



OPEN ACCESS

EDITED BY

Mehdi Pirooznia,
Johnson and Johnson, United States

REVIEWED BY

Ammar Husami,
Cincinnati Children's Hospital Medical Center,
United States
Oscar Haigh,
Commissariat à l'Énergie Atomique et aux
Énergies Alternatives (CEA), France

*CORRESPONDENCE

Thien-Phuc Nguyen Hoang,
✉ phucnguyen1@genesolutions.vn
Lan N. Tu,
✉ lantu@genesolutions.vn

†These authors have contributed equally to
this work

RECEIVED 07 November 2023

ACCEPTED 26 January 2024

PUBLISHED 09 February 2024

CITATION

Nguyen Hoang T-P, Nguyen TA, Tran NHB,
Nguyen Hoang V-A, Thi Dao HT, Tran V-U,
Nguyen YN, Nguyen AT, Nguyen Thi CT,
Do Thi TT, Nguyen DS, Nguyen H-N, Giang H
and Tu LN (2024), Analytical validation and
clinical utilization of K-4CARE™: a
comprehensive genomic profiling assay with
personalized MRD detection.
Front. Mol. Biosci. 11:1334808.
doi: 10.3389/fmolb.2024.1334808

COPYRIGHT

© 2024 Nguyen Hoang, Nguyen, Tran, Nguyen
Hoang, Thi Dao, Tran, Nguyen, Nguyen, Nguyen
Thi, Do Thi, Nguyen, Nguyen, Giang and Tu. This
is an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Analytical validation and clinical utilization of K-4CARE™: a comprehensive genomic profiling assay with personalized MRD detection

Thien-Phuc Nguyen Hoang^{1,2*†}, Tien Anh Nguyen^{1,2†},
Nam H. B. Tran^{1,2†}, Van-Anh Nguyen Hoang^{1,2},
Hong Thuy Thi Dao^{1,2}, Vu-Uyen Tran^{1,2}, Yen Nhi Nguyen^{1,2},
Anh Tuan Nguyen^{1,2}, Cam Tu Nguyen Thi^{1,2}, Thanh Thuy Do Thi¹,
Duy Sinh Nguyen^{1,2}, Hoai-Nghia Nguyen^{1,2}, Hoa Giang^{1,2} and
Lan N. Tu^{1,2*}

¹Medical Genetics Institute, Ho Chi Minh City, Vietnam, ²Gene Solutions, Ho Chi Minh City, Vietnam

Background: Biomarker testing has gradually become standard of care in precision oncology to help physicians select optimal treatment for patients. Compared to single-gene or small gene panel testing, comprehensive genomic profiling (CGP) has emerged as a more time- and tissue-efficient method. This study demonstrated in-depth analytical validation of K-4CARE, a CGP assay that integrates circulating tumor DNA (ctDNA) tracking for residual cancer surveillance.

Methods: The assay utilized a panel of 473 cancer-relevant genes with a total length of 1.7 Mb. Reference standards were used to evaluate limit of detection (LOD), concordance, sensitivity, specificity and precision of the assay to detect single nucleotide variants (SNVs), small insertion/deletions (Indels), gene amplification and fusion, microsatellite instability (MSI) and tumor mutational burden (TMB). The assay was then benchmarked against orthogonal methods using 155 clinical samples from 10 cancer types. In selected cancers, top tumor-derived somatic mutations, as ranked by our proprietary algorithm, were used to detect ctDNA in the plasma.

Results: For detection of somatic SNVs and Indels, gene fusion and amplification, the assay had sensitivity of >99%, 94% and >99% respectively, and specificity of >99%. Detection of germline variants also achieved sensitivity and specificity of >99%. For TMB measurement, the correlation coefficient between whole-exome sequencing and our targeted panel was 97%. MSI analysis when benchmarked against polymerase chain reaction method showed sensitivity of 94% and specificity of >99%. The concordance between our assay and the TruSight Oncology 500 assay for detection of somatic variants, TMB and MSI measurement was 100%, 89%, and 98% respectively. When CGP-informed mutations were used to personalize ctDNA tracking, the detection rate of ctDNA in liquid biopsy was 79%, and clinical utility in cancer surveillance was demonstrated in 2 case studies.

Conclusion: K-4CARE™ assay provides comprehensive and reliable genomic information that fulfills all guideline-based biomarker testing for both targeted therapy and immunotherapy. Integration of ctDNA tracking helps clinicians to further monitor treatment response and ultimately provide well-rounded care to cancer patients.

KEYWORDS

comprehensive genomic profiling, tumor mutational burden, microsatellite instability, minimal residual disease, ctDNA (circulating tumor DNA), targeted therapy, immunotherapy, immune checkpoint inhibitor (ICI)

Introduction

Gene mutation-targeted therapy was first approved by the U.S. Food and Drug Administration (FDA) in 2001 (Waarts et al., 2022). Many ground-breaking developments since then have revolutionized cancer treatment and transformed the field of cancer diagnosis. Such progress is tightly linked with the advancement in genetic sequencing technologies, which have uncovered the mutational landscapes of cancer and hence revealed new actionable alterations (Waarts et al., 2022). Once a targeted therapy is approved, it is crucial to identify patients with the right genomic alterations predictive of response to the treatment. Therefore, biomarker testing has been recommended as the standard of care to empower precision oncology. As the number of therapies increases, the list of biomarkers also expands, not just in number but also in the complexity. Moving beyond single genomic alterations, some complex biomarkers have emerged, such as tumor mutational burden (TMB) and microsatellite instability (MSI). For the use of immune checkpoint inhibitors (ICIs) in advanced cancer, TMB and MSI have both been approved by the FDA as independent predictive biomarkers in multiple solid tumors, besides programmed death-ligand 1 (PD-L1) positivity (Buttner et al., 2019).

In real-world practice, biomarker testing is often run as single-gene, hotspot or small panel assays due to their low cost, simple workflow and quick turnaround time. Particularly in developing countries, popular methods are based on polymerase chain reaction (PCR) such as digital droplet PCR to detect single point mutations; or immunohistochemistry (IHC) and fluorescent *in situ* hybridization to detect gene amplification or rearrangement (Pennell et al., 2019; Radich et al., 2022). Next-generation sequencing (NGS) is increasingly popular but still limited for pre-selected gene panels. The main disadvantage of this practice is the inadequate coverage of all companion biomarkers and failure to analyze complex genomic signatures such as TMB and MSI, which could ultimately lead to patients missing treatment opportunities. For instance, it was reported that for lung carcinoma, the rate of *EGFR* testing was the highest (>80%) while the rates of *ALK* and *ROS1* testing were remarkably lower (Pennell et al., 2019), despite the fact that *ALK* and *ROS1* mutations could occur in up to 10% of the cases (Song et al., 2017; Dang et al., 2020). Therefore, comprehensive genomic testing (CGP) that analyzes hundreds of genes in a single assay could be a more efficient solution. CGP consolidates results for multiple biomarkers for both approved and emerging therapies while conserving precious biopsy samples. Knowing all

treatment options upfront could help physicians select or change therapy promptly when the disease progresses.

The major drawbacks of CGP are the cost and necessity of qualified equipment and highly skilled labor. Hence, CGP assays are either not available or not affordable for majority of patients in resource-restricted countries. Moreover, the quality of formalin-fixed paraffin-embedded (FFPE) tumor tissues, the primary resource for CGP, remains a big challenge in these countries. DNA and particularly RNA from FFPE could be heavily degraded due to problems in formalin fixation and storage condition. Limitation in both sample quantity and quality thus requires optimization, and standardized CGP protocols developed for FFPE samples of high quality might not be applicable. Furthermore, CGP is criticized to detect many mutations that are not treatment-meaningful (Christofyllakis et al., 2022). If this mutation profile could be leveraged to guide detection of circulating tumor DNA (ctDNA) in the plasma, this would then facilitate real-time monitoring of tumor burden and patient's response to treatment. Meaningful clinical benefits of ctDNA monitoring to detect minimal residual disease (MRD) have consistently been documented (Chin et al., 2019).

In this study, we established K-4CARE™ (K4C), a tissue-based CGP assay utilizing a panel of 473 cancer-associated genes to detect single nucleotide variants (SNVs), small insertion/deletions (Indels), gene amplification and fusion, as well as TMB and MSI assessment. The assay was thoroughly validated using reference standards and optimized for clinical samples of cancer patients in Vietnam. For selected types of cancer, the K4C assay also integrated ctDNA detection in the liquid biopsy for cancer surveillance.

Materials and methods

Sample collection

Genomic DNA from formalin-fixed paraffin-embedded (FFPE) tumor samples and matching white blood cell (WBC) samples of cancer patients were obtained from the Medical Genetics Institute, Ho Chi Minh city, Vietnam. All samples were de-identified and stored for less than 2 years. In total, 282 FFPE samples from more than 10 types of cancer were included in the study. The full list and allocation of samples for different experiments are in [Supplementary Table S1](#). The list of commercial reference standards used to assess the performance of the K4C assay is in [Supplementary Table S2](#). For ctDNA detection, plasma samples of 40 patients that had either colorectal cancer (CRC) or breast cancer were used in the analysis ([Supplementary Table S1](#)).

Library preparation and sequencing for FFPE and WBC

Genomic DNA was already isolated from FFPE and WBC samples as previously described [(Dang et al., 2020; Nguyen Hoang et al., 2023; Nguyen et al., 2022; Tran et al., 2021)]. DNA fragmentation and library preparation for paired FFPE and WBC samples were performed using the NEBNext Ultra II FS DNA library prep kit (New England Biolabs, USA) following the manufacturer's instructions. Libraries were pooled together and hybridized with predesigned probes for a panel of 473 cancer-relevant genes (Integrated DNA Technologies, USA). This gene panel was curated from large public cancer databases, high-impact cancer genomic studies, and particularly in-house database of prevalent mutations from >10,000 cancer patients. The panel included entire exons of 473 genes, promoter region of *TERT*, and selected introns of 6 common fusion genes (Supplementary Table S3). Massive parallel sequencing of DNA libraries was performed on the DNBSEQ-G400 sequencer (MGI, China) with an average depth of 500X for FFPE and 500X for WBC samples.

In the experiment to validate TMB, DNA libraries were also hybridized to the xGen Exome Hyb Panel v2 (IDT, USA) that covered the entire human exome. Sequencing was performed on the DNBSEQ-G400 sequencer (MGI, China), with the average depth for FFPE and WBC samples of 250X and 175X, respectively. In the experiments to compare K4C assay with the commercial TruSight Oncology 500 (Illumina, USA), the library preparation and target enrichment for FFPE samples were performed following the manufacturer's protocol. DNA libraries were sequenced on the NextSeq 550 (Illumina, USA) in paired-end 150-cycle runs, with the average depth of 1000X.

Bioinformatics analysis

For quality control, we used Illumina DRAGENTM Bio-IT Platform (v3.10) to trim low-quality bases and adapter sequences in fastq files and to remove duplicated reads from the bam files. FFPE samples that passed quality control (QC) of the K4C assay must have a mean depth greater than or equal to 150X and a percent coverage at 100X exceeding or equal to 75%. The failure rate was 10.4% and samples that failed QC were removed from analysis.

Somatic variants were called from bam files by the DRAGENTM Somatic pipeline, using matched tumor-normal pairs. To reduce the number of artifacts related to formalin treatment and tissue sample processing, post somatic calling filtering was performed including ignoring soft-clip bases and removing clustered events. VEP (version 105) (McLaren et al., 2016) was used to predict the effects of variants and annotate them against dbSNP (Sherry et al., 2001), ClinVar (Landrum et al., 2016), and COSMIC (Tate et al., 2019) databases. For gene amplification, DRAGENTM Somatic CNV Calling pipeline was utilized. Ninety WBC samples were used to build a reference baseline, from which the normalized copy ratio value was calculated. The gain or loss of the targeted genes was called, with bin segment = 100 bp. Gene amplification was detected when the copy ratio of segment mean ≥ 2.0 . For gene fusion, the aligned bam files from DRAGENTM with the marked soft-clipped reads were analyzed using Factera (v1.4.4) (Newman et al., 2014) with default

parameters. Fusions were detected when minimal 2 breakpoint-spanning reads were reported. The OncoKB database (Chakravarty et al., 2017) (last updated 10/02/2023) was then used to identify actionable alterations eligible for FDA-approved drugs (level 1–2).

Germline variants were called from bam files of WBC samples by the DRAGENTM Germline pipeline. Variants were then classified according to the guidelines of The American College of Medical Genetics and Genomics (ACMG) as previously described [(Tran et al., 2021; Nguyen et al., 2023)]. Pathogenic and likely pathogenic variants were reported.

TMB calculation was performed using our in-house developed script to divide the total number of eligible somatic variants by the size of the interrogated panel. An eligible somatic variant must meet all of the following criteria: (1) pass filtering parameters of the variant calling pipeline, (2) not likely a germline variant as filtered by the dbSNP database (Sherry et al., 2001), (3) locate within the coding region, (4) not a synonymous mutation, (5) have VAF $\geq 5\%$, allele depth $\geq 5X$, total depth $\geq 15X$. The panel size was counted for the bases within coding regions with the minimal total depth $\geq 15X$. The threshold for TMB-High (TMB-H) was determined as 10 mutations/Mb. For MSI status, unstable microsatellite loci were detected by MSIsensor-pro (v1.2.0) (Jia et al., 2020) in matched tumor-normal mode. If the proportion of unstable loci among all detected microsatellite loci was at least 20%, the sample was determined as MSI-High (MSI-H).

For samples analyzed using the TruSight Oncology 500 kit, the TSO500 analysis pipeline was executed utilizing the DRAGENTM TSO500 (v2.1) software on the Illumina DRAGENTM Bio-IT Platform (v3.10). One sample (1/40) that failed the recommended QC metrics according to the TSO500 manufacturer's guidelines was excluded from the analysis.

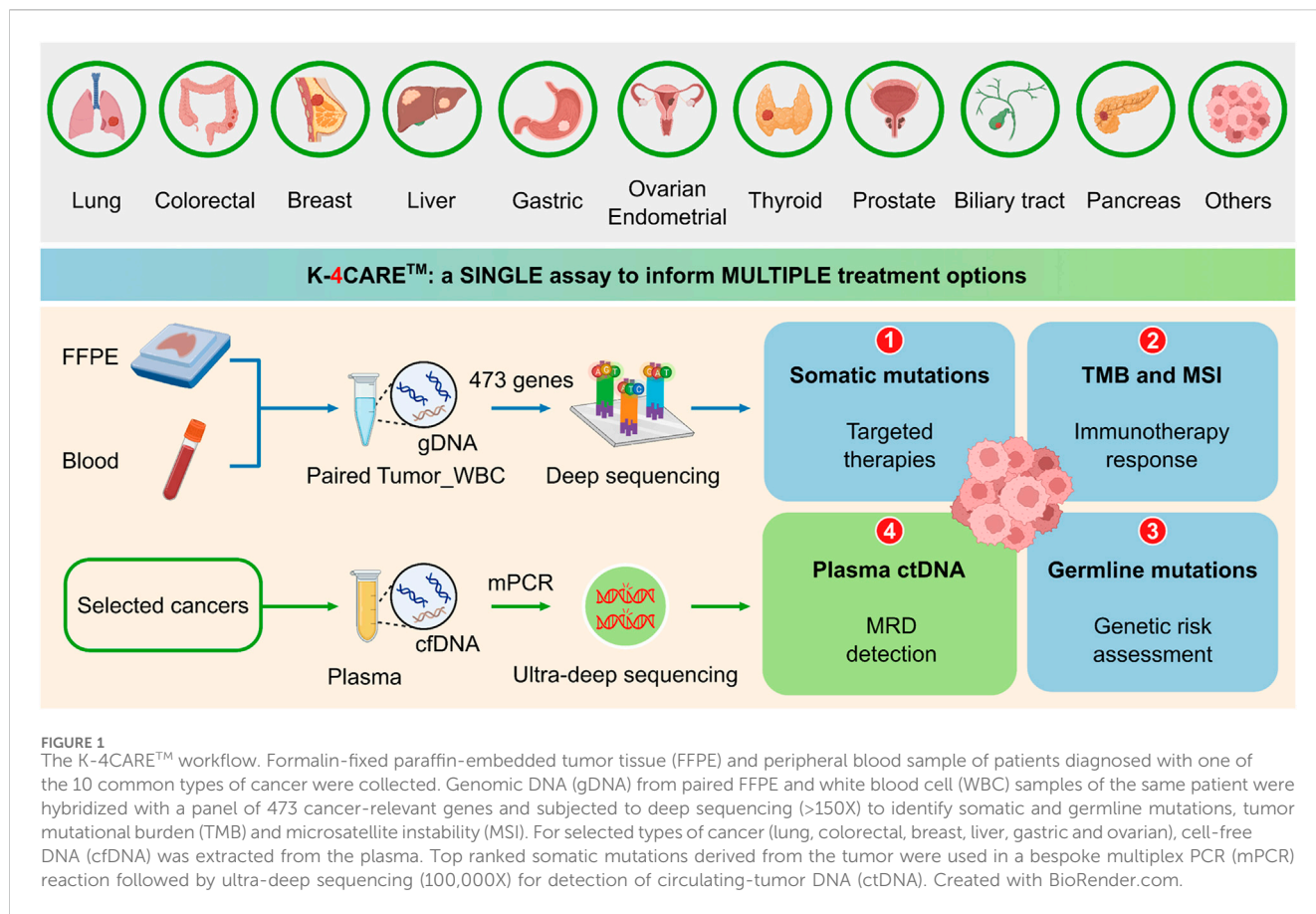
MSI analysis by PCR

DNA from paired FFPE and WBC samples were subjected to fluorescent multiplex PCR reactions using the MSI Analysis System, Version 1.2 (Promega, USA) to examine 5 mononucleotide repeat markers: BAT-25, BAT-26, MONO-27, NR-21 and NR-24. Samples were run on the SeqStudioTM genetic analyzer (ThermoFisher, USA). Those with at least 2 altered markers were classified as MSI-H, while others were classified as microsatellite stable (MSS).

Plasma ctDNA detection

All FFPE-derived mutations were ranked using our proprietary scoring algorithm as previously described (Nguyen et al., 2022; Nguyen Hoang et al., 2023). The main criteria included VAF in FFPE, predicted pathogenicity and deleteriousness, mutation type, reported frequency in large cancer databases, and internal validation as a tumor-derived mutation. The top ranked mutations specific to each patient were then used for ctDNA tracking in the plasma samples.

cfDNA was already extracted from plasma samples as previously described (Nguyen et al., 2022; Nguyen Hoang et al., 2023). The minimum amount of cfDNA required was ≥ 0.15 ng/ μ L or total of ≥ 3 ng. cfDNA fragments with the selected mutation sites were



amplified using KAPA HiFi DNA Polymerase (Roche, USA) and targeted primer pairs (PhuSa Biochem, Vietnam). Amplified cfDNA fragments were sequenced on the NextSeq 2000 system (Illumina, USA) with an average depth of 100,000X per amplicon. Amplicons with less than 10,000X coverage were considered unsuccessful. Variant calling was performed using mpileup from Samtools (v1.11) as previously described (Nguyen et al., 2022; Nguyen Hoang et al., 2023). A plasma sample that had at least one mutation with VAF above 0.05% was defined as ctDNA positive.

Statistical analysis

The degree of linearity between K4C results and reference standards or benchmarked methods was assessed using R2 values derived from the “lm” linear model method within the R v4.3.0 Stats package. All graphical representations were generated using Rstudio server version 2022.07.02 (Rstudio Team, USA).

Results

Assay workflow

The K4C curated gene panel of 473 cancer-relevant genes included the entire exons, selected promoter and intron regions, with the total genomic size of 1.7 Mb (Supplementary Table S3). FFPE and blood

samples were required for the analysis (Figure 1). Genomic DNA from tumor and matching WBC samples were sequenced to profile all somatic and germline variants in 473 genes. The results could inform physicians of 1) actionable and resistant somatic mutations predictive of response to FDA-approved targeted therapies; 2) TMB and MSI status to guide immunotherapy decision, 3) pathogenic or likely pathogenic germline variants for genetic risk assessment. For selected 6 types of cancer, all tumor-derived mutations were further ranked by our proprietary algorithm and the top mutations were used to detect ctDNA presence in plasma using a bespoke mPCR assay. This result was used to 4) detect MRD and monitor treatment response for cancer patients.

Analytical validation was first performed using reference standards to determine the accuracy and reproducibility of the K4C assay. Clinical samples from more than 10 types of cancer were then used to further evaluate the performance of the assay, especially when benchmarked against other orthogonal methods. The overall performance of the K4C assay is summarized in Table 1.

Detection of somatic and germline variants

The reference standard, OncoSpan gDNA, had 228 verified variants (211 SNVs, 17 INDELS) within interrogated region of our K4C gene panel. Sensitivity of K4C assay was estimated from 11 replicates: two to three replicates/run for 4 independent runs. For both SNV and Indel detection, sensitivity was estimated to be >99%

TABLE 1 Overall performance of K-4CARE™ assay.

		Sensitivity	Specificity (%)
Targeted therapies (FFPE)	Somatic variants (at VAF ≥5%)*		
	SNV	>99%	>99%
	INDEL	>99%	>99%
	Amplification	>99%	>99%
	Fusion	94%	>99%
Hereditary cancers (WBC)	Germline variants		
	SNV	>99%	>99%
	INDEL	>99%	>99%
Immunotherapy (FFPE)	MSI	94%	>99%
	TMB	Correlation R-squared	
		97%	
Minimal residual disease (Plasma—Selected cancers)	ctDNA	Limit of detection	
		VAF ≥0.05%	

*For somatic SNVs, and INDELS, variants at 1% ≤ VAF <5% were detected with sensitivity 82%, specificity >99%.

Abbreviations: ctDNA: circulating tumor DNA, FFPE: formalin fixed paraffin-embedded tissue, Indel: insertion/deletion, MSI: microsatellite instability, SNV: single nucleotide variation, TMB: tumor mutational burden, VAF: variant allele frequency, WBC: white blood cells.

Bold names on the left are different categories of genetic markers.

for VAF ≥5% (Supplementary Table S4). There was high correlation ($R^2 > 0.90$) between the expected VAFs and VAFs determined by the K4C assay (Figure 2A, Supplementary Table S5). Specificity was established at >99% for somatic SNV and Indel detection using TruQ 0 standard (Supplementary Table S6). Limit of Detection (LOD) of the assay was defined as the minimum VAF detectable in at least 95% of technical replicates using the same input material. Using the OncoSpan standard, LOD for SNVs and Indels was determined at 5% VAFs (Supplementary Table S7). Variants at 1% ≤ VAF <5% were detected with sensitivity of 82% and specificity of >99% (Supplementary Tables S6, S7). The high repeatability and reproducibility of the assay using the OncoSpan standard was shown in Supplementary Table S8.

Next, we compared the performance to detect SNVs and Indels of the K4C assay with the TruSight Oncology 500 (TSO500) using 39 clinical samples. Concordance between the K4C and TSO500 assays to detect actionable SNV and Indel variants was 100%. The correlation of VAFs determined by two assays was 0.93 (Figure 2B). The high reproducibility of VAFs measured by the K4C assay was also demonstrated in clinical samples measured in 4 independent replicates (Figure 2C, Supplementary Table S9).

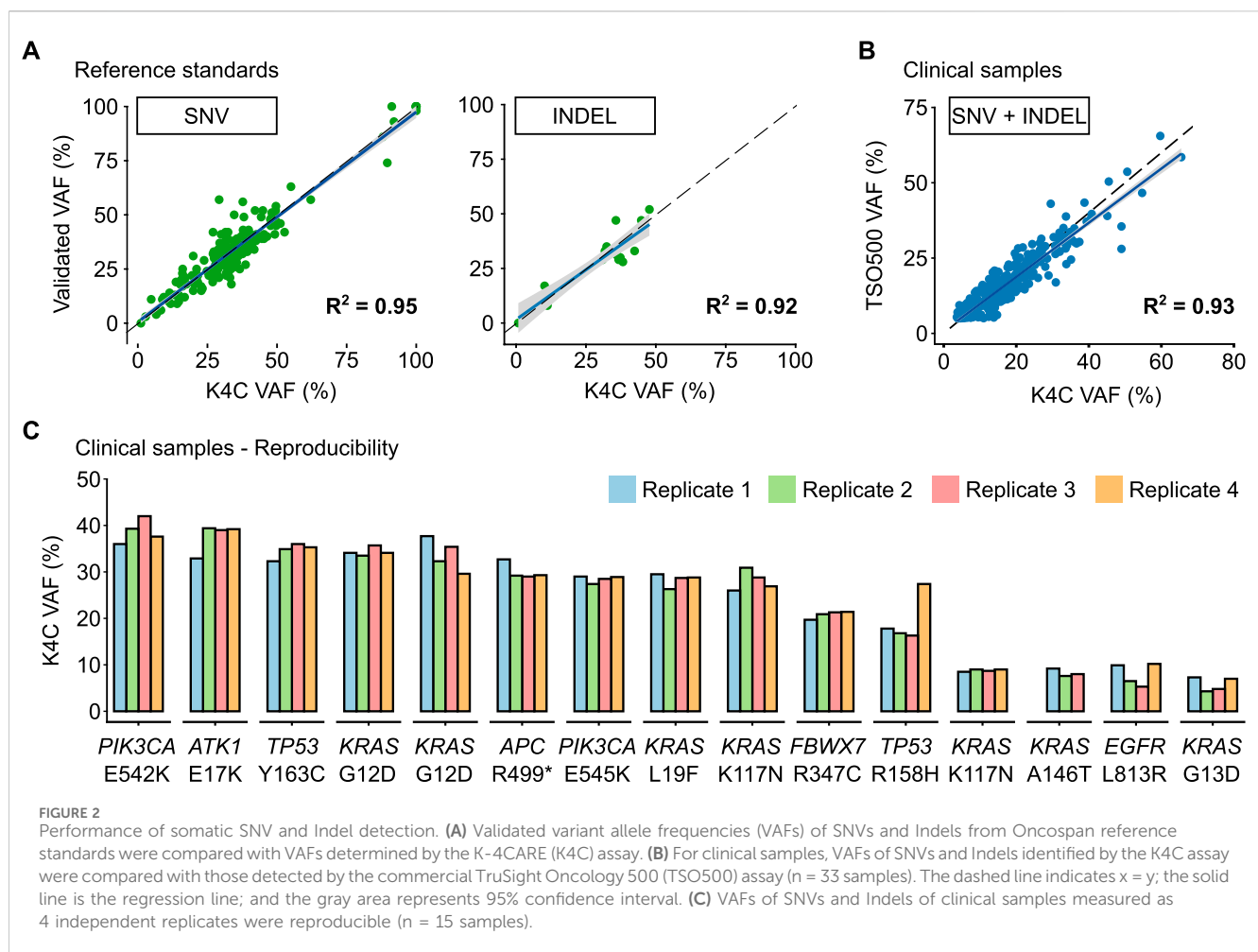
For amplification, we used Breast CNV Mix and Lung and Brain CNV Mix standards comprising of 6 verified CNVs (ERBB2, FGFR3, MYC, EGFR, MET, MYCN) with 3 additional copies each. The K4C assay was able to identify amplification alterations in all six genes while it did not detect amplification in negative reference standards (Supplementary Table S10). For gene fusion, Pan Cancer 6 Fusion standard that had six types of fusions: TPM3-NTRK1, QKI-NTRK2, ETV6-NTRK3, EML4-ALK, CCDC6-RET, and SLC34A2-ROS1 was used. K4C assay detected 94% of fusion events in 3 independent replicates, while it did not detect any in negative reference standards (Supplementary Table S11).

For germline variants, the K4C assay detected both SNV and Indel variants from the reference standard, BRCA germline I, with sensitivity and specificity of >99%; and with high correlation between expected VAFs and VAFs determined by the K4C assay (Supplementary Table S12).

Tumor mutational burden (TMB)

K4C assay determined the TMB of SeraSeq® TMB-7 and -13 standards as 7.7 and 14.2 respectively, which were within the recommended range from the manufacturer. Since these standards have a tumor fraction of 100%, which rarely applies to clinical samples, we next evaluated the stability of TMB values at lower tumor fractions and different sequencing performance. Fastq data of tumor standards were mixed with WBC standards at the ratios of 3:1, 1:1, 2:3, 3:7, 1:4, and 1:9 to create simulated tumor fractions ranging from 75% to 10%; followed by random read sampling to vary median depth and percentage coverage at 100X across different tumor fractions. We found that TMB values remained stable when tumor fraction was at least 30%, mean depth of at least 150X, and a percent coverage at 100X exceeding or equal to 75% (Figure 3A).

We then determined TMB values of 127 clinical samples using either K4C targeted panel or WES, which is the gold standard for TMB measurement. The correlation co-efficiency of TMB values between the 2 methods was 0.97 for all clinical samples, and 0.94 for samples with TMB <60 (Figure 3B, Supplementary Table S13). When benchmarked against the TSO500 assay using 39 clinical samples, the concordance to determine TMB status of the K4C assay was 89% (Figure 3C, Supplementary Table S14). TMB scores were highly reproducible in clinical samples analyzed in 4 independent runs (Figure 3D).



Microsatellite instability (MSI)

MSI status was first determined for 158 colorectal and gastric cancers, using PCR and fragment analysis for 5 microsatellite markers recommended by the National Cancer Institute (Boland et al., 1998). K4C assay was performed for randomly selected 31 samples, including both MSI-H and MSS samples. The concordance of K4C assay to determine MSI status was 97% compared to the PCR method (Figure 4A, Supplementary Table S15). When benchmarked against the TSO500 assay using 39 clinical samples, the concordance between K4C and TSO500 for MSI measurement was 98% (Figure 4B, Supplementary Table S14). MSI scores determined by the K4C assay were highly reproducible in clinical samples analyzed in 4 independent runs (Figure 4C).

Clinical utilization

In order to illustrate the clinical usefulness of the K4C assay, we identified actionable alterations predictive of treatment response to FDA-approved drugs (OncoKB level 1 and 2 only). The percentage of samples with at least 1 actionable mutation was 68.4%, 57.1%, 50%, 47.1%, 9.1% and 7.1% in thyroid, breast, lung, CRC, ovarian and cervical cancers, respectively (Figure 5A, Supplementary Table

S16). The clinical evaluation of TMB showed that TMB score was highest in CRC, gastric and lung cancers (Figure 5B). For MSI, the prevalence of MSI-H was 16.9%, 9.1%, 5.3% and 3.4% in CRC, ovarian, thyroid and gastric cancers (Figure 5C). We also examined the relationship between TMB and MSI status in clinical samples (Supplementary Table S17). Majority of the samples with concurrent TMB-H and MSI-H were CRC and gastric cancer. Several samples had TMB-H but MSS status, including lung and liver cancers (Figure 5D).

Detection of plasma ctDNA

In 40 patients with either CRC or breast cancer, the total number of tumor-specific mutations detected were 4257 variants. Using in-house ranking algorithm, we selected the top two to five variants unique for each patient and amplified them in bespoke mPCR assays to detect ctDNA in the plasma. Out of total 127 variants selected from tumors, 100 variants (78.7%) were detected in plasma samples (Figure 6A). There was no correlation between VAF of a variant in FFPE tissue and its VAF or detection status in the plasma (Figure 6B). When the same set of mutations were used to monitor ctDNA dynamics in longitudinal plasma samples, we could detect MRD and molecular relapse ahead of clinical diagnosis. The 2 case studies were presented to illustrate the

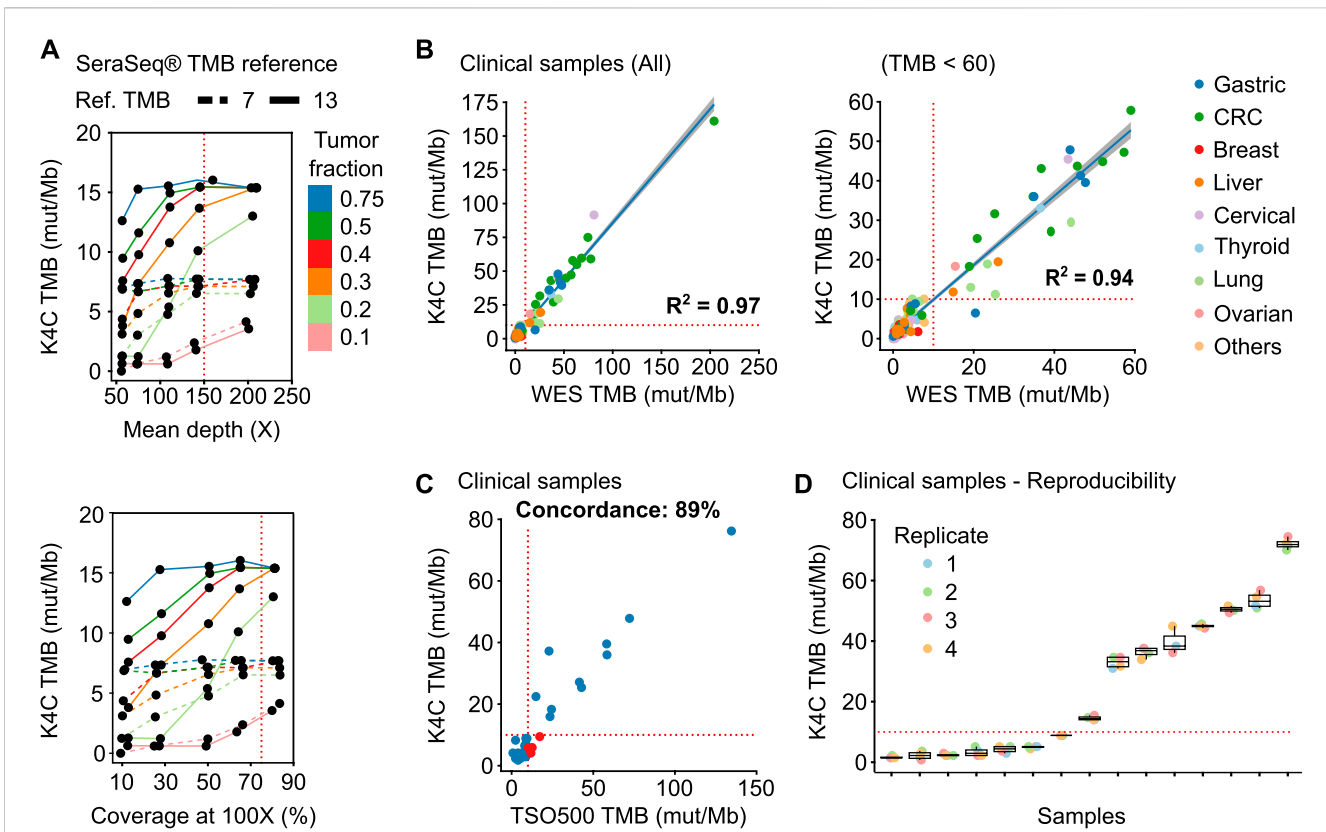


FIGURE 3 Performance of TMB measurement. (A) TMB values of the SeraSeq® TMB reference standards were measured by K4C assay when the following parameters varied: tumor fraction, mean depth and percent coverage at 100X. (B) Correlation of TMB scores determined using the K4C gene panel to those determined by whole exome sequencing (WES) in all clinical samples (n = 127 samples), or samples with TMB < 60 (n = 121 samples). (C) Concordance of TMB in clinical samples as determined by the K4C assay and the commercial TruSight Oncology 500 (TSO500) assay (n = 35 samples). (D) TMB scores of clinical samples measured as 4 independent replicates by the K4C assay were reproducible (n = 15 samples).

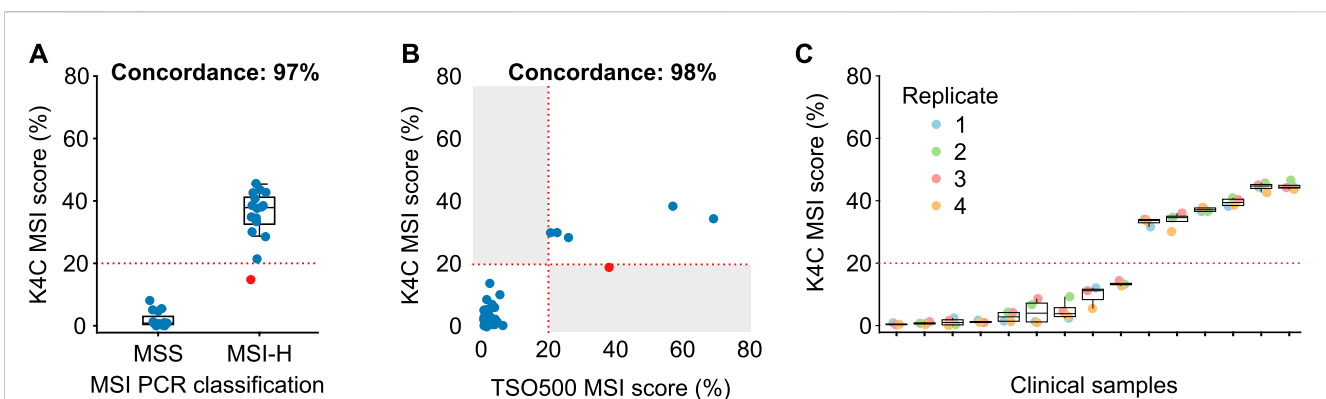
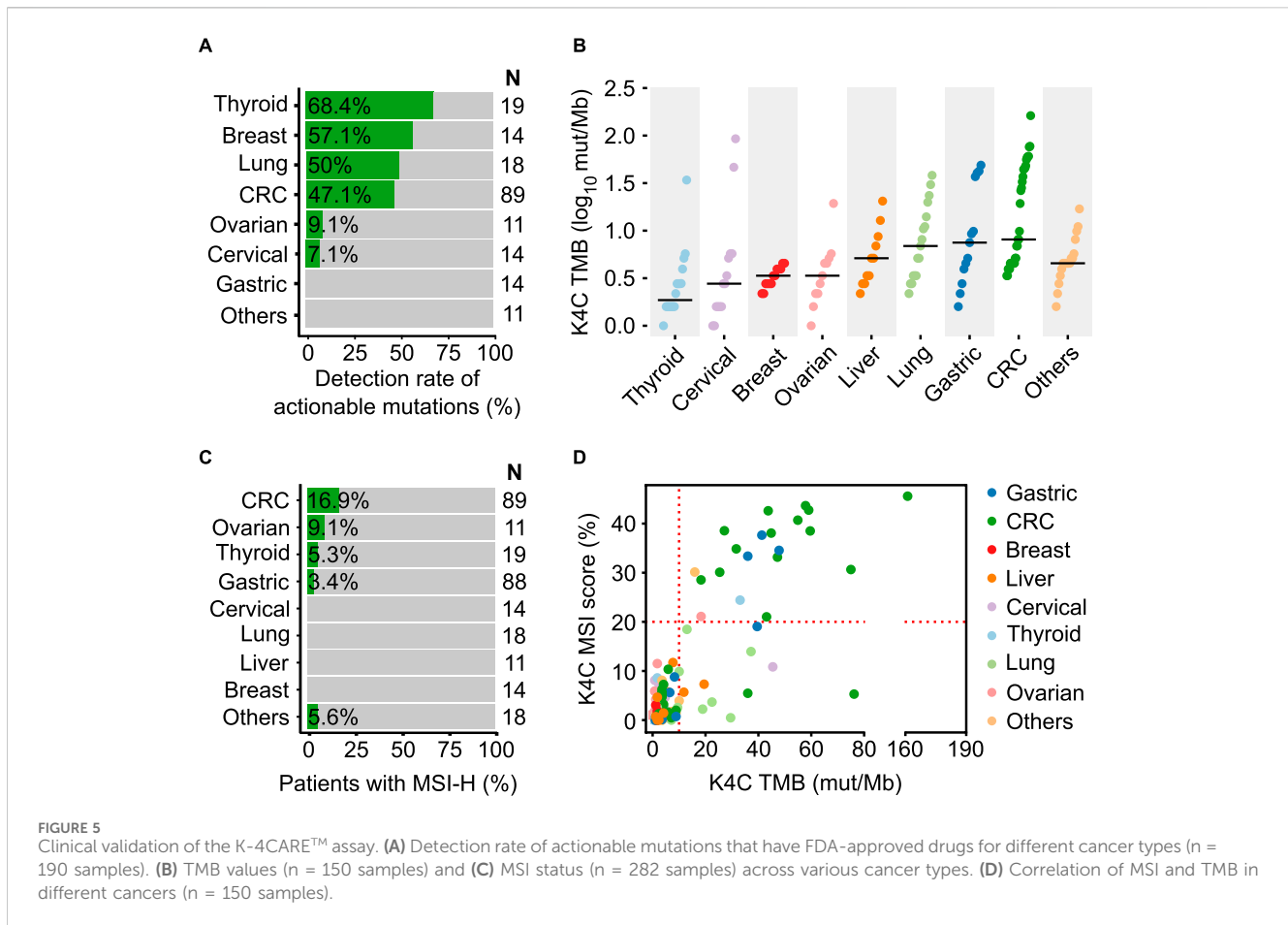


FIGURE 4 Performance of MSI measurement. (A) Concordance of MSI status determined by the K4C assay and by PCR-fragment analysis using gastric and colorectal cancers (n = 31 samples). (B) Concordance of MSI status determined by the K4C assay and the commercial TruSight Oncology 500 (TSO500) assay for clinical samples (n = 39 samples). (C) MSI scores of clinical samples measured as 4 independent replicates by the K4C assay were reproducible (n = 15 samples).

clinical utility of the test. First, patient K4C012 was diagnosed with stage I, triple-negative breast cancer. ctDNA was detected in the pre-operative plasma, which was cleared after surgery but returned positive after adjuvant chemoradiotherapy (CRT). Clinical examination only confirmed cancer relapse 10 months after

ctDNA was found positive (Figure 6C). Second, patient K4C077 was diagnosed with T4aN2M0 colon cancer. After surgery, VAF of ctDNA drastically declined but the patient remained ctDNA-positive, indicating presence of residual cancer cells. Despite CRT, he was diagnosed with liver metastasis



10.5 months later. Levels of carcinoembryonic antigen (CEA), the current biomarker used to monitor CRC patients, were normal both before and after surgery, and only rised at 6 months after ctDNA was found positive (Figure 6D).

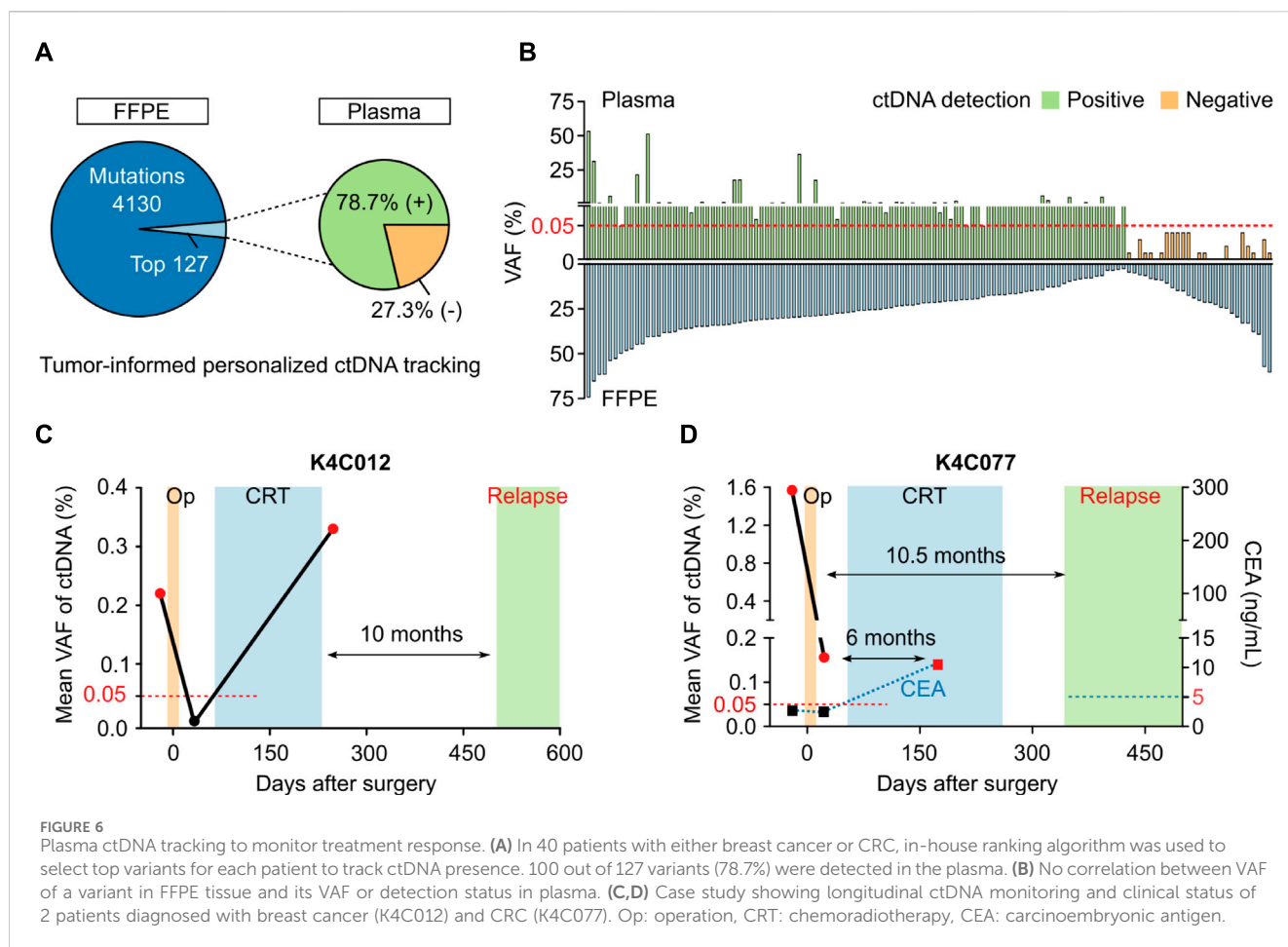
Discussion

In this study, we presented the analytical validation and clinical utilization of K-4CARE™, a tumor-based CGP assay that enables the detection of somatic variants, germline variants, TMB and MSI measurement, as well as ctDNA monitoring, all within a single platform. Performance metrics of the assay have been rigorously established and validated using both reference standards and orthogonal methods.

Firstly, the K4C assay could detect 4 types of somatic variants with high sensitivity and specificity, using tumor DNA only. Unlike other CGP assays that used RNA to detect gene rearrangement (White et al., 2021; Ng et al., 2022; Saldivar et al., 2023), we designed probes to target selected intron regions of commonly fused genes, and DNA libraries were used to hybridize with these probes. Our bioinformatic pipeline to detect gene fusion was robust and similar to those developed in previous studies (Wang et al., 2023; Xia et al., 2023; Ye et al., 2023). Since this eliminates the necessity to process tumor RNA, it expedites the result and makes the assay more cost-effective and feasible, particularly for developing countries like Vietnam where the quality of FFPE-derived

RNA is often not sufficient for NGS (in-house data). Furthermore, we demonstrated clinical usefulness of our assay to identify at least 1 actionable mutations (OncoKB level 1–2) in 38.9% of the clinical samples (74/190). This rate seemed to be higher than other CGP assays applied in other cohorts, such as 12.0% in Ng et al (Ng et al., 2022), 10.7% in Aoyagi et al (Aoyagi et al., 2022) and 23.7% in Karol et al (Karol et al., 2021). It was likely attributed to the distribution of samples across different cancer types in each cohort. Our analysis included more CRC (46.8%) and thyroid cancer (10.0%), both of which have high prevalence of actionable mutations, compared to other cohorts. Particularly for papillary thyroid cancer, we identified the prevalence of BRAF V600E mutations in 68.4% of the samples, which was at the high end of the reported range (32%–65%) (Xing et al., 2013). Since the sample size in each cancer type was small, this result was solely used to illustrate the clinical utilization of the assay, rather than providing representative mutational spectrum or true level of actionability. More comprehensive analysis of actionable mutation profiles in substantially larger cohorts of lung, breast and colorectal cancer was reported in our previous studies (Dang et al., 2020; Nguyen et al., 2022; Nguyen Hoang et al., 2023).

For TMB measurement, we included non-synonymous mutations in a genomic region of approximately 1.7 Mb, larger than the recommended 1 Mb to ensure accuracy and stability of TMB (Buttner et al., 2019; Bevins et al., 2020). The validation process has been aligned with current standards and previous CGP assays (Merino et al., 2020; White et al., 2021; Ng et al., 2022; Saldivar et al., 2023),



showing high correlation between K4C panel and the gold standard WES ($R^2 = 0.97$). The concordance of TMB measurement between K4C (tumor-normal) and TSO500 (tumor-only) assays was 89%, slightly lower than the concordance among tumor-only panels (>90%) (Bevins et al., 2020). We observed that all non-concordant samples had higher TSO500 TMB compared to the K4C value, which was also reported for a different tumor-normal CGP assay (Saldivar et al., 2023). Such discrepancy was likely due to the absence of matched normal sample and the inclusion of synonymous mutations in the TMB calculation algorithm of the TSO500 assay (Saldivar et al., 2023). In addition, among our clinical samples, cancer types with the highest TMB scores were CRC, gastric and lung cancers. This result was consistent with previous studies (Lawrence et al., 2013; Wu et al., 2022; Saldivar et al., 2023) but different from the Ng et al study reporting the highest TMB in liver and breast cancers in an Asian cohort (Ng et al., 2022).

For MSI analysis, IHC and PCR-based assay examining 5 mononucleotide repeat markers remain the first choices in cancers belonging to the spectrum of Lynch syndrome such as CRC and endometrial cancer (Luchini et al., 2019). MSI evaluation by NGS is also recommended as an option for these cancers for the additional benefit of examining TMB, another biomarker for ICIs (Luchini et al., 2019). For all cancers not related to Lynch syndrome, NGS is the method of choice to determine MSI as the reliability of IHC and the 5 markers by PCR has not been proven (Luchini et al., 2019). In our analysis using CRC and gastric cancer samples, we

demonstrated excellent concordance of MSI measured by PCR for 5 repeat markers and NGS examining ~800 repeat loci, indicating the reliability and accuracy of the NGS technique. The prevalence of MSI-H was 16.9% in CRC, similar to the rate reported previously (Bonneville et al., 2017; Middha et al., 2017). The prevalence of MSI-H in our gastric cancer samples (3.4%) was lower than that in Bonneville et al (19.9%) (Bonneville et al., 2017) but comparable to other large cohorts (2.5–3.4%) (Middha et al., 2017; Trabucco et al., 2019). The surprisingly high incidence of MSI-H in ovarian and thyroid cancers in our cohort needs to be further investigated with a larger sample size.

Since PD-L1, MSI and TMB have been approved as independent biomarkers for ICIs, the relationship among these three has been extensively examined in previous studies. PD-L1 was found to have low concordance with TMB and MSI, while the relationship between TMB and MSI varied significantly by cancer type (Luchini et al., 2019; Wu et al., 2022). The highest concordance of TMB and MSI was found in CRC and gastric cancer, and almost none in lung cancer (Luchini et al., 2019; Wu et al., 2022). This completely agreed with our observation in this study. Moreover, similar to those studies, we also found TMB-H could occur in MSS samples across different cancer types, while the incidence of MSI-H in TMB-L samples was rare. This again emphasizes the importance of comprehensive testing for multiple biomarkers to maximize the treatment opportunity for patients. It should also be noted that the interpretation of these biomarkers has to be cancer type-specific, as

the cut-off for TMB-H at 10 mut/Mb, for example, has been shown to be only applicable for lung carcinoma (Samstein et al., 2019).

Lastly, a unique application of the K4C assay is the ability to use personalized mutations derived from tumor to track for the presence of ctDNA in plasma. The analytical validation of the mPCR technique was already elaborated in our previous studies (Nguyen et al., 2022; Nguyen Hoang et al., 2023). At the LOD of 0.05%, the assay robustly and accurately identified ctDNA in the plasma using mutations ranked by our in-house algorithm. The prognostic and predictive value of ctDNA in routine monitoring of cancer patients during and after treatment have been consistently demonstrated in several clinical trials (Chin et al., 2019). Specifically for ICIs, ctDNA clearance after 2 cycles of treatment was shown as a strong indicator of long-lasting response in multiple solid tumors (Cabel et al., 2017; Bratman et al., 2020).

Taken together, the data presented in this study highlights the accuracy and reliability of the K4C assay in providing comprehensive genomic alterations and signatures to inform targeted- and immuno- therapies for cancer patients. The assay could also detect ctDNA in the plasma, that can later be used to routinely monitor treatment response and detect residual cancer.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1035299/>.

Ethics statement

The studies involving humans were approved by the institutional ethics committee of the University of Medicine and Pharmacy, Ho Chi Minh City (approval number 164/HDDD). The use of plasma samples for ctDNA tracking was under the studies approved by the ethics committee of the University of Medicine and Pharmacy, Ho Chi Minh City (approval number 14/GCN-HDDD for colorectal cancer, 300/HDDD for breast cancer). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

T-PN: Writing–original draft, Writing–review and editing, Conceptualization, Formal Analysis, Investigation, Methodology, Visualization. TN: Data curation, Formal Analysis, Methodology, Software, Writing–review and editing. NT: Data curation, Visualization, Writing–review and editing, Formal Analysis,

Methodology, Writing–original draft. V-AN: Data curation, Visualization, Writing–review and editing. HT: Data curation, Writing–review and editing. V-UT: Data curation, Writing–review and editing. YN: Data curation, Writing–review and editing. AN: Data curation, Writing–review and editing. CN: Data curation, Writing–review and editing. TD: Formal Analysis, Writing–review and editing. DN: Formal Analysis, Writing–review and editing. H-NN: Formal Analysis, Writing–review and editing, Supervision. HG: Formal Analysis, Software, Writing–review and editing, Supervision. LT: Conceptualization, Formal Analysis, Investigation, Methodology, Project administration, Supervision, Writing–original draft, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was funded by Gene Solutions, Vietnam. The funder did not have any role in the study design, data collection and analysis, or preparation of the manuscript.

Acknowledgments

We thank Hoa Phan, Medical Genetics Institute, Vietnam for assisting in manuscript preparation.

Conflict of interest

Authors T-PN, TN, NT, V-AN, AN, HT, V-UT, YN, AN, CN, DN, H-NN, HG, and LT was employed by Gene Solutions.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2024.1334808/full#supplementary-material>

References

Aoyagi, Y., Kano, Y., Tohyama, K., Matsudera, S., Kumaki, Y., Takahashi, K., et al. (2022). Clinical utility of comprehensive genomic profiling in Japan: result of PROFILE-F study. *PLoS One* 17, e0266112. doi:10.1371/journal.pone.0266112

Bevins, N., Sun, S., Gaieb, Z., Thorson, J. A., and Murray, S. S. (2020). Comparison of commonly used solid tumor targeted gene sequencing panels for estimating tumor mutation burden shows analytical and prognostic concordance within the cancer

- genome atlas cohort. *J. Immunother. Cancer* 8, e000613. doi:10.1136/jitc-2020-000613
- Boland, C. R., Thibodeau, S. N., Hamilton, S. R., Sidransky, D., Eshleman, J. R., Burt, R. W., et al. (1998). A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.* 58, 5248–5257.
- Bonneville, R., Krook, M. A., Kautto, E. A., Miya, J., Wing, M. R., Chen, H. Z., et al. (2017). *Landscape of microsatellite instability across 39 cancer types*. *JCO precision oncology 2017*. United States: Lippincott Williams and Wilkins Ltd.
- Bratman, S. V., Yang, S. Y. C., Iafolla, M. A. J., Liu, Z., Hansen, A. R., Bedard, P. L., et al. (2020). Personalized circulating tumor DNA analysis as a predictive biomarker in solid tumor patients treated with pembrolizumab. *Nat. cancer* 1, 873–881. doi:10.1038/s43018-020-0096-5
- Buttner, R., Longshore, J. W., Lopez-Rios, F., Merkelbach-Bruse, S., Normanno, N., Rouleau, E., et al. (2019). Implementing TMB measurement in clinical practice: considerations on assay requirements. *ESMO Open* 4, e000442. doi:10.1136/esmoopen-2018-000442
- Cabel, L., Riva, F., Servois, V., Livartowski, A., Daniel, C., Rampanou, A., et al. (2017). Circulating tumor DNA changes for early monitoring of anti-PD1 immunotherapy: a proof-of-concept study. *Ann. Oncol.* 28, 1996–2001. doi:10.1093/annonc/mdx212
- Chakravarty, D., Gao, J., Phillips, S. M., Kundra, R., Zhang, H., Wang, J., et al. (2017). *OncokB: a precision oncology knowledge base*. *JCO precision oncology 2017*. United States: Lippincott Williams and Wilkins Ltd.
- Chin, R. I., Chen, K., Usmani, A., Chua, C., Harris, P. K., Binkley, M. S., et al. (2019). Detection of solid tumor molecular residual disease (MRD) using circulating tumor DNA (ctDNA). *Mol. diagnosis Ther.* 23, 311–331. doi:10.1007/s40291-019-00390-5
- Christofyllakis, K., Bittenbring, J. T., Thurner, L., Ahlgrim, M., Stilgenbauer, S., Bewarder, M., et al. (2022). Cost-effectiveness of precision cancer medicine-current challenges in the use of next generation sequencing for comprehensive tumour genomic profiling and the role of clinical utility frameworks (Review). *Mol. Clin. Oncol.* 16, 21. doi:10.3892/mco.2021.2453
- Dang, A. H., Tran, V. U., Tran, T. T., Thi Pham, H. A., Le, D. T., Nguyen, L., et al. (2020). Actionable mutation profiles of non-small cell lung cancer patients from Vietnamese population. *Sci. Rep.* 10, 2707. doi:10.1038/s41598-020-59744-3
- Jia, P., Yang, X., Guo, L., Liu, B., Lin, J., Liang, H., et al. (2020). MSIsensor-pro: fast, accurate, and matched-normal-sample-free detection of microsatellite instability. *Genomics, proteomics Bioinforma.* 18, 65–71. doi:10.1016/j.gpb.2020.02.001
- Karol, D., McKinnon, M., Mukhtar, L., Awan, A., Lo, B., and Wheatley-Price, P. (2021). The impact of foundation medicine testing on cancer patients: a single academic centre experience. *Front. Oncol.* 11, 687730. doi:10.3389/fonc.2021.687730
- Landrum, M. J., Lee, J. M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., et al. (2016). ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic acids Res.* 44, D862–D868. doi:10.1093/nar/gkv1222
- Lawrence, M. S., Stojanov, P., Polak, P., Kryukov, G. V., Cibulskis, K., Sivachenko, A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499, 214–218. doi:10.1038/nature12213
- Luchini, C., Bibeau, F., Ligtenberg, M. J. L., Singh, N., Nottegar, A., Bosse, T., et al. (2019). ESMO recommendations on microsatellite instability testing for immunotherapy in cancer, and its relationship with PD-1/PD-L1 expression and tumour mutational burden: a systematic review-based approach. *Ann. Oncol.* 30, 1232–1243. doi:10.1093/annonc/mdz116
- McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., et al. (2016). The ensembl variant effect predictor. *Genome Biol.* 17, 122. doi:10.1186/s13059-016-0974-4
- Merino, D. M., McShane, L. M., Fabrizio, D., Funari, V., Chen, S. J., White, J. R., et al. (2020). Establishing guidelines to harmonize tumor mutational burden (TMB): *in silico* assessment of variation in TMB quantification across diagnostic platforms: phase I of the Friends of Cancer Research TMB Harmonization Project. *J. Immunother. Cancer* 8, e000147. doi:10.1136/jitc-2019-000147
- Middha, S., Zhang, L., Nafa, K., Jayakumaran, G., Wong, D., Kim, H. R., et al. (2017). Reliable pan-cancer microsatellite instability assessment by using targeted next-generation sequencing data. *JCO Precis. Oncol.* 2017, 1–17. doi:10.1200/PO.17.00084
- Newman, A. M., Bratman, S. V., Stehr, H., Lee, L. J., Liu, C. L., Diehn, M., et al. (2014). FACTERA: a practical method for the discovery of genomic rearrangements at breakpoint resolution. *Bioinforma. Oxf. Engl.* 30, 3390–3393. doi:10.1093/bioinformatics/btu549
- Ng, C. C., Lim, S., Lim, A. H., Md Nasir, N. D., Zhang, J., Rajasegaran, V., et al. (2022). A comprehensive next generation sequencing tissue assay for Asian-prevalent cancers-Analytical validation and performance evaluation with clinical samples. *Front. Mol. Biosci.* 9, 963243. doi:10.3389/fmolb.2022.963243
- Nguyen, H. T., Lu, Y. T., Tran, D. H., Tieu, B. L., Le, K. T., Pham, T. N., et al. (2023). Prevalence and genetic spectrum associated with hereditary colorectal cancer syndromes, the need to improve cancer risk awareness, and family cascade testing in Vietnam. *Fam. cancer* 22, 449–458. doi:10.1007/s10689-023-00344-1
- Nguyen, H. T., Nguyen, T. V., Nguyen Hoang, V. A., Tran, D. H., Le Trinh, N. A., Le, M. T., et al. (2022). Tumor genomic profiling and personalized tracking of circulating tumor DNA in Vietnamese colorectal cancer patients. *Front. Oncol.* 12, 1069296. doi:10.3389/fonc.2022.1069296
- Nguyen Hoang, V. A., Nguyen, S. T., Nguyen, T. V., Pham, T. H., Doan, P. L., Nguyen Thi, N. T., et al. (2023). Genetic landscape and personalized tracking of tumor mutations in Vietnamese women with breast cancer. *Mol. Oncol.* 17, 598–610. doi:10.1002/1878-0261.13356
- Pennell, N. A., Arcila, M. E., Gandara, D. R., and West, H. (2019). Biomarker testing for patients with advanced non-small cell lung cancer: real-world issues and tough choices. *Annu. Meet.* 39, 531–542. doi:10.1200/EDBK_237863
- Radich, J. P., Briercheck, E., Chiu, D. T., Menon, M. P., Sala Torra, O., Yeung, C. C. S., et al. (2022). Precision medicine in low- and middle-income countries. *Annu. Rev. pathology* 17, 387–402. doi:10.1146/annurev-pathol-042320-034052
- Saldívar, J. S., Harris, J., Ayash, E., Hong, M., Tandon, P., Sinha, S., et al. (2023). Analytic validation of NeXT Dx™, a comprehensive genomic profiling assay. *Oncotarget* 14, 789–806. doi:10.18632/oncotarget.28490
- Samstein, R. M., Lee, C. H., Shoushtari, A. N., Hellmann, M. D., Shen, R., Janjigian, Y. Y., et al. (2019). Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat. Genet.* 51, 202–206. doi:10.1038/s41588-018-0312-8
- Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., et al. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic acids Res.* 29, 308–311. doi:10.1093/nar/29.1.308
- Song, Z., Zheng, Y., Wang, X., Su, H., Zhang, Y., and Song, Y. (2017). ALK and ROS1 rearrangements, coexistence and treatment in epidermal growth factor receptor-wild type lung adenocarcinoma: a multicenter study of 732 cases. *J. Thorac. Dis.* 9, 3919–3926. doi:10.21037/jtd.2017.09.79
- Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., et al. (2019). COSMIC: the catalogue of somatic mutations in cancer. *Nucleic acids Res.* 47, D941–d947. doi:10.1093/nar/gky1015
- Trabucco, S. E., Gowen, K., Maund, S. L., Sanford, E., Fabrizio, D. A., Hall, M. J., et al. (2019). A novel next-generation sequencing approach to detecting microsatellite instability and pan-tumor characterization of 1000 microsatellite instability-high cases in 67,000 patient samples. *J. Mol. diagnostics JMD* 21, 1053–1066. doi:10.1016/j.jmoldx.2019.06.011
- Tran, V. T., Nguyen, S. T., Pham, X. D., Phan, T. H., Nguyen, V. C., Nguyen, H. T., et al. (2021). Pathogenic variant profile of hereditary cancer syndromes in a Vietnamese cohort. *Front. Oncol.* 11, 789659. doi:10.3389/fonc.2021.789659
- Waarts, M. R., Stonestrom, A. J., Park, Y. C., and Levine, R. L. (2022). Targeting mutations in cancer. *J. Clin. investigation* 132. doi:10.1172/jci154943
- Wang, J., Shi, Y., Chen, J., Liu, J., Zhao, X., Pang, J., et al. (2023). Genomic profiling, prognosis, and potential interventional targets in young and old patients with cholangiocarcinoma. *Cancer Biol. Ther.* 24, 2223375. doi:10.1080/15384047.2023.2223375
- White, T., Szlinger, S., LoBello, J., King, A., Aldrich, J., Garinger, N., et al. (2021). Analytic validation and clinical utilization of the comprehensive genomic profiling test, GEM ExTra®. *Oncotarget* 12, 726–739. doi:10.18632/oncotarget.27945
- Wu, L., Yao, H., Chen, H., Wang, A., Guo, K., Gou, W., et al. (2022). Landscape of somatic alterations in large-scale solid tumors from an Asian population. *Nat. Commun.* 13, 4264. doi:10.1038/s41467-022-31780-9
- Xia, H., Zhang, J., Chen, T., Wang, M., Chen, D., Si, T., et al. (2023). Molecular characterization of MET fusions from a large real-world Chinese population: a multicenter study. *Cancer Med.* 12, 14015–14024. doi:10.1002/cam4.6047
- Xing, M., Alzahrani, A. S., Carson, K. A., Viola, D., Elisei, R., Bendlova, B., et al. (2013). Association between BRAF V600E mutation and mortality in patients with papillary thyroid cancer. *Jama* 309, 1493–1501. doi:10.1001/jama.2013.3190
- Ye, R., Yu, Y., Zhao, R., Han, Y., and Lu, S. (2023). Comprehensive molecular characterizations of stage I-III lung adenocarcinoma with tumor spread through air spaces. *Front. Genet.* 14, 1101443. doi:10.3389/fgene.2023.1101443