이학박사 학위논문

장기간 보관된 낮은 품질의 액체 생검에서도 비소세포폐암 핫스팟 변이의 고감도 검출이 가능한 검사법 개발에 대한 연구

Revolutionizing NSCLC Diagnosis: Ultra-High-Sensitive ctDNA Analysis for detecting hotspot mutations with long-term stored Plasma

울 산 대 학 교 대 학 원

의 과 학 과

이 지 영

장기간 보관된 낮은 품질의 액체 생검에서도 비소세포폐암 핫스팟 변이의 고감도 검출이 가능한 검사법 개발에 대한 연구

지 도 교 수　　장 세 진 , 천 성 민

이 논문을 이학박사 학위 논문으로 제출함

2023 년　8 월

울 산 대 학 교 대 학 원

의　과　학　과

이　지　영

이지영의 이학박사학위 논문을 인준함

심사위원장	성 창 옥	(인)
심사위원	김 지 선	(인)
심사위원	김 지 훈	(인)
심사위원	오 용 택	(인)
심사위원	천 성 민	(인)

울 산 대 학 교 대 학 원

2023 년 8 월

감사의 글

어떻게 생각해 보면 학위과정이 PCR 과정과도 같았다는 생각이 문득 들었습니다. 정말 보잘 것 없는 제 자신이라는 Template에서 의미 있는 무언가를 찾아내기 위한 여정이, Template 하나로만 수행될 수 없는 PCR 과정과도 같았음을 비로소 깨닫습니다. 제 삶의 PCR 과정에 다양한 형태로 도움을 주신 분들께 이 자리를 빌려 감사의 인사를 드립니다.

먼저, 제 연구실 생활에 있어 Taq polymerase, MSP와 같은 역할로 지도해 주신 장세진 교수님, 천성민 교수님께 고개 숙여 감사드립니다. Low-quality template이자 만학도인 저를 받아들여 주시고, 인내해 주신 덕분에 이만큼 올 수 있었습니다. 제 인생 3분의 1을 교수님들과 함께 하면서 교수님들은 늘 저에게 성실함을 보여주셨습니다. 저 또한 그 스승에 그 제자답게 늘 성실하고 노력하는 사람이 되도록 하겠습니다.

그리고 바쁘신 시간을 할애하여 논문을 심사해 주신 김지선 교수님, 김지훈 교수님, 성창옥 교수님, 오용택 부장님께도 감사의 인사를 드립니다. 각자의 자리에서 최선을 다하시는 교수님들과 부장님을 본받아 제게 주어진 것에 감사하며 열심히 연구에 정진하겠습니다.

실험실 동무이자 동지인 여러 선생님들께도 감사드립니다. 늘 열과 성의를 다해 배우고자 하시는 우리 선생님들이 있었기에 저 또한 그에 발맞춰 따라가고자 노력할 수 있었던 것 같습니다. 앞으로 연구를 할 때마다 선생님들을 떠올리며 초심을 잃지 않도록 하겠습니다.

가끔은 삶이 힘들고 답답할 때, 많은 것을 함께 나누며 진심으로 위로해 준 제이오(第二吾)라 칭하기도 미안한 한지원, 한지현 친구들에게도 매우 고맙다는 말을 전하고 싶습니다. 우리 친구들처럼 상대를 헤아릴 줄 알고, 잘 나누는 사람이 되도록 하겠습니다.

마지막으로, 이 세상에서 아무 조건 없이 저를 사랑해 주시고 항상 제 편이 되어 주시는 부모님께 감사드립니다. 부모님의 희생과 사랑 덕분에 이 과정을 잘 마무리할 수 있었습니다. 받은 것을 나눌 줄 아는 사람이 되겠습니다.

일일이 말씀드리지는 못했지만 저를 응원해 주신 많은 분들께 감사드리며 더 나은 연구자가 되도록 노력하겠습니다. 감사합니다.

Abstract

Background: Circulating cell-free DNA (cfDNA) has enormous potential in clinical applications, particularly in cancer patients. The prognostic and predictive values of cfDNA in non-small cell lung cancer (NSCLC) have been previously reported, with *EGFR*, *KRAS*, and *BRAF* mutations in tumor-derived cfDNAs acting as biomarkers for the early stages of tumor progression and recurrence. However, effective applications of cfDNA require a reliable approach to detect low mutant fractions of <1%, considering the limitations of extremely low tumor-derived DNA proportions. Herein, we developed an ultra-high sensitivity lung version 1 (ULV1) panel targeting *EGFR*, *KRAS*, and *BRAF* hotspot mutations and evaluated the detection ability of hotspot mutations using low-quality plasma samples from NSCLC patients.

Methods: The ULV1 panel comprised two multiplexing pools capable of detecting 14 hotspots. We utilized positive samples corresponding to each mutation to evaluate the optimization and performance of the ULV1 panel and samples from 104 NSCLC patients with stages I–IV to verify the potential of the ULV1 panel as a high-sensitivity diagnostic tool. The *EGFR* mutation in tissue using targeted NGS or diverse diagnostic platforms alone was tested at the Asan Medical Center, Seoul, South Korea. We compared the mutational concordance in tumor tissue samples with matched plasma samples and evaluated the performance between ULV1 and CT-ULTRA to detect *EGFR* mutation status in plasma cfDNA.

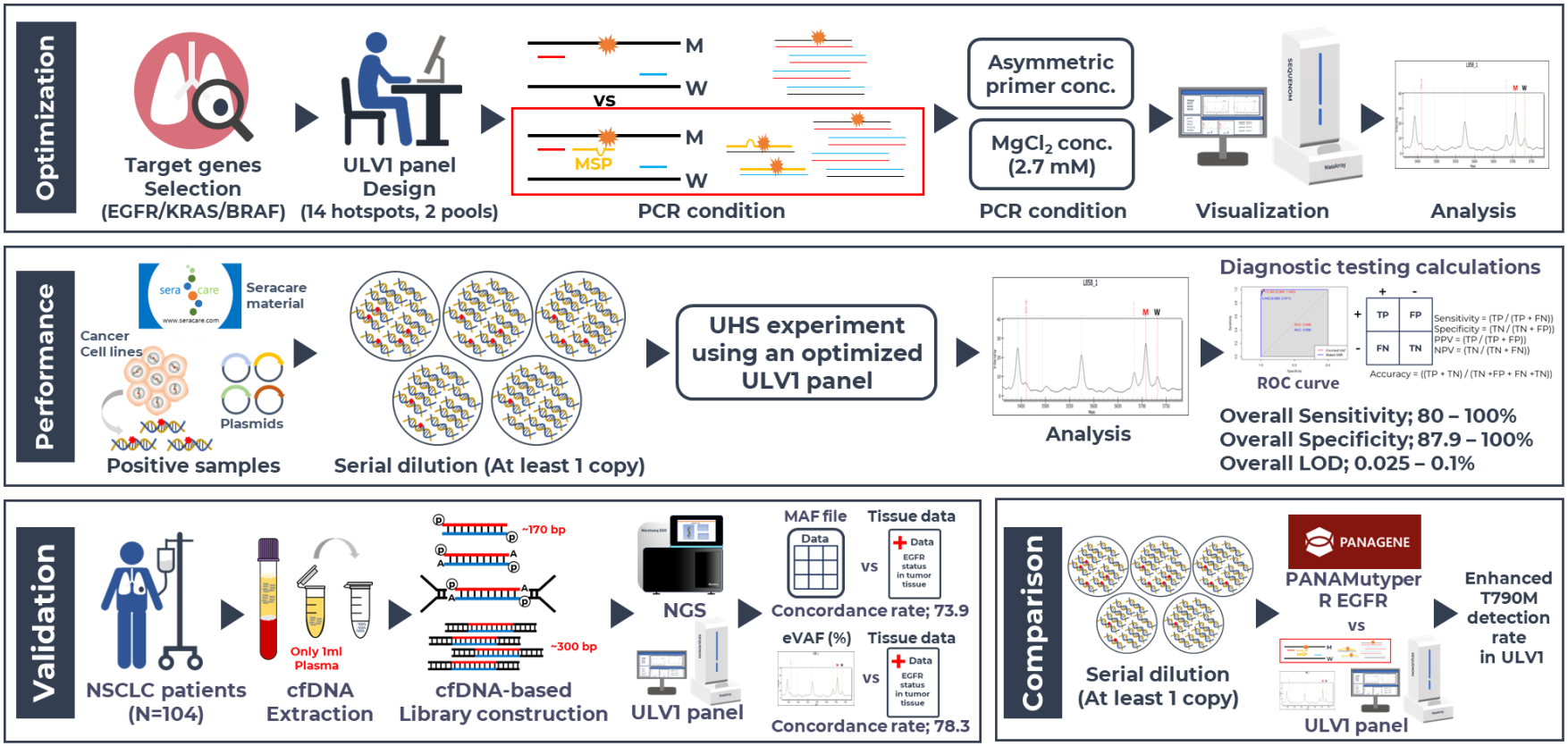
Results: The performance of the ULV1 panel using positive samples demonstrated overall sensitivity (80–100%) and specificity (87.9–100%) with a limit of detection of 0.025–0.1%. Focusing on *EGFR* hotspot mutations included in the ULV1 panel in stage IV, we identified 23 hotspot mutations. Of these, 18 (78.3%) and 17 (73.9%) mutations demonstrated the same mutation pattern as the matched tumor tissue DNA in cfDNA samples using ULV1 and CT-ULTRA, respectively. Interestingly, additional mutations were identified in only three of 80 (3.75%) cfDNAs using ULV1 panel analysis. Considering the limited volume (1 mL) and long-term storage (12–50 months) before cfDNA extraction, the impressive detectability of ULV1

panel is expected to increase with the cfDNA input amount.

Conclusion: The results of our study demonstrate that the ULV1 panel has clinical potential as a reliable detection tool for identifying genetic mutations in low-quality plasma samples from NSCLC patients. This not only enables the monitoring of treatment responsiveness but also validates NGS data.

Keywords: Circulating cell-free DNA; Non-small cell lung cancer; *EGFR*; low-quality plasma; Ultra-high sensitivity lung version 1; MassARRAY

Graphical Abstract



Contents

List of Figures

List of Tables

Introduction

Assessing the targetable driver and treatment-resistance mutations in non-small cell lung cancer (NSCLC) can help guide therapeutic decision-making[1–4], including the real-time monitoring of treatment responses and disease progression in patients[5]. Moreover, it can aid in verifying the absence or presence of minimal residual disease (MRD) post tumor resection surgery[6]. Therefore, accurately identifying *EGFR* mutations in NSCLC patients is the first step in treatment decision-making and can be utilized as a critical patient health indicator. Within the context of detecting oncogenic alterations, tumor tissue is considered the gold standard method owing to its reliable mutation information. However, not only are the corresponding procedures invasive but it is also challenging to collect tumor tissue from locations that are inaccessible via biopsy[7]. Moreover, tissue biopsy does not reflect intra-tumor heterogeneity, resulting in resistance to effective targeted therapy[8].

Recently, liquid biopsies have emerged as an alternative to tissue biopsies, with advantages including minimally invasive procedures, low complications risk, and better tumor heterogeneity representation[9]. Liquid biopsies can target circulating cell-free DNA (cfDNA), which originates from the plasma or serum and is released into the bloodstream from dying cells through apoptosis and necrosis[10]. Notably, circulating tumor DNA (ctDNA) released from cancer patients can be utilized for real-time monitoring of DNA mutations and tumor burden and can provide useful information for predicting the condition of cancer patients[11,12]. However, using cfDNA for disease diagnosis and prognosis has several limitations. The excess amount of wild DNA derived from hematopoietic and normal cells can interfere with tumor-derived DNA detection, making it difficult to identify mutant DNA in cancer diagnostic tests. Next, cfDNA concentration is associated with clinical features, including disease or cancer types[13], tumor stages[14], tumor burden[15], and response to treatment[16]. Further, it has been reported that the fraction of ctDNA in total cfDNA can vary markedly from 0.1–90%[17]. Thus, clinical samples with low variant allelic frequencies (VAFs) as low as ≤1% are unsuitable for

detecting mutant alleles using methods with a low detection limit, such as low coverage and Sanger sequencing.

Several studies have reported a variety of PCR-based technologies, including peptide nuclei acid-locked nucleic acid (PNA-LNA) PCR clamp[18], droplet digital PCR[19], and allele-specific amplification (AS-PCR)[20], to detect somatic mutations in cfDNA. Although these techniques are robust in detecting low-level mutations, the complicated workflow, high-cost analysis, and limitations of multiplexing potential present notable challenges.

In our previous work, we developed an improved approach, the Onco Ultra-High Sensitivity (OncoUHS) assay, based on a combination of the MassARRAY platform and the amplification-refractory mutation system (ARMS)-PCR. By selectively amplifying shorter mutant-specific amplicons through the combination of a common outer primer set and additional mutant allele-specific primer (MSP), the resulting amplified products demonstrated a substantial increase in the mutant allele proportions compared to the original mutant allele frequencies observed in the sample. The OncoUHS assay was previously used to detect hotspot mutations in colorectal cancer patients using plasma samples with extremely low variant allele fractions[21].

In this study, a new panel was designed, namely the ultra-high sensitivity lung version 1 (ULV1) panel, which comprised frequently mutated genes in NSCLC patients. *EGFR* mutations, which play a critical role in targeted therapy (*EGFR*-TKI), are present in approximately 40–50% of NSCLC cases in Asian populations[22], predominantly occurring in the tyrosine kinase domain of the receptor (exon 18–21). Furthermore, *KRAS* and *BRAF*, which are involved in downstream *EGFR* signaling, are mutated in lung cancer, with *KRAS* being associated with *EGFR*-TKI resistance[23]. Using the ULV1 panel, we evaluated its performance in cfDNA analysis from 104 NSCLC patients. Our findings suggest that the ULV1 panel is a simple, rapid, highly sensitive, and cost-efficient approach for detecting somatic mutations in NSCLC patients.

Material and Methods

Preparation of positive controls and plasma samples

For evaluating the analytical performance of the ULV1 panel, we used the ctDNA Complete Mutation mix AF5% (#0710-0528; SeraCare, Massachusetts, USA), seven cell lines (AsPC1, H1975, HCC827, HCT-15, HT-29, SW900, and PC9), and four synthetic plasmids were used as positive controls. Also, HL-60 cell lines expressing wild-type *BRAF*, *KRAS*, and *EGFR* were used as a negative control. Cell line-derived gDNA was extracted using a NEXprep FFPE tissue kit (#NexK-9000; Genes Laboratories Inc, Gyeonggi-do, Republic of Korea), following the manufacturer's instructions, excluding the deparaffinization step. Extracted gDNAs were quantified using the Qubit™ dsDNA HS Assay kit (Thermo Fisher Scientific, Waltham, MA, USA) and stored them at −20°C until further use.

The plasma cohort consisted of 104 samples collected from patients with NSCLC at Asan Medical Center between 2011 and 2018, representing various stages. cfDNA was extracted from only 1 mL of plasma using an STB cell-free DNA kit (Syntekabio, Seoul, Republic of Korea), following the manufacturer's instructions. Extracted cfDNA was qualified and quantified using Bioanalyzer High Sensitivity DNA Analysis (Agilent, Santa Clara, USA) and stored at −80°C until further use. We used data on *EGFR* mutations from matched tissue samples obtained from the same patients, which were previously analyzed using targeted NGS assay or different diagnostic platforms at Asan Medical Center, to assess the concordance rate of *EGFR* mutations detected in cfDNA with the ULV1 panel or targeted NGS **(Figure 1)**. The protocol of this study was approved by the Ethics Committee at Asan Medical Center (approval no. 2016-0692).

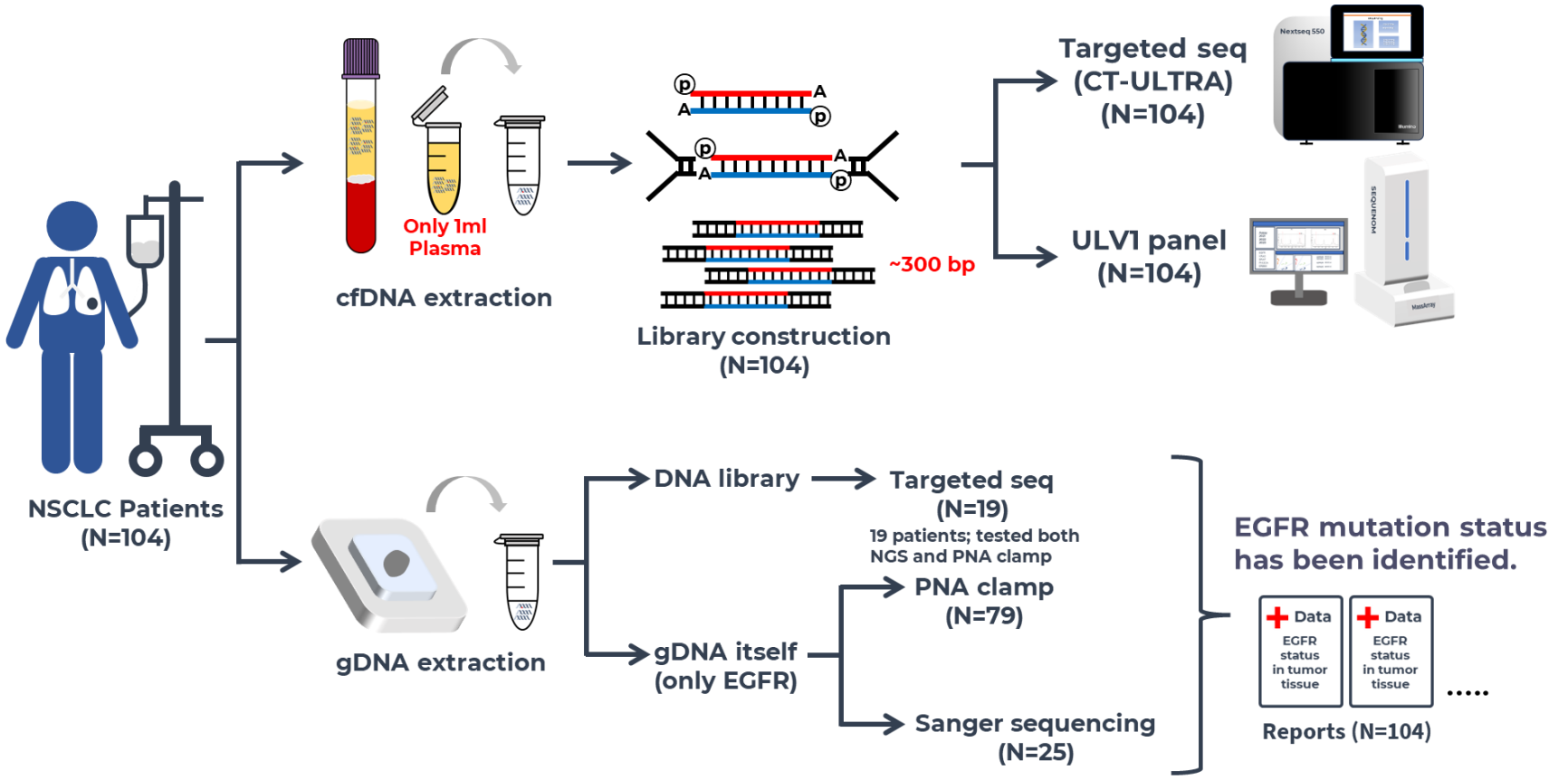


Figure 1. Schematic of the experimental procedure for identifying mutational concordance between the tumor tissue and plasma cfDNA.

Nineteen patients were tested for both targeted sequencing and PNA clamp in tissue samples.

cfDNA library construction

To address the issue of limited DNA quantity, both the ctDNA Complete Mutation mix and cfDNA samples from 104 patients were transformed into DNA libraries. The construction of DNA libraries with various input cfDNA amounts ranging from 1.9 - 10 ng and 10 ng of ctDNA.

Complete Mutation mix AF5% was performed using a SureSelectXT Reagent kit (Agilent Technologies, Santa Clara, CA, USA). The process involved a series of steps that included end repair, A-tailing, ligation with a TruSeq adapter, and enrichment of adapter-ligated libraries, which were used as templates for targeted NGS and ULV1 experiments. **Table 1** listed the library yield according to the input cfDNA amount from 104 patients.

Positive sample dilution

Each mutant allele positive DNA was serially diluted with HL-60 (wild-type control) genomic DNA to generate the following mutant-to-wild-type allele fractions: 100%, 10%, 5%, 1%, 0.5%, 0.25%, 0.1%, 0.05%, 0.025%, and 0%. Next, these DNA series were used as templates to determine the detection limits of each assay and to compare the performance of the different methods.

Table 1. cfDNA library yield according to the input cfDNA amount from 104 patients

Patient ID	Stage	Extracted cfDNA yield (ng / 1mL plasma)	cfDNA input, ng (for library)	Total library yield, ng
AMC-001	I	6.0	5.5	585
AMC-002	I	6.1	5.6	822
AMC-003	I	3.6	3.3	516
AMC-004	I	5.3	4.9	771
AMC-005	I	9.8	9.1	951
AMC-006	I	5.9	5.5	792
AMC-007	I	4.5	4.2	657
AMC-008	I	8.9	8.3	660
AMC-009	I	7.6	7.1	540
AMC-010	I	8.0	7.5	669
AMC-011	I	7.5	7.1	672
AMC-012	I	2.1	2.0	271
AMC-013	I	10.1	9.5	861
AMC-014	I	5.5	5.1	669
AMC-015	I	4.3	4.0	639
AMC-016	I	5.3	4.9	651
AMC-017	I	6.2	5.7	738
AMC-018	I	169.0	10.1	666
AMC-019	I	5.6	5.3	651
AMC-020	I	6.5	6.1	816
AMC-021	I	9.8	9.2	819
AMC-022	I	6.7	6.3	450
AMC-023	I	4.1	3.8	663
AMC-024	II	9.0	8.3	924
AMC-025	II	11.0	10.0	915
AMC-026	II	3.9	3.7	480

Patient ID	Stage	Extracted cfDNA yield (ng / 1mL plasma)	cfDNA input, ng (for library)	Total library yield, ng
AMC-027	II	9.3	8.7	894
AMC-028	II	14.3	10.0	690
AMC-029	II	18.7	10.0	825
AMC-030	II	4.2	3.9	492
AMC-031	II	9.0	8.4	891
AMC-032	II	5.2	4.8	963
AMC-033	II	10.8	10.0	894
AMC-034	II	4.0	3.7	636
AMC-035	II	18.1	10.0	978
AMC-036	II	2.5	2.4	414
AMC-037	III	4.8	4.5	654
AMC-038	III	2.1	2.0	272
AMC-039	III	14.4	10.0	1020
AMC-040	III	9.6	8.9	696
AMC-041	III	7.8	7.3	669
AMC-042	III	2.5	2.3	408
AMC-043	III	6.8	6.3	744
AMC-044	III	7.0	6.5	816
AMC-045	III	7.2	6.6	645
AMC-046	III	9.2	8.5	783
AMC-047	III	13.6	10.0	648
AMC-048	III	5.6	5.3	669
AMC-049	III	4.4	4.1	630
AMC-050	III	4.9	4.6	462
AMC-051	IV	2.0	1.9	408
AMC-052	IV	4.2	3.9	429
AMC-053	IV	2.2	2.0	306

Patient ID	Stage	Extracted cfDNA yield (ng / 1mL plasma)	cfDNA input, ng (for library)	Total library yield, ng
AMC-054	IV	9.4	8.7	1158
AMC-055	IV	5.8	5.4	735
AMC-056	IV	5.1	4.7	756
AMC-057	IV	3.9	3.6	666
AMC-058	IV	12.0	10.0	1143
AMC-059	IV	6.2	5.7	822
AMC-060	IV	9.5	8.8	1269
AMC-061	IV	6.2	5.7	846
AMC-062	IV	9.9	9.1	792
AMC-063	IV	16.8	15.5	1281
AMC-064	IV	6.4	5.9	834
AMC-065	IV	5.4	5.0	570
AMC-066	IV	5.7	5.3	606
AMC-067	IV	3.4	3.1	531
AMC-068	IV	4.2	3.8	405
AMC-069	IV	4.6	4.3	471
AMC-070	IV	11.0	10.0	831
AMC-071	IV	8.2	7.6	900
AMC-072	IV	30.4	10.0	492
AMC-073	IV	20.9	10.0	1101
AMC-074	IV	7.4	6.9	714
AMC-075	IV	36.3	10.0	1008
AMC-076	IV	15.6	10.0	972
AMC-077	IV	6.0	5.6	708
AMC-078	IV	16.5	10.0	828
AMC-079	IV	11.7	10.0	882
AMC-080	IV	10.4	9.8	1014

Patient ID	Stage	Extracted cfDNA yield (ng / 1mL plasma)	cfDNA input, ng (for library)	Total library yield, ng
AMC-081	IV	5.5	5.1	600
AMC-082	IV	13.6	10.0	1152
AMC-083	IV	17.1	10.0	930
AMC-084	IV	17.0	10.0	822
AMC-085	IV	15.1	10.0	990
AMC-086	IV	10.3	9.6	1017
AMC-087	IV	4.9	4.6	690
AMC-088	IV	6.6	6.2	702
AMC-089	IV	95.3	10.0	819
AMC-090	IV	13.0	10.0	816
AMC-091	IV	6.9	6.4	537
AMC-092	IV	8.6	8.1	453
AMC-093	IV	6.9	6.5	927
AMC-094	IV	6.3	5.9	696
AMC-095	IV	19.9	10.0	804
AMC-096	IV	10.9	10.2	906
AMC-097	IV	16.1	10.0	954
AMC-098	IV	13.3	10.0	951
AMC-099	IV	8.2	7.7	924
AMC-100	IV	7.4	6.9	654
AMC-101	IV	33.3	10.0	858
AMC-102	IV	11.0	10.3	1290
AMC-103	IV	7.1	6.6	909
AMC-104	IV	4.9	4.6	570

Analysis of hotspot mutations using the ULV1 panel

The ULV1 panel was designed to detect 14 hotspot mutations of *BRAF, KRAS*, and *EGFR* genes that are common in patients with NSCLC. This was achieved via a two-pool multiplexing method based on the ultrahigh sensitive (UHS) assay, combining additional MSP and common outer primer sets. The outer PCR primer sets and single base extension primers were designed using the Assay Designer of MassARRAY Typer 4.0 software (Agena Bioscience), following the manufacturer's instructions. MSPs containing mismatched bases between the first and fourth base positions from the 3'-termini of primer were manually designed in the opposite direction to the single base extension primer for each mutant locus. As cfDNA is a template, the panel was designed to generate short amplicons of 51 to 129 bp. The primer sequences used in this study are listed in **Table 2**.

PCR amplification was performed using 34.6 ng of input DNA (equivalent to human DNA 10,000 copies), 0.5 U of Taq polymerase (Qiagen, Germany), 1.25× PCR buffer (including 1.875 mM $MgCl_2$), 0.815 mM $MgCl_2$, 500 μM deoxynucleotide triphosphates, and an optimal concentration of primers per pool. The following PCR program was used for amplification: 94°C for 15 min, followed by 45 cycles of (94°C for 20 s, 60°C for 30 s, and 72°C for 1 min), and a final extension step at 72°C for 3 min. After multiplex PCR amplification, residual deoxynucleotides were inactivated by treatment with shrimp alkaline phosphatase (Agenabio, USA) at 37°C for 40 min and 85°C for 5 min. Next, single-base extension reactions were performed in a total reaction volume of 9 μL, containing 0.222× iPLEX buffer plus, 0.5× iPLEX termination mix, extension primer mix (1.86~2.67 μM), and 0.5× iPLEX enzyme (Thermo Sequenase) using the following nested thermocycler programs: 94°C for 30 s, followed by 40 cycles of 94°C for 5 s, 52°C for 5 s, and 80°C for 5 s. The annealing and extension steps were repeated five times within the 40 cycles program (i.e., 40 × 5 = 200 short cycles), before a final extension step of 3 min at 72°C. After spotting the desalted product onto a 384-format SpectroCHIP II, the spectrum profiles generated by matrix-assisted laser

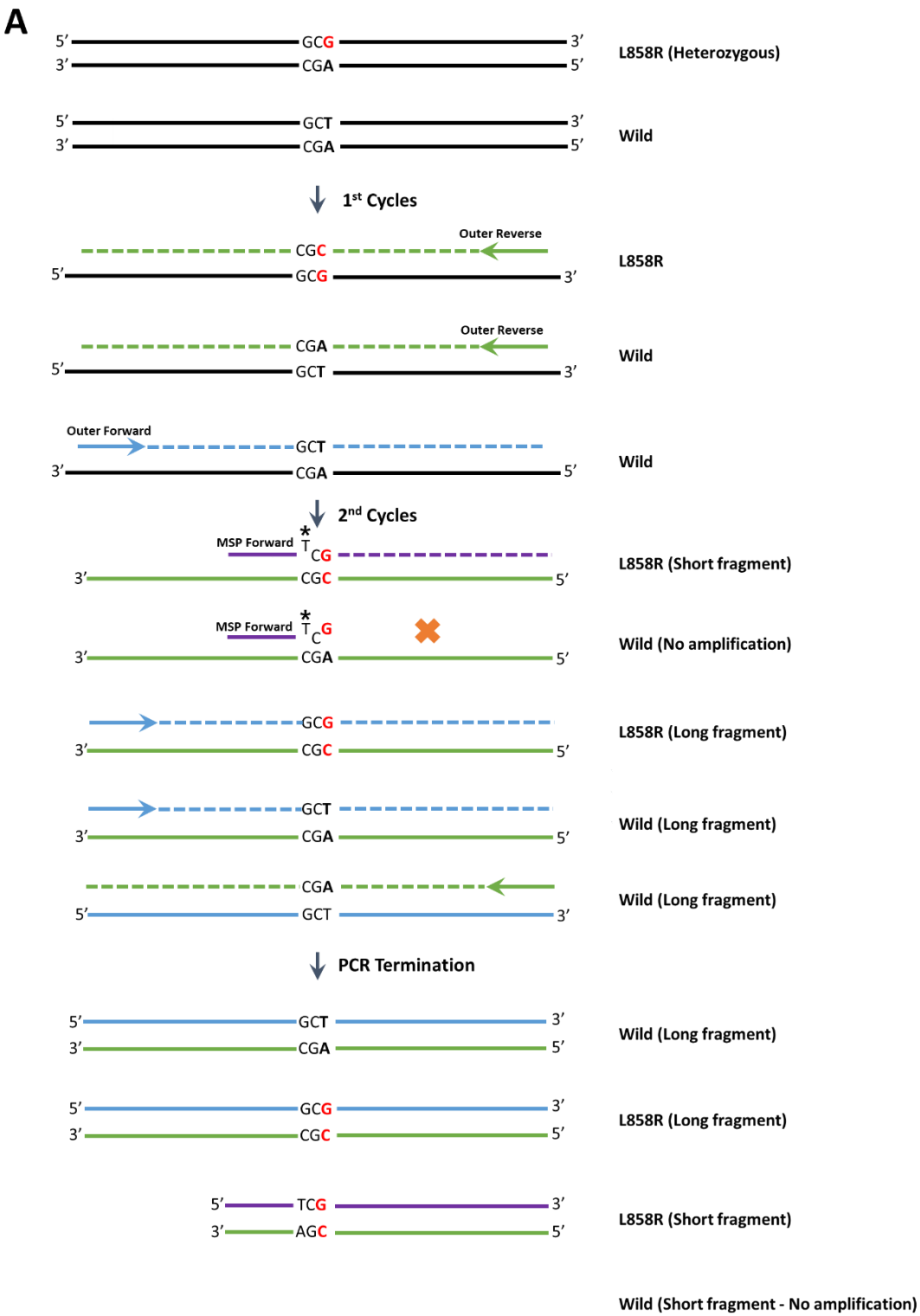
desorption/ionization time-of-flight mass spectrometry were acquired and interpreted using Typer 4.0 software. The experimental process using the ULV1 panel is presented in **Figure 2**.

Table 2. Primer sequences of the ULV1 panel

ULV1 Pool	Assay	Forward (5' → 3')	Reverse (5' → 3')	MSP (5' → 3')[a]	UEP (5' → 3')
Pool 1	***EGFR* E709K**	ACGTTGGATG AGGGACCTTACCTTATACACC	ACGTTGGATG AGTGGAGAAGCTCCCAACCA	ACGTTGGATG GCTCTCTTGAGGATCTTGA**g**GA	TTGATCTTTTTGAATT CAGTTT
	***EGFR* G719A**	ACGTTGGATG AGGGACCTTACCTTATACACC	ACGTTGGATG AGTGGAGAAGCTCCCAACCA	AATAAATCATAA CGTGCCGAACGCACCGG**g**GG	cAAAAGATCAAAGTG CTGG
	***EGFR* G719D**	ACGTTGGATG AGGGACCTTACCTTATACACC	ACGTTGGATG AGTGGAGAAGCTCCCAACCA	AATAAATCATAA CGTGCCGAACGCACCG**ag**GT	cAAAAGATCAAAGTG CTGG
	***EGFR* E746_A750delinsK**	ACGTTGGATG AGCAGAAACTCACATCGAGG	ACGTTGGATG GATCCCAGAAGGTGAGAAAG	AATAAATCATAA CCTTGTTGGCTTTCGGAGATG**c**TT	AAAATTCCCGTCGCT ATCAA
	***EGFR* L747_P753>S**	ACGTTGGATG AGCAGAAACTCACATCGAGG	ACGTTGGATG GATCCCAGAAGGTGAGAAAG	ACGTTGGATG CCCGTCGCTATCAAGGA**t**TC	GATTTCCTTGTTGGC TTTCG
	***EGFR* C797S**	ACGTTGGATG TACTGGGAGCCAATATTGTC	ACGTTGGATG CAGCTCATCACGCAGCTCAT	ACGTTGGATG TCACGCAGCTCATGCCCTTCG**t**CA	GGACATAGTCCAGGA GGC
	***KRAS* G13D**	ACGTTGGATG TAGCTGTATCGTCAAGGCAC	ACGTTGGATG TAAGGCCTGCTGAAAATGAC	ACGTTGGATG TGTGGTAGTTGGAGCTGG**g**GA	AGGCACTCTTGCCTA CG

ULV1 Pool	Assay	Forward (5' → 3')	Reverse (5' → 3')	MSP (5' → 3')[a]	UEP (5' → 3')
	***KRAS* G13V**	ACGTTGGATG TAGCTGTATCGTCAAGGCAC	ACGTTGGATG TAAGGCCTGCTGAAAATGAC	ACGTTGGATGTGTGGTAGTTGGAGCT GG**g**GT	AGGCACTCTTGCCTA CG
Pool 2	***EGFR* E746_A750del**	ACGTTGGATG TCGAGGATTTCCTTGTTGGC	ACGTTGGATG GATCCCAGAAGGTGAGAAAG	ACGTTGGATG AATTCCCGTCGCTATCA**c**GA	TTGGCTTTCGGAGAT GT
	***EGFR* T790M**	ACGTTGGATG ATCTGCCTCACCTCCACCGT	ACGTTGGATG TGTTCCCGGACATAGTCCAG	AGCCGAAGGGCATGAGCTG**a**A	CACCGTGCAGCTCAT CA
	***EGFR* L858R**	ACGTTGGATG AGCCAGGAACGTACTGGTGA	ACGTTGGATG AAAGCCACCTCCTTACTTTGC	ACGTTGGATG GTTCAAGATCACAGATTTTGG**t**CG	GCACCCAGCAGTTTG GCC
	***BRAF* V600E**	ACGTTGGATG TTCATGAAGACCTCACAGTAAAAA	ACGTTGGATG AGCCTCAATTCTTACCATCCA	CT GTGATTTTGGTCTAGCTAC**g**GA	CCCACTCCATCGAGA TTTC
	***KRAS* G12D**	ACGTTGGATG TTTATTATAAGGCCTGCTGAAAATG	ACGTTGGATG ATTGTTGGATCATATTCGTCCAC	ACTTGTGGTAGTTGGAGC**g**GA	CACTCTTGCCTACGC CA
	***KRAS* G12V**	ACGTTGGATG TTTATTATAAGGCCTGCTGAAAATG	ACGTTGGATG ATTGTTGGATCATATTCGTCCAC	ACTTGTGGTAGTTGGAGC**a**GT	CACTCTTGCCTACGC CA

MSP, Mutant-specific primer; UEP, unextended primer, [a] The lower case in the MSP sequence indicates 3' terminal mismatch.


A
5′ GCG 3′
3′ CGA 5′
L858R (Heterozygous)
5′ GCT 3′
3′ CGA 5′
Wild
1st Cycles
Outer Reverse
CGC
5′ GCG 3′
L858R
Outer Reverse
CGA
5′ GCT 3′
Wild
Outer Forward
GCT
3′ CGA 5′
Wild
2nd Cycles
MSP Forward
* T
CG
3′ CGC 5′
L858R (Short fragment)
MSP Forward
* T
C G
3′ CGA 5′
Wild (No amplification)
GCG
3′ CGC 5′
L858R (Long fragment)
GCT
3′ CGA 5′
Wild (Long fragment)
CGA
5′ GCT 3′
Wild (Long fragment)
PCR Termination
5′ GCT 3′
3′ CGA 5′
Wild (Long fragment)
5′ GCG 3′
3′ CGC 5′
L858R (Long fragment)
5′ TCG 3′
3′ AGC 5′
L858R (Short fragment)
Wild (Short fragment - No amplification)

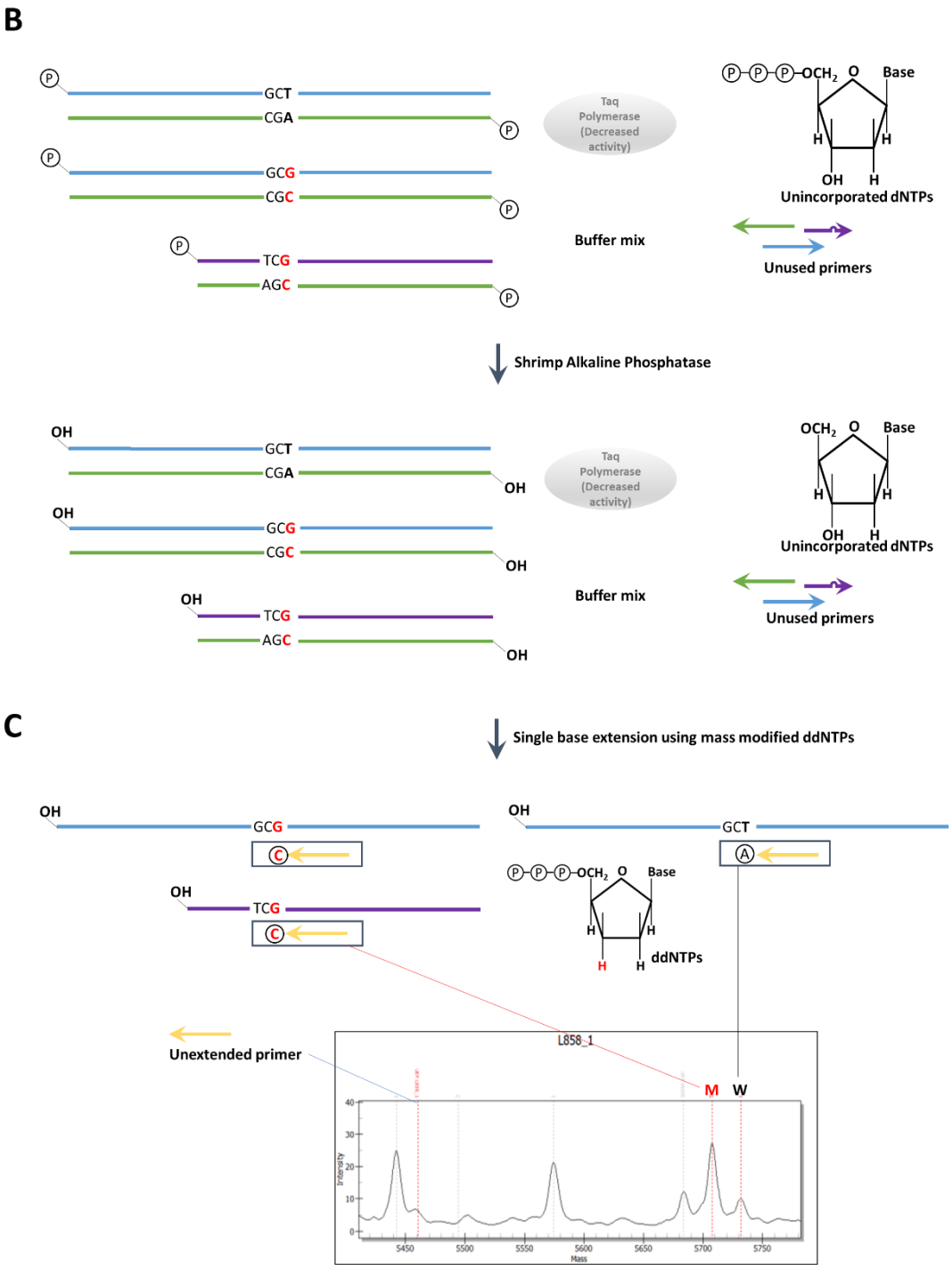


Figure 2. Schematic of PCR amplification steps for the ULV1 panel. A, The process of initial PCR and generated PCR products. Asterisk (*) indicates MSP mismatch. B, Treatment of SAP enzyme to remove any unincorporated dNTPs. C, A single base extension reaction by dideoxynucleotides at 3′ ends of the unextended primers.

Targeted NGS analysis of cfDNA using CT-ULTRA panel

To evaluate the somatic mutational profile of each cfDNA, NGS analysis was performed using the CT-ULTRA panel, which was developed specifically for specimens with low variant allele fractions. Each library was denoted with a sample-specific unique sequencing barcode (6 bp) and quantified using the Qubit™ dsDNA HS Assay kit before pooling 12-20 libraries (yielding a total of 750 ng) for target capture using the Agilent SureSelectXT custom kit (CT-ULTRA, RNA bait, 0.26 Mb; Agilent Technologies).

A CT-ULTRA panel was designed to target a total 118 genes, including the entire exons of 88 genes, the partial introns of four genes that are often rearranged in solid cancer, and an additional small-sized (9,412 bp) specific SNP loci for genetic fingerprinting **(Table 3)**. The captured libraries were enriched by limited PCR (10 cycles), followed by measurement using the Qubit™ kit. DNA libraries that passed the quality checks were then sequenced using the NextSeq 550 platform (Illumina, San Diego, CA, USA) in paired-end mode. Sequenced reads were aligned to the human reference genome (NCBI build 37) using Burrows-Wheeler Aligner (0.5.9) in default mode, and PCR de-duplication was performed using a Picard's MarkDuplicates package. After the initial alignment process, de-duplicated reads were re-aligned at known indel positions with the GATK4 BaseRecalibrator tool (version 4.1.3.0). Then, the base quality was recalibrated using the GATK4 ApplyBQSR tool and used as final BAM for variant calling.

The Mutect2 tool was used for the somatic variant calling of single nucleotide variants (SNV) and short indels. Germline variants from the somatic variant candidates were filtered out using the Single Nucleotide Polymorphism database (dbSNP, build 141; found in >1% of samples), Exome Aggregation Consortium database (ExAC; r0.3.1, threshold frequency 0.001), Korean Reference Genome database (KRGDB), and an in-house panel of normal controls. After additional filtering using GATK4 FilterMutectCalls tools, final somatic variants were annotated using the Variant Effect Predictor (version 86), which were then converted to the

Mutation Annotation Format (MAF) file format using vcf2maf. The manual curation of SNV and indel alterations was performed carefully using the Integrative Genomics Viewer (IGV).

Table 3. Gene list of CT-ULTRA

CT-ULTRA panel includes the complete exonic sequence of 88 genes, introns of 4 genes, and 31 hotspots.
Gene List: Entire Exonic Sequence for Detection of Base Substitution, Insertions/Deletions, and Copy Number Alterations *AKT1, ALK, APC, AR, ARAF, ARID1A, ATM, AURKA, AXIN1, BCL2, BRAF, BRCA1, BRCA2, CCND1, CCND2, CCNE1, CD274, CDH1, CDK4, CDK6, CDKN2A, CRKL, CTNNB1, DDR2, EGFR, ERBB2, ERG, ESR1, EZH2, FBXW7, FGFR1, FGFR2, FGFR3, FLT3, GATA3, GNA11, GNAQ, GNAS, HNF1A, HRAS, IDH1, IDH2, JAK2, JAK3, KIT, KRAS, MAP2K1, MAP2K2, MAPK1, MAPK3, MDM2, MET, MLH1, MPL, MTOR, MYC, MYCN, NF1, NFE2L2, NOTCH1, NPM1, NRAS, NTRK1, NTRK3, PDGFRA, PIK3CA, PTEN, PTPN11, RAF1, RB1, RET, RHEB, RIT1, ROS1, RSPO2, RUNX1, SMAD2, SMAD4, SMO, SRC, STK11, TCF19, TERT, TMPRSS2, TOP1, TP53, TSC1, VHL*
Gene List: Partial Intronic Sequence for Detection of Rearrangements *ALK, NTRK1, RET, ROS1*
Gene List: Hot spot for Detection of Base Substitution, Insertions/Deletions *ACVR2A, ADNP, AK9, BICC1, BTK, CLOCK, CROT, DCAKD, DOCK1, EYS, FRG2B, IL10RB, ITGA9, LINC00299, NBAS, PAH, PLCL2, PPPSR3B, PRDM2, RAD51B, RBBP8, RGS17, RHOA, RIPK3, SEMA6D, SLC23A2, SYN3, TEAD2, TERT, U2AF1, ZNRF3*

LOD comparison with PANAMutyper test for three hotspot mutations

The PANAMutyper R *EGFR* kit (Panagene, Daejeon, Korea) was used to verify that ULV1 had comparable performance to other platforms in terms of *EGFR* hotspot mutations, based on real-time PCR analysis. *EGFR* assays were performed according to the manufacturer's instructions. For PCR amplification, 5 μL of DNA template, including 34.6 ng of DNA, was added to 19 μL of each master mix and 1 μL of Taq DNA polymerase. PCR reactions were performed using the CFX96 real-time PCR detection system (Bio-Rad, USA) following the thermal cycling program. Amplification and melting curves analysis for each fluorescent dye were generated, and the genotype of each sample was determined based on each assay threshold and the melting temperature range. Sample analysis was repeated at least three times, and ULV1 tests were conducted simultaneously using the same samples for direct comparison with PANAMutyper. Cohen's kappa value was used to assess the agreement between ULV1 and PANAMutyper.

Statistical analysis

Raw data which was produced by the MassARRAY system was further analyzed by assessing the signal-to-noise ratio (SNR) value for extension analytes. To clearly distinguish between mutant samples, the enriched VAF was calculated using the following equation:

Enriched VAF (%) = (Mutant peak SNR / (Mutant peak SNR + Wild-type peak SNR)) × 100.

To evaluate characteristics of the diagnostic potential, the receiver operating characteristic (ROC) curve and the area under the ROC curve were determined for the highest sensitivity, specificity, and cutoff value using the pROC package in R. Additionally, three unique replicates were extracted from the five replicates to consider the variability of cutoff value in each assay, and 10 possible combinations were confirmed. Variability was assessed by calculating the mean and SD value of the cutoff value identified from 10 combinations.

The concordance rate of *EGFR* hotspot mutations was measured by dividing the number of

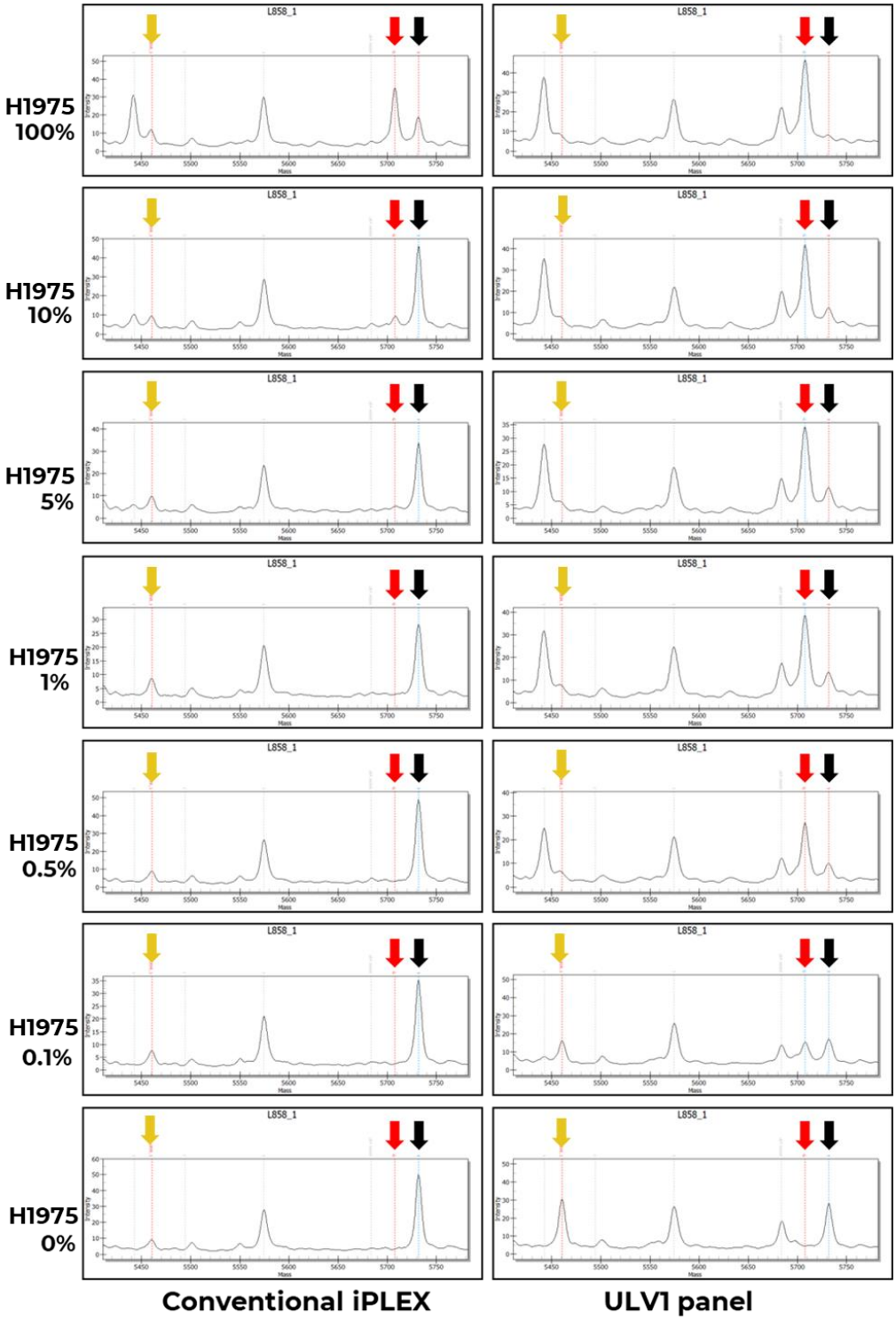
EGFR hotspot mutations in cfDNA identified in ULV1 and CT-ULTRA by the number of *EGFR* hotspot mutations identified in tissues. The concordance rate of oval *EGFR* mutations was also calculated by dividing the number of *EGFR* mutations of cfDNA identified in ULV1 and CT-ULTRA by the total number of *EGFR* mutations identified in matched tumor tissue DNA.

Results

Optimization for mutant allele selective amplification using the ULV1 panel

The optimal primer conditions for the ULV1 panel were determined based on the asymmetric primer concentration between the outer and MSP primers in a multiplexing manner. The detection limit of *EGFR* L858R mutation using a conventional PCR product amplified with an outer primer set was approximately 5%, which is consistent with the diagnostic sensitivity of iPLEX chemistry disclosed by the manufacturer. No mutation signal was detected in the positive samples with low mutant fractions (<5%). However, the ULV1 product amplified by incorporating the appropriate MSP concentration for selective enrichment of the mutant allele demonstrated a clear mutation signal, even in a 0.1% positive sample **(Figure 3)**. Furthermore, the mutation signal increased with the ratio of MSP compared with that of the outer primer in the same direction. For the *EGFR* L858R mutation, the ratio of the outer forward primer (OF: the same direction as MSP) to MSP was 0.2:1 (the final concentration of OF and MSP was 0.024 and 0.12 μM, respectively). In the wild-type control sample, no mutant signal was detected under any primer conditions **(Figure 4)**.

Considering that the ULV1 panel is based on multiplex PCR, the effect of the $MgCl_2$ concentration in diverse ranges (3.5, 3, 2.7, 2.3, and 1.875 mM) was investigated using a ctDNA Complete Mutation mix with a 0.5% mutant allele fraction. Although mutant signals were clearly detected in all mutations at an $MgCl_2$ concentration of 2.3 to 3.5 mM, signals were undetected at low $MgCl_2$ concentrations (1.875 mM) owing to amplification failure **(Figure 5)**. In this regard, we adopted the optimum conditions that maintain the balance of wild-type amplification while having the highest mutant signal intensity without non-specificity as the final experimental conditions **(Table 4)**.



GCACCCAGCAGTTTGGCC (reverse); 5460.6 Da - **UEP**

GCACCCAGCAGTTTGGCC+**G** (reverse) ; 5707.7 Da - **Mutant**

GCACCCAGCAGTTTGGCC+T (reverse) ; 5731.8 Da - **Wild**

L858R ; c.2573T>G (p.Leu858Arg)

Figure 3. Direct comparison of detection sensitivity between the ULV1 panel and conventional MassARRAY using serially diluted H1975 gDNA. Among the assays comprising ULV1 panel, L858R assay demonstrated strong intensity of mutant signal (red arrow) compared with that of conventional MassARRAY in the same sample. Black and gold arrows indicate wild and UEP signal intensities, respectively.

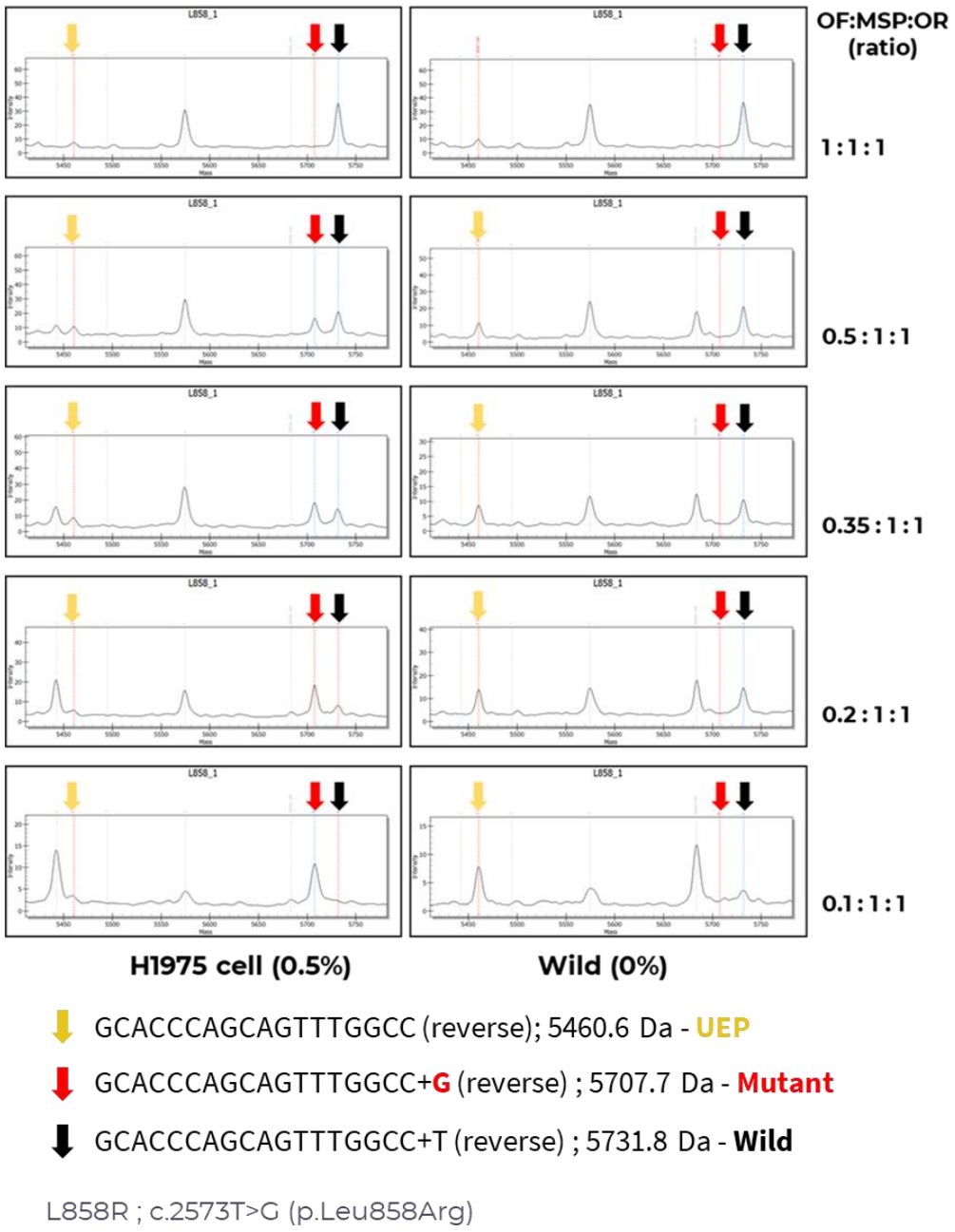
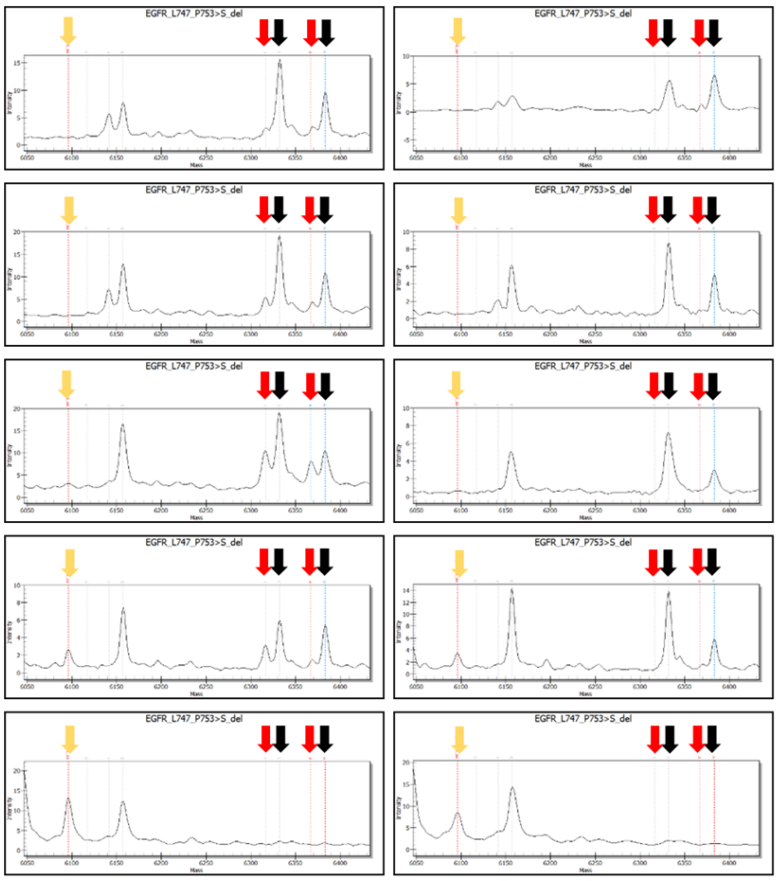


Figure 4. Optimization of PCR conditions depending on the unbalanced primer concentration using diluted H1975 gDNA. Among the assays comprising ULV1 panel, L858R assay gradually demonstrated a strong mutant signal (red arrow) as it decreased the concentration of the outer primer located in the same direction as L858R MSP. Black and gold arrows indicate wild and UEP signal intensities, respectively. OF, outer forward primer; MSP, mutation-specific primer; OR, outer reverse primer

A



GATTTCCTTGTTGGCTTTCG (forward); 6096 Da - **UEP**

GATTTCCTTGTTGGCTTTCG +**A** (forward) ; 6367.2 Da - **Mutant**

GATTTCCTTGTTGGCTTTCG +G (forward) ; 6383.2 Da - **Wild**

L747_P753>S c.2240_2257del (p.Leu747_Pro753delinsSer)

AAAATTCCCGTCGCTATCAA +**A** (reverse) ; 6316.2 Da - **Mutant**

AAAATTCCCGTCGCTATCAA +G (reverse) ; 6332.2 Da - **Wild**

EGFR_E746_A750del_c.55209959_55209973del (p.Glu746_Ala750delinsLys)

B

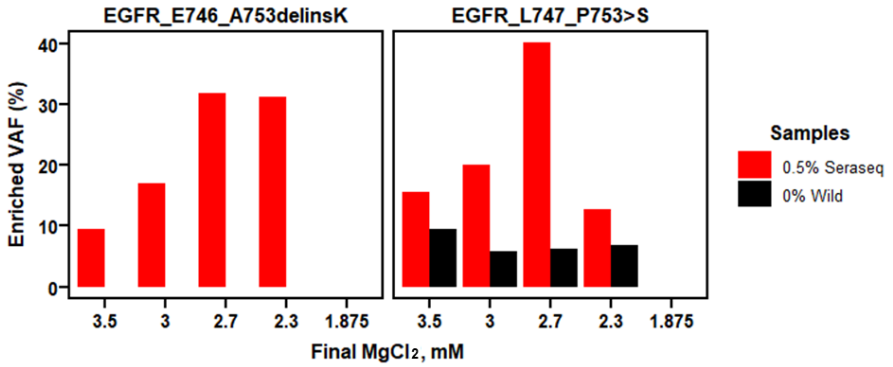


Figure 5. Optimization of PCR conditions depending on $MgCl_2$ concentration using 0.5% Seraseq material. A, All exon 19 assays included in Pool 1 showed the highest mutant signal at 2.7 mM $Mgcl_2$ concentration (red arrow). All exon 19 del hotspots were shown in one window. The mutant and wild-type masses in E746_A750delinsK assay are 6316.2 Da and 6332.2 Da. The mutant and wild-type masses in L747_P753>S assay are 6367.2 Da and 6383.2 Da. B, The bar plot represented the enriched VAF for each concentration.

Table 4. The optimal condition of the ULV1 Panel

ULV1	Direction of assay primers	Final conc. μM (each)	Ratio (OF : MSP : OR)	$MgCl_2$
Poo 1	**E709K/G719X F**	0.12	1 : 1 : 0.35	2.75 mM
	E709K/G719X R	0.04		
	E709K MSP R	0.12		
	G719A MSP F	0.12		
	G719D MSP F	0.12		
	E746_A750delinsK/L747_P753>S F	0.04	0.1 : 1 : 0.1	
	E746_A750delinsK/L747_P753>S R	0.04		
	E746_A750deinsK MSP F	0.36		
	L747_P753>S MSP R	0.36		
	C797S F	0.12	1 : 1 : 0.35	
	C797S R	0.04		
	C797 MSP R	0.12		
	***KRAS* G13D/V F**	0.24	1 : 1 : 0.5	
	***KRAS* G13D/V R**	0.12		
	***KRAS* G13D MSP R**	0.24		
	***KRAS* G13V MSP R**	0.24		
Pool 2	**E746_A750del F**	0.12	1 : 1 : 0.2	
	E746_A750del R	0.02		
	E746_A750del MSP R	0.12		
	T790M F	0.12	1 : 1 : 0.2	
	T790M R	0.02		
	T790M MSP R	0.12		

ULV1	Direction of assay primers	Final conc. μM (each)	Ratio (OF : MSP : OR)	$MgCl_2$
	L858R F	0.024	0.2 : 1 : 1	2.75 Mm
	L858R R	0.12		
	L858R MSP F	0.12		
	***BRAF* V600E F**	0.04	0.35 : 1 : 1	
	***BRAF* V600E R**	0.12		
	***BRAF* V600E MSP F**	0.12		
	***KRAS* G12D/V F**	0.06	0.5 : 1 : 1	
	***KRAS* G12D/V R**	0.12		
	***KRAS* G12D MSP F**	0.12		
	***KRAS* G12V MSP F**	0.12		

OF, Outer Forward primer; MSP, Mutation-specific Primer; OR, Outer Reverse primer

Performance of the ULV1 panel with high analytical sensitivity and specificity

To evaluate the diagnostic performance of the ULV1 panel, a total of 180–250 DNA samples were tested, comprising 35–70 positive with mutant allele proportions of 5%, 1%, 0.5%, 0.25%, 0.1%, 0.05%, and 0.025% (five replicates for each positive), including 145–215 negative controls **(Table 5)** for each target hotspot mutation. With these results, the cutoff values for each mutation had been determined using ROC curve analysis. The UHS assay resulted in a markedly increased mutant signal intensity, which was referred to as the enriched VAF. The enriched VAF was computed by dividing the mutant signal by the sum of mutant and wild-type signals. The cutoff values of enriched VAF for identifying each positive mutant differed for each mutation. The UHS assay exhibited slightly different efficiencies in mutant allele selective enrichment between target mutations; however, all assays excluding *KRAS* G12D demonstrated high sensitivity (85.7–100%) and specificity (87.9–100%) with a limit of detection (LOD) of 0.025–0.1%. In the case of *KRAS* G12D, the sensitivity and specificity were 80 and 91.7%, respectively, with an LOD of 0.1% **(Table 6)**.

The amplification efficiency was evaluated between the initial VAF of the positive samples and the enriched VAF of the UHS products for all 14 hotspot mutations in the ULV1 panel **(Figure 6)**. Owing to the selective enrichment of mutant alleles via the UHS assay, all positive samples with a mutant fraction greater than 1% exhibited saturated enriched VAF **(Figure 7)**.

Table 5. Summary of positive controls for evaluating the analytical performance of the ULV1 panel

ULV1	Gene	Hotspots	Nucleotide Change	Positive samples	Tested range	Replicates
Pool 1	***EGFR***	E709K	c.2125G>A	E709K/G719A Plasmid	5% ~ 0.025% (7 ranges)	5
	EGFR	E709K	c.2125G>A	E709K/G719D Plasmid	5% ~ 0.025% (7 ranges)	5
	EGFR	G719A	c.2156G>C	E709K/G719A Plasmid	5% ~ 0.025% (7 ranges)	5
	EGFR	G719D	c.2156G>A	E709K/G719D Plasmid	5% ~ 0.025% (7 ranges)	5
	EGFR	E746_A750delinsK	c.2235_2249del15	PC9 gDNA	5% ~ 0.025% (7 ranges)	5
	EGFR	E746_A750delinsK	c.2235_2249del15	Seraseq ctDNA reference	5% ~ 0.025% (7 ranges)	5
	EGFR	L747_P753>S	c.2240_2257del18	Seraseq ctDNA reference	5% ~ 0.025% (7 ranges)	5
	EGFR	C797S	c.2389T>A	C797S Plasmid	5% ~ 0.025% (7 ranges)	5
	KRAS	G13D	c.38G>A	HCT15 gDNA	5% ~ 0.025% (7 ranges)	5
	KRAS	G13V	c.38G>T	G13V Plasmid	5% ~ 0.025% (7 ranges)	5
Pool 2	***EGFR***	E746_A750del	c.2236_2250del15	HCC827 gDNA	5% ~ 0.025% (7 ranges)	5
	EGFR	T790M	c.2369C>T	H1975 gDNA	5% ~ 0.025% (7 ranges)	5
	EGFR	L858R	c.2573T>G	H1975 gDNA	5% ~ 0.025% (7 ranges)	5
	BRAF	V600E	c.1799T>A	HT29 gDNA	5% ~ 0.025% (7 ranges)	5
	KRAS	G12D	c.35G>A	AsPC1 gDNA	5% ~ 0.025% (7 ranges)	5
	KRAS	G12V	c.35G>T	SW900 gDNA	5% ~ 0.025% (7 ranges)	5

Table 6. Performance of *EGFR* assays in the ULV1 panel

	EGFR								
Parameters	**E709K**	**G719A**	**G719D**	**E746_A750delinsK**	**E746_A750del**	**L747_P753>S**	**T790M**	**C797S**	**L858R**
Sensitivity (%)	100	100	94.3	95.7	94.3	85.7	88.6	100	100
(95% CI)	(94.87 ~ 100)	(90 ~ 100)	(80.84 ~ 99.3)	(87.98 ~ 99.11)	(80.84 ~ 99.3)	(69.74 ~ 95.19)	(73.26 ~ 96.8)	(90 ~ 100)	(90 ~ 100)
Specificity (%)	97.2	98.6	98.1	98.9	99.3	95.3	98.6	100	96.6
(95% CI)	(93.64 ~ 99.09)	(95.98 ~ 99.71)	(95.31 ~ 99.49)	(96.04 ~ 99.87)	(96.22 ~ 99.98)	(91.61 ~ 97.75)	(95.11 ~ 99.83)	(98.3 ~ 100)	(92.14 ~ 98.87)
AUC (%)	99.3	99.8	98.8	98.9	96.9	91	96.9	100	99.8
Cutoff point (%)	4.9	17.2	17.2	1.8	4.8	7.4	10.6	38.0	13.5
The limit of detection (%)	0.025	0.025	0.05	0.025	0.05	0.05	0.1	0.025	0.025
PPV (%)	93.33	92.11	89.19	97.1	97.06	75	93.94	100	87.5
(95% CI)	(85.51 ~ 97.08)	(79.13 ~ 97.29)	(75.69 ~ 95.62)	(89.4 ~ 99.25)	(82.37 ~99.57)	(61.74 ~ 84.8)	(79.57 ~ 98.41)		(74.74 ~ 94.31)
NPV (%)	100	100	99.06	98.34	98.63	97.62	97.28	100	100
(95% CI)			(96.49 ~ 99.75)	(95.15 ~ 99.45)	(94.94 ~ 99.64)	(94.79 ~ 98.93)	(93.43 ~ 98.9)		
Accuracy (%)	98	98.8	97.6	98	98.33	94	96.67	100	97.22
(95% CI)	(95.39 ~ 99.35)	(96.53 ~ 99.75)	(94.85 ~ 99.11)	(95.39 ~ 99.35)	(95.21~ 99.65)	(90.3 ~ 96.6)	(92.89 ~ 98.77)	(98.54 ~ 100)	(93.64 ~ 99.09)

AUC, Area Under the Curve; PPV, Positive Predictive Value; NPV, Negative Predictive Value; 95% CI, 95% Confidence Interval

Table 6. Performance of *BRAF* and *KRAS* in the ULV1 panel (continued)

	BRAF	***KRAS***			
Parameters	**V600E**	**G12D**	**G12V**	**G13D**	**G13V**
Sensitivity (%)	97.1	80	94.3	97.1	91.4
(95% CI)	(85.08 ~ 99.93)	(63.06 ~ 91.56)	(80.84 ~ 99.3)	(85.08 ~ 99.93)	(76.94 ~ 98.2)
Specificity (%)	97.9	91.7	99.3	89.8	87.9
(95% CI)	(94.07 ~ 99.57)	(85.99 ~ 95.65)	(96.22 ~ 99.98)	(84.92 ~ 93.48)	(82.78 ~91.95)
AUC (%)	99.4	92.2	98.5	98.3	92.8
Cutoff point (%)	3.8	3.1	13.3	6.0	5.4
The limit of detection (%)	0.025	0.1	0.025	0.025	0.05
PPV (%)	91.89	70	97.06	60.71	55.17
(95% CI)	(78.69 ~ 97.21)	(56.97 ~ 80.44)	(82.37 ~ 99.57)	(50.88 ~ 69.75)	(45.84 ~ 64.15)
NPV (%)	99.3	95	98.63	99.48	98.44
(95% CI)	(95.36 ~ 99.9)	(90.72 ~ 97.36)	(94.94 ~ 99.64)	(96.55 ~ 99.93)	(95.52 ~ 99.47)
Accuracy (%)	97.78	89.44	98.33	90.8	88.4
(95% CI)	(94.41 ~ 99.39)	(84.01 ~ 93.52)	(95.21 ~ 99.65)	(86.52 ~ 94.08)	(83.77 ~ 92.09)

AUC, Area Under the Curve; PPV, Positive Predictive Value; NPV, Negative Predictive Value; 95% CI, 95% Confidence Interval

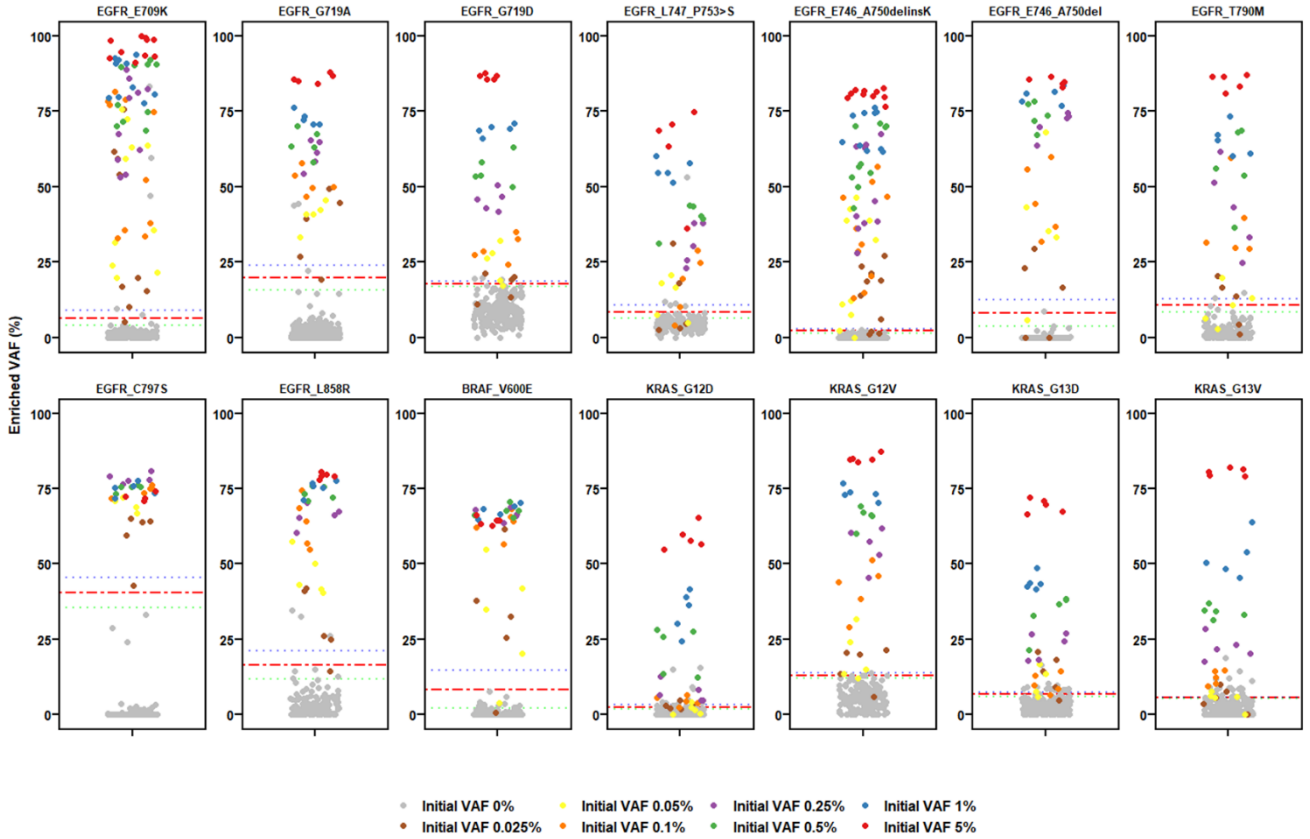


Figure 6. **Amplification efficiency and sensitivity of the ULV1 panel.** All assays have different cutoff values depending on amplification efficiency in a multiplexing manner. The red dashed line on the horizontal is marked with the mean of a cutoff value. The blue and green dotted lines on the horizontal are marked with mean +SD and −SD, indicating the acceptable ranges of cutoff.

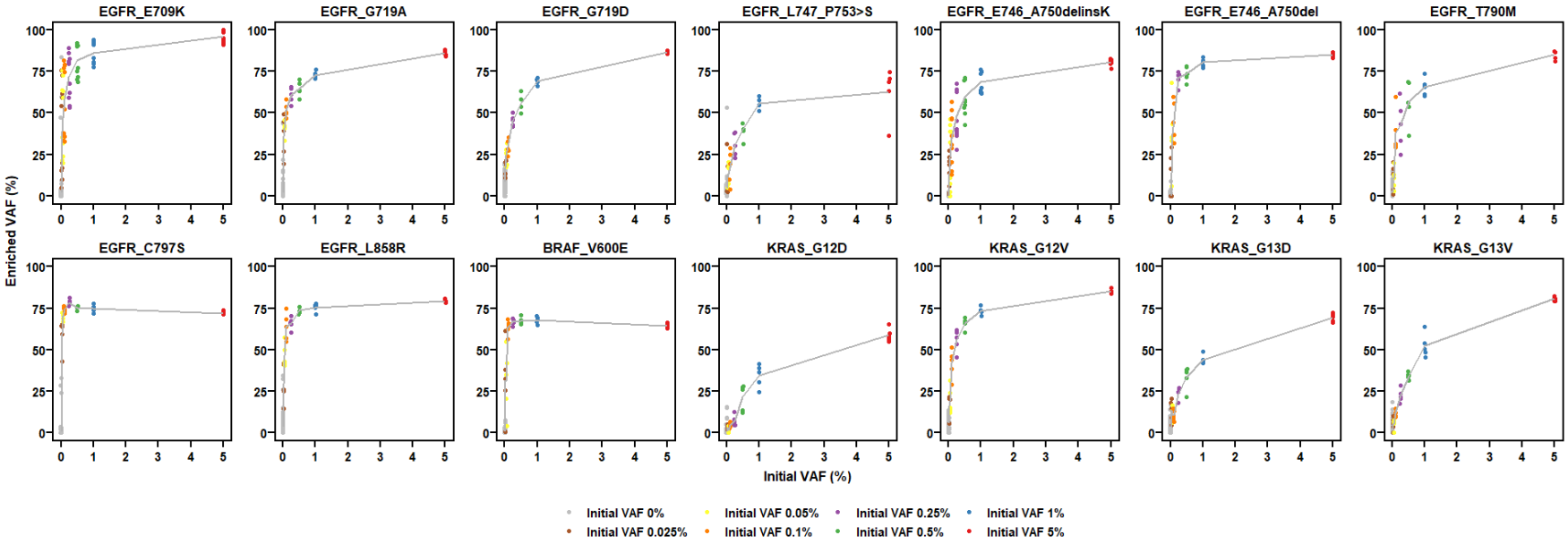


Figure 7. Amplification curve shape of the enriched VAF against the initial VAF. Initial and enriched VAF in each positive sample are depicted on the x-axis and y-axis as a percentage. All assays depict saturation of enriched VAF in more than 1% of the initial VAF.

Application of the ULV1 panel using cfDNA from NSCLC patients

We evaluated the potential diagnostic capabilities of the ULV1 panel in clinical setting, focusing on achieving high sensitivity. In this regard, we evaluated the mutational patterns of cfDNAs samples from 104 NSCLC patients using the ULV1 panel. We compared the results to mutations identified in the DNA of matched tumor tissue for each patient. Additionally, we conducted targeted NGS using a CT-ULTRA panel for the 104 cfDNA samples and computed the concordance rate of mutations identified using ULV1 panel. The demographic and clinical characteristics of all 104 patients are represented in **Table 7**.

Table 7. Characteristics of 104 patients for the ULV1 and CT-ULTRA test

Characteristic	Number	Percentage (%)
Age (years)		
Median age	65	
Range	39-83	
Gender		
Female, n (%)	48	46.2
Male, n (%)	56	53.8
Pathologic Stage, n (%)		
I	23	22.1
II	13	12.5
III	14	13.5
IV	54	51.9
Activating *EGFR* mutation		
Exon 18	3	2.9
Exon 19	36	34.6
Exon 20	3	2.9
Exon 21	37	35.6
Exon19 + Exon 20	1	1.0
Wild	24	23.1

The matched tumor tissue DNA of the enrolled patients was previously screened for *EGFR* somatic mutations using diverse platforms, including NGS, PNA-clamping, and Sanger sequencing. A total of 81 *EGFR* mutations were identified in 80 (76.9%) of 104 tumor tissue DNA samples comprising 37 cases of exon 19 deletions (45.7%), 34 of L858R (42%), three of G719X (3.7%), one of T790M (1.2%), and six of other *EGFR* mutations (7.4%). Overall, ULV1 and CT-ULTRA panels detect *EGFR* mutations in cfDNA from a total of 23/81 (28.4%) and 26/81 cases (32.1%), respectively. The detection rate of *EGFR* mutations in cfDNA was significantly different depending on the tumor stage of NSCLC patients from which cfDNA was extracted, regardless of the mutation detection method used or the mutation location. As the tumor stage progressed, the detection rates of *EGFR* mutations in cfDNA using the ULV1 or CT-ULTRA panels gradually increased. In the initial stages (stage I or II), majority of cfDNA samples had few mutations detected by neither method (blue boxes). Conversely, in the late stages (stage III or IV), the number of red (detected using both methods), yellow (detected using ULV1 alone), and orange boxes (detected using CT-ULTRA alone) increased **(Figure 8A)**. The concentration of cfDNA extracted from the plasma of 104 NSCLC patients had an average value of 11.4 ng/mL (range 2–169 ng/mL), and there was a statistically significant difference in cfDNA concentration only between stage I and IV ($p < 0.05$). No differences in cfDNA concentration were observed among the other stages **(Figure 8B)**. The median VAF of the *EGFR* mutation in cfDNA detected using the CT-ULTRA panel was 5.3% (range 0.8–50.2%) and a statistically significant increase in VAF values was observed in stage IV patients alone compared to other stages **(Figure 8C)**.

When we limited our analysis to the hotspot mutations included in the ULV1 panel, the concordance rates of the mutations detected in cfDNA using the ULV1 and CT-ULTRA panels with the *EGFR* mutations detected in the patient's tumor tissues were 51.1% (23/45) and 44.4% (20/45), respectively. As the tumors progressed from stages I to IV, the *EGFR* hotspot mutations detection rates in cfDNA using the ULV1 panel gradually increased to 8.33% (1/12),

25% (1/4), 50% (3/6), and 78.3% (18/23), respectively. Although a similar trend was observed in the CT-ULTRA analysis, it was relatively low than that in the ULV1 analysis. In tumor samples at stages II and III, ULV1 analysis demonstrated a statistically higher concordance rate than that of CT-ULTRA analysis **(Figure 8D)**. Additionally, the concordance rate of *EGFR* overall mutations between the patient's tumor tissues and matched plasma samples demonstrated a gradual increase in the ULV1 panel compared to CT-ULTRA according to stage, whereas the concordance rate in stage IV was higher in CT-ULTRA **(Figure 8E)**.

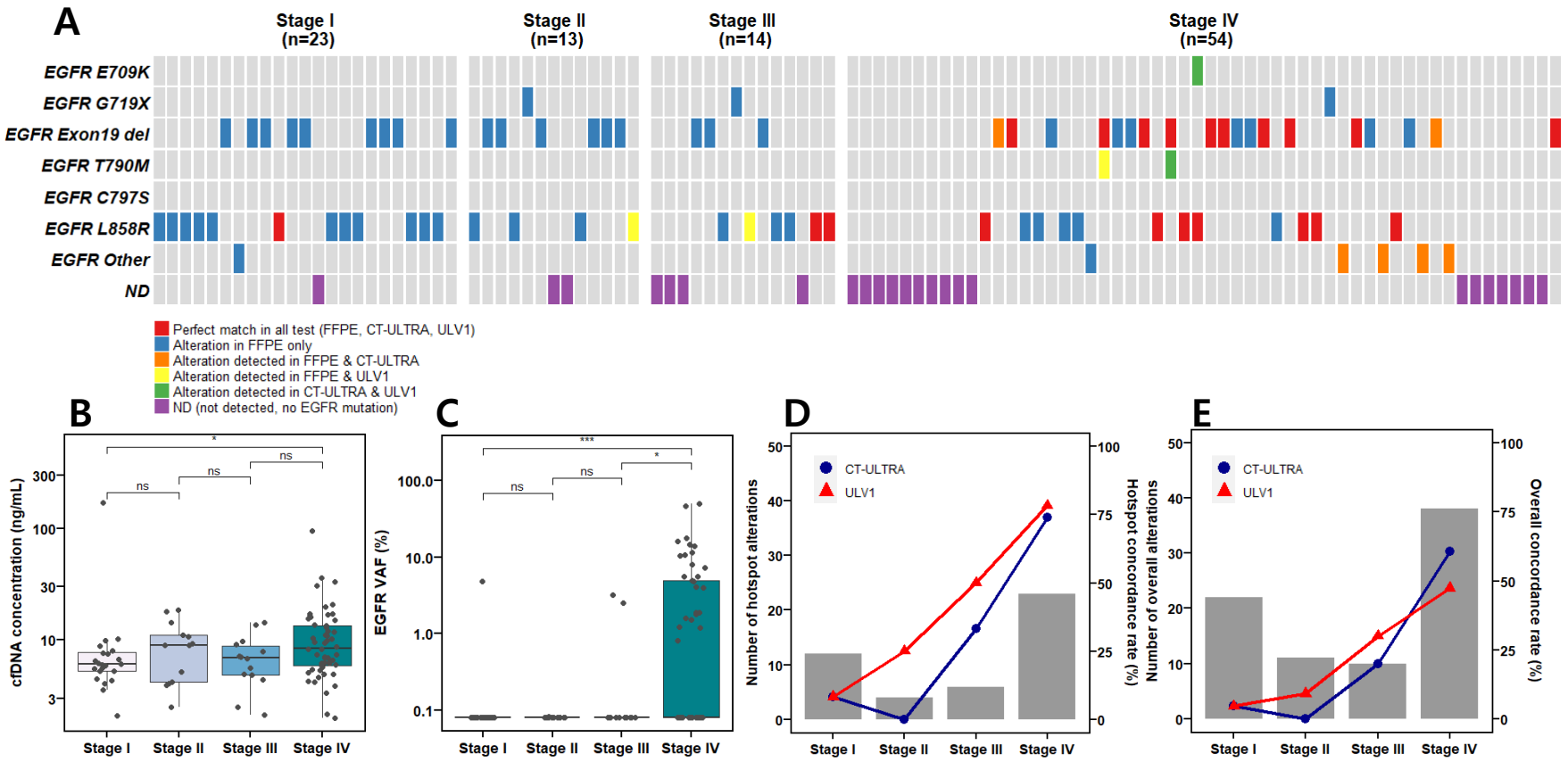
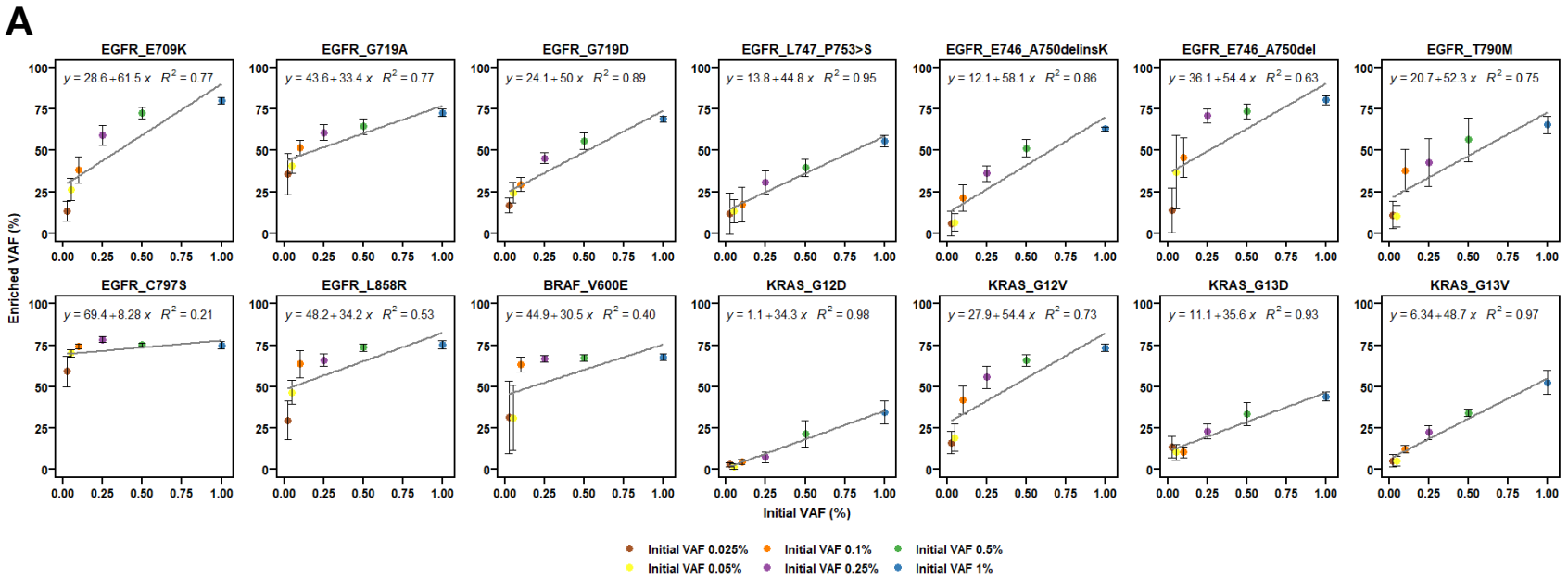


Figure 8. Validation of the ULV1 panel using 104 patient samples. A, Mutational profiling and concordance between tumor tissue and matched plasma in 104 patient samples. The mutational concordance between the tissue and plasma of each patient, including the concordance between both methods in plasma samples is depicted as a heatmap.

B, Comparison of cfDNA concentration between stages. High cfDNA concentration is associated with advanced stages of cancer. C, Comparison of *EGFR* VAF between stages. High VAF of cfDNA is associated with advanced stages of cancer. D, Comparison of concordance rate between tissue and matched plasma in *EGFR* hotspot mutation. Gray bars represent the number of *EGFR* hotspot mutations identified in the tissue. The red and blue dotted lines demonstrate the concordance rate of *EGFR* hotspot mutation identified using ULV1 and CT-UTRA compared to that of the tissue. E, Gray bars represent the number of *EGFR* overall mutations identified in the tissue. The red and blue dotted lines depict the concordance rate of *EGFR* overall mutation identified using ULV1 and CT-UTRA compared to that of the tissue.

Feasibility of semi-quantitative analysis of the ULV1 panel

Positive samples with low mutant allele fractions <1% demonstrated a relatively good correlation between the initial and enriched VAF, with almost linear phased enrichment **(Figure 9A)**. These results indicate the feasibility of semi-quantitative analysis using the ULV1 panel for clinical samples with low mutation fractions, including cfDNA extracted from the plasma of cancer patients. In fact, the ULV1 panel analysis detected three *EGFR* mutations, including two L858R mutations and one T790M mutation, that were not detected by the CT-ULTRA panel. **Table 8** provides a comparison of *EGFR* hotspot mutations detected using ULV1 and CT-ULTRA. Additionally, the enriched VAF in most samples with <1% VAF identified in CT-ULTRA was distributed within the semi-quantitative ranges **(Figure 9B)**.


A
EGFR_E709K
y = 28.6+61.5 x R² = 0.77
EGFR_G719A
y = 43.6+33.4 x R² = 0.77
EGFR_G719D
y = 24.1+50 x R² = 0.89
EGFR_L747_P753>S
y = 13.8+44.8 x R² = 0.95
EGFR_E746_A750delinsK
y = 12.1+58.1 x R² = 0.86
EGFR_E746_A750del
y = 36.1+54.4 x R² = 0.63
EGFR_T790M
y = 20.7+52.3 x R² = 0.75
EGFR_C797S
y = 69.4+8.28 x R² = 0.21
EGFR_L858R
y = 48.2+34.2 x R² = 0.53
BRAF_V600E
y = 44.9+30.5 x R² = 0.40
KRAS_G12D
y = 1.1+34.3 x R² = 0.98
KRAS_G12V
y = 27.9+54.4 x R² = 0.73
KRAS_G13D
y = 11.1+35.6 x R² = 0.93
KRAS_G13V
y = 6.34+48.7 x R² = 0.97
Enriched VAF (%)
Initial VAF (%)
Initial VAF 0.025%
Initial VAF 0.05%
Initial VAF 0.1%
Initial VAF 0.25%
Initial VAF 0.5%
Initial VAF 1%

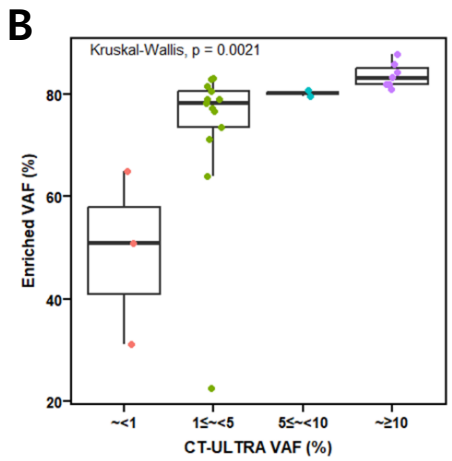

B
Kruskal-Wallis, p = 0.0021
Enriched VAF (%)
~<1
1≤~<5
5≤~<10
~≥10
CT-ULTRA VAF (%)

Figure 9. Feasibility of semi-quantitative analysis. A, Semi-quantitative detection range. The linear relation between the initial and enriched VAF identified by positive samples was in the range of 0.025–1%. B, Amplification curve shape between the CT-ULTRA VAF and the enriched VAF in patient samples. The CT-ULTRA and enriched VAF in each patient sample are depicted on the x-axis and y-axis as a percentage. The saturation of enriched VAF is depicted in patient samples with CT-ULTRA VAF of >1%.

Table 8. Variants detected in cfDNA by the ULV 1 panel only

Patients	Stage	cfDNA amount, ng (for DNA library)	Gene	Mutation	Observed VAF, % (CT-ULTRA)	Observed enriched VAF, % (ULV1)	Semi-Quantitative Dynamic range, %	Expected VAF, %
AMC036	**II**	**2.4**	***TP53***	**p.A161Pfs*9**	3.1	NA	NA	NA
			EGFR	**p.L858R**	1.1[a]	82	29.7 - 75.4	1< ~ ≤5
AMC044	**III**	**6.5**	***MTOR***	**p.A134V**	1.4	NA	NA	NA
			EGFR	**p.L858R**	0.33[a]	51	29.7 - 75.4	0.08
AMC070	**IV**	**10**	***EGFR***	**p.E746_A750delinsK**	1.2	64	5.8 - 62.9	1< ~ ≤5
			EGFR	**p.T790M**	0.23[a]	31	11.2 - 65.5	0.20

NA, Not Available; [a] The manual review results through each BAM file load into the IGV are shown.

Comparison of the ULV1 and PANAMutyper™ R *EGFR* performances

To evaluate the diagnostic performance of the ULV1 panel and the Korean Food and Drug Administration-approved PANAMutyper kit, a total of 21 DNA samples were tested. These included 18 positive samples (three replicates for each positive with mutant allele proportions of 1%, 0.5%, 0.25%, 0.1%, 0.05%, and 0.025%) and three negative samples (three replicates for each negative with mutant allele proportion of 0%) for each hotspot. The ULV1 panel was tested simultaneously with the same samples to ensure an accurate comparative analysis. (**Table 9**). Assays for exon 19 deletions and L858R demonstrated comparable detection rates regardless of the method used, whereas the detection rate of T790M was enhanced at low mutation fractions in ULV1. Cohen's κ coefficient demonstrated a nearly perfect concordance between ULV1 and PANAMutyper™ R *EGFR* for *EGFR* detection ($\kappa = 0.84$).

Table 9. Direct comparison between the ULV1 and PANAMutyper

Hotspot	Diluted samples	ULV1	PANAMutyper
EGFR	1%	3/3 (100%)	3/3 (100%)
Exon 19 del	0.5%	3/3 (100%)	3/3 (100%)
	0.25%	3/3 (100%)	3/3 (100%)
	0.1%	3/3 (100%)	3/3 (100%)
	0.05%	3/3 (100%)	3/3 (100%)
	0.025%	3/3 (100%)	3/3 (100%)
	0%	0/3 (0%)	0/3 (0%)
EGFR	1%	3/3 (100%)	3/3 (100%)
T790M	0.5%	3/3 (100%)	3/3 (100%)
	0.25%	3/3 (100%)	3/3 (100%)
	0.1%	3/3 (100%)	3/3 (100%)
	0.05%	**3/3 (100%)**	**2/3 (66.7%)**
	0.025%	**2/3 (66.7%)**	**0/3 (0%)**
	0%	0/3 (0%)	0/3 (0%)
EGFR	1%	3/3 (100%)	3/3 (100%)
L858R	0.5%	3/3 (100%)	3/3 (100%)
	0.25%	3/3 (100%)	3/3 (100%)
	0.1%	3/3 (100%)	3/3 (100%)
	0.05%	3/3 (100%)	3/3 (100%)
	0.025%	3/3 (100%)	3/3 (100%)
	0%	0/3 (0%)	0/3 (0%)

Discussion

In this study, a ULV1 panel was developed for highly sensitive detection of frequent hotspot mutations in NSCLC patients, and its clinical applicability was demonstrated through performance analysis using cfDNA samples with extremely low fractions of tumor-derived mutations. From a technical perspective, the most critical step of this study was selectively amplifying the minor mutant alleles from excess wild-type, which was addressed using the ARMS method **(Figure 10)** with the optimization of a ratio of additional MSP to the outer primer **(Figure. 4, 11)** and $MgCl_2$ concentration **(Figure. 5)**. Consequently, the product had a higher mutant allele fraction than that of the original, suggesting that enriched mutant products provide better sensitivity, as indicated by the red arrow in **Figure 3**.

Although several approaches with high sensitivity have been developed, the detection of mutations in cfDNA is often limited by low cfDNA levels[24]. To overcome this limitation, all limited DNA samples, including clinical samples and positive materials, were constructed as DNA libraries ligated with an NGS adaptor. Constructing cfDNA libraries offers several advantages, including preserving the limited number of clinical samples as amplified amounts and improving the detection rates of positive samples using a larger amount of DNA in the experiment. Moreover, it enables further analysis, such as NGS. Herein, the constructed cfDNA library was used as a template for the ULV1 panel and targeted NGS to compare the performance of NGS and ULV1 panel. Sufficient cfDNA libraries were constructed in the range of 271–1290 ng without bias to wild-type and mutant alleles using small amount of cfDNAs (1.9–10 ng) extracted from the plasma of 104 patients **(Table 1)**. The relatively low amounts of cfDNAs used in this study compared with previous other studies may be because only 1 mL of plasma was used for cfDNA extraction, including the long-term storage of the plasma (>12 months) prior to cfDNA extraction. Indeed, the storage term of plasma samples in stage IV tends to be shorter than that of plasma samples in stages I–III with low mutational concordance rates **(Figure 12)**. This result supports that long-term storage of both extracted

cfDNA and plasma samples induced DNA fragmentation, leading to reduced cfDNA yield[25]. High sensitivity and accurate detection of somatic alterations in samples with few copies of the target mutation requires an optimal cutoff value to accurately determine the presence of somatic mutations. ROC curve analysis was utilized to measure the efficiency of the ULV1 panel as a diagnostic tool, and different cutoff values were established for each mutation.

The major factors that determined the cutoff value of each mutation in this study were the limitations of multiplex PCR, including uneven amplification efficiency, generation of non-specific products caused by mispriming, and interference between primers targeting mutations located in the same or adjacent positions. These features were identified in *KRAS* assays, except for the G12V assay. The *KRAS* G12D assay demonstrated gentle amplification curve and the lowest enriched VAF in all ranges owing to lowest amplification efficiency **(Figure 7)**, which may hinder a clear distinction in enriched VAF between positive samples with low mutant fractions and negative samples. G13D and G13V assays demonstrated relatively low specificity. The G13D assay identified sporadic non-specific products that were not biased toward specific samples in false positives, whereas the G13V assay demonstrated that the G13D-positive sample contributed to the non-specific product, resulting in a low specificity. Furthermore, non-specific products were generated for *EGFR* E746_A750delinsK positive samples with mutant fractions of >5% in the E746_A750del assay. This non-specific amplification appears to be the result of mispriming of E746_A750del MSP to E746_A750delinsK, a similar mutation in the immediately adjacent region. Therefore, in cases where the mutation fraction of *EGFR* E746_A750delinsK is >5%, there is a risk of detecting false-positive signals for the E746_A750del mutation, along with a true-positive signal for the E746_A750delinsK mutation. However, as E746_A750delinsK mutant signal is much higher, and these two mutations rarely coexist, this would not present challenges in distinguishing the types of exon-19 deletions (data not shown).

To determine the clinical applicability of using actual patient samples, we evaluated the

correspondence of hotspot mutations between cfDNA and tumor tissue DNA in 104 patients at diverse stages of NSCLC using targeted NGS and ULV1 panel analyses. Several studies have evaluated the feasibility of using cfDNA as a diagnostic sample from cancer patients by evaluating the concordance of somatic alterations between tumor tissues and matched cfDNA using targeted NGS. These studies predominantly evaluated the concordance between cfDNA and the primary or metastatic tissues of NSCLC patients [26]. A recent study also evaluated concordance in early- and late-stage NSCLC[27]. Notably, Guo et al. observed positive concordance rates of 44% and 71.4% in early- and late-stage patients for somatic mutations between cfDNA and matched tissues, respectively, using a parallel NGS panel. In comparison, in this study, the somatic mutation concordance was 3.03% (1/33) in early-stage (I–II) and 53.2% (25/47) in late-stage (III–IV) patients, which represents a relatively low concordance rate compared to that reported by Guo et al. This study has several limitations. As mentioned previously, the initial amount of cfDNA used in this study was low, which explains the reason for mutation concordance rate between tumor tissue and cfDNA was lower than that reported in previous studies. Given that ctDNA levels are <0.1% VAF in most patients with stage I lung cancer[28], it is estimated that there are approximately 2–3 copies of mutant alleles in cfDNA of <10 ng. Herein, a limited sample of 1 mL was used, which particularly affected the plasma samples in the early stage, which were stored relatively long compared to those at stage IV. Additionally, such small ctDNA copies are likely to be lost during different clean-up steps in the DNA library construction process. This is likely the reason for the discordance in the early stages, regardless of the approach used. Generally, the higher input of cfDNA and advanced cancer stages result in increased detection rates[28,29]. Therefore, an increase in plasma volume may be required to improve the sensitivity of mutation detection using cfDNA from early-stage patients. A read depth of at least 1000× is required to detect mutations with 0.1% VAF. However, a relatively low sequencing depth of mean target coverage of 300× was used herein, compared with previous studies, resulting in false-negative outcomes for ctDNA identified

below the NGS detection limit. Only 19 out of 104 tissue samples were tested using the NGS platform, whereas the others were tested for *EGFR*-specific mutations alone by employing diverse diagnostic platforms. Notably, additional T790M and E709K mutations were detected in AMC075 and AMC077 patients in stage IV using both methods with cfDNA, whereas neither mutation was detected in the tissue samples **(Figure 8A, green)**. Although previous studies used the same technique to determine mutation concordance, our study verified that discrepancies in detection methods between tissue and cfDNA affect the concordance rate.

Interestingly, in analyses using the same cfDNA library, the ULV1 panel demonstrated a slightly higher concordance rate with mutations in the tumor tissue than targeted NGS **(Figure 8A, yellow).** Comparing the concordance rates for all *EGFR* mutations, the ULV1 panel demonstrated better results than those of the targeted NGS in stages II and III, and all the identified mutations were L858R, which is an oncogenic driver. Alternatively, the ULV1 panel concordance rate was lower than that of the targeted NGS in stage IV, which was expected given that some mutations were excluded from the ULV1 hotspot panel **(Figure 8E)**. Therefore, we attempted to determine the concordance by selecting only the hotspot mutations included in the ULV1 panel. Thus, it was verified that the ULV1 panel resulted in a slightly higher concordance rate than that of the targeted NGS in stage IV, along with stages II and III **(Figure 8D)**. Notably, in three patient samples (AMC036, 044, and 070) that contributed to the concordance rate, each *EGFR* mutation consistent with tumor tissue was detected in the ULV1 panel but not in the targeted NGS **(Figure 8A, yellow)**. To identify the ctDNA levels in these samples, the VAFs of the mutations observed using the automatic pipeline or expected by manual curation from targeted NGS were evaluated **(Table 8)**. Most somatic alterations detected in cfDNA have been reported to have much lower VAFs than those detected in matched solid tumor tissues[30,31] owing to the extremely low tumor-derived DNA proportion in cfDNA. Additionally, as ctDNA is derived from the entire tumor of cancer patients, unlike tissue biopsy, the presence of subclones that have a relatively low VAF compared to major

clones can be identified. Consistent with this, only one alteration with a VAF of <5% was observed using the automatic pipeline in each of the three patient samples **(Table 8)**. Furthermore, owing to the manual review of loading each BAM file into the IGV, it was verified that the VAFs of the *EGFR* mutations detected using the ULV1 panel alone were extremely low (0.23–1.1%). These results indicate that the *EGFR* mutations detected using the ULV1 panel alone may be subclonal mutations, implying a difference in detection sensitivity between cfDNA targeted sequencing and the ULV1 panel.

In this study, we demonstrated the potential of a semi-quantitative analysis of ULV1 panel results. This possibility, along with the detection of *EGFR* hotspot mutations, allowed for an estimation of the ctDNA fraction containing these mutations. To estimate the approximate fraction of three false-negative *EGFR* mutations, L858R (n = 2) and T790M (n = 1) in cfDNA targeted NGS, linear regression analyses were used. The R^2 value indicated that the suitability of each assay was for semi-quantitative analysis. The L858R assay had a relatively low R^2 value compared to the T790M assay. Using the ULV1 panel, the enriched VAFs of *EGFR* mutations detected in AMC044 (L858R) and AMC070 (T790M) were within the dynamic range of semi-quantification using linear regression, yielding 0.08% and 0.2% of the input VAF, respectively. Alternatively, the enriched AMC036 (L858R) sample VAF was found to be above the maximum value of the dynamic range, verifying that the input VAF alone was >1%. The *EGFR* C797S, *EGFR* L858R, and *BRAF* V600E assays had relatively low R^2 values. Owing to the high amplification efficiency, >60% of the enriched VAFs were created in 0.1% positive samples. The absence of a gradual correlation between the initial and enriched VAF may have been the reason for the less R^2 value. As ULV1 is a qualitative and semi-quantitative method, it may be more appropriate for use as a qualitative method in assays with a high amplification efficiency.

Additionally, we evaluated ULV1 performance through a direct comparison with the commercial kits. PANAMutyper combines PANAClamp and PANARealTyper and induces

mutant-specific amplification while suppressing wild-type DNA amplification of the target region. Therefore, a positive result from PANAMutyper could be easily determined using the threshold and melting temperature range for each assay, whereas a negative result was indirectly determined. However, ULV1 selectively amplifies mutant-type DNA from excess wild-type DNA, which could be more conducive to determining sample genotypes. As a result of the direct comparison between positive samples using these two methods, we observed that the detection rate was higher in samples with low mutant fractions in ULV1. This suggests that ULV1 can be utilized to screen multiple *EGFR* mutations.

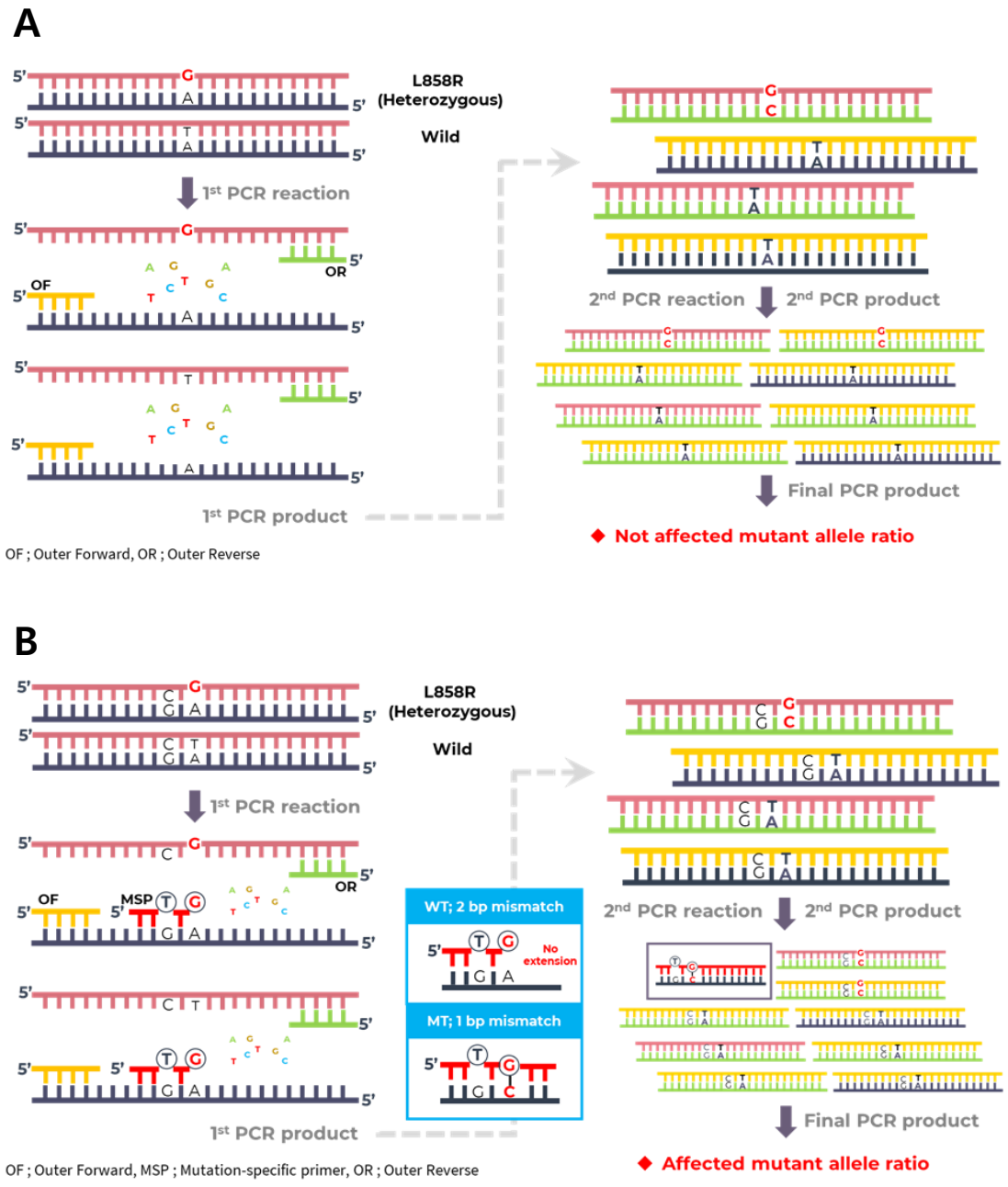


Figure 10. Schematic of the difference between the conventional PCR and ULV1 panel. A, The conventional PCR method utilizes forward and reverse primers to amplify the target, thereby maintaining the original mutant-to-wild type ratio present in the sample during amplification. B, The ULV1 method incorporates an MSP in addition to forward and reverse primers, resulting in preferential amplification of smaller-sized mutants. Thus, the

amplification in ULV1 leads to a higher mutant-to-wild type ratio than the original ratio present in the sample.As a result, the amplification in ULV1 leads to a higher mutant-to-wild type ratio than the original ratio present in the sample.

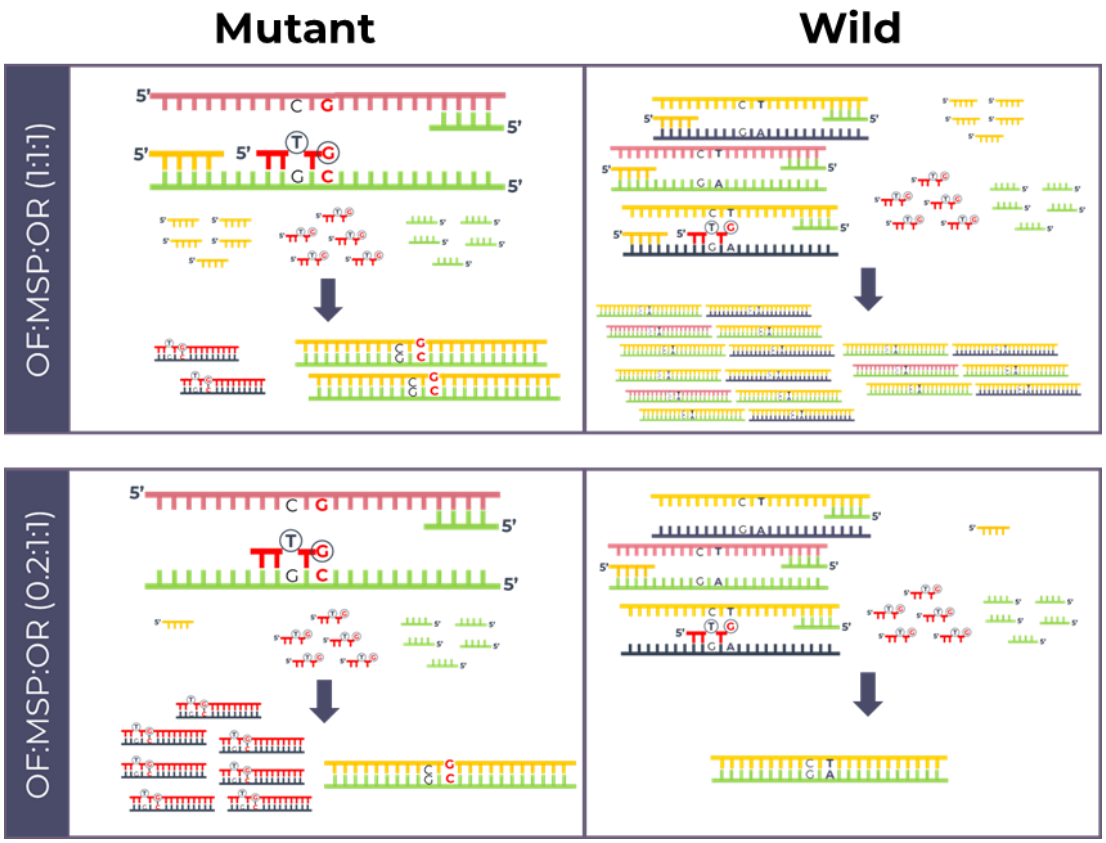


Figure 11. Schematic of the amplification efficiency depending on unbalanced primer concentration. As the concentration of the outer primer in the same direction as MSP decreases, amplification of smaller-sized targets mediated by the outer primer in the opposite direction to MSP is enhanced.

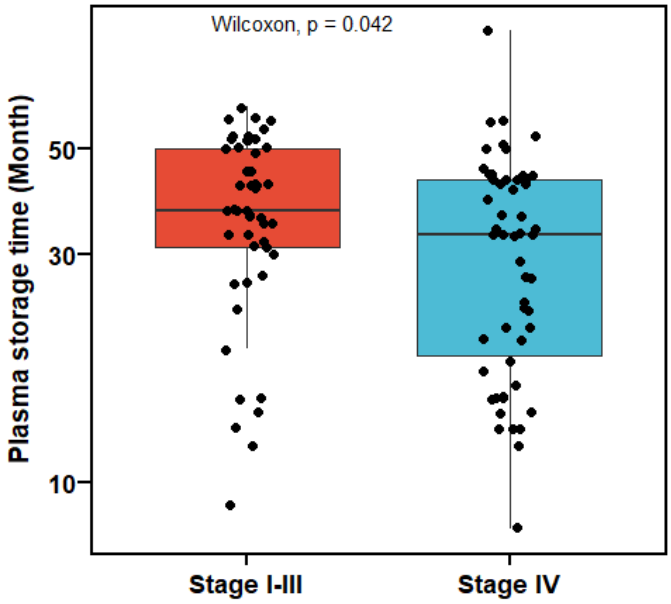


Figure 12. Distribution of plasma storage duration between stages in the cohort.

Plasma samples from advanced cancer patients demonstrated that the plasma storage time was relatively short.

Conclusion

In conclusion, Onco-UHS is a simple, highly sensitive, and cost-effective multiplex profiling method for detecting driver mutations in various types of carcinomas. The ULV1 panel may be useful for identifying somatic alterations in plasma samples from NSCLC patients, enabling the monitoring of the responsiveness of patients treated with *EGFR*-TKI, including NGS validation.

Bibliography

1. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004;350:2129–2139.

2. Yuan M, Huang L-L, Chen J-H, et al. The emerging treatment landscape of targeted therapy in non-small-cell lung cancer. Signal Transduction and Targeted Therapy 2019;4:1–14.

3. Wang S, Cang S, Liu D. Third-generation inhibitors targeting *EGFR* T790M mutation in advanced non-small cell lung cancer. Journal of Hematology & Oncology 2016;9:34.

4. Wang S, Tsui ST, Liu C, et al. *EGFR* C797S mutation mediates resistance to third-generation inhibitors in T790M-positive non-small cell lung cancer. J Hematol Oncol 2016;9:59.

5. Taus Á, Camacho L, Rocha P, et al. Dynamics of *EGFR* Mutation Load in Plasma for Prediction of Treatment Response and Disease Progression in Patients With *EGFR*-Mutant Lung Adenocarcinoma. Clinical Lung Cancer 2018;19:387-394.e2.

6. Zhang J-T, Liu S-Y, Gao W, et al. Longitudinal Undetectable Molecular Residual Disease Defines Potentially Cured Population in Localized Non–Small Cell Lung Cancer. Cancer Discovery 2022;12:1690–1701.

7. Temilola DO, Wium M, Coulidiati TH, et al. The Prospect and Challenges to the Flow of Liquid Biopsy in Africa. Cells 2019;8. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6721679/ [Accessed October 31, 2020].

8. Filho OM, Viale G, Stein S, et al. Impact of HER2 heterogeneity on treatment response of early-stage HER2-positive breast cancer: phase II neoadjuvant clinical trial of T-DM1 combined with pertuzumab. Cancer Discov 2021;11:2474–2487.

9. Russano M, Napolitano A, Ribelli G, et al. Liquid biopsy and tumor heterogeneity in

metastatic solid tumors: the potentiality of blood samples. Journal of Experimental & Clinical Cancer Research 2020;39:95.

10. Heitzer E, Auinger L, Speicher MR. Cell-Free DNA and Apoptosis: How Dead Cells Inform About the Living. Trends in Molecular Medicine 2020;26:519–528.

11. Bronkhorst AJ, Ungerer V, Holdenrieder S. The emerging role of cell-free DNA as a molecular marker for cancer management. Biomolecular Detection and Quantification 2019;17:100087.

12. Yi X, Ma J, Guan Y, et al. The feasibility of using mutation detection in ctDNA to assess tumor dynamics. Int J Cancer 2017;140:2642–2647.

13. Schwarzenbach H, Hoon DSB, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. Nat Rev Cancer 2011;11:426–437.

14. Noguchi T, Iwahashi N, Sakai K, et al. Comprehensive Gene Mutation Profiling of Circulating Tumor DNA in Ovarian Cancer: Its Pathological and Prognostic Impact. Cancers 2020;12:3382.

15. Xu X, Yu Y, Shen M, et al. Role of circulating free DNA in evaluating clinical tumor burden and predicting survival in Chinese metastatic colorectal cancer patients. BMC Cancer 2020;20:1006.

16. Peled M, Agassi R, Czeiger D, et al. Cell-free DNA concentration in patients with clinical or mammographic suspicion of breast cancer. Sci Rep 2020;10:14601.

17. Corcoran RB, Chabner BA. Application of Cell-free DNA Analysis to Cancer Treatment. N Engl J Med 2018;379:1754–1765.

18. Fujita K, Nakayama M, Sata M, et al. Highly sensitive detection of driver mutations from cytological samples and cfDNA in lung cancer. Cancer Med 2021;10:8595–8603.

19. Jerič Kokelj B, Štalekar M, Vencken S, et al. Feasibility of Droplet Digital PCR

Analysis of Plasma Cell-Free DNA From Kidney Transplant Patients. Front Med (Lausanne) 2021;8:748668.

20. Feng W, Gu W, Zhao N, et al. Comparison of the SuperARMS and Droplet Digital PCR for Detecting *EGFR* Mutation in ctDNA From NSCLC Patients. Translational Oncology 2018;11:542–545.

21. Shin S-J, Chun S-M, Kim T-I, et al. Feasibility of multiplexed gene mutation detection in plasma samples of colorectal cancer patients by mass spectrometric genotyping. PLoS ONE 2017;12.

22. Midha A, Dearden S, McCormack R. *EGFR* mutation incidence in non-small-cell lung cancer of adenocarcinoma histology: a systematic review and global map by ethnicity (mutMapII). Am J Cancer Res 2015;5:2892–2911.

23. Mao C, Qiu L-X, Liao R-Y, et al. *KRAS* mutations and resistance to *EGFR*-TKIs treatment in patients with non-small cell lung cancer: a meta-analysis of 22 studies. Lung Cancer 2010;69:272–278.

24. Elazezy M, Joosse SA. Techniques of using circulating tumor DNA as a liquid biopsy component in cancer management. Comput Struct Biotechnol J 2018;16:370–378.

25. Sozzi G, Roz L, Conte D, et al. Effects of prolonged storage of whole plasma or isolated plasma DNA on the results of circulating DNA quantification assays. J Natl Cancer Inst 2005;97:1848–1850.

26. Jahangiri L, Hurst T. Assessing the Concordance of Genomic Alterations between Circulating-Free DNA and Tumour Tissue in Cancer Patients. Cancers (Basel) 2019;11. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6966532/ [Accessed November 5, 2020].

27. Guo Q, Wang J, Xiao J, et al. Heterogeneous mutation pattern in tumor tissue and circulating tumor DNA warrants parallel NGS panel testing. Mol Cancer 2018;17:131.

28. Jiang J, Adams H-P, Yao L, et al. Concordance of Genomic Alterations by Next-Generation Sequencing in Tumor Tissue versus Cell-Free DNA in Stage I–IV Non–Small Cell Lung Cancer. The Journal of Molecular Diagnostics 2020;22:228–235.

29. Bettegowda C, Sausen M, Leary RJ, et al. Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies. Sci Transl Med 2014;6:224ra24.

30. Kujala J, Hartikainen JM, Tengström M, et al. Circulating Cell-Free DNA Reflects the Clonal Evolution of Breast Cancer Tumors. Cancers 2022;14:1332.

31. Kujala J, Hartikainen JM, Tengström M, et al. High mutation burden of circulating cell-free DNA in early-stage breast cancer patients is associated with a poor relapse-free survival. Cancer Medicine 2020;9:5922–5931.

국문 요약

암은 전 세계적으로 여전히 높은 사망률에 기여하고 있는 질병으로서, 암에 대한 치료는 종양 조직의 프로파일링에 의존하고 있다. 일반적으로, 치료의 단서가 될 종양 조직을 얻기 위해서는 환자의 건강 상태와 종양의 접근성을 고려해야 하는데, 일부 환자는 종양의 위치상 접근이 용이하지 않아 조직의 확보가 어려운 것으로 알려져 있다. 그리고, 종양은 클론진화로 인해 발생한 종양 내 이질성의 성향을 띄고 있기 때문에 FFPE section처럼 종양 조직의 일부만을 포함하는 검체에는 치료의 실마리가 될 클론이 포함되지 않을 가능성이 있고, 이로 인해 치료의 단서를 놓치는 결과로 이어질 수 있다. 또한, 암으로 진단받은 환자는 종양의 진행과 재발 및 치료의 반응성을 추적하기 위해서 주기적인 모니터링이 필요하다. 그러나 때마다 반복적인 조직의 채취는 쉽지 않음은 물론 영상학적인 테크닉을 통해 암의 관리가 가능하기는 하지만 비용적인 문제 및 초기단계의 종양이나 매우 작은 전이성 종양의 발견은 여전히 어려운 것으로 알려져 있다.

이러한 조직 수집 및 종양 내 이질성에 대한 문제를 극복할 대안으로서 액체생검이 최근 급부상하고 있다. 하지만 액체 생검의 가장 큰 한계는 종양 유래의 DNA가 매우 낮은 농도로 존재하기 때문에, 이를 탐지할 수준의 초고감도 기법 개발이 필요하다. 실제 암 환자의 혈액 내 ctDNA의 비율은 0.1% 수준의 매우 낮은 농도부터 90%까지 다양한 것으로 알려져 있다. ctDNA 농도는 암종, 종양의 크기 및 위치, 그리고 치료의 반응성과 같은 다양한 요인이 복합적으로 작용하여 영향을 미친다. 그러므로 low-coverage sequencing이나 Sanger sequencing과 같은 저감도 검출 한계를 가진 테크닉을 사용하여, ctDNA의 비율이 낮은 검체로부터 질병의 진단 및 예후에 활용되는 변이를 검출하는 것은 적합하지 않다.

본 연구는 기존 MassARRAY 플랫폼에 AMRS primer를 추가하여 유전자 변이의 검출 감도를 높인 OncoUHS assay를 활용하고자 하였고, 비소세포폐암 환자에서 빈번하게 발생하는 유전자 변이를 선별한 후 Ultra-High Sensitivity Lung Version 1 (ULV1) panel을 디자인하고, 이를 1년 이상 장기 보관된 혈장을 이용하여 ctDNA 비율이 낮은 검체에서도 유전자 변이를 검출할 수 있는 수준의 검사법임

을 확인하고자 하였다.

우리는 먼저 비소세포폐암의 진단, 치료의 반응성 및 예후를 예측할 수 있는데 중요한 역할을 하는 *BRAF*, *KRAS*, *EGFR* 유전자에 한하여 hotspot mutations 선정하였고, Multiplexing 적용이 가능한 MassARRAY의 이점을 이용하여 선정된 14 hotspot mutations을 2개의 pool로 나누고, 고감도 검출을 위해 반드시 필요한 Mutant-specific primer (MSP) 및 Primer set, 그리고 Unextended extend primer (UEP)를 디자인하고 이를 Ultra-high sensitivity lung version 1 (ULV1)으로 명명하였다. ULV1 panel의 target site를 포함한 target region의 증폭을 극대화하기 위해, Primers와 $MgCl_2$ 농도의 변화를 주어 최적의 농도 조건으로 최적화 한 후 표준 물질, cell lines 유래의 gDNA 및 hotspot을 포함하는 plasmid를 구축한 다음 이들을 단계별로 희석하여 다양한 농도의 샘플로 만든 후 ULV1 panel의 성능을 평가하였고, 그 결과 80 – 100%의 민감도와 87.9 – 100%의 특이도 그리고 0.025 – 0.1% 수준의 LOD를 확인하였다.

다음으로 ULV1 panel을 이용하여 진단의 정확성을 확인하고자, 장기간 보관된 104명의 소세포폐암 환자로부터 얻은 혈장 1 mL에서 cfDNA를 추출하고 이를 Adapter-ligated DNA library 형태로 제작하여 ULV1과 CT-ULTRA 실험에 사용하였다. 104명의 환자는 다양한 테크닉을 통해 종양 조직에서의 *EGFR* mutational status가 이미 확인된 sample cohort로 이 결과를 *EGFR* 진단의 정답지로 간주하고, 앞서 library 형태로 구축한 cfDNA를 사용하여 ULV1과 CT-ULTRA에서 확인된 *EGFR* mutation의 결과를 각각 비교 분석하였다. 그 결과 Stage II – IV에서 CT-ULTRA에 비해 더 나은 mutational concordance를 확인하였다.

또한, 단계별로 희석된 양성 샘플들을 이용하여 enrichment performance를 확인하고 특정 구간에서의 linearity phase가 확인됨으로써, 1% 이하의 낮은 mutant allele frequency를 가진 검체의 대략적인 ctDNA fraction을 유추할 수 있을 뿐만 아니라 낮은 수준에서의 ctDNA fluctuation을 파악할 수 있는 Semi-quantitative analysis의 가능성도 확인되었다.

마지막으로, 상용화된 PANAMutyper R *EGFR* kit와 ULV1 panel과의 성능을 직접적으로 비교함으로써 유사제품의 성능에 못지 않은 효율성을 확인하였다. 종합해보면, ULV1 panel은 비소세포폐암 환자의 혈장 검체를 이용하여 유전자 변이를 식별하는데 유용할 수 있으며, 이를 통해 치료의 반응성을 낮은 수준에서도 모니터링 할 수 있을 뿐만 아니라 NGS data의 validation에도 활용될 수 있는 tool로써 이용 가능성을 제시한다.

핵심어: 세포유리 핵산; 비소세포폐암; *EGFR*; 저품질 혈장; Ultra-high sensitivity lung version 1; MassARRAY