Establishment and application of a multiplex genetic mutation-detection method of lung cancer based on MassARRAY platform

Hong-Xia Tian, Xu-Chao Zhang, Zhen Wang, Jian-Guang Chen, Shi-Liang Chen, Wei-Bang Guo, Yi-Long Wu
Medical Research Center, Guangdong Lung Cancer Institute, Guangdong General Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, China

ABSTRACT

Objective: This study aims to establish a method for highly parallel multiplexed detection of genetic mutations in Chinese lung cancer samples through Agena iPLEX chemistry and matrix-assisted laser desorption ionization time-of-flight analysis on MassARRAY mass spectrometry platform.

Methods: We reviewed the related literature and data on lung cancer treatments. We also identified 99 mutation hot spots in 13 target genes closely related to the pathogenesis, drug resistance, and metastasis of lung cancer. A total of 297 primers, composed of 99 paired forward and reverse amplification primers and 99 matched extension primers, were designed using Assay Design software. The detection method was established by analyzing eight cell lines and six lung cancer specimens. The proposed method was then validated through comparisons by using a LungCarta™ kit. The sensitivity and specificity of the proposed method were evaluated by directly sequencing EGFR and KRAS genes in 100 lung cancer cases.

Results: The proposed method was able to detect multiplex genetic mutations in lung cancer cell lines. This finding was consistent with the observations on previously reported mutations. The proposed method can also detect such mutations in clinical lung cancer specimens. This result was consistent with the observations with LungCarta™ kit. However, an FGFR2 mutation was detected only through the proposed method. The measured sensitivity and specificity were 100% and 96.3%, respectively.

Conclusions: The proposed MassARRAY technology-based multiplex method can detect genetic mutations in Chinese lung cancer patients. Therefore, the proposed method can be applied to detect mutations in other cancer tissues.

KEYWORDS

Lung neoplasms; driver genes; mutation; multigene testing; MassARRAY

Introduction

Lung cancer is the leading cause of cancer-related deaths worldwide, and its rate of occurrence has been increasing. Precision medicine or personalized medicine based on genomic changes is a new breakthrough that allows more targeted therapy than traditional radiation and chemotherapy. Some of these mutations have been identified as predictive of clinical responses to targeted therapeutic agents in non-small cell lung cancer (NSCLC), such as the improved response to gefitinib or erlotinib in patients with certain epidermal growth factor receptor (EGFR) mutations and the response of crizotinib in patients with echinoderm microtubule-associated protein-like 4-anaplastic lymphoma kinase (EML4-ALK) fusion mutation. Given that many clinically relevant biomarkers continue to be identified, and targeted drugs are developed for personalized treatment, multigene mutation screening will be a requirement in routine clinical practice.

Genetic mutations can be detected by employing iPLEX mass-modified single-base extension technology and matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry on MassARRAY platform. The target sequences are amplified with amplification primers and extension primers, which are located adjacent to the interrogated loci, and they are used for single-base extension reactions to hybridize and elongate the extension primers at the nucleotide position of interest. Mutations are distinguished via detection and resolution of the mass-modified extension bases through MALDI-TOF. MassARRAY platform is an ideal platform for the mutation screening of multiple genes. LungCarta™ Panel can
analyze 214 mutations in 26 oncogenes. However, this method presents several limitations. For instance, the panel is expensive and does not include the recently identified molecular targets of interest. Therefore, a multigene detection method that is suitable for Chinese lung cancer patients should be established.

Briefly, the motivations of developing this new method are as follows: (1) MALDI-TOF method can simultaneously interrogate multiple gene mutation loci, which are screened for in-routine clinical practice. (2) If all the genes are analyzed individually, most analyses are certainly not possible because of the complete depletion of tissues and nucleic acid. However, multiplexing, which is employed in this method, reduces the amount of the clinical sample used. (3) Some novel molecular targets can be enrolled in the new method. (4) The custom panel in the new method is more cost-effective than the commercial panel because the former is established by ourselves.

Materials and methods

Patient specimens and cell lines

We obtained a cohort of 106 lung cancer specimens from the Guangdong Lung Cancer Institute of Guangdong General Hospital between January 2014 and December 2014. An informed consent was obtained from each patient, which was approved by the ethics committee of Guangdong General Hospital. All samples, which were stored at –80 °C after being frozen in liquid nitrogen, were assessed by two pathologists to ensure that >50% of each sample consisted of tumor tissue. We used eight NSCLC cell lines (i.e. H460, PC9, H1650, H1975, A549, GLC82, L78, and HCC827), which were purchased from the cell bank of the Chinese Academy of Sciences in Shanghai.

Reagents and instruments

The following materials were used in this study: QIAsymphony DNA Mini Kit (Qiagen, Valencia, Germany); LungCarta kit, PCR Accessory Set, iPLEX Pro Reagent Kit, and SpectroCHIP® (Agena Bioscience, San Diego, USA); H₂O (Sigma-Aldrich, USA); QIAsymphony SP (Qiagen, Valencia, Germany); Thermo NanoDrop 1000 (Thermo, MA, USA); MassARRAY® Nanodispenser and MassARRAY® Analyzer (Agena Bioscience, San Diego, USA); ABI 3730xl Sequencing Machine (Life Technologies, New York, USA); and polymerase chain reaction (PCR) machine (Life Technologies, New York, USA).

Preparation of polygenic primer panel

Determination of related driver genes of lung cancer

We reviewed the related literature, data on lung cancer treatments, and gene spectrum characteristics of the Chinese lung cancer population. A total of 13 target oncogenes (i.e. EGFR, KRAS, ALK, FGFR1, FGFR2, FGFR3, PIK3CA, BRAF, PTEN, MET, ERBB2, AKTI, and STK11), which were closely related to the pathogenesis, drug resistance, and metastasis of lung cancer and associated with relevant transduction pathways, were enrolled in the polygenic primer panel[5–12].

Selected mutation hot spots of driver genes

We determined the Catalogue of Somatic Mutation in Cancer (COSMIC) identifier numbers of the following 13 genes by browsing the COSMIC database: EGFR(ENST000002757493), KRAS(ENST00000256078), ALK(ENST00000389048), FGFR1(ENST00000447712), FGFR2(ENST00000358487), FGFR3(ENST00000440486), PIK3CA(ENST00000263967), BRAF(ENST00000288602), PTEN(ENST00000371953), MET(ENST00000318493), ERBB2(ENST00000269571), AKTI(ENST00000349310), and STK11(ENST00000326873). In accordance with the mutation frequencies of 13 oncogenes in lung cancer, 99 mutation hot spots associated with lung cancer were added to the polygenic primer panel (Table 1).

Designed polygenic primer panel

The genome sequence numbers of the following 13 target genes were identified in GenBank: EGFR(NG_007726.3), KRAS(NG_007524.2), ALK(NG_00945.1), FGFR1(NG_007729), FGFR2(NG_012449.1), FGFR3(NG_012632.1), PIK3CA(NG_012113.2), BRAF(NG_007783.3), PTEN(NG_007466.2), MET(NG_008996.1), ERBB2(NG_007503.1), AKTI(NG_012188.1), and STK11(NG_007460.2). We marked mutant loci in genomic DNA (gDNA) sequences in accordance with the mutation label and format requirements of the MassARRAY platform. The relevant parameters of Assay Design Software (ADS) were adjusted to include all 99 loci in the polygenic primer panel. The highest number of loci that can be detected simultaneously was set to 10 mutant loci. A total of 99 loci were randomly distributed in 12 wells through ADS on the basis of primer design (avoidance of the formation of dimers/mismatches). A total of 297 primers, including 99 paired forward and reverse amplification primers and 99 matched single-base extension primers, were designed. The target sequences were amplified by
amplification primers, and the extension primers were located on the base before the mutation sites, which were complementary to the amplification products and performed single-base extension reactions.

**Configuration of polygenic primer panel**

The primers were synthesized by Shanghai Sangon Biological Engineering Technology and Service Co., Ltd. The polygenic primer panel was configured as follows: (1) The amplification primers were first diluted to 10 μM, and the working liquid was a mixture of all primers in each well, including 0.5 μM of each primer. In accordance with the parameters of ADS, the forward and reverse amplification primers were distributed into 12 pipes. The forward primer P1-F and reverse amplification primer P1-R comprised group 1, the forward primer P2-F and reverse amplification primer P2-R comprised group 2, and so on. (2) The extension primers were first diluted to 500 μM and then mixed on the basis of individual molecular weights. The extension primers of E1 comprised group 1, the extension primers of E2 comprised group 2, and so on (E1–E12). A total of 12 amplification primers corresponded to 12 extension primers (e.g. P1 corresponded to E1, etc.).

**Establishment of detection method**

The detection method was verified by analyzing 8 cell lines (i.e. H460, PC9, H1650, H1975, A549, GLC82, HCC827, and H1299) and 6 lung cancer specimens. The proposed method was then validated through comparisons by using a LungCarta™ kit or previously reported results. The gDNA of a healthy person (from foreskin tissue; with informed consent) was used as a negative control, and H₂O was used as a blank control. Each sample required 120 ng gDNA for 12 wells (10 ng/well).

The procedure was conducted in a MassARRAY system platform. The experiments using the LungCarta™ kit were performed in accordance with the manufacturer’s protocol. The polygenic primer kit was used in the detection method as follows.

**PCR**

gDNA was extracted from cell lines and patient specimens in accordance with the manufacturer’s protocol. Afterward, gDNA was quantified using a NanoDrop ND-1000 spectrophotometer. gDNA was then amplified using a PCR Accessory Set. The thermocycling cocktail comprised 0.5 μL of PCR buffer (10×), 0.4 μL of MgCl₂ (25 mM), 0.1 μL of dNTPs (25 mM), 0.2 μL of PCR enzyme (5 U/μL), 1 μL of amplification primer mix (P1-P12), 1 μL of gDNA (10 ng/μL), and H₂O to a total volume of 5 μL. The thermocycling conditions were 94 °C for 2 min; followed by 45 cycles of 94 °C for 30 s, 56 °C for 30 s, and 72 °C for 1 min; and a final extension step of 72 °C for 5 min.

**Shrimp alkaline phosphatase (SAP) reaction**

dNTPs in the PCR products were removed via SAP. For this reaction, 0.3 μL of SAP (1.7 U/μL) and 0.17 μL of SAP buffer
(10×) were added to step 1 PCR products, and H₂O was added to a total volume of 7 μL. The reaction conditions were 37 °C for 40 min and 85 °C for 5 min.

**Extension reaction**
A single-base extension reaction was performed with iPLEX Pro Reagent Kit to hybridize and elongate the extension primers at the nucleotide position of interest. For the single-base extension, 0.2 μL of TypePlex buffer (10×), 0.2 μL of TypePlex termination mix (10×), 0.041 μL of TypePlex thermosequenase enzyme (33 U/μL), and 0.804 μL of extension primers (E1–E12) were mixed with step 2 products, and H₂O was added to a total volume of 9 μL. The reaction conditions were 94 °C for 30 s; followed by 5 cycles of 94 °C for 5 s, 52 °C for 5 s, and 80 °C for 5 s, with this step repeated 35 times; and a final extension at 72 °C for 3 min.

**Desalination**
For desalination, 41 μL of H₂O and 15 mg of Clean Resin (96-well microplates) were added to step 3 extension products, and the plate was rotated for 30–60 minutes. The plate was then centrifuged at 3,200 g for 5 min.

**Spotter and analysis**
The supernatant from the step 4 extension products was spotted onto matrix-precoated SpectroCHIP® through MassARRAY® Nanodispenser and scanned using MassARRAY® Analyzer. The results were analyzed using MassArray® Workstation Software (v.4.0). The mutation was distinguished with TOF mass spectrometry on the basis of different molecular weights. The peaks in the mass spectrum were identified as mutations.

**Analysis of sensitivity and specificity**
The newly established methods were evaluated by Sanger sequencing of EGFR and KRAS genes in 100 lung cancer cases. EGFR and KRAS mutations were detected through Sanger sequencing in accordance with a previously published protocol. The sequencing data were analyzed using Sequencing Analysis Software v5.2 (Applied Biosystems).

**Results**

**Polygenic primer panel containing mutation sites**
We tested 99 mutation hot spots in 13 target genes (i.e. EGFR, KRAS, ALK, FGFR1, FGFR2, FGFR3, PIK3CA, BRAF, PTEN, MET, ERBB2, AKT1, and STK11), which were closely related to the pathogenesis, drug resistance, and metastasis of lung cancer and associated with relevant transduction pathways. The detailed protein mutation sites of genes are listed in Table 1.

**Establishing a detection method in cell lines**
The detection method used with the polygenic primer kit was established by analyzing 8 cell lines of lung cancer. The results from the newly established method were consistent with the previously reported mutations in cell lines. All the cell lines were confirmed by Sanger sequencing, and the sequencing results from A549 are shown in Figure 1. The detailed mutation sites within genes and proteins are listed in Table 2. The negative control and blank control did not detect mutations.

**Established detection method in lung cancer specimens**
The detection method used with the polygenic primer kit was established by analyzing 6 lung cancer tissue specimens, and this method was validated through comparison by using a LungCarta™ kit (Table 3). An FGFR2 gene mutation was
found in a sample of lung cancer tissue (K1747T) through the newly established method. However, the LungCarta™ kit cannot detect the mutation (Figure 2). Unfortunately, other methods cannot perform validation because of the depletion of the material. With respect to other clinical lung cancer specimens, the findings were consistent with the observations with LungCarta™ kit (Figure 3).

**Sensitivity and specificity of the established method**

For 100 samples, 99 mutation hot spots in 13 targeted genes were analyzed using the established method. For comparison, EGFR and KRAS mutations were detected via Sanger sequencing. With Sanger sequencing as the gold standard, the sensitivity of the established method was 100%, and its specificity was 96.3% (positive predictive value, 95.8%);

---

**Table 2** The result of established method was validated in cell lines

<table>
<thead>
<tr>
<th>Cell line</th>
<th>ATCC previously reported</th>
<th>Result of established method</th>
</tr>
</thead>
<tbody>
<tr>
<td>H460</td>
<td>KRAS mutation PIK3CA mutation</td>
<td>KRAS_Q61H (c.183A&gt;T) PIK3CA_E545K (c.1633G&gt;A )</td>
</tr>
<tr>
<td>PC9</td>
<td>EGFR_ Exon19 deletion</td>
<td>p.E746_A750delELREA (c.2235-2249del15)</td>
</tr>
<tr>
<td>H1650</td>
<td>EGFR_ Exon19 deletion</td>
<td>p.E746_A750delELREA (c.2235-2249del15)</td>
</tr>
<tr>
<td>H1975</td>
<td>EGFR_L858R EGFR_T790M</td>
<td>EGFR_L858R (c.2573T&gt;G) EGFR_T790M (c.2369C&gt;T)</td>
</tr>
<tr>
<td>A549</td>
<td>KRAS mutation</td>
<td>KRAS_G12S (c.34G&gt;A)</td>
</tr>
<tr>
<td>GLC82</td>
<td>EGFR_L858R</td>
<td>EGFR_L858R (c.2573_2574TG&gt;G)</td>
</tr>
<tr>
<td>HCC827</td>
<td>EGFR Exon19 deletion</td>
<td>p.E746_A750delELREA (c.2236_2250del15)</td>
</tr>
<tr>
<td>H1299</td>
<td>EGFR/ALK/KRAS negative</td>
<td>No mutation</td>
</tr>
</tbody>
</table>

**Table 3** The result of established method was validated by comparing with LungCarta™ in lung cancer tissue samples

<table>
<thead>
<tr>
<th>Sample number</th>
<th>Result of LungCarta™</th>
<th>Result of established method</th>
</tr>
</thead>
<tbody>
<tr>
<td>K1736T</td>
<td>EGFR_L858R MET_N375S</td>
<td>EGFR_L858R (c.2573T&gt;G) MET_N375S (c.1124A&gt;G)</td>
</tr>
<tr>
<td>K1744T</td>
<td>EGFR_DEL</td>
<td>EGFR_p.E746_A750delELREA (c.2235-2249del15)</td>
</tr>
<tr>
<td>K1748T</td>
<td>STK11_F354L</td>
<td>STK11_F354L (c.1062C&gt;G)</td>
</tr>
<tr>
<td>K1745T</td>
<td>No mutation</td>
<td>No mutation</td>
</tr>
<tr>
<td>K1746T</td>
<td>No mutation</td>
<td>No mutation</td>
</tr>
<tr>
<td>K1747T</td>
<td>No mutation</td>
<td>FGFR2_N549H (c.1645A&gt;C)</td>
</tr>
</tbody>
</table>

**Figure 2** FGFR2 gene mutation was detected by the newly developed method in K1747T tissue samples of lung cancer. C: mutant allele; A: wild-type allele.
Coexistence of L858R and N375S was detected by the newly developed method and LungCarta™ in K1736T tissue sample of lung cancer. (A) EGFR L858R was detected by LungCarta™. (B) MET N375S was detected by LungCarta™. (C) EGFR L858R was detected by the newly developed method. (D) MET N375S was detected by the newly developed method.
negative predictive value, 100%). A detailed comparison of the results is shown in Table 4.

Table 4  The result of the newly developed method compared with Sanger sequencing in 100 cases of lung cancer samples

<table>
<thead>
<tr>
<th>Result of new method</th>
<th>Result of Sanger sequencing</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mutation</td>
<td>Wild type</td>
</tr>
<tr>
<td>Mutation</td>
<td>46</td>
<td>2</td>
</tr>
<tr>
<td>Wild type</td>
<td>0</td>
<td>52</td>
</tr>
<tr>
<td>Total</td>
<td>46</td>
<td>54</td>
</tr>
</tbody>
</table>

Discussion

Quantitative PCR, Sanger sequencing, fluorescence in situ hybridization, and immunohistochemistry are the main single-gene detection methods for the molecular classification of lung cancer. These methods present low throughput, and they are also time consuming and expensive. Moreover, these methods significantly restrict the screening of ROS1, RET, MET, HER2, NRTK, and NRG in clinical practice with the exception of EGFR, ALK, and KRAS oncogenes. Patients with newly diagnosed lung cancer who are undergoing surgery (major surgery or minor surgical biopsy) exhibit considerable cancer tissues, but only a small sample can be usually acquired from advanced patients through percutaneous biopsy, bronchoscopy biopsy, and endobronchial ultrasound biopsy. As such, any tissue barely remains for characterization of other important molecules and exploratory research on drug resistance mechanisms after diagnosis and molecular detection of individual genes. Considering the complexity of tumor driver genes and the limitations of single-gene detection techniques, we need to establish a multiplex genetic mutation-detection method for lung cancer in clinical practice and translational medicine.

Next-generation sequencing (NGS) is currently one of the most important approaches to precise treatment employed by studies on the domain of life sciences, but the complexities of NGS technologies and data analysis are slowing their wide spread availability in the diagnostics laboratories. So a robust genotyping platform which can carry-on daily routine testing is need. Other multiplex genetic mutation-detection methods have been reported, such as Cancer Personalized Profiling by deep sequencing and SNaPShot. Molecular characterization via MALDI-TOF on MassARRAY platform, which is characterized by high throughput, high sensitivity, and a simple operation, can be implemented by combining single-base extension and mass spectrometry technology. This approach also minimizes the weaknesses of traditional single-gene tests, such as high cost and their time-consuming and tedious procedures. Given the short segments of the PCR amplification products, not only fresh tissue specimens but also formalin-fixed paraffin-embedded (FFPE) specimens, pleural effusion, and biopsy samples can be used for effective detection. LungCarta™ Panel, which is commercially available, can facilitate comprehensive screening of relevant biomarkers of lung cancer. However, this method presents several limitations; for example, the panel is confidential. The LungCarta™ Panel is also expensive, and it does not include recently identified molecular targets of interest.

According to the review of the related literature and data on lung cancer treatments, 99 mutation hot spots in 13 oncogenes of lung cancer were enrolled into the polygenic primer panel, such as MET mutations, which are associated with targeting the therapy resistance of EGFR tyrosine kinase inhibitors (TKIs). The mutations in multiple loci of the ALK gene are also associated with the resistance of crizotinib; for instance, the recently reported C1156Y, L1196M, F1174L, and G1269S mutations in ALK gene are identified as resistance mutations following crizotinib treatment in lung cancer. Some mutations are highly sensitive to the structurally unrelated ALK inhibitor TAE684. In recent years, FGFR signaling pathways have been shown to demonstrate abnormal activation in various tumors, and one FGFR2 gene mutation was detected in one clinical specimen through the newly established method. However, this mutation cannot be detected using the LungCarta™ kit. Unfortunately, the mutation cannot be confirmed by another method because of the depletion of the material. This observation demonstrates that the multiplex genetic mutation-detection method may provide accurate variant information for patients. MALDI-TOF mass spectrometry also presents advantages because of its high accuracy and sensitivity compared with the direct sequencing method. At present, Sanger sequencing is generally regarded as the gold standard for detecting somatic mutations in tumor specimens. However, Sanger sequencing suffers from limited sensitivity for low-level mutant alleles, particularly in FFPE specimens; this sequencing also takes a slow turnaround time. Given that MALDI-TOF mass spectrometry established detection limits of 1%–5% mutant alleles, the specificity of the in-house panel compared with that of Sanger sequencing is 96.3%, which may be associated with the high sensitivity of mass spectrometry and may compensate for the low sensitivity of direct sequencing.

The proposed MALDI-TOF multiplex mutation-detection method demonstrates the following advantages: (1) The novel method examines oncogenes and multiple drug-resistant loci in parallel, which are associated with lung-
cancer-targeted therapy. The screening results optimize the molecular characterization of the disease and provide information regarding targeted therapy. (2) The polygenic primer panel can reduce the patient’s molecular testing costs. Following clinical and translational applications, the proposed method provides direct economic benefits and optimizes clinical resources. (3) Moreover, the proposed method shows high throughput, and it can simultaneously detect 99 loci of 13 oncogenes in 12 wells. Thus, a sensitive highly parallel approach is provided, and the patient diagnostic samples are used efficiently for an accurate molecular classification. Furthermore, we can freely enroll updated mutations to the in-house panel based on the flexibility of the MassARRAY platform. Thus, the newly developed method is extensible. However, the newly developed method also presents limitations. For instance, this method is based on a panel of preselected sites. Therefore, the mutational patterns in tumor suppressors, such as P53 gene\(^{25}\), which is highly random and usually disperse across the gene, cannot be comprehensively detected with the newly developed method. For these mutation patterns, Sanger sequencing or NGS may represent a good alternative. In conclusion, successful establishment of the novel detection method, which demonstrates a great potential for lung cancer treatment, can significantly reduce the costs of biomarker tests.

**Acknowledgements**

This work was supported by the Special Fund for Research in the Public Interest from the National Health and Family Planning Commission of PRC (Grant No. 201402031), the Key Lab System Project of the Guangdong Science and Technology Department (Grant No. 2012A061400006), and the Special Fund for Research in the Public Interest and Capacity Building from the Guangdong Science and Technology Department (Grant No. 2014A020212225). This article was published originally in Chinese Journal of Clinical Oncology 2015; 42(17): 856-61 (in Chinese).

**Conflict of interest statement**

No potential conflicts of interest are disclosed.

**References**


