Targeted sequencing reveals distinct pathogenic variants in Chinese patients with lung adenocarcinoma brain metastases

YANCHUN MA1*, KUN CHEN1*, ZHENHUA YANG2 and MING GUAN1,3

1Department of Laboratory Medicine, Huashan Hospital North, Fudan University, Shanghai 201907;
2Clinical Laboratory, Baoshan District Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai 201999;
3Central Laboratory, Huashan Hospital, Shanghai Medical School, Fudan University, Shanghai 200000, P.R. China

Received March 12, 2017; Accepted December 6, 2017

DOI: 10.3892/ol.2018.7859

Abstract. Lung cancer is the most common type of malignancy to metastasize to the brain, with the median survival time of patients being 6-11 months. In the present study, the aim was to compare the actionable gene mutation profiles of primary lung adenocarcinoma (LC) samples and LC brain metastasis (LCBM) samples through targeted sequencing. Next generation sequencing (NGS) of 13 formalin-fixed, paraffin-embedded LC samples and 15 LCBM samples was performed using a customized OncoAim™ cancer panel and OncoAim™ RNA fusion panel on the MiSeq platform. The OncoAim™ cancer panel pipeline and OncoAim™ RNA fusion panel pipeline were used for bioinformatic analysis. Together, 43 variants were observed in 7 genes from the 28 cancer samples. The mutated genes of LCBM were tumor protein (TP)53, epidermal growth factor receptor (EGFR), catenin β1, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit α, mothers against decapentaplegic homolog 4, Kirsten rat sarcoma viral oncogene homolog (KRAS) and proto-oncogene B-Raf, which were exhibited in 10/15 (66.7%), 6/15 (40.0%), 3/15 (20.0%), 2/15 (13.3%), 2/15 (13.3%), 1/15 (6.7%) and 1/15 (6.7%) of samples, respectively. The mutated genes of LC were TP53, EGFR and KRAS, which were exhibited in 11/13 (84.6%), 5/13 (38.5%) and 2/13 (18.2%) of samples, respectively. echinoderm microtubule associated protein like 4-anaplastic lymphoma kinase rearrangements were present in 1 LCBM sample. For 2 LC samples and 1 LCBM sample, no genetic alterations were observed. The NGS data also revealed a novel 4-codon deletion of TP53 (p.V166_H169del) and a novel TP53 splice site mutation (7577157 -63del TACTCAG). Further potentially actionable mutations were detected in LCBM, indicating a high degree of genetic heterogeneity between the LC and LCBM samples that were analyzed. The present study demonstrated that NGS provides an improved approach for the discovery of potentially actionable mutations and the understanding of the mechanisms underlying tumor progression and evolution.

Introduction

Brain metastases (BM), particularly those from non-small cell lung cancer (NSCLC), constitute an increasing proportion of the global cancer burden. It is estimated that 50% of all patients with BM will experience a primary lung malignancy (1). In patients with central nervous system (CNS) metastases, the administration of standard chemotherapies or targeted agents is limited due to low CNS penetration of antineoplastic agents across the blood-brain barrier (2).

However, in the last decade, several biological agents targeting specific molecular alterations, including gefitinib or erlotinib targeting epithelial growth factor receptor (EGFR) mutations and crizotinib targeting anaplastic lymphoma kinase (ALK) translocations, have been approved for clinical use and used in personalized treatment regimes (3-7). Selected molecular-targeted therapy for BM of lung cancer has been notably effective, demonstrated by response rates, progression-free survival and quality of life in small studies and case reports (8-11); however, there are limited data available regarding the differential genetic aberrations between lung adenocarcinoma BM (LCBM) and primary lung cancer (LC) (12).

Next-generation sequencing (NGS) technologies provide an unprecedented opportunity to identify the clinically actionable mutations relevant to personalized cancer care, and are increasingly used for the mutational analysis of tumors for clinical and research applications (13-15). Targeted NGS permits the concurrent deep sequencing of hundreds of cancer genes, and has emerged as a sensitive
and efficient tool to detect complex and heterogeneous gene mutations (16-18).

In the present study, targeted NGS was conducted to detect potentially actionable genetic aberrations in LCBM as well as primary LC, to provide target gene mutation profiles and an improved understanding of the genetic mechanism underlying lung cancer metastasis. The aim of the present study was to identify novel and potentially beneficial therapeutic targets for patients with BM.

Materials and methods

Patients. A total of 28 formalin-fixed, paraffin-embedded (FFPE) tumor samples, including 13 cases of LC and 15 cases of LCBM, were retrospectively collected from the archives of the Pathology Laboratory in Huashan Hospital North in Fudan University (Shanghai, China) between January 2013 and December 2015. There were 10 unstained FFPE tissue sections (10 µm) that were cut for DNA and RNA extraction. The pathologist identified the lesional area and the percentage tumor cell content (minimum of 50%) for each case, which were the areas microdissected for extraction. Assessment of BM was concluded using all of the radiology reports, which contained brain imaging or pathology.

DNA and RNA extraction from FFPE samples. Genomic DNA was extracted from FFPE slices of solid tumor using the Qiagen®QIAamp DNA FFPE kit (Qiagen, Inc., Valencia, CA, USA), according to the manufacturer's protocol. DNA was eluted with 40 µl elution buffer (10 mM Tris-Cl, pH 8.3; 0.1 mM EDTA; 0.04% NaN₃). The concentration was measured using the Qubit HS kit (Thermo Fisher Scientific, Inc., Waltham, MA, USA). Total cellular RNA was extracted using the Qiagen®RNeasy FFPE kit, according to the manufacturer's protocol with minor modifications as specified below (Qiagen, Inc.). A total of 1 ml xylene was added to the 1.5 ml tube containing the FFPE tissue. The tube was vortexed vigorously for 10 sec and centrifuged at 21,130 x g for 2 min at 25˚C. The supernatant was carefully removed using a pipette. Then 1 ml ethanol (96-100%) was added into the tube, followed by vigorous vortexing followed by centrifugation at 21,130 x g for 2 min at 25˚C. The above step was repeated once. Then, the ethanol was carefully removed. The tube was kept open at room temperature (15-25˚C) for 10 min until all residual ethanol evaporated. Then, the manufacturer's protocol (Qiagen, Inc.) was followed, adding 150 µl Buffer PKD and following the remaining steps in order to obtain RNA. The RNA concentration was measured with the Qubit™ RNA kit (Thermo Fisher Scientific, Inc.), and RNA integrity was evaluated using RNA gel electrophoresis (1% agarose gel in 0.5% TBE buffer, running for 35 min at a 120 v voltage; with a ratio of 28S rRNA to 18S rRNA of ~2:1). Reverse transcription of RNA to cDNA was conducted using the SuperScript™ VILO™ cDNA synthesis kit (Thermo Fisher Scientific, Inc.), and was followed by library construction.

Sequencing library construction. The OncoAim™ cancer panel (Singlera Genomics, Inc., Shanghai, China) was used with the recommended DNA input (20 ng for FFPE samples) to generate sequencing libraries compatible with MiSeq system Dx (DX-410-1001; Illumina, Inc., San Diego, CA, USA). The panel included 400 amplicons targeting cancer hotspots, the specific genetic loci of which were depicted on the website of the Singlera Genomics, Inc. (Shanghai, China; http://www.singlera.com.cn/product/index.aspx?pid=1). The targeted regions were amplified using the customized OncoAim™ cancer panel followed by the addition of P5/P7 dual barcode adapters (Illumina, Inc.).

In addition, 20 ng reverse transcribed cDNA was amplified with the OncoAim™ RNA fusion panel (Singlera Genomics, Inc.) to generate sequencing libraries compatible with MiSeq. The panel included 170 amplicons targeting fusion events for ALK, Ret proto-oncogene, proto-oncogene c-Ros1 and neurotrophic receptor tyrosine kinase, along with five genes (hydroxymethylbilane synthase, integrin subunit β 7, lamin A/C, MYC proto-oncogene, bHLH transcription factor and TATA-box binding protein) as internal expression controls. Fusion variants, if present, were enriched by the customized OncoAim™ RNA fusion panel followed by a P5/P7 addition (Illumina, Inc.).

LabChip® GX Touch (PerkinElmer, Inc., Waltham, MA, USA) was used to confirm successful library amplifications to a desired length (~220 bp), and library concentration was assessed using the Qubit™ HS kit (Thermo Fisher Scientific, Inc.). No detectable DNA in the library prepared from non-template control (negative control) was used to indicate that no DNA contamination was introduced during library preparation. There were 12-16 barcoded libraries that were pooled. The ratios were confirmed and quantified by qPCR using KAPA Library Quantification kit (Kapa Biosystems, Inc., Wilmington, MA, USA) following the manufacturer's protocol, using the Bio-Rad CFX96 qPCR system (Bio-Rad Laboratories, Inc., Hercules, CA, USA). The KAPA Library Quantification kit uses the novel KAPA SYBR FAST DNA Polymerase, engineered through a process of directed evolution for high-performance SYBR Green I-based qPCR (Kapa Biosystems, Inc., Wilmington, MA, USA). The qPCR primers were 5'-AATGATACGCGCCACCACCGA-3' and 5'-CAAGCAGACAGCTACAGCA-3'. The PCR cycling conditions were: initial denaturation: 95˚C for 5 min; cycling: 35 cycles of 95˚C for 30 sec and 60˚C for 45 sec. The library concentration was calculated against the standard curve generated from 6 standards using qPCR software on the Bio-Rad CFX96 qPCR system (CFX Maestro software 12004110; Bio-Rad Laboratories, Inc.).

Data processing. A bioinformatic analysis pipeline was created using a corresponding OncoAim™ cancer panel pipeline (OncoAim™ version 7.2) and OncoAim™ RNA fusion panel pipeline (OncoAim™ RNA version 3.1). Briefly, raw reads (fastq files) were quality-filtered with FastQC version 0.9.5 (Babraham Bioinformatics, Cambridge, UK) and base quality score recalibration was performed. Clean reads were assembled and aligned against the reference genome hg19, or the targeted fusion reference to detect sequence/structure variations. The Burrow-Wheeler Aligner algorithm (https://github.com/lh3/bwa; version 0.7.12-r1039; Dec 2015) was implemented into the procedure for alignment, using default parameters. Insertions and deletions in sequence alignment files were left-aligned, and local realignment was...
Table I. Baseline characteristics of all patients (n=28).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>LC (n=13)</th>
<th>LCBM (n=15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>61 (50-74)</td>
<td>55 (35-65)</td>
</tr>
<tr>
<td>Sex</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td></td>
<td>9 (3)</td>
<td>4 (2)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>7 (4)</td>
<td>8 (5)</td>
</tr>
<tr>
<td>Smoking history</td>
<td>Never</td>
<td>Have smoked</td>
</tr>
<tr>
<td></td>
<td>9 (6)</td>
<td>4 (3)</td>
</tr>
<tr>
<td></td>
<td>Smoked</td>
<td>Total</td>
</tr>
<tr>
<td></td>
<td>4 (3)</td>
<td>7 (5)</td>
</tr>
<tr>
<td>TNM stage</td>
<td>Ia-Ilb</td>
<td>IIa-IIlb</td>
</tr>
<tr>
<td></td>
<td>3 (2)</td>
<td>10 (8)</td>
</tr>
<tr>
<td></td>
<td>IIIa-IIIb</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>10 (8)</td>
<td>15 (12)</td>
</tr>
<tr>
<td>Pathological diagnosis</td>
<td>LC</td>
<td>LCBM</td>
</tr>
<tr>
<td></td>
<td>12 (9)</td>
<td>13 (10)</td>
</tr>
<tr>
<td>BM</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>12 (9)</td>
<td>13 (10)</td>
</tr>
</tbody>
</table>

LC, lung adenocarcinoma; LCBM, lung adenocarcinoma brain metastasis; TNM, tumor node metastasis; BM, brain metastasis.

Results

Patient characteristics. Characteristics of the 13 patients with LC and 15 patients with LCBM are presented in Table I. The median age of patients diagnosed with LC and LCBM was 61 years (range, 50-74 years) and 55 years (range, 35-65 years), respectively. Overall, 57% of patients (16 cases) were male, and the majority of patients (24 cases) did not have a history of smoking. The majority of patients (89.3%, 25/28) were diagnosed in advanced (tumor-node-metastasis III and IV) stages (19).

NGS. A total of 28 samples were successfully sequenced. Table II presents the results of NGS of 28 samples with the OncoAim™ cancer panel and the presence of each sample gene mutation can be observed in Fig. 1. The mutated genes in these 28 samples were TP53, EGFR, Kirsten rat sarcoma viral oncogene homolog (KRAS), CTNNB1, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit α (PIK3CA), mothers against decapentaplegic homolog 4 (SMAD4) and proto-oncogene B-Raf (BRAF). The number of mutant genes are summarized in Fig. 2A, and were identified in 21/28 (75.0%), 11/28 (39.3%), 3/28 (10.7%), 3/28 (10.7%), 2/28 (7.1%), 2/28 (7.1%) and 1/28 (3.6%) of samples, respectively. Overall, 43 variants in the 7 genes were observed. These 43 genomic variant aberrations included 36 base substitutions and 7 short deletions. A total of 5 samples (5/28, 18%) exhibited 3 potentially biologically significant variants, 8 samples (8/28, 29%) exhibited 2 variants, 11 samples (11/28, 39%) exhibited only a single variant, and 4 samples (4/28, 14%) exhibited no alterations (Fig. 2B). The average number of variants was 1.5/tumor. The DNA fusion panel tests revealed only echinoderm microtubule associated protein like 4 [EML4 (13)]-ALK (20) in 1 sample of LCBM (Fig. 3).

Comparison of gene mutations between LC and LCBM. The mutated genes in LCBM tissues were TP53, EGFR, CTNNB1, PIK3CA, SMAD4, KRAS and BRAF; which were identified in 10/15 (66.7%), 6/15 (40.0%), 3/15 (20.0%), 2/15 (13.3%), 2/15 (13.3%), 1/15 (6.7%) and 1/15 (6.7%) of samples, respectively. The mutated genes of LC were TP53, EGFR and KRAS, which were exhibited in 11/13 (84.6%), 5/13 (38.5%) and 2/13 (18.2%) of samples, respectively. EML4-ALK rearrangement existed in 1 LCBM sample. A total of 2 LC samples and 1 LCBM sample exhibited no genetic alterations (Fig. 4). No significant differences were identified in the number of mutations or the frequency of TP53, EGFR or KRAS between patients with LC (n=13) and patients with LCBM (n=15; P>0.05).
Novel TP53 gene mutations in LCBM and mutation verification by Sanger sequencing. The LCBM-7 and LCBM-11 samples exhibited a novel TP53 mutation form p.V166_H169del and a new splice site mutation for 7577157-63del TACTCAG, respectively, located in the non-coding regions. Fig. 5A depicts the Sanger sequencing chromatograms of p.V166_H169del compared with the wild-type (Fig. 5B). Fig. 5C depicts the Sanger sequencing chromatograms of 7577157-63del TACTCAG compared with the wild-type (Fig. 5D).

Discussion

Lung cancer is the most common cause of cancer-associated mortality globally, and 80-85% of these cases are NSCLC (20). In total, 10-50% of patients with NSCLC develop BM (21). Treatment options are palliative, and usually include a combination of steroids, radiation and surgery (22). Despite advances in cancer therapy, the median survival time for patients with lung cancer BM is only 6-11 months (23). The poor prognosis associated with disseminated disease that has metastasized to the brain is not addressed by existing therapies. Efforts to develop improved personalized therapies and prognostic biomarkers must be driven by the understanding of the mechanisms underlying this disease. In the present study, a genetic alteration profile of the BM was produced, as well as for primary LC, with the intention of identifying potential biomarkers associated with brain metastasis from lung cancer. Currently, limited numbers of biomarkers possess the potential to be used clinically for BM, however a large number of promising candidate biomarkers are being evaluated in current studies, and may soon become relevant for clinical practice (12).

The targeted NGS analysis of 15 tumor samples from patients with LCBM as well as 13 patients with primary LC revealed that 14/15 (93.3%) patients exhibited ≥1 somatic molecular activating mutation and 11/13 (84.6%) had ≥1 molecular aberration. The most frequent genomic alterations in these types of cancer were TP53 (11 for LC, 10 for LCBM), KRAS (1 for LC, 2 for LCBM) and EGFR (5 for LC, 6 for LCBM). Notably, mutations in CTNNB1, BRAF, SMAD4 and PIK3CA were identified only in LCBM, although their frequencies were low.

In the present study, 60.7% (17/28) of the lung cancer samples analyzed exhibited a TP53 mutation, which is higher than that reported in a number of studies from Western countries (24,25). This discrepancy may be explained by relatively few early-stage samples being analyzed in the present cohort. Lung cancer tissues exhibiting EGFR mutations have demonstrated success in responding to EGFR tyrosine kinase inhibitors (26,27). An equal distribution of EGFR mutations in exon 19 (5/28) and the L858R point mutation in exon 21 (5/28) were identified in the present study, followed by the p.G719A point mutation in exon 8 (1/28).

KRAS mutations were observed in patients with LC in a previous study; KRAS and other driver mutations, including EGFR, human epidermal growth factor receptor 2 and BRAF, being mutually exclusive, as KRAS mutations may represent negative predictive biomarkers for tumor responses in patients with NSCLC treated with EGFR-TKIs (28). Whereas EGFR mutations were identified more frequently in never-smokers,
KRAS mutations were first identified among patients who reported greater tobacco use (29). KRAS mutations were identified in 20–30% of patients with lung cancer in North American populations (30). The prevalence of KRAS mutations in the present study was less than that observed in the Caucasian group. Similarly, a recent study also identified KRAS mutations in 8.3% (113/1,368) of a patient cohort with LC (31). In the present study, BRAF and PIK3CA mutations were identified in a small proportion of patients with LCBM, however were not exhibited in patients with LC. However, these

Figure 3. RNA fusion panel tests revealed only EML4(13)-ALK(20) in one sample of lung adenocarcinoma brain metastasis. EML4, echinoderm microtubule associated protein like 4; ALK, anaplastic lymphoma kinase. The blue bar represents the Cris, and the pink bar represents the Watson strand of DNA in the reference genome (hg19).

Figure 4. Variant frequency detected by next generation sequencing in 28 formalin-fixed, paraffin-embedded samples sub-grouped by LC and LCBM. CTNNB1, PIK3CA, SMAD4, BRAF and EML4-ALK mutations were present in the metastatic lesions in 3/15, 2/15, 2/15, 2/15 and 1/15 LCBM cases, but not in primary LC, LCBM, lung adenocarcinoma brain metastasis; LC, lung adenocarcinoma; TP53, tumor protein 53; EGFR, epidermal growth factor receptor; CTNNB1, catenin-β1; KRAS, Kirsten rat sarcoma viral oncogene; PIK3CA, phosphatidylinositol-4,5-bisphosphate 3-kinase gene; BRAF, proto-oncogene B-Raf; SMAD4, mothers against decapentaplegic homolog 4.

KRAS mutations were first identified among patients who reported greater tobacco use (29). KRAS mutations were identified in 20–30% of patients with lung cancer in North American populations (30). The prevalence of KRAS mutations in the present study was less than that observed in the Caucasian group. Similarly, a recent study also identified KRAS mutations in 8.3% (113/1,368) of a patient cohort with LC (31). In the present study, BRAF and PIK3CA mutations were identified in a small proportion of patients with LCBM, however were not exhibited in patients with LC. However, these
results do not imply that these mutations are associated with BM in patients with lung cancer as mutations of BRAF (11) and PIK3CA (32) have been identified in a variety of brain tumors and BM; however, to the best of our knowledge, SMAD4 and CTNNB1 mutations have not previously been identified in brain tumors. Mutations of SMAD4 have been identified primarily in pancreatic and colorectal cancer, and among other types of cancer only occasionally (≤7% in lung cancer) (33). SMAD4 L533R is a pathogenic mutation, which was first identified in juvenile polyposis (34). Inactivation of SMAD4 was considered to serve an important function in impairing transforming growth factor (TGF)-β signaling. Yu et al (35) reported that activation of the TGF-β/SMAD signaling pathway promotes LC metastasis by enhancing the ability of LC cells to extravasate into distant organs. It is reasonable to consider that SMAD4 is also involved in LCBM.

Mutational analysis of exon 3 of CTNNB1 in the present study identified p.S45F, p.S37C and p.D32V mutations. These mutations caused amino acid substitutions in 3/15 LCBM samples (Ser33Phe, Ser33Cys and Ser37Cys). The mutations in exon 3 of CTNNB1 affected glycogen synthase kinase-3β phosphorylation sites of the degradation-targeting box of CTNNB1 and resulted in nuclear CTNNB1 protein accumulation (36). Notably, CTNNB1 mutations were detected in BM from primary LC, in the present study. This finding is consistent with a recent study by Mäki-Nevala et al (37), which demonstrated that CTNNB1 mutations located at codon 37 were detected in two specimens of LCBM. If codons 33 and 37 are a possible location for a number of CTNNB1 mutations in LCBM, this would indicate that CTNNB1 mutations may be important for a secondary event during cancer progression in a carcinogen-specific manner.

There were two novel types of TP53 deletion mutations in LCBM identified in the present study, the TP53 c.495_506del 12 (p.V166_H169del VTAH) and TP53 intron region (7577157-63del TACTCAG), which were validated by Sanger sequencing. Han et al (38) revealed a deletion of two bases, which altered glutamine 167 to alanine, and the resulting frame-shift produced an in-frame stop codon located at amino acid 179 in a human colon cancer cell line (HCC 278). This mutation may be associated with “slippage” or misalignment during DNA replication, due to the occurrence of the CA dinucleotide four times within a stretch of 12 bases (codons 166-169). A novel TP53 gene splice variant was observed in the present study, a 7-base deletion at position 7577157-63 at the splice site of intron 7, which may affect the splice site of exon 8. Although splice sites in TP53 are not typical mutation sites, there is evidence that TP53 splicing mutations lead to exon dropping, indicating a biological relevance (39). Whether these two novel mutations serve a function in BM of lung cancer should be subject to further investigation.

There are several limitations to the present study. The sample size was small, and the retrospective design hindered the ability to study more clinically relevant outcomes, including personalized treatment success, due to the rarity of BM specimens. The NGS approach is designed to detect actionable mutations, including single nucleotide variants, insertions and deletions; however, copy number variants, including gene amplifications and structural variants/rearrangements, are detected less frequently. Therefore, the data acquired may underestimate the mutation burden for the cases. There was no opportunity to investigate matched primary tumors or to analyze in detail intra-patient differences in the mutation status of these genes between tumor sites.

Figure 5. Sanger sequencing chromatograms. (A) TP53-p.V166_H169del and (B) TP53-wild-type; (C) TP53-7577157-63del TACTCAG and (D) TP53-wild-type. TP53, tumor protein 53.
In conclusion, NGS demonstrated a high-throughput in mutational analysis for the patients with LCBM, calculated by detecting the molecular alterations that have a potential clinical relevance. Such findings may assist in clinical decision-making regarding therapeutic intervention for individual patients and provide improved diagnosis or prognosis.

Acknowledgements

The present study was supported by the Research Fund for Shanghai Leading Talents (2015-2017).

Competing interests

The authors declare that they have no competing interests.

References


