Development of an optimal protocol for molecular profiling of tumor cells in pleural effusions at single-cell level

メタデータ	
	言語: English
	出版者:
	公開日: 2020-03-20
	キーワード (Ja):
	キーワード (En):
	作成者: 中村, 育子
	メールアドレス:
	所属:
URL	https://jair.repo.nii.ac.jp/records/2002443

ORIGINAL ARTICLE

Development of an optimal protocol for molecular profiling of tumor cells in pleural effusions at single-cell level

Ikuko Takeda Nakamura[1,2] | Masachika Ikegami[1,3] | Nobuhiko Hasegawa[1,4] |
Takuo Hayashi[5] | Toshihide Ueno[1] | Masahito Kawazu[1] | Shigehiro Yagishita[6] |
Yasushi Goto[7] | Yuki Shinno[7] | Yuki Kojima[8] | Kazuya Takamochi[9] |
Fumiyuki Takahashi[2] | Kazuhisa Takahashi[2] | Hiroyuki Mano[1] | Shinji Kohsaka[1]

[1]Division of Cellular Signaling, National Cancer Center Research Institute, Tokyo, Japan

[2]Department of Respiratory Medicine, Graduate School of Medicine, Juntendo University, Tokyo, Japan

[3]Department of Orthopedic Surgery, Faculty of Medicine, The University of Tokyo, Tokyo, Japan

[4]Department of Orthopedic Surgery, Graduate School of Medicine, Juntendo University, Tokyo, Japan

[5]Department of Human Pathology, Graduate School of Medicine, Juntendo University, Tokyo, Japan

[6]Division of Molecular Pharmacology, National Cancer Center Research Institute, Tokyo, Japan

[7]Department of Thoracic Oncology, National Cancer Center Hospital, Tokyo, Japan

[8]Department of Breast and Medical Oncology, National Cancer Center Hospital, Tokyo, Japan

[9]Department of General Thoracic Surgery, Graduate School of Medicine, Juntendo University, Tokyo, Japan

Correspondence
Shinji Kohsaka and Hiroyuki Mano, Division of Cellular Signaling, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan.
Emails: skohsaka@ncc.go.jp; hmano@ncc.go.jp

Abstract

Liquid biopsy analyzes the current status of primary tumors and their metastatic regions. We aimed to develop an optimized protocol for single-cell sequencing of floating tumor cells (FTCs) in pleural effusion as a laboratory test. FTCs were enriched using a negative selection of white blood cells by a magnetic-activated cell sorting system, and CD45-negative and cytokeratin-positive selection using a microfluidic cell separation system with a dielectrophoretic array. The enriched tumor cells were subjected to whole-genome amplification (WGA) followed by genome sequencing. The FTC analysis detected an *EGFR* exon 19 deletion in Case 1 (12/19 cells, 63.2%), and *EML4-ALK* fusion (17/20 cells, 85%) with an alectinib-resistant mutation of *ALK* (p.G1202R) in Case 2. To eliminate WGA-associated errors and increase the uniformity of the WGA product, the protocol was revised to sequence multiple single FTCs individually. An analytical pipeline, accurate single-cell mutation detector (ASMD), was developed to identify somatic mutations of FTCs. The large numbers of WGA-associated errors were cleaned up, and the somatic mutations detected in FTCs by ASMD were concordant with those found in tissue specimens. This protocol is applicable to circulating tumor cells analysis of peripheral blood and expands the possibility of utilizing molecular profiling of cancers.

KEYWORDS
circulating tumor cells, floating tumor cells, liquid biopsy, lung adenocarcinoma, molecular profiling

1 | INTRODUCTION

A liquid biopsy refers to blood-based cancer testing and involves detailed molecular analysis. Human blood samples contain several sources of cancer-derived materials, such as circulating tumor cells (CTCs), circulating tumor DNA (ctDNA), and exosomes.[1] Analyses of these components enable clinicians to repeatedly and noninvasively evaluate the current status of primary tumors and their metastatic regions in cancer patients.[2]

CTCs in lung cancer can be detected using CTC-isolating technologies such as the CELLSEARCH system (Silicon Biosystems, Bologna, Italy) and filtration techniques.[3-5] The amount of CTCs in peripheral blood has been shown to be of prognostic value and to change depending on treatment response.[6] The DEPArray system (Silicon Biosystems) is a semi-automatic microfluidics apparatus that allows the isolation of rare cells, such as CTCs,[7] that may be subjected to further molecular profiling.[8]

In addition to liquid biopsy of blood, other body fluids, such as pleural effusion, urine, and saliva, may contain genomic materials from tumors as well.[9,10] Malignant pleural effusion is a devastating complication of lung cancer, and sometimes the only source available for genetic profiling analysis. However, comprehensive molecular profiling of floating tumor cells (FTCs) in malignant pleural effusion has not been fully investigated.[11]

Although the clinical utility of ctDNA analyses has been recently reported in lung cancer,[12,13] a clinical platform for DNA sequencing using CTCs has not yet been developed. In this study, we applied DEPArray technology to detect and sort tumor cells from malignant pleural effusion and established a protocol for molecular profiling.

2 | MATERIALS AND METHODS

2.1 | Patients and sample collection

As a pilot cohort study, we enrolled 2 patients with lung adenocarcinoma from the Department of Respiratory Medicine at Juntendo University Hospital in Tokyo. As an additional cohort, we enrolled 9 patients with either lung adenocarcinoma or another thoracic malignancy from the Department of Respiratory Medicine at Juntendo University Hospital in Tokyo or the Department of Thoracic Oncology at National Cancer Center Hospital in Tokyo. All study participants gave written informed consent. The study was approved by the Ethics Committee of National Cancer Center Research Institute (No. 2015-202) and Juntendo University School of Medicine (No. 2014176). The malignant pleural effusion samples were collected by clinicians and stored on ice during sample processing. The enrichment process was initiated within 2 h after specimen collection.

2.2 | Tumor cell enrichment with ferromagnetic beads

The collected samples were centrifuged at $400 \times g$ for 10 min to separate blood cells, including FTCs, from the fluid. Following centrifugation, the supernatant was carefully removed, and the remaining cells were incubated with anti-CD45 antibody ferromagnetic beads (Miltenyi Biotec, Bergisch Gladbach, Germany) for 15 min at 4°C. Afterward, CD45-negative cells were selected by immunomagnetic separation using an autoMACS Pro Separator (Miltenyi Biotec) in accordance with the manufacturer's protocol. The negative fractions were collected and proceeded to cell lysis by incubation with RBC Lysis Buffer (Tonbo Biosciences, San Diego, CA) for 15 min at room temperature. The CD45-negative cells were then incubated with anti-CD235a (glycophorin A) antibody conjugated to phycoerythrin (PE) (BioLegend, San Diego, CA) followed by incubation with anti-PE ferromagnetic beads (Miltenyi Biotec, Cologne, Germany) for 10 min at 4°C, respectively. Selection of CD235a-negative cells was performed using magnetic-activated cell sorting on an autoMACS. The number of enriched cells was counted using a Countess II FL Automated Cell Counter (Thermo Fisher Scientific, Waltham, MA).

2.3 | Immunocytochemistry

The cells were fixed with 4% paraformaldehyde for 20 min at room temperature and stained with anti-CD45 antibody conjugated to allophycocyanin (APC) (Miltenyi Biotec) for 30 min at room temperature. Then, the cells underwent permeabilization using an Inside Stain Kit (Miltenyi Biotec). Intracellular staining was carried out for 10 min at 4°C using anti-pan cytokeratin (CK) antibody conjugated to PE (BioLegend). Nuclear staining was performed using Hoechst staining solution for 5 min at room temperature.

2.4 | Single-cell isolation using DEPArray

The immunolabeled cells were washed twice with SB115 buffer (Silicon Biosystems) and loaded onto a DEPArray cartridge (Silicon Biosystems). Chip scanning was performed using an automated fluorescence microscope to generate an image gallery on the DEPArray system. FTCs were selected in accordance with their morphology and staining pattern (Hoechst positive, CK positive, and CD45 negative) and automatically recovered into 200 μL tubes. All samples were washed in phosphate-buffered saline (PBS) and then preserved at −80°C with 1 μL PBS.

2.5 | Whole-genome amplification

Single cells underwent whole-genome amplification (WGA) using a SMARTer PicoPLEX WGA Kit (TaKaRa Bio Inc, Shiga, Japan) in accordance with the manufacturer's instructions, and the amplified products were purified with a MinElute polymerase chain reaction (PCR) Purification Kit (QIAGEN, Hilden, Germany). The DNA concentration was measured with a NanoDrop One UV-Vis spectrophotometer (Thermo Fisher Scientific).

2.6 | Library preparation for next-generation sequencing

The DNA samples were prepared for analysis using 3 platforms: the AmpliSeq for Illumina Comprehensive Cancer Panel, the Todai OncoPanel (TOP),[14] and whole-genome sequencing (WGS). AmpliSeq is an amplicon-based cancer panel designed to detect somatic variants from whole exons of 409 genes associated with cancer. Library preparation was conducted following the manufacturer's instructions except for the number of PCR cycles, which was optimized for single-cell sequencing. The TOP DNA panel is a hybridization capture-based panel developed to detect single nucleotide variations (SNVs) and insertions/deletions of 464 genes. Library preparation of TOP was conducted as previously described.[14] WGS samples were prepared using a NEBNext Ultra DNA Library Prep Kit for Illumina (New England BioLabs, Ipswich, MA). Library concentration was measured using a Qubit 2.0 Fluorometer (Thermo Fisher Scientific). Library quality was considered to be sufficient if a single peak around 250-350 bp was detected using the Agilent 2200 TapeStation system (Thermo Fisher Scientific). Sequencing analysis was performed on an Illumina HiSeq 2500 system (Illumina, San Diego, CA).

2.7 | DNA extraction from leukocyte and formalin-fixed paraffin-embedded tissues

The CD45-positive fractions isolated with autoMACS were analyzed as paired normal leukocytes. Genomic DNA was then extracted using a QIAamp DNA Mini Kit (QIAGEN). Genomic DNA of primary tumors was extracted from formalin-fixed paraffin-embedded (FFPE) tissues with a GeneRead DNA FFPE Kit (QIAGEN). Extracted DNA from the FFPE samples was analyzed using the TOP panel. FFPE specimen of Case 1 was a cell block of pleural effusion. FFPE specimen of Case 2 was collected from mediastinal lymph nodes. Both specimens were provided as treatment naïve cancer.

2.8 | Sanger sequencing for *EGFR* exon 19 deletion

For capillary sequencing with a 3130xl Genetic Analyzer (Thermo Fisher Scientific), 10 ng of template DNA were used to amplify *EGFR* exon 19 with GoTaq G2 Hot Start Green Master Mix (Promega, Madison, WI) in accordance with the manufacturer's instructions. The primers used were 5′-CATGTGGCACCATCTCACA-3′ and 5′-CCACACAGCAAAGCAGAAAC-3′.

2.9 | Identification of *EML4-ALK* breakpoint

Next-generation sequencing (NGS) reads spanning the *EML4-ALK* fusion points were obtained by WGS, and to confirm the fusion point, 3 forward primers (f1, f2, and f3) at *EML4* intron 6 and three reverse primers (r1, r2, and r3) at *ALK* intron 19 were designed.

PCR analysis was performed in 2 FTC groups (each group consisted of whole-genome amplified DNA from 10 single FTCs) with 5 primer sets: f1 and r1 (primer set 1), f1 and r2 (primer set 2), f1 and r3 (primer set 3), f2 and r3 (primer set 4), and f3 and r3 (primer set 5). The PCR products were analyzed by Sanger sequencing to identify the fusion sequence of *EML4-ALK*. The primer sequences were: f1: 5′-TAAATCCTGTGCCACGTCCC-3′, f2: 5′-GCTTTTTAAACTTTGGTGAAC-3′, f3: 5′-TAATATAGTGAGCACCTG-3′, r1: 5′-AGCTTCCGTTTTGGCTTGG-3′, r2: 5′-TGAGGTGCAGAATCAGGG-3′, and r3: 5′-TTCACCATCGTGATGGACAC-3′.

2.10 | Data analysis

The paired-end reads were independently aligned to the human reference genome (hg38) using the Burrows-Wheeler Alignment tool,[15] Bowtie 2 (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml), and NovoAlign (http://www.novocraft.com/products/novoalign/). Somatic mutations were called using MuTect (http://www.broadinstitute.org/cancer/cga/mutect), SomaticIndelDetector, and VarScan (http://varscan.sourceforge.net). Mutations were discarded if: (a) the read depth was <20 or the variant allele frequency (VAF) was <0.3, (b) they were present in paired leukocytes and whole-genome amplified leukocyte DNA, or (c) they were present in normal human genomes in either the 1000 Genomes Project dataset (http://www.internationalgenome.org/) or our in-house database. Gene mutations were annotated by SnpEff (http://snpeff.sourceforge.net).

2.11 | Accurate single-cell mutation detector (ASMD) pipeline

To accurately detect somatic mutations from FTCs by removing the WGA errors, we developed a four-step analytical pipeline. First, the somatic mutations of single-cell FTCs and leukocytes were identified and compared with those detected in bulk leukocytes from the same individual. Second, mutations were discarded if: (a) the read depth was <100, (b) the variant was not annotated as pathogenic by the ClinVar database and its VAF was <0.3, (c) the variant was annotated as pathogenic by the ClinVar database and its VAF was <.05, and (d) they were present in whole-genome amplified DNA of non-paired single-cell leukocytes isolated with DEPArray. Third, mutations that were detected in only 1 of the FTCs from the same individual were excluded. Non-pathogenic mutations detected in more than 1 case were also removed.

3 | RESULTS

3.1 | Patient profiles of the 2 pilot cases

The 2 lung adenocarcinoma patients enrolled in this study were admitted to the Department of Respiratory Medicine at Juntendo

University Hospital in Tokyo. Case 1 was a 35-y-old woman diagnosed with lung adenocarcinoma with *EGFR* exon 19 deletion. Before commencing first-line chemotherapy, 10 mL of pleural effusion was obtained. Case 2 was a 79-y-old male diagnosed with lung adenocarcinoma with *EML4-ALK* fusion. Crizotinib, an anaplastic lymphoma kinase (ALK) and ROS1 kinase inhibitor, were administered for 7 mo. Then, due to disease progression, the drug was changed to alectinib, another ALK inhibitor. At 7 mo later, the patient relapsed with massive pleural effusion, 15 mL of which was obtained for this study (Figure 1A).

3.2 | Isolation and sequencing of FTCs from malignant pleural effusion

Malignant pleural effusion was collected and centrifuged to separate the cellular components (Figure 1B, Step 1). Blood cell depletion was carried out using CD45-microbeads and CD235a-microbeads using autoMACS (Figure 1B, Step 2-1). The CD45 antigen is expressed on all cells of hematopoietic origin except erythrocytes, platelets, and their precursor cells, and the CD235a antigen, a single-pass transmembrane glycoprotein, is expressed on mature erythrocytes and erythroid precursor cells. Red blood cell lysis was performed using RBC Lysis Buffer. Then, the remaining cells were fixed and stained with anti-CD45 antibody-APC, anti-pan-CK antibody-PE, and Hoechst stain solution (Figure 1B, Step 2-2). Subsequently, the samples were loaded onto a DEPArray cartridge, and chip scanning was performed (Figure 1B, Step 2-3). FTCs were sorted based on their morphology and staining pattern and recovered with 10 μL of elution buffer into 200 μL tubes (Figure 1B, Step 2-4).

WGA was performed on single FTCs using PicoPLEX (Figure 1B, Step 3). The amplified DNA was subjected to Sanger sequencing, NGS analyses with Illumina Comprehensive Cancer Panel, TOP panel, and WGS (Figure 1B, Step 4).

3.3 | Analysis of DEPArray and genotyping using Sanger sequencing

After FTC enrichment using autoMACS, 1812 and 7350 cells in Cases 1 and 2, respectively, were isolated with the dielectrophoretic cage. CD45-negative and CK-positive cells were further selected as FTCs; the FTC counts in Case 1 and 2 were 114 and 3787 cells, respectively (Figure 2). In total, 51 cells (the maximum capacity for the DEPArray system) from each case were collected into 200 μL tubes. Leukocytes from Case 1 (CD45-positive, CK-negative, and Hoechst-positive cells) were also isolated as a negative control for the subsequent NGS analysis (Figure 2B).

WGA of 20 single FTCs from each patient was performed using PicoPLEX. The genomes of 19 and 20 FTCs for Cases 1 and 2, respectively, were successfully amplified up to 2 μg. *EGFR* del L747_A750 was confirmed in 12/19 (63.2%) samples from Case 1 using Sanger sequencing (Figure 2A).

3.4 | Whole-genome sequencing

Because the 12 *EGFR* mutation-positive samples from Case 1 were identified as tumor cells, they were subjected to further investigation using WGS. All 20 samples from Case 2 proceeded to WGS
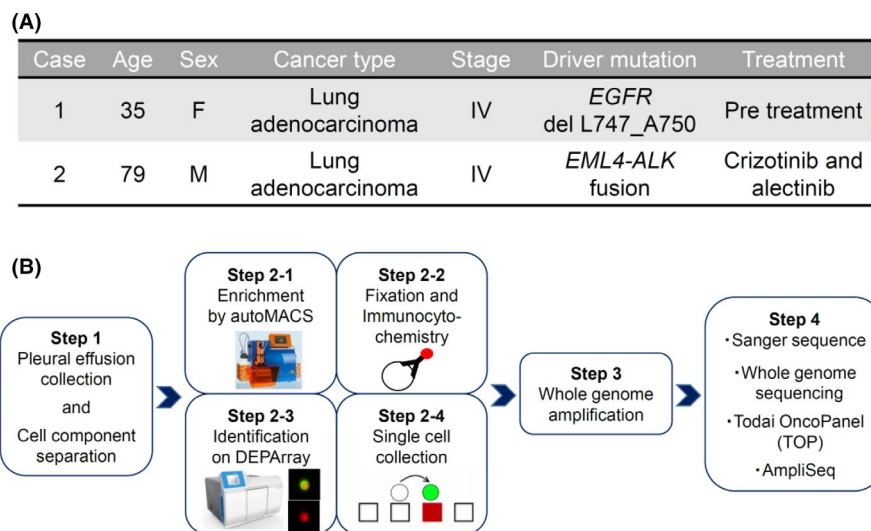

FIGURE 1 Patient profile and overview of FTC analysis. A, Patient general information, somatic mutation, and treatment history of the 2 pilot cases. B, Schematic of FTC isolation from malignant pleural effusion and sequencing. The pleural effusion specimen was collected, and cellular components were separated by centrifugation (Step 1). Using CD45 microbeads and CD235a microbeads, leukocytes and erythrocytes were removed using autoMACS. Then, the cells were fixed and stained with anti-CD45 antibody-APC, anti-pan-CK antibody-PE, and Hoechst. FTC-enriched samples were loaded onto a DEPArray cartridge, and chip scanning was performed. The immunolabeled cells were selected by their morphology and staining pattern and recovered into 200 μL tubes (Step 2). WGA was performed on single FTCs using PicoPLEX (Step 3). The amplified genome was analyzed by Sanger sequencing, AmpliSeq, the TOP, and WGS (Step 4). F, female; M, male
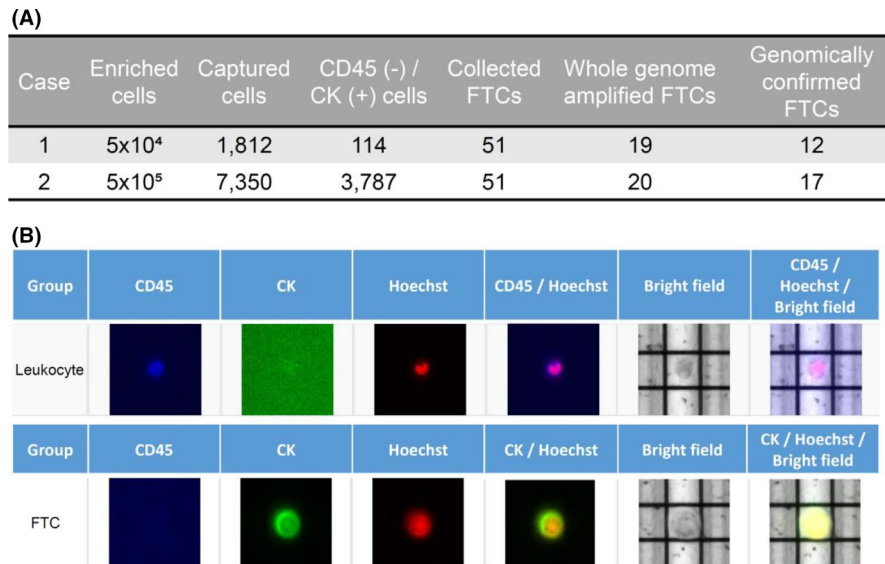
(A)

Case	Enriched cells	Captured cells	CD45 (-) / CK (+) cells	Collected FTCs	Whole genome amplified FTCs	Genomically confirmed FTCs
1	5×10^4	1,812	114	51	19	12
2	5×10^5	7,350	3,787	51	20	17

(B)

FIGURE 2 Enumeration and image analysis of FTCs on DEPArray. A, The number of cells enriched using autoMACS and captured by dielectrophoretic cages was counted. CD45-negative and CK-positive cells were selected as FTCs and recovered (51 was the maximum capacity for the DEPArray system). Twenty single FTCs from each case underwent WGA, and genomic DNA of 19 and 20 cells from Cases 1 and 2, respectively, were successfully amplified. *EGFR* exon 19 of FTCs from Case 1 was analyzed by Sanger sequencing, which detected *EGFR* exon 19 deletion in 12 cells. FTCs from Case 2 were analyzed with *EML4-ALK* fusion breakpoint sequencing, which detected *EML4-ALK* fusion in 17 cells. B, Representative images of leukocytes (CD45 positive, CK negative, and Hoechst positive) are shown in the upper row, and images of FTCs (CD45 negative, CK positive, and Hoechst positive) are in the lower row. CK, cytokeratin; FTC, floating tumor cell
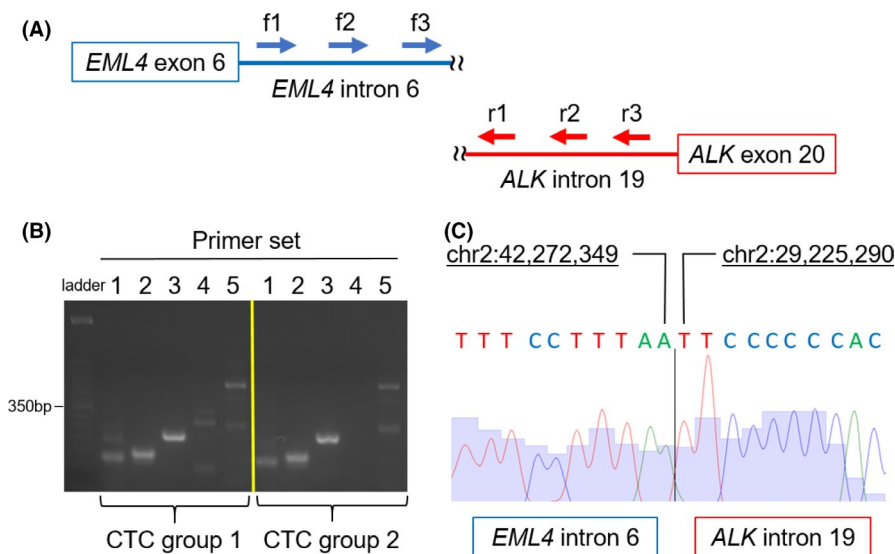

FIGURE 3 Identification of the *EML4-ALK* breakpoint sequence in FTCs. A, WGS of Case 2 indicated a span read of *EML4-ALK* fusion. To determine the breakpoint, 3 forward primers (f1, f2, and f3) and 3 reverse primers (r1, r2, and r3) were prepared. B, PCR was performed for each FTC group, which consisted of the amplified genome of 10 single FTCs. The results of electrophoresis of the PCR products are shown. The following 5 primer sets were used: primer set 1 (f1 and r1), primer set 2 (f1 and r2), primer set 3 (f1 and r3), primer set 4 (f2 and r3), and primer set 5 (f3 and r3). C, The breakpoint sequence of *EML4* intron 6 to *ALK* intron 19 was identified using Sanger sequencing

because no SNV information was available for Case 2. Library preparation for WGS was conducted for each single-cell DNA sample individually.

Approximately 5-20 million NGS reads were obtained from each single-cell DNA sample (Figure S1A, B). However, because of uneven amplification of the genome, the mean coverage of each cell was extremely low (Figure S1B). The concentration of whole-genome amplified single cells was approximately 100 ng/μL while that of the negative control (water) was less than 10 ng/μL, suggesting that the genome amplification was performed without contamination from environmental bacteria. The low mapped reads ratio was due to the low quality of base call as our analytical pipeline discarded the reads containing multiple masked bases.

TABLE 1 Read summary and mutational analysis from each platforms

Case 1										
Sample type	FFPE	FTC					FTC	Leukocyte		FTC
Platform	TOP	WGS	AmpliSeq		TOP		AmpliSeq	AmpliSeq	TOP	TOP
Fixed/live	Fixed	Fixed	Fixed				Live	Live	Live	Live
Sample set	-	-	T1	T2	T1	T2	-	-	-	-
Whole-genome amplification	N	Y	Y	Y	Y	Y	Y	N	N	Y
Raw read number	5.3×10^7	2.0×10^8	9.7×10^7	8.8×10^7	4.6×10^7	3.4×10^7	9.1×10^7	1.1×10^8	9.4×10^7	8.4×10^6
Mapped read number	4.9×10^7	4.1×10^7	8.1×10^7	7.4×10^7	3.6×10^7	2.7×10^7	7.8×10^7	9.5×10^7	8.7×10^7	7.1×10^6
Mean coverage	123	0.85	5050	4561	293	229	4783	5892	740	87.9
Breadth of coverage % 1× / 20× (%)	99.8/99.2	23.9/0.2	90.6/82.9	90.1/82.7	90.4/67.6	89.0/64.6	91.0/82.6	96.3/95.5	99.9/99.7	83.9/50.0
Somatic mutation	EGFR del L747_A750 FGF4 E154L	ND	EGFR del L747_A750 ERCC2 P475L SYNE1 D397N KMT2C D348N		EGFR del L747_A750 FGF4 E154L		-	-	-	-

Case 2										
Sample type	FFPE	FTC							Leukocyte	
Platform	TOP	WGS	AmpliSeq			TOP			AmpliSeq	TOP
Fixed/live	Fixed	Fixed	Fixed			Fixed			Live	Live
Sample set	-	-	T1	T2	T3	T1	T2	T3	-	-
Whole-genome amplification	N	Y	Y	Y	Y	Y	Y	Y	N	N
Raw read number	2.9×10^7	2.0×10^8	9.1×10^7	1.0×10^8	3.6×10^7	9.4×10^7	3.1×10^7	4.8×10^7	9.6×10^7	7.6×10^7
Mapped read number	2.7×10^7	5.0×10^7	7.5×10^7	8.6×10^7	2.9×10^7	7.8×10^7	2.5×10^7	3.8×10^7	8.1×10^7	7.0×10^7
Mean coverage	10.6	1.14	4586	5280	248	4826	232	262	4919	947
Breadth of coverage % 1× / 20× (%)	99.3/5.3	27.9/0.3	91.6/84.8	92.8/86.6	91.1/68.1	93.1/87.9	92.7/71.0	92.7/71.0	96.2/95.6	99.8/99.5
Somatic mutation	ND	EML4-ALK fusion	EML4-ALK fusion ALK G1202R			EML4-ALK fusion ALK G1202R			-	-

Note: Sequence read summary and detected mutations of FTCs, leukocytes, and tumor FFPE specimens.

Abbreviations: FFPE, formalin-fixed paraffin-embedded specimen; FTC, floating tumor cell; TOP, Todai OncoPanel; WGS, whole-genome sequencing; n, not performed; y, performed; ND, not detected.
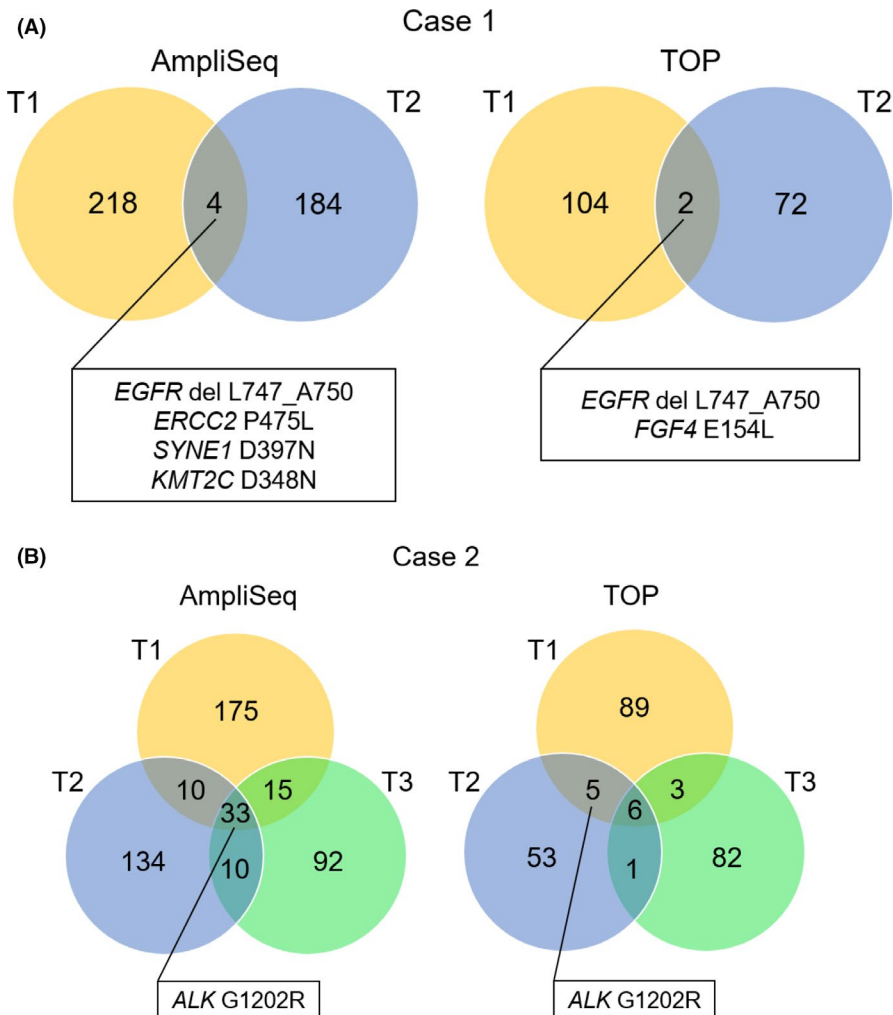
FIGURE 4 SNV distribution among each FTC set for Cases 1 and 2. A, Number of identified mutations in each FTC set of Case 1 analyzed by AmpliSeq and TOP. B, Number of identified mutations in each FTC set of Case 2 analyzed by AmpliSeq and TOP

Merging the NGS data of the single cells from the same case improved the coverage, suggesting that the amplified regions in the genome were random among the samples (Figure S1C, D). Amplification bias and coverage uniformity of merged data were assessed by Lorenz curves and Gini's coefficient. The Gini's coefficients were 0.21 (case 1) and 0.17 (case 2) (Figure S1E), which suggested that merging the data of multiple FTCs may be a reasonable approach to make the coverage uniform. Uneven WGA was also observed in a single-cell analysis of H2228 lung cancer cells (Figure S2), suggesting that amplification bias is due to WGA and not acquired in the process of clinical sample preparation.

EGFR del L747_A750 was not detected using WGS in Case 1, because no reads were obtained around the *EGFR* exon 19 locus. In contrast, span reads of the *EML4-ALK* fusion from merged NGS data of Case 2 were detected, despite the low number of mapped reads. To determine the genomic breakpoint of the fusion gene, we performed PCR using a primer set specific to the *EML4-ALK* breakpoint (Figure 3A). The expected size of the PCR products was observed with primer sets 2 and 3 (Figure 3B), and the breakpoint sequence of *EML4-ALK* was confirmed using Sanger sequencing (Figure 3C). *EML4-ALK* fusion was further confirmed by PCR amplification in 17/20 (85%) isolated single cells from Case 2.

3.5 | Target sequencing by hybridization capture-based and amplicon-based methods

We merged 5 or 6 WGA samples to establish 2 or 3 sets of pooled DNAs (T1 and T2 for Case 1, and T1-T3 for Case 2, respectively) for target sequencing. Target sequencing with a hybridization capture-based method (TOP panel) and an amplicon-based method (AmpliSeq) was performed on all pooled DNA sets (Table 1).

The TOP panel analyses yielded a mean sequence coverage of 20× for 70% of the captured regions in the FTCs, while the same analyses in bulk white blood cells covered >99% of the target regions (Table 1 and Figure S3). The PCR cycles of AmpliSeq were optimized to increase the detection frequency of low amplification regions in WGA. We compared the sequencing coverage among samples prepared with 12, 16, 23, and 30 PCR cycles. The 20× coverage area was widest (>80%) after 23 cycles (Figure S4).

To eliminate WGA errors, sequence reads of bulk leukocytes and single leukocytes from the same individuals were compared. The mutations identified specifically in single leukocytes were regarded as WGA errors. The VAF of mutations only identified in single leukocytes was <0.3 (Figure S5). Therefore, mutations only found in single leukocytes were removed from mutation analyses for FTCs (Figure 4).

EGFR exon 19 deletion was detected in T1 and T2 of Case 1 by both AmpliSeq and TOP (Figures 4A and S6A), whereas >200 sample-specific or panel-specific mutations were also identified. *ALK* p.G1202R, a known alectinib-resistant mutation, was the only mutation identified throughout multiple samples from Case 2 by both panels (Figures 4B and S6B, and Table S1).

3.6 | Development of the ASMD pipeline and validation in 9 cases

Considering the number of WGA-associated errors and the low uniformity of the WGA product, it is preferable to sequence single CTCs individually. An analytical pipeline, ASMD was constructed to identify somatic mutations of CTCs, which were supported by multiple reads of different clones (Figure 5A). The FTC sequences of 9 additional cases were conducted to evaluate the utility of ASMD

(Tables 2 and S2). For each case, except for Case 5, 5 FTCs were sequenced, which were performed using WGA and analyzed with TOP panels individually. The coverage was improved by merging the data of multiple FTCs (Figure 5B and Table S3). The large numbers of WGA-associated errors were cleaned up, and the somatic mutations detected in FTCs by ASMD were concordant with those found in FFPE specimens, although some additional or depleted mutations were observed in FTCs, suggesting clonal evolution of the tumors during metastasis (Figure 5C and Table 3).

4 | DISCUSSION

This study is the first study to perform NGS of single cells in pleural effusions of multiple cases. The combination of autoMACS and DEPArray was applied to detect and enrich FTCs in pleural effusion. The enrichment method for only 10 mL of pleural effusion enables
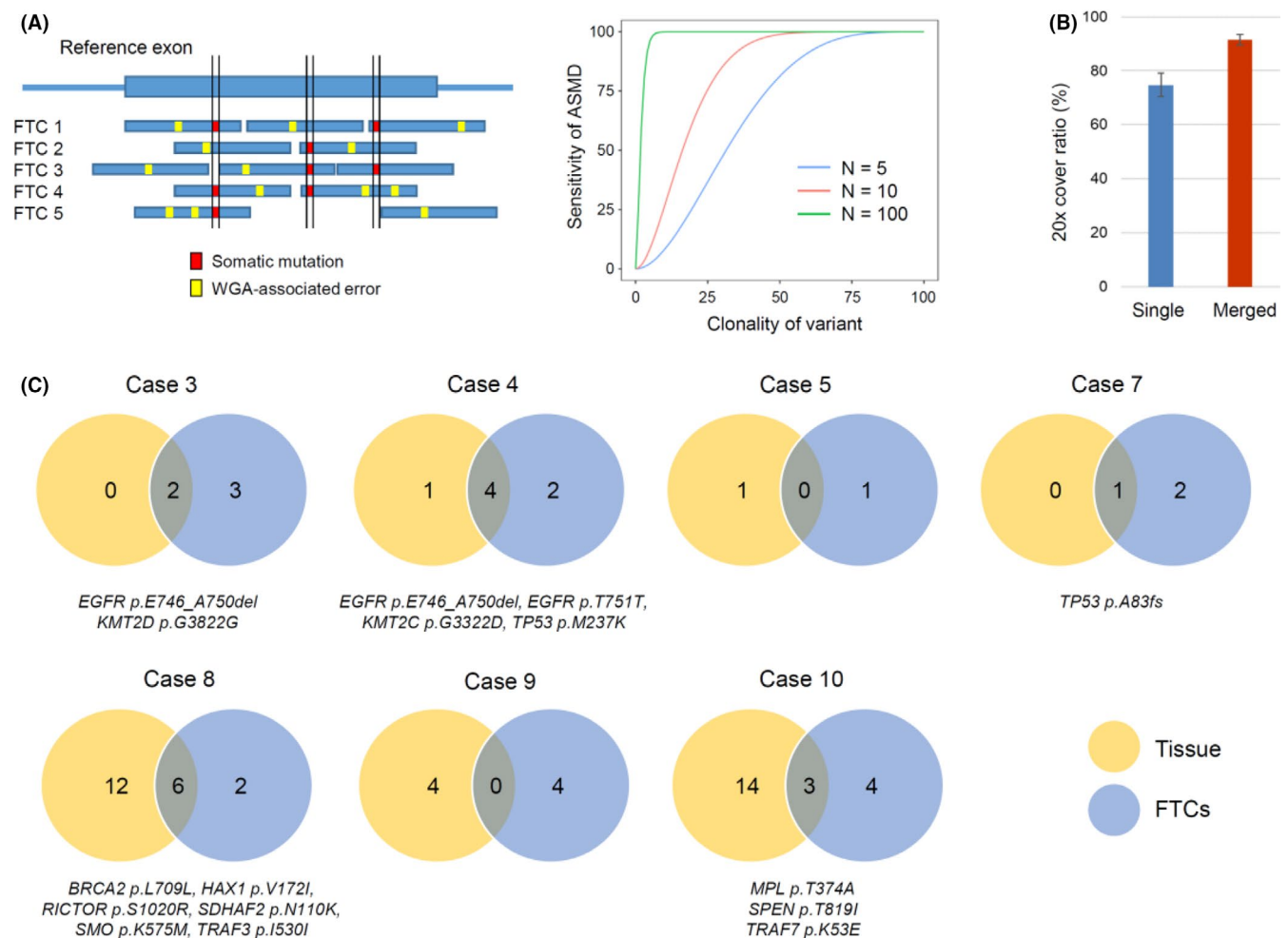

FIGURE 5 Precise analysis of somatic mutations in FTCs using the ASMD pipeline. A, The schema on the left shows the overview of the accurate single-cell mutation detector (ASMD) pipeline. The ASMD identifies somatic mutations of FTCs, which are supported by multiple reads of different clones. The unique mutation identified in 1 sample and mutations with low VAF were discarded as WGA-associated errors. The figure on the right shows that the sensitivity of the ASMD is dependent on the clonality of the variant (the ratio of the clone harboring the variant in the tumor) and the number (N) of FTC sequences. B, The average 20× cover ratio of single-cell FTCs or the merged data of FTCs is shown. The coverage was improved by merging the data of multiple FTCs. C, The number of somatic mutations identified in tissue specimens or FTCs are indicated. The variants detected in both specimens are revealed in the Venn diagrams

TABLE 2 Patient profiles of 9 additional cases

Case no.	Age	Sex	Cancer type	Genomic aberration at diagnosis	Analysis	Treatment	Current status
3	81	M	LUAD	*EGFR* ex19 del	Cobas EGFR Mutation Test v2	Op	PD
4	41	M	LUAD	*EGFR* ex19 del	Cobas EGFR Mutation Test v2	Afatinib; CDDP + PEM; Clinical trial	PD
5	36	M	NUT midline carcinoma	*BRD4-NUT*	IHC, FISH	Pre-treatment	Pre-treatment
6	68	M	LUAD	*EGFR* E709K/G719A	Oncomine	Afatinib	PD
7	77	F	LUAD	*EGFR* L858R	Cobas EGFR Mutation Test v2	Op; Osimertinib	PD
8	50	F	MPM	*ETV6* deletion, NF2 448-1G > C, BAP1 D408fs*17, SF3B1 Y623C, PERM1 R921fs*87	FoundationOne CDx	CDDP + PEM; Nivolumab; RT	PD
9	60	M	LUAD	*EGFR* (−), *ALK* (−),*ROS1* (−), PD-L1 50%	Cobas EGFR Mutation Test v2	Pembro; RT; CDDP + PEM	PD
10	69	F	LUAD	*ROS1* (+)	Cobas EGFR Mutation Test v2, OncoGuide AmoyDx	Pre-treatment	Pre-treatment
11	67	F	LUAD	*EGFR* (−), *ALK* (−),*ROS1* (−)	NA	CBDCA + PEM+Bev; Docetaxel; TS-1; Atezolizumab	PD

Note: Patient general information, genomic aberrations and analysis, treatment history, and current status of each case.

Abbreviations: Bev, bevacizumab; CDDP, cisplatin; F, female; FISH, fluorescence in situ hybridization; ICH, immunohistochemistry; LUAD, lung adenocarcinoma; M, male; MPM, malignant pleural mesothelioma; Op, operation; PD, progressive disease; PEM, pemetrexed; RT, radiation therapy.

the diagnosis of carcinomatous pleuritis and the performance of comprehensive molecular profiling. In contrast with this method, the conventional techniques require a considerable amount of pleural effusion specimen for diagnosis.[11]

The *ALK* p.G1202R mutation identified in Case 2 is located at the solvent front of the ALK domain and is recognized as an acquired mutation following ALK-tyrosine kinase inhibitor (TKI) therapy.[16,17] Lorlatinib is known as effective against this mutation.[18,19] Therefore, if clinicians had chosen lorlatinib as the third-line treatment, they might have controlled disease progression in Case 2. Unfortunately, the patient could not receive loratinib because his performance status was too poor due to the accumulation of pleural effusion.

Notably, in this study, targetable *EGFR* driver mutations were identified in 4 cases. Furthermore, potentially targetable variants, such as *BRCA1* truncating mutation (case #9) and *RAF1* oncogenic mutation (case #11) were identified. These may be acquired mutations that arose in response to treatment.

Genomic mutations in peripheral blood CTCs were first detected in non-small-cell lung cancer (NSCLC) patients in 2008 using the CTC-Chip.[18,20] This study revealed EGFR TKI-resistant mutations in some patients who experienced a faster progression compared with those without these mutations. Several other groups also detected *EGFR* mutations in CTCs.[21]

NGS of single CTCs remains challenging because single cells do not contain a sufficient quantity of DNA for molecular analysis. Therefore, WGA is required for NGS analysis of CTCs, although the analyses of the first 2 cases indicated that the DNA amplification process is often error prone. Because the VAF of WGA error was <0.3, library construction was performed individually in single FTCs of the additional 9 cases to keep the VAF of somatic mutation in each cell high.

CTCs are highly heterogeneous, even within the same individuals.[22] Therefore, multiple single FTC sequencing should be performed to evaluate intratumor heterogeneity. Such data would provide insight into the origin of tumors and the molecular mechanism of disease progression and resistance acquisition.[23] It is estimated that ASMD analysis of 5 FTCs per case is capable of detecting mutations, which 68% of tumor cells harbor, at 90% sensitivity. Sequencing more clones can improve the sensitivity of the ASMD method, and the sample indexing technique before pooling for capture might be an efficient way to reduce the cost for capture probes.

Other issues to be overcome in the sequencing of WGA samples include a high allelic drop-out rate, artifacts, and uneven genome amplification. Methods have been developed to minimize WGA amplification bias by reducing the limiting volume for multiple displacement amplification to avoid exponential preferential amplification[24,25] or looping of the amplicons to induce quasi-linear amplification (MALBAC).[26] Despite these advances, WGA remains far from uniform.

There are several studies that compare the methods for WGA.[27-30] Optimization of library preparation using samples that have undergone

TABLE 3 Somatic mutations of 9 additional cases

Case no.	Gene	Variant	VAF						ClinVar	COSMIC	ASMD
			Clone 1	Clone 2	Clone 3	Clone 4	Clone 5	Tissue			
3	ACVR2A	p.Met1fs	55.9	–	42.5	69.8	52.4	–	–	–	Y
	SMARCA4	p.Pro281Pro	30.8	–	30.1	–	–	–	–	–	Y
	TSC1	p.Ser526fs	63.2	71.7	24.5	49.7	72	–	–	–	Y
	EGFR	p.Glu746_Ala750del	–	–	11.4	14.6	11.9	17.3	Drug_response	1163	Y
	KMT2D	p.Gly3822Gly	100	–	–	100	99.6	15.1	–	–	Y
4	CHEK2	p.Arg117Gly	–	5.9	9.1	11.1	–	–	Pathogenic/Likely_pathogenic	1	Y
	PTEN	c.802-2A > G	–	7.4	5.7	5.9	–	–	Pathogenic/Likely_pathogenic	1	Y
	FAT1	p.Val2071Ala	–	–	–	–	99.1	28.7	–	–	N
	TP53	p.Met237Lys	100	99.6	100	99.8	100	24.8	Likely_pathogenic	18	Y
	KMT2C	p.Gly3322Asp	82.4	26.4	15.9	68.1	20	19.1	–	1	Y
	EGFR	p.Thr751Thr	41.7	–	–	48	45.2	18.8	–	–	Y
	EGFR	p.Glu746_Ala750del	21.5	13.2	–	26	29.4	17.6	Drug_response	1163	Y
5	DICER1	p.Asp1810Gly	5.6	13.7	–	NA	NA	–	Pathogenic	–	Y
	JAK1	p.Arg360Gln	–	–	–	NA	NA	28.2	–	–	N
6	EGFR	p.Glu709Lys	84	95.1	69.6	48.9	77.8	NA	Drug_response	36	Y
	EGFR	p.Gly719Ala	83.8	97.6	70.6	48.2	78.5	NA	Pathogenic	75	Y
	IKZF1	p.His127Asn	56.8	21.8	93	32.5	35.7	NA	–	–	Y
	MST1	p.Arg461His	38.3	40.9	8.4	9.7	–	NA	–	2	Y
	PMS1	c.583-1G > A	–	–	–	100	100	NA	–	1	Y
	TP53	p.Arg342*	–	–	100	99.6	100	NA	Pathogenic	379	Y
7	TP53	p.Ala83fs	93.6	93.8	91.8	93.8	94.7	25.1	Pathogenic_HGVS.p	–	Y

(Continues)

TABLE 3 (Continued)

Case no.	Gene	Variant	VAF Clone 1	Clone 2	Clone 3	Clone 4	Clone 5	Tissue	ClinVar	COSMIC	ASMD
8	ACTN2	p.Met361Val	44.6	88.8	36.2	36.6	75.8	-	Uncertain_significance	-	Y
	CDKN2A	p.Arg80*	79.7	100	98.9	-	99.9	-	Pathogenic	293	Y
	TRAF3	p.Ile530Ile	36.2	33.3	-	40.9	37.1	50	-	-	Y
	BRCA2	p.Leu709Leu	41.5	38.7	42.5	48.6	43.3	49.7	Benign	-	Y
	SMO	p.Lys575Met	98.7	66.8	-	46	-	49.1	-	4	Y
	SDHAF2	p.Asn110Lys	67.5	-	-	53.6	57.4	48.9	Uncertain_significance	-	Y
	SOCS1	p.Phe113Phe	-	-	-	-	-	48.1	-	-	N
	RICTOR	p.Ser1020Arg	23.9	-	68.6	38.7	5.7	47.6	-	-	Y
	FGFR4	p.Gly528Ser	-	-	-	-	-	47.5	-	-	N
	ARID1B	p.Arg1789Gln	-	-	-	-	-	47.3	-	-	N
	HAX1	p.Val172Ile	64	69.5	58.8	25.5	63.5	46.7	Benign/Likely_benign	-	Y
	ERBB4	p.Val83Val	-	-	-	-	-	46.4	-	-	N
	KLF5	p.Val357Val	-	-	-	-	-	44.9	-	-	N
	ZNF503	p.Ser16_Gly17dup	-	-	-	-	-	42.2	-	-	N
	MEF2A	p.Gln421_Pro423dup	-	-	15	35.6	13.7	40.3	-	-	N
	ATF7IP	p.Thr175_Ala179del	-	-	-	-	-	37.6	-	-	N
	KLHL14	p.Asp83Gly	-	-	-	-	-	35.8	-	-	N
	MCC	p.Ser25_Ser28dup	-	-	-	-	-	26.2	-	-	N
	SRP14	p.Pro117_Thr126del	-	-	-	-	-	18.7	-	-	N
	DNMT1	p.Thr161Ala	-	-	-	-	-	11.3	-	-	N
9	BRCA1	p.Gln759*	6.8	-	6	-	-	-	Pathogenic	-	Y
	KARS	p.Leu568Phe	51.6	99.4	-	58.6	-	-	-	-	Y
	MELK	p.Ile578Val	62.5	98.8	100	-	5	-	-	-	Y
	UGT1A5	p.Ala158Ala	77.6	9.3	5.8	98.7	82.4	-	-	4	Y
	INSR	p.Arg7Arg	-	-	-	-	-	81.9	-	-	N
	MRE11A	p.Met157Val	-	-	-	-	-	80.2	Conflicting_interpretations	1	N
	RECQL4	p.Asn406Lys	-	-	-	-	100	57.4	-	-	N
	NOTCH1	p.Cys725Cys	-	-	-	-	-	26	-	-	N

(Continues)

TABLE 3 (Continued)

Case no.	Gene	Variant	VAF					Tissue	ClinVar	COSMIC	ASMD
			Clone 1	Clone 2	Clone 3	Clone 4	Clone 5				
10	BRCA2	p.Glu425Glu	33.4	12.3	84.3	61.2	41	–	Benign	1	Y
	FAT1	p.Val476Met	–	–	46	38.9	–	–	–	–	Y
	PTEN	c.802-2A > G	–	8.3	–	–	6.7	–	Pathogenic/Likely_pathogenic	1	Y
	TP63	p.Leu562Pro	–	7.8	–	9	–	–	Likely_pathogenic	–	Y
	TRAF7	p.Lys53Glu	39	6.3	7.5	75.2	38.5	68.8	–	–	Y
	SPEN	p.Thr819Ile	–	53.4	–	66	44.7	61	–	–	Y
	CEBPA	p.Thr265Thr	–	–	–	–	–	54.4	Benign	0	N
	MPL	p.Thr374Ala	–	57.9	100	36.1	38.6	49.1	Benign/Likely_benign	1	Y
	SH2B3	p.Ala36Ala	–	–	–	–	–	47.8	–	–	N
	FOXA1	p.Tyr394Cys	–	–	–	–	–	45.4	–	–	N
	MET	p.Leu674Leu	–	–	–	–	49.6	45.1	–	–	N
	KLF5	p.Glu419Gln	–	–	–	–	–	39.2	–	5	N
	FGFR4	p.Pro356Leu	99.9	–	18.3	28.9	–	38	–	–	N
	ATXN1	p.Gln225del	–	–	–	31.5	–	16.6	–	–	N
	STAT3	p.Ile339Ile	–	–	–	–	–	15.9	–	–	N
	HAX1	p.Leu6Phe	–	–	–	–	–	15.6	–	–	N
	SETDB1	p.Ser983*	–	–	–	–	–	14.7	–	–	N
	SMARCA4	p.Ile1045Ile	–	–	–	–	–	13.8	–	–	N
	STAT3	p.Gln469*	–	–	–	–	–	13.7	–	–	N
	SMARCA4	p.Phe1234Leu	–	–	–	–	–	13	–	3	N
	NUP93	p.Glu14Lys	–	–	–	–	–	10.4	–	37	N
11	FOXP1	p.Ala645Ala	–	–	–	93.3	79	NA	–	–	Y
	IGF1R	p.Ile601Val	–	66.9	45.2	–	–	NA	–	–	Y
	NF2	p.Trp184*	9	–	14	–	–	NA	Likely_pathogenic	3	Y
	RAF1	p.Ser259Pro	16.5	–	30.7	–	–	NA	Likely_pathogenic	5	Y

Note: The detected somatic mutations of FTCs (clone 1-5) and tumor tissue specimens. The variant allele frequency (VAF) in each sample, the clinical significance of the variants annotated by ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), and the variants' count number reported in the COSMIC database v91 (https://cancer.sanger.ac.uk/cosmic) are indicated. The ASMD (accurate single-cell mutation detector) column shows whether ASMD detected the variant as a somatic mutation in FTCs or not (Y/N).

Abbreviation: NA, not analyzed.

WGA to ensure sufficient coverage of the target regions is required for the reliable molecular profiling of single cells. The ASMD pipeline improved the cover ratio by mutually compensating the regions that frequently dropped out. The concept of an ASMD pipeline is similar to 'census-based variant detection' developed by Zhang et al.[31] There are 2 significant differences in ASMD from that method. First, ASMD uses sequence data from single-cell and bulk leukocytes from the same individual to call somatic mutations. In particular, mutation identified non-paired single-cell leukocytes were used to exclude false-positive mutations, which are errors that arise during WGA. Second, ASMD uses information on allele frequency combined with clinical annotation of variants to accurately call true-positive mutations.

The CTC-Chip was the first specific microfluidic device for CTC isolation. However, it captured many false-positive results from the blood of healthy individuals.[32] Using a filtration technique and fluorescence in situ hybridization, Pailler et al detected *ALK* rearrangements in CTCs from NSCLC patients.[33] However, false-positive results were also obtained due to contamination by normal endothelial cells and rare hematologic cells such as megakaryocytes.[34] Potential limitations of this study are as follows:

(1) The utility of single-cell compared with bulk sequencing of pleural fluid is insufficiently well confirmed in this study. FTCs were concentrated at least 15 times in Case 1 through the sequential enrichment, considering that 114 candidate FTCs were sorted from 1812 captured cells on DEPArray after CD45 negative selection. In contrast, because the ratio of FTCs to blood cells in Case 2 was relatively high (approximately half of the captured cells were CK positive), the bulk analysis without enrichment might have also identified the *ALK* mutations. The actual utility of this enrichment technology depends on the FTC concentration in pleural effusion. The enrichment method must ensure that tumor cells can be efficiently recovered, even from a pleural effusion with a few tumor cells.

(2) The TOP panel basically cannot detect fusion genes as the panel does not include probes for fusion genes. Capture probes targeting the intronic region of fusions should be added to the TOP panel to cover fusion genes to strengthen the clinical utility of this method.

To capture tumor cells efficiently, further optimization with additional immunostaining may be necessary. Lustgarten et al reported that large numbers of epithelial cells were observed in benign effusions, and selection with an antibody to Claudin-4 was required for epithelial cell adhesion molecule-based CTC detection with the CELLSEARCH system.[35]

An advantage of DNA sequencing for single tumor cells over ctDNA assay might be that it is not affected by patient-related factors such as inflammatory conditions and autoimmune disorders that might contribute to the release of cell-free DNA and could be confounding factors in ctDNA assays.[36] Furthermore, because ctDNA is often detected at very low VAF even in patients with advanced cancer,[37-40] sequencing of tumor cells in body fluids might be complementary to ctDNA assays.

It is notable that deep-sequencing of *EGFR* using ctDNA from Case 1 could not detect the *EGFR* exon 19 deletion and that WGS with ctDNA from Case 2 could not identify the supporting reads for *EML4-ALK* fusion (data not shown). A large-scale study comparing the concordance among assays using tissues, ctDNA, FTCs, and CTCs is necessary to establish the clinical validity of molecular profiling for FTCs and CTCs.

Overall, we established an effective enrichment method of tumor cells from pleural effusion specimens using the autoMACS system and DEPArray. The optimized protocol is the first report of *EML4-ALK* fusion detection from a single FTC. Furthermore, the novel ASMD pipeline dramatically improved the accuracy of the detection of mutations of FTCs and might also be applied to CTC analysis of peripheral blood.

We believe that this method expands the possibility of utilizing molecular profiling in primary and metastatic tumors, aids clinicians in their choice of appropriate drugs, and provides patients with more opportunities for personalized medicine.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

ORCID

Hiroyuki Mano (iD) https://orcid.org/0000-0003-4645-0181
Shinji Kohsaka (iD) https://orcid.org/0000-0001-8651-6136

REFERENCES

1. Crowley E, Di Nicolantonio F, Loupakis F, Bardelli A. Liquid biopsy: monitoring cancer-genetics in the blood. *Nat Rev Clin Oncol*. 2013;10:472-484.
2. Haber DA, Velculescu VE. Blood-based analyses of cancer: circulating tumor cells and circulating tumor DNA. *Cancer Discov*. 2014;4:650-661.
3. Rolfo C, Mack PC, Scagliotti GV, et al. Liquid biopsy for advanced non-small cell lung cancer (NSCLC): A statement paper from the IASLC. *J Thorac Oncol*. 2018;13:1248-1268.
4. Zhu WF, Li J, Yu LC, et al. Prognostic value of EpCAM/MUC1 mRNA-positive cells in non-small cell lung cancer patients. *Tumour Biol*. 2014;35:1211-1219.
5. Zhang Z, Shiratsuchi H, Lin J, et al. Expansion of CTCs from early stage lung cancer patients using a microfluidic co-culture model. *Oncotarget*. 2014;5:12383-12397.
6. Fusi A, Metcalf R, Krebs M, Dive C, Blackhall F. Clinical utility of circulating tumour cell detection in non-small-cell lung cancer. *Curr Treat Options Oncol*. 2013;14:610-622.

7. Fuchs AB, Romani A, Freida D, et al. Electronic sorting and recovery of single live cells from microlitre sized samples. *Lab Chip*. 2006;6:121-126.

8. Peeters DJ, De Laere B, Van den Eynden GG, et al. Semiautomated isolation and molecular characterisation of single or highly purified tumour cells from Cell Search enriched blood samples using dielectrophoretic cell sorting. *Br J Cancer*. 2013;108:1358-1367.

9. Siravegna G, Marsoni S, Siena S, Bardelli A. Integrating liquid biopsies into the management of cancer. *Nat Rev Clin Oncol*. 2017;14:531-548.

10. Wu X, Zhu L, Ma PC. Next-Generation Novel Noninvasive Cancer Molecular Diagnostics Platforms Beyond Tissues. *Am Soc Clin Oncol Educ Book*. 2018;38;964-977.

11. Liu X, Lu Y, Zhu G, et al. The diagnostic accuracy of pleural effusion and plasma samples versus tumour tissue for detection of EGFR mutation in patients with advanced non-small cell lung cancer: comparison of methodologies. *J Clin Pathol*. 2013;66:1065-1069.

12. McCoach CE, Blakely CM, Banks KC, et al. Clinical Utility of Cell-Free DNA for the Detection of ALK Fusions and Genomic Mechanisms of ALK Inhibitor Resistance in Non-Small Cell Lung Cancer. *Clin Cancer Res*. 2018;24:2758-2770.

13. Zugazagoitia J, Ramos I, Trigo JM, et al. Clinical utility of plasma-based digital next-generation sequencing in patients with advance-stage lung adenocarcinomas with insufficient tumor samples for tissue genotyping. *Ann Oncol*. 2019;30:290-296.

14. Kohsaka S, Tatsuno K, Ueno T, et al. Comprehensive assay for the molecular profiling of cancer by target enrichment from formalin-fixed paraffin-embedded specimens. *Cancer Sci*. 2019;110:1464-1479.

15. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25:1754-1760.

16. Ignatius Ou SH, Azada M, Hsiang DJ, et al. Next-generation sequencing reveals a Novel NSCLC ALK F1174V mutation and confirms ALK G1202R mutation confers high-level resistance to alectinib (CH5424802/RO5424802) in ALK-rearranged NSCLC patients who progressed on crizotinib. *J Thorac Oncol*. 2014;9:549-553.

17. Lin JJ, Zhu VW, Yoda S, et al. Impact of EML4-ALK Variant on Resistance Mechanisms and Clinical Outcomes in ALK-Positive Lung Cancer. *J Clin Oncol*. 2018;36:1199-1206.

18. Zou HY, Friboulet L, Kodack DP, et al. PF-06463922, an ALK/ROS1 inhibitor, overcomes resistance to first and second generation ALK inhibitors in preclinical models. *Cancer Cell*. 2015;28:70-81.

19. Shaw AT, Solomon BJ, Besse B, et al. ALK resistance mutations and efficacy of Lorlatinib in advanced anaplastic lymphoma kinase-positive non–small-cell lung cancer. *J Clin Oncol*. 2019;37:1370-1379.

20. Maheswaran S, Sequist LV, Nagrath S, et al. Detection of mutations in EGFR in circulating lung-cancer cells. *N Engl J Med*. 2008;359:366-377.

21. Marchetti A, Del Grammastro M, Felicioni L, et al. Assessment of EGFR mutations in circulating tumor cell preparations from NSCLC patients by next generation sequencing: toward a real-time liquid biopsy for treatment. *PLoS One*. 2014;9:e103883.

22. Powell AA, Talasaz AH, Zhang H, et al. Single cell profiling of circulating tumor cells: transcriptional heterogeneity and diversity from breast cancer cell lines. *PLoS One*. 2012;7:e33788.

23. Polzer B, Medoro G, Pasch S, et al. Molecular profiling of single circulating tumor cells with diagnostic intention. *EMBO Mol Med*. 2014;6:1371-1386.

24. Gole J, Gore A, Richards A, et al. Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells. *Nat Biotechnol*. 2013;31:1126-1132.

25. Wang J, Fan HC, Behr B, Quake SR. Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. *Cell*. 2012;150:402-412.

26. Zong C, Lu S, Chapman AR, Xie XS. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science*. 2012;338:1622-1626.

27. Normand E, Qdaisat S, Bi W, et al. Comparison of three whole genome amplification methods for detection of genomic aberrations in single cells. *Prenat Diagn*. 2016;36:823-830.

28. Babayan A, Alawi M, Gormley M, et al. Comparative study of whole genome amplification and next generation sequencing performance of single cancer cells. *Oncotarget*. 2017;8:56066-56080.

29. Borgstrom E, Paterlini M, Mold JE, Frisen J, Lundeberg J. Comparison of whole genome amplification techniques for human single cell exome sequencing. *PLoS One*. 2017;12:e0171566.

30. Huang L, Ma F, Chapman A, Lu S, Xie XS. Single-Cell Whole-Genome Amplification and Sequencing: Methodology and Applications. *Annu Rev Genomics Hum Genet*. 2015;16:79-102.

31. Zhang CZ, Adalsteinsson VA, Francis J, et al. Calibrating genomic and allelic coverage bias in single-cell sequencing. *Nat Commun*. 2015;6:6822.

32. Nagrath S, Sequist LV, Maheswaran S, et al. Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature*. 2007;450:1235-1239.

33. Pailler E, Adam J, Barthélémy A, et al. Detection of circulating tumor cells harboring a unique ALK rearrangement in ALK-positive non–small-cell lung cancer. *J Clin Oncol*. 2013;31:2273-2281.

34. El-Heliebi A, Kroneis T, Zohrer E, et al. Are morphological criteria sufficient for the identification of circulating tumor cells in renal cancer? *J Transl Med*. 2013;11:214.

35. Lustgarten DES, Thompson J, Yu G, et al. Use of circulating tumor cell technology (CELLSEARCH) for the diagnosis of malignant pleural effusions. *Annal Am Thoracic Soc*. 2013;10:582-589.

36. Merker JD, Oxnard GR, Compton C, et al. Circulating Tumor DNA Analysis in Patients With Cancer: American Society of Clinical Oncology and College of American Pathologists Joint Review. *J Clin Oncol*. 2018;36:1631-1641.

37. Newman AM, Bratman SV, To J, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat Med*. 2014;20:548-554.

38. Bettegowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med*. 2014;6:224ra224.

39. Cohen JD, Javed AA, Thoburn C, et al. Combined circulating tumor DNA and protein biomarker-based liquid biopsy for the earlier detection of pancreatic cancers. *Proc Natl Acad Sci U S A*. 2017;114:10202-10207.

40. Cohen JD, Li L, Wang Y, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science*. 2018;359:926-930.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.