (12/12)</td>
<td>100% (12/12)</td>
</tr>
<tr>
<td></td>
<td>Exon 21 p.L858R</td>
<td>95.7% (22/23)</td>
<td>100% (12/12)</td>
<td>91.7% (11/12)</td>
</tr>
<tr>
<td></td>
<td>Exon 21 p.L861Q</td>
<td>100% (23/23)</td>
<td>100% (12/12)</td>
<td>100% (12/12)</td>
</tr>
<tr>
<td></td>
<td>Exon 19 deletions</td>
<td>95.7% (22/23)</td>
<td>91.7% (11/12)</td>
<td>91.7% (11/12)</td>
</tr>
<tr>
<td>TKIs resistant EGFR mutation</td>
<td>Exon 20 p.T790M</td>
<td>91.3% (21/23)</td>
<td>91.7% (11/12)</td>
<td>100% (12/12)</td>
</tr>
</tbody>
</table>

EGFR, epidermal growth factor receptor; NGS, next-generation sequencing; TKI, tyrosine kinase inhibitor.

### Table 2

Testing tumour biomarker genes (other than epidermal growth factor receptor and TP53) variants in cell-free DNA by the next-generation sequencing test

<table>
<thead>
<tr>
<th>SNV gene</th>
<th>Variant</th>
<th>Coding change</th>
<th>Variant description</th>
<th>Allele fraction (%)</th>
<th>Variant depth</th>
<th>Molecular mutant per mL (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTRK1</td>
<td>Exon 17 p.I737I</td>
<td>c.2211C&gt;T</td>
<td>Synonymous variant</td>
<td>2.97</td>
<td>119/4007</td>
<td>225</td>
</tr>
<tr>
<td>RB1</td>
<td>Exon 19 p.V654M</td>
<td>c.1960G&gt;A</td>
<td>Missense variant</td>
<td>7.91</td>
<td>286/3616</td>
<td>598</td>
</tr>
<tr>
<td>ROS1</td>
<td>Exon 41 p.L2157V</td>
<td>c.6469C&gt;G</td>
<td>Missense variant</td>
<td>0.28</td>
<td>4/1417</td>
<td>2.1</td>
</tr>
<tr>
<td>PMS2</td>
<td>Exon 11 p.A660A</td>
<td>c.1980C&gt;T</td>
<td>Synonymous variant</td>
<td>44.08</td>
<td>901/2044</td>
<td>425</td>
</tr>
<tr>
<td>PMS2</td>
<td>Exon 3 p.D70G</td>
<td>c.209A&gt;G</td>
<td>Missense variant</td>
<td>47.13</td>
<td>345/732</td>
<td>505</td>
</tr>
<tr>
<td>PTCH1</td>
<td>Exon 17 p.N929N</td>
<td>c.2787C&gt;T</td>
<td>Synonymous variant</td>
<td>46.50</td>
<td>1789/3847</td>
<td>4490</td>
</tr>
<tr>
<td>PTCH1</td>
<td>Exon 22 p.A1247A</td>
<td>c.3741G&gt;A</td>
<td>Synonymous variant</td>
<td>1.57</td>
<td>32/2034</td>
<td>22.5</td>
</tr>
<tr>
<td>BRAF</td>
<td>Exon 15 p.V600E</td>
<td>c.1799T&gt;A</td>
<td>Missense variant</td>
<td>0.08</td>
<td>2/2561</td>
<td>1.3</td>
</tr>
<tr>
<td>GNAS</td>
<td>Exon 8 p.R202C</td>
<td>c.604C&gt;T</td>
<td>Missense variant</td>
<td>1.32</td>
<td>35/2655</td>
<td>21.2</td>
</tr>
<tr>
<td>ERBB2</td>
<td>Exon 23 p.K937R</td>
<td>c.2810A&gt;G</td>
<td>Missense variant</td>
<td>0.21</td>
<td>5/2390</td>
<td>4.5</td>
</tr>
<tr>
<td>ERBB2</td>
<td>Exon 26 p.R1111Q</td>
<td>c.3332G&gt;A</td>
<td>Missense variant</td>
<td>0.61</td>
<td>12/1961</td>
<td>13.1</td>
</tr>
<tr>
<td>CTNNB1</td>
<td>Exon 3 p.S33Y</td>
<td>c.98C&gt;A</td>
<td>Missense variant</td>
<td>0.07</td>
<td>3/4597</td>
<td>63</td>
</tr>
<tr>
<td>PTEN</td>
<td>Exon 5 p.H93Y</td>
<td>c.277C&gt;T</td>
<td>Frameshift variant</td>
<td>4.32</td>
<td>157/3633</td>
<td>162</td>
</tr>
<tr>
<td>PTEN</td>
<td>Exon 8 p.N323fs</td>
<td>c.968dupA</td>
<td>Frameshift variant</td>
<td>0.12</td>
<td>1/841</td>
<td>0.9</td>
</tr>
<tr>
<td>PTEN</td>
<td>Exon 8 p.N323fs</td>
<td>c.968dupA</td>
<td>Frameshift variant</td>
<td>0.11</td>
<td>1/892</td>
<td>1.2</td>
</tr>
</tbody>
</table>
for MET gene variants was 9.19–20.35 with a median value of 18.31 (figure 4). Twenty-five per cent (6/24) cfDNA samples were positive for the EGFR copy number amplification mutation, including 50% (3/6) cfDNA samples were also positive for MET amplification mutation. Fifty per cent (3/6) of EGFR amplification positive cfDNA samples and 66.7% (2/3) MET amplification positive cfDNA samples showed EGFR p.T790M negative results.

**DISCUSSION**

EGFR gene is the most important tumour biomarker to be tested for the targeted treatment of NSCLC because EGFR-TKIs treatment is the first-line therapy method for patients with NSCLC with sensitising EGFR mutations. It can be observed that some cfDNA samples showed a low EGFR variant allele fraction which can be lower than 0.1%, the testing of baseline EGFR p.T790M variant is critical for clinical NSCLC treatment.30 31 EGFR testing method with LOD lower than 0.1% should be applied for sensitive detection of EGFR variants in the plasma cfDNA of patients with NSCLC.

As shown in online supplemental table S4, the cfDNA samples which showed mismatch results from different assays were investigated. EGFR sensitising mutations with variant allele fractions less than 0.2% were easily missed in the MassArray-based test. And EGFR exon 20 p.T790M mutation with variant allele fraction less than 5% were missed by both the MassArray and qPCR-based assays. Though the variant allele fraction (EGFR exon 20 p.T790M) at 0.42% normally can still be caught by the MassArray and qPCR-based assays, the low copy number of variant alleles (4.45 copies/mL plasma) may be the reason for missing detected (online supplemental table S4). Besides the variant allele fraction, the copy number of gene variants can also be an important factor that can impact the sensitivity of assays on testing tumour biomarker genes in plasma cfDNA.

As a high-throughput testing method, the NGS assay has the advantages of simultaneous testing of multiple tumour biomarker genes in plasma cfDNA.
biomarker genes in one test, interrogating both known and unknown gene variants. This makes it superior to qPCR and MassArray-based assays on gene detection efficiency.\(^{19, 30}\) It was observed that the NGS-based assay showed high sensitivity in testing tumour biomarker genes in plasma cfDNA as gene allele variant fraction of less than 0.1% and the copy number of gene allele variant of less than 1 copy/mL plasma were also successfully detected (figures 1 and 3). This confirms other studies in which the deep sequencing approach showed higher detection sensitivity in testing tumour biomarker genes in cfDNA than qPCR and MassArray-based assays.\(^{30, 32, 33}\)

Although EGFR exon 20 p.T790M mutation is the most possible reason for the resistance of NSCLC to the first-generation and second-generation EGFR TKIs, by approximately 50% of resistance cases are caused by other molecular mechanisms such as the amplification of ERBB2, MET, and the mutation of BRAF, PI3K and KRAS.\(^{34-36}\) Also, other EGFR mutations such as exon 20 p.C797S, oncogenic fusions such as FGFR3-TACC3, and oncogene amplification such as MET and ERBB2 can also cause the resistance of NSCLC to the third-generation TKIs.\(^{37}\) This study found that the detection of oncogenic amplification mutations can help clarify the resistance of NSCLC cases to the EGFR-TKIs by 12.5% (figure 4). Besides monitoring the EGFR status, it is necessary to comprehensively test other TKIs treatment-associated genes for choosing the most appropriate targeted treatment method. In this study, 4.2% (1/24) sample showed BRAF/exon 15 p.V600E positive but EGFR p.T790M negative which helped clarify the EGFR independent TKIs resistance in NSCLC (table 2).\(^{37}\) 12.5% (3/24) and 4.2% (1/24) samples showed mutations in PTEN and ERBB2, respectively (table 2), which may contribute to the resistance of NSCLC to TKIs as well, because ERBB2 amplification and PTEN loss have been found to activate PI3K/Akt pathway that is involved NSCLC progression.\(^{34}\) It was observed that the PTEN mutation (exon 8 p.N323fs) was only detected in patients with NSCLC without EGFR mutations. However, further studies need to be performed to confirm the relationship between PTEN and ERBB2 mutations and TKIs resistance.

**CONCLUSION**

All three methods (NGS, MassArray and qPCR-based assays) can effectively detect EGFR variants in the plasma cfDNA of patients with NSCLC and the concordance of these assays on testing EGFR variants was 83.3% (10/12). Based on the findings from this study, cfDNA testing methods with the LOD (lower than 2% and less than 10\(^{6}\) copies/mL plasma) for target gene variants can detect more than 50% of EGFR mutation-positive plasma samples from patients with NSCLC. Compared with the MassArray and qPCR tests, the AVENIO cfDNA expanded kit (Roche Diagnostics) possesses the highest sensitivity and throughput on testing tumour biomarker genes in plasma cfDNA which shows high competitiveness to be clinically used for targeted therapy of NSCLC. The targeted capture sequencing assay is an excellent tool for comprehensive molecular analysis of NSCLC targeted therapy associated genes, which can help improve the precision treatment of NSCLC by TKIs. Considering the sample size used in this study is limited, further studies are needed to confirm the findings.

**Contributors** LZ: methodology, project administration, writing original draft and editing. JC: resources and supervision. KF: methodology and resources. QC: resources, writing original draft and editing, responsible for the overall content as a guarantor. IV: conceptualisation, funding acquisition, methodology, project administration, resources, supervision, writing original draft and editing, responsible for the overall content as a guarantor.

**Funding** This study was supported by a research grant from Hoffmann - La Roche and AstraZeneca.

**Competing interests** It has received honorarium for consultation from AstraZeneca, Roche, Bayer. The other authors of this study declare that they have no competing interests.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants and was approved by the Ethics Committee of the Health Research Ethics Board of Universities of Alberta (HREBA.CC-18-0092). Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data relevant to the study are included in the article or uploaded as supplementary information.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) 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, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**REFERENCES**

35 Nagano T, Tachihara M, Nishimura Y. Mechanism of resistance to epidermal growth factor receptor-tyrosine kinase inhibitors and a potential treatment strategy. Cells 2018;7:212.
Application of liquid biopsy-based targeted capture sequencing analysis to improve the precision treatment of non-small cell lung cancer by tyrosine kinase inhibitors

Lei Zhang¹², John Coffin¹, Kim Formenti¹, Quincy Chu², Iyare Izevbaye¹ *

¹Department of Laboratory Medicine and Pathology, University of Alberta
²Cross Cancer Institute, Alberta Health Services

*Corresponding author: Dr. Iyare Izevbaye, 4B1.21 Walter C MacKenzie Health Sciences Centre, Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, Alberta, T6G 2B7, Canada. E-mail address: izevbaye@ualberta.ca
Appendix S1-Methods

Testing EGFR variants in cfDNA by qPCR

Cobas EGFR Mutation Test v2 (Roche Diagnostics, QC, Canada) a real-time qPCR-based assay, was used for testing EGFR variants in plasma ctDNA from NSCLC patients. It is an in vitro diagnostic product that detects known mutations in the exons 18, 19, 20, and 21 of EGFR through PCR amplification. The Cobas cfDNA Sample Preparation Kit (Roche Diagnostics) was used for the extraction of cfDNA from 2 mL plasma samples. Briefly, plasma samples were treated with a mixture of protease K and lysis/binding buffer for releasing cfDNA. After binding to the glass fiber filter, cfDNA was purified through the serial of washing steps by using the provided washing buffer. Plasma cfDNA was then collected into the provided elusion buffer. The extracted plasma cfDNA was used for EGFR testing by the Cobas EGFR Mutation Test kit v2 (Roche Diagnostics). Three qPCR amplification detection systems were prepared for separately testing different EGFR variants present in one plasma cfDNA sample. Following the manufacturer’s instructions, the qPCR detection of target EGFR variants in cfDNA was performed on the Cobas z 480 analyzer (Roche Diagnostics).

Testing pan-cancer gene mutations in cfDNA by NGS

AVENIO ctDNA expanded kit (Roche Diagnostics) that includes the reagents for sequencing test and the post-sequencing analysis software was used for testing somatic mutations in the cfDNA of NSCLC patients. The AVENIO ctDNA expanded kit (Roche Diagnostics) is a hybridization capture sequencing-based 77 genes pan-cancer assay, and target tumour biomarker genes covered by the panel are listed in table S1. AVENIO cfDNA Isolation Kit (Roche Diagnostics) was used to extract cfDNA from plasma according to the user’s manual. With a high pure extender assembly (HPEA)
unit, the cfDNA extraction kit was used to isolate cfDNA from 4 mL (2-5 mL) plasma samples. The quality of extracted plasma cfDNA samples was checked before library preparation. Agilent High Sensitivity DNA Analysis Kit (Agilent Technologies, CA, USA) was used to determine the length distributions of different cfDNA fragments for assessing the genomic DNA contamination. The concentration of cfDNA was also tested by using the Qubit® dsDNA HS Assay Kit on the Qubit 3.0 fluorometer (ThermoFisher Scientific, CA, USA).

After cfDNA extraction, the AVENIO ctDNA Expanded Kit (Roche Diagnostics) was used to prepare the sequencing library based on a hybridization capture-based technology. Briefly, with 10-50 ng input of cfDNA, the first step was adapter ligation. After post-ligation cleanup, PCR was performed to amplify the adapter-ligated cfDNA template. Hybridization was performed by using biotin-labeled probes to specifically capture the target genes fragments which were then enriched through binding to streptavidin beads. The expanded panel included in this kit was used for identifying and characterizing 77 genes (e.g. BRAF, EGFR, PIK3CA, et al.) that are associated with solid tumors (Table S1). After post-hybridization washes, a second PCR was performed to amplify the enriched library. The quality of the enriched library was assessed again by the Agilent High Sensitivity DNA Analysis Kit (Agilent Technologies) and the Qubit® dsDNA HS Assay Kit on the Qubit 3.0 fluorometer (ThermoFisher Scientific). With 50 ng input cfDNA, the typical concentration of enriched library is expected from 0.5 ng/μL to 30 ng/μL. The ideal peak size of the library is approximately 300 bp in the Bioanalyzer detection profile.
Prepared libraries were sequenced on the NextSeq 500 500/550 High Output Kit V2 (300 cycles) on Illumina NextSeq sequencing platform according to manufacturer’s instructions, (Illumina, CA, USA). A 15% PhiX Sequencing Control V3 (Illumina) pike-in sequencing control was included. AVENIO ctDNA Analysis Software (Roche Diagnostics) was used for sequencing data analysis and generating diagnostic reports. Specifically, the AVENIO ctDNA Analysis Software uses integrated digital error suppression (iDES) strategies, combining molecular bar codes within silico error suppression techniques to call variants (1). According to the manufacture’s introduction, the limit of detection of this kit has been estimated to reach down to 0.1%, and it has also been reported that the sequencing kit can detect mutant allele fraction down to 0.02% (2), which is consistent with the results from this study. By using the high-quality control process that is the iDES strategies, combining molecular bar codes within silico error suppression techniques, there is no cut-off value available for different called gene variants. The mutant allele fraction that is lower than 0.1% can still be reported with high confidence. The run metrics of this study are shown in figure S2 and table S2.

Testing **EGFR** mutations in cfDNA by MassArray

The MassARRAY detection system is based on matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) technology. It is a sensitive method for testing low allele frequency biomarker gene variants (3). A clinically validated UltraSEEK Lung Cancer Panel (Agena Bioscience, CA, USA) was the reference method in this study.

QIAamp Circulating Nucleic Acid Kit (Qiagen, CA, USA) was used for plasma cfDNA extraction. Briefly, purified cfDNA from 4 mL of plasma samples were prepared
through four steps (lysis, binding, washing, and elution) of extraction procedures. The Qubit® dsDNA HS Assay Kit (ThermoFisher Scientific) was used for testing the concentration of extracted plasma cfDNA. Using 10 ng input cfDNA, a multiplex PCR was performed to amplify target genes in cfDNA, and followed by a mutation-specific single based extension reaction for the MassArray UltraSEEK Lung Cancer Panel (Agena Bioscience), according to the manufacturer’s instructions. The extension products were enriched through a streptavidin-biotin interaction-based hybridization capture process and then transferred to a SpectroCHIP Array (Agena Bioscience) for the following detection through the MassArray Analyzer (Agena Bioscience). Generally, the enriched extension products (DNA sequences) were ionized and accelerated into a flight tube towards a detector. The extended probes were separated based on the differences of time-of-flight needed, and a mass spectrum for each sample was generated with relative intensity on the y-axis and mass/charge on the x-axis. The acquired data were further processed by the MassARRAY Typer (Agena Bioscience) software for generating a detailed detection report of the detected gene variants.

Appendix S2-Figures and tables
Figure S1. Electrophoresis files of typical good quality cfDNA (A, B) and enriched library (C, D) tested by the high-sensitivity DNA assay of Agilent Bioanalyzer.
Figure S2. Performance of testing tumour biomarkers in cfDNA by NGS-based Avenio expanded panel. A: Average of the total mapped reads, on-target rate, theoretical sensitivity and coverage uniformity of the NGS assay for testing plasma cfDNA samples. B: Typical coverage uniformity for each base of the cfDNA tested by the NGS assay.
Table S1. Target tumour biomarker genes covered by the Avenio Expand Panel

<table>
<thead>
<tr>
<th>List of target genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABL1, AKT1, AKT2, ALK, APC, AR, ARAF, BRAF, BRCA1, BRCA2, CCND1, CCND2, CCND3, CD274, CDK4, CDK6, CDKN2A, CSF1R, CTNNB1, DDR2, DPYD, EGFR, ERBB2, ESRI, EZH2, FBXW7, FGFR1, FGFR2, FGFR3, FLT1, FLT3, FLT4, GATA3, GNAI1, GNAQ, GNAS, IDH1, IDH2, JAK2, JAK3, KDR, KEAP1, KIT, KRAS, MAP2K1, MAP2K2, MET, MLH1, MSH2, MSH6, MTOR, NF2, NFE2L2, NRAS, NTRK1, PDCD1LG2, PDGFA, PDGFB, PIK3CA, PIK3R1, PMS2, PTCH1, PTEN, RAF1, RB1, RET, RNF43, ROS1, SMAD4, SMO, STK11, TERT, TP53, TSC1, TSC2, UGT1A1, VHL</td>
</tr>
</tbody>
</table>

Table S2. The metrics of testing tumour biomarker genes in cfDNA by the Avenio panel

<table>
<thead>
<tr>
<th>Number of read pairs (M)</th>
<th>Sequencing depth (Median)</th>
<th>Unique depth (Median)</th>
<th>Error rate</th>
<th>Fragment length (bp) (Median)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>25</td>
<td>9004</td>
<td>2870</td>
<td>1.14E-05</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>5</td>
<td>1543</td>
<td>1094</td>
<td>5.33E-06</td>
</tr>
</tbody>
</table>

Table S3. Results of testing EGFR from plasma cfDNA by NGS, MassARRAY and qPCR methods

<table>
<thead>
<tr>
<th>Mutations</th>
<th>Variants</th>
<th>NGS-based Avenio Expanded Panel</th>
<th>MassARRAY-based UltraSEEK Lung Panel</th>
<th>qPCR-based Cobas EGFR Mutation Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFR sensitizing mutations</td>
<td>Exon 18 p.G719A</td>
<td>4.2% (1/24)</td>
<td>4.3% (1/23)</td>
<td>8.3% (1/12)</td>
</tr>
<tr>
<td></td>
<td>Exon 20 p.S768I</td>
<td>4.2% (1/24)</td>
<td>4.3% (1/23)</td>
<td>8.3% (1/12)</td>
</tr>
<tr>
<td></td>
<td>Exon 21 p.L858R</td>
<td>20.8% (5/24)</td>
<td>17.4% (4/23)</td>
<td>16.7% (2/12)</td>
</tr>
<tr>
<td></td>
<td>Exon 21 p.L861Q</td>
<td>8.3% (2/24)</td>
<td>8.7% (2/23)</td>
<td>8.3% (1/12)</td>
</tr>
<tr>
<td></td>
<td>Exon 19 Dels</td>
<td>33.3% (8/24)</td>
<td>26.1% (6/23)</td>
<td>41.7% (5/12)</td>
</tr>
<tr>
<td>TKIs resistant EGFR mutation</td>
<td>Exon 20 p.T790M</td>
<td>33.3% (8/24)</td>
<td>34.8% (8/23)</td>
<td>25.0% (3/12)</td>
</tr>
</tbody>
</table>
Table S4. List of EGFR variants showed mismatched testing results by different assays

<table>
<thead>
<tr>
<th>Mutations</th>
<th>Variants</th>
<th>NGS-based Avenio Expanded Panel</th>
<th>MassARRAY-based UltraSEEK Lung Panel</th>
<th>qPCR-based Cobas EGFR Mutation Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>VAF</td>
<td>CNV (Copies/mL)</td>
<td>Test result</td>
</tr>
<tr>
<td>EGFR sensitizing mutations</td>
<td>Exon 21 p.L858R</td>
<td>0.17%</td>
<td>1.79</td>
<td>Negative</td>
</tr>
<tr>
<td>EGFR mutation</td>
<td>Exon 19 Dels</td>
<td>21.00%</td>
<td>1590</td>
<td>Inconclusive</td>
</tr>
<tr>
<td>TKIs resistant</td>
<td>Exon 19 Dels</td>
<td>0.11%</td>
<td>89.1</td>
<td>Negative</td>
</tr>
<tr>
<td>EGFR mutation</td>
<td>Exon 19 Dels</td>
<td>0.00%</td>
<td>0</td>
<td>Negative</td>
</tr>
<tr>
<td></td>
<td>Exon 20 p.T790M</td>
<td>0.42%</td>
<td>4.45</td>
<td>Negative</td>
</tr>
<tr>
<td></td>
<td>Exon 20 p.T790M</td>
<td>0.00%</td>
<td>0</td>
<td>Positive</td>
</tr>
</tbody>
</table>

References

