

RESEARCH ARTICLE

Analytical and clinical validation of an amplicon-based next generation sequencing assay for ultrasensitive detection of circulating tumor DNA

Jonathan Poh¹, Kao Chin Ngeow¹, Michelle Pek², Kian-Hin Tan², Jing Shan Lim¹, Hao Chen¹, Choon Kiat Ong³, Jing Quan Lim³, Soon Thye Lim³, Chwee Ming Lim⁴, Boon Cher Goh⁵, Yukti Choudhury^{1*}

1 Lucence Diagnostics Pte Ltd, Singapore, Singapore, **2** Lucence Health Inc, Palo Alto, California, United States of America, **3** National Cancer Centre Singapore, Singapore, Singapore, **4** Department of Otolaryngology-Head and Neck Surgery, Singapore General Hospital, Duke-NUS Medical School, Singapore, Singapore, **5** National University Hospital, Singapore, Singapore

* yukti.c@lucence.com



OPEN ACCESS

Citation: Poh J, Ngeow KC, Pek M, Tan K-H, Lim JS, Chen H, et al. (2022) Analytical and clinical validation of an amplicon-based next generation sequencing assay for ultrasensitive detection of circulating tumor DNA. PLoS ONE 17(4): e0267389. <https://doi.org/10.1371/journal.pone.0267389>

Editor: Jeffrey Chalmers, The Ohio State University, UNITED STATES

Received: August 11, 2021

Accepted: April 7, 2022

Published: April 29, 2022

Copyright: © 2022 Poh et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting Information](#) files.

Funding: This research was funded by Lucence Diagnostics Pte Ltd and Lucence Health Inc. The funders provided support in the form of salaries for authors JP, KCN, MP, KT, JSL, HC, and YC but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of

Abstract

Next-generation sequencing of circulating tumor DNA presents a promising approach to cancer diagnostics, complementing conventional tissue-based diagnostic testing by enabling minimally invasive serial testing and broad genomic coverage through a simple blood draw to maximize therapeutic benefit to patients. LiquidHALLMARK® is an amplicon-based next-generation sequencing assay developed for the genomic profiling of plasma-derived cell-free DNA (cfDNA). The comprehensive 80-gene panel profiles point mutations, insertions/deletions, copy number alterations, and gene fusions, and further detects oncogenic viruses (Epstein-Barr virus (EBV) and hepatitis B virus (HBV)) and microsatellite instability (MSI). Here, the analytical and clinical validation of the assay is reported. Analytical validation using reference genetic materials demonstrated a sensitivity of 99.38% for point mutations and 95.83% for insertions/deletions at 0.1% variant allele frequency (VAF), and a sensitivity of 91.67% for gene fusions at 0.5% VAF. In non-cancer samples, a high specificity ($\geq 99.9999\%$ per-base) was observed. The limit of detection for copy number alterations, EBV, HBV, and MSI were also empirically determined. Orthogonal comparison of epidermal growth factor receptor (*EGFR*) variant calls made by LiquidHALLMARK and a reference allele-specific polymerase chain reaction (AS-PCR) method for 355 lung cancer specimens revealed an overall concordance of 93.80%, while external validation with cobas® *EGFR* Mutation Test v2 for 50 lung cancer specimens demonstrated an overall concordance of 84.00%, with a 100% concordance rate for *EGFR* variants above 0.4% VAF. Clinical application of LiquidHALLMARK in 1,592 consecutive patients demonstrated a high detection rate (74.8% circulating tumor DNA (ctDNA)-positive in cancer samples) and broad actionability (50.0% of cancer samples harboring alterations with biological evidence for actionability). Among ctDNA-positive lung cancers, 72.5% harbored at least one biomarker with a guideline-approved drug indication. These results establish the high sensitivity, specificity,

these authors are articulated in the 'author contributions' section.

Competing interests: These studies were funded by Lucence Diagnostics Pte Ltd or Lucence Health Inc. JP, KCN, MP, KT, JSL, HC, and YC are employees and/or share-holders of Lucence Diagnostics Pte Ltd or Lucence Health Inc, the funder of this study. Patents relating to the technology described in this manuscript have been filed (WO2020005159) by YC and HC and the study describes the validation of a product developed by Lucence Diagnostics Pte Ltd. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials, as detailed online in the guide for authors.

accuracy, and precision of the LiquidHALLMARK assay and supports its clinical application for blood-based genomic testing.

Introduction

The diagnosis and subclassification of cancers, fundamental to any cancer treatment algorithm, has historically been achieved through the study of the primary tumor tissue. The availability of high-quality diagnostic material suitable for high-end genomic technologies, including multiplexed tissue-based genotyping, is often limiting and associated with high risks and costs [1]. Indeed, for diseases such as advanced-stage, non-squamous, non-small-cell lung cancers (NSCLCs), approximately 30% of tumor samples yield insufficient or inadequate tissue for successful molecular genotyping [2,3]. Such limitations have positioned minimally invasive 'liquid biopsy'-based diagnostics as an attractive alternative for cancer management. Profiling of cell-free DNA (cfDNA) in blood to detect circulating tumor DNA (ctDNA) [4,5] has tremendous potential in overcoming the limits of sampling frequency and tumor accessibility, as well as the requirement of clinically overt disease imposed by traditional tumor biopsies. Such advances are reflected in clinical practice upgrades, such as the National Comprehensive Cancer Network (NCCN) guidelines for NSCLC now encouraging the use of liquid biopsies in cases where tissue biopsies may not be feasible [6], particularly as many patients harbor biomarkers that can be targeted by current or emerging therapies [7]. Precisely due to its non-invasive nature, the utility of cfDNA profiling extends to every stage of clinical management, from diagnosis, selection of targeted treatments, minimal residual disease (MRD) detection to identifying acquired drug resistance, in the background of spatial and temporal tumor heterogeneity [8,9]. Recent years have seen a number of approvals for cfDNA-based tests in the companion diagnostic setting based on polymerase chain reaction (PCR), such as epidermal growth factor receptor (*EGFR*) in the cobas® *EGFR* Mutation Test v2 and phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (*PIK3CA*) in the theascreen® *PIK3CA* RGQ PCR Kit [10], and very recently based on next-generation sequencing (NGS) [11,12].

Broad molecular coverage offered by NGS-based panel tests for ctDNA profiling allows the evaluation of multiple markers from the same plasma cfDNA sample, improving the detection of ctDNA in a sample below the technical lower limit of detection of the assay. Despite this, alterations identified through routine analysis of tumor tissue are detected in comprehensive ctDNA analysis with a sensitivity of 80–90% [13–16]. Successful ctDNA detection is known to be affected by the anatomical site of disease, tumor burden, and the sensitivity/specificity profile of the assay [9]. For utility in multiple clinical settings, cfDNA-based tests should ideally be able to detect ctDNA fractions <1%, against the setting of inherent error rates of NGS platforms of 0.1–1% [17]. The technical limits of detection of ctDNA platforms have been tremendously improved by ultra-deep sequencing of targeted variants, error-suppression methods, and statistical error-modeling [18,19]. Still, the challenges of detecting low-level ctDNA accurately, particularly for low disease burden (correlated to overall ctDNA amount) hamper concordance between tissue and cfDNA tests, leading to significant discordances between multiple cfDNA platforms at <1% ctDNA levels [20], and limit the generalizability of any one ctDNA biomarker. Negative ctDNA results are typically reflexed to tissue genotyping for NSCLC [21] and testing for *PIK3CA* in hormone receptor-positive (HR+) breast cancers [22].

Both biological and technical factors work in concert to determine the ability to detect ctDNA accurately, and overcoming technical limitations presents near-term opportunities to

enhance this ability. Two basic approaches to broadly capture cfDNA are amplicon (PCR)-based and hybrid capture-based methods [23]. Broad cfDNA panel tests that have been validated are presently largely based on the latter approach [15,24], with amplicon-based assays being relatively limited in panel breadth [25]. In principle, the amplicon-based target capture utilizing a pair of primers per target site, as opposed to a single hybridization probe, affords a greater degree of specificity for the enrichment of desired target regions. In turn, this enables interrogation of limiting amounts of ctDNA (as in low tumor burden samples), and enriches the density of informative sequencing reads which are on-target. Analytical sensitivities of amplicon-based methods are reportedly superior [25,26] compared to hybrid capture-based methods [15,24] for single nucleotide variants (SNVs) and insertion/deletion events (INDELs), in the biologically relevant sub-1% ctDNA range. While multiple analytical and post-analytical factors differ between individual methods, including variant calling algorithms, better sensitivity can at least be partially attributed to the superior on-target capture rate of amplicon-based enrichment methods. At present, there is a paucity of comprehensive validation studies for NGS-based cfDNA tests, particularly those based on amplicon-based enrichment, which possess technical advantages in terms of simpler workflow, effectiveness with limited input DNA, and higher on-target capture rates.

In this work, we present the analytical validation of an amplicon-based NGS assay, LiquidHALLMARK[®], with broad target coverage for clinically actionable genes. The assay incorporates two degrees of error-correction, through the use of unique molecular identifiers (UMI) introduced at target capture, and position-specific background error suppression through statistical modeling of noise inherent to samples and their processing. In addition to the profiling of SNVs, INDELs, copy number alterations (CNAs), and gene fusions (known and novel), the assay is the first NGS-based cfDNA panel test to incorporate the detection of cancer-causing viruses, including Epstein-Barr virus (EBV) and hepatitis B virus (HBV), and detection of microsatellite instability (MSI) in cfDNA, validated through multiple orthogonal comparisons. We demonstrate the clinical applicability of LiquidHALLMARK by quantitatively presenting the diagnostic findings in consecutive samples received for routine clinical testing at our Clinical Laboratory Improvement Amendment (CLIA)-certified, College of American Pathologists (CAP)-accredited laboratory. LiquidHALLMARK is applicable for ultra-sensitive detection of ctDNA in multiple cancer types, with a limit of detection of 0.1% VAF for SNVs and INDELs, with which clinically actionable alterations can be identified and monitored.

Results

Assay design/specifications

LiquidHALLMARK is a targeted NGS-based sequencing diagnostic assay that profiles 80 cancer-related genes using cfDNA isolated from stabilized peripheral whole blood-derived plasma of cancer patients (Fig 1). The assay detects SNVs and INDELs in 77 genes, with nearly full exonic coverage in five genes and carefully curated inclusion of hotspot and critical regions in select oncogenes and tumor suppressors, spanning 58.8 kb of genome coverage. In addition, CNAs in 18 genes, and structural rearrangements in 10 genes, including the detection of both novel and pre-determined gene fusion events, are captured by the assay. Further, the assay detects MSI, as well as the presence of HBV and EBV, which are key oncogenic signatures in colorectal cancer, nasopharyngeal cancer, and liver cancer, respectively. Although less comprehensive than conventional 500-gene tissue-based NGS panels, rational inclusion of key cancer-specific and pan-cancer driver genes, non-genetic biomarkers, as well as targets for which FDA-approved drugs, investigational drugs, or clinical trials are available (S1 Fig), has provided a diagnostic yield of 74.8% in all cancer samples processed to date (n = 1521).

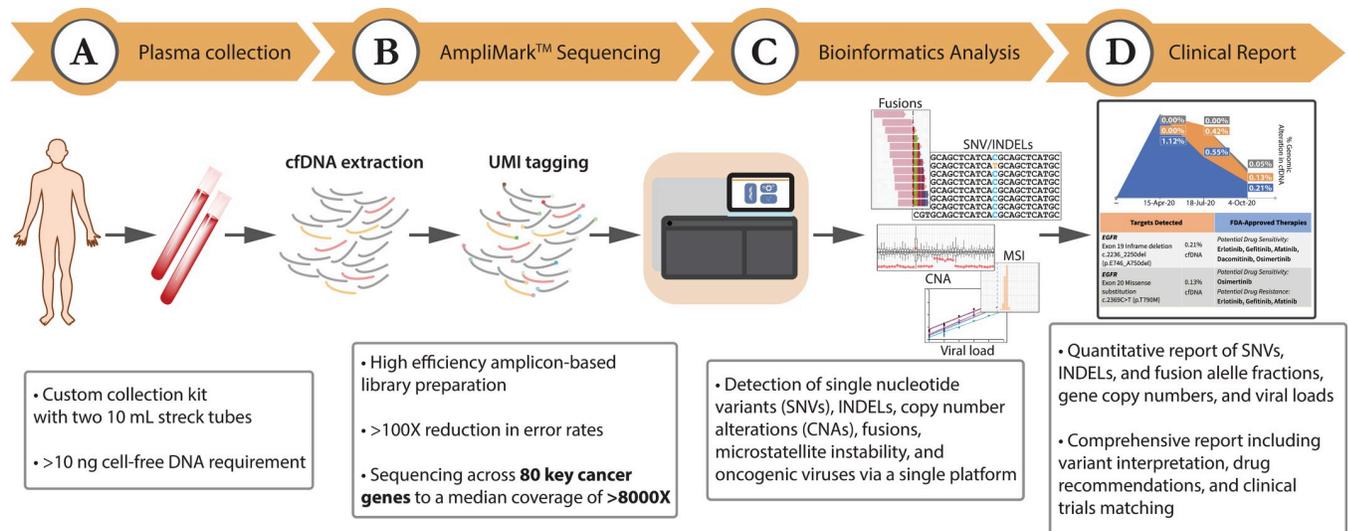


Fig 1. Overview of the LiquidHALLMARK assay workflow. (A) Stabilized whole blood is collected using custom collection kits and cell-free DNA (cfDNA) isolated from plasma. Up to 20 ng of cfDNA is used for library preparation. (B) Targeted genomic regions are tagged with oligonucleotide barcodes and enriched via high efficiency amplicon-based (AmpliMark) library preparation before being sequenced to a median coverage of >8000X. UMI, unique molecular identifier. (C) Sequencing data is bioinformatically deconvoluted, with molecular barcodes enabling error suppression rates above 100X. (D) Somatic variants and other biomarkers are quantitative reported, with comprehensive interpretation encompassing drug actionability and clinical trial eligibility.

<https://doi.org/10.1371/journal.pone.0267389.g001>

LiquidHALLMARK utilizes a highly optimized amplicon-based library preparation method (AmpliMark™, Lucence Diagnostics) combined with UMI labeling to efficiently recover unique molecules from just 20 ng input cfDNA, sequencing to a median unique coverage of 8354X (range 42.0–98.4% recovery of unique DNA molecules). With a median amplicon length of 166 base pairs (bp), which approximates the median length of cfDNA [27], the assay generates highly uniform sequencing profiles; on average 97.1% of the panel has a depth no less than 0.2X the mean panel coverage and 79.0% of the panel has a depth no less than half the mean panel coverage. In addition, UMI sequencing combined with statistical background error-modeling approaches enable error suppression of more than 100X, allowing the confident calling of SNVs/INDELS above noise levels at as low as 0.01% variant allele frequency (VAF; S2 Fig).

SNV/INDEL detection performance

To determine the analytical sensitivity, precision, and accuracy of the LiquidHALLMARK assay in detecting SNVs and INDELS, the reference standard sets HD780 (Fig 2A) and diluted Tru-Q (Fig 2B; Horizon Discovery), comprising a total of 39 unique clinically relevant SNVs and three INDELS in 19 genes across 5% (3.6–6.5%), 1% (1.0–1.3%), and 0.1% (0.10–0.13%) VAF verified by droplet digital PCR (ddPCR), were run in triplicates across multiple operators, two reagent lots, and two NextSeq 550 sequencers. In addition, the HD786 reference standard (Horizon Discovery) comprising six SNVs and four INDELS at 5% (5.0–5.6%) and one INDEL at 2.5% was run by two different operators. Across a total of 674 SNV/INDEL observations, only one INDEL (EGFR E746_A750del) and three SNVs (KRAS G12D, Q61H, Q61L) were missed at 0.1% VAF, and one SNV at 1% VAF (KRAS Q61H) (S1 Table). Because the KRAS Q61 locus consistently gave a depth <6% of the mean panel coverage (average depth at locus <850X) specifically in the Tru-Q standards, the KRAS Q61H and KRAS Q61L variants covered by this locus were excluded from all sensitivity analyses. Hence, these results yield a

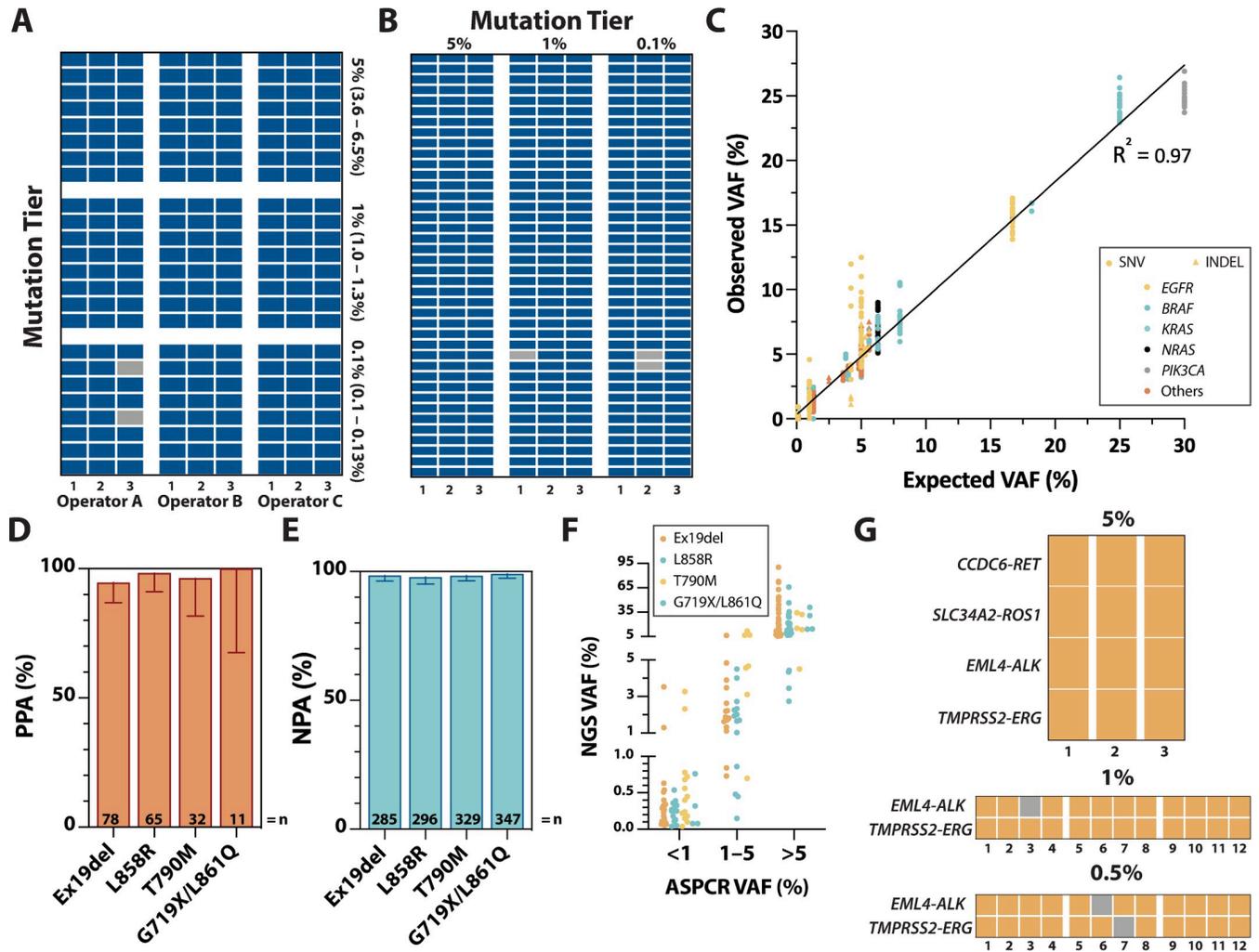


Fig 2. Analytical performance for (A-F) SNV/INDEL and (G) fusion detection. Hit/miss observations across different single nucleotide variant (SNV)/insertion-deletion (INDEL) mutation tiers and operators/repeats in (A) HD780 and (B) Tru-Q reference standard materials. Each row represents a unique SNV/INDEL variant. Detected SNV/INDEL variants are highlighted in blue and missed variants in gray. (C) Correlation between expected and observed SNV/INDEL variant allele frequencies (VAF) based on linear regression. Circles represent SNVs and triangles INDELS. Each color represents a unique gene. (D) Positive percent agreement (PPA), (E) negative percent agreement (NPA) and (F) positive VAF concordance of LiquidHALLMARK *EGFR* variant calls relative to allele-specific PCR (AS-PCR) in 355 clinical lung cancer samples. Error bars in D and E represent 95% CI. Each color in (F) represents a unique class of *EGFR* variant; ex19del, exon 19 deletion. (G) Hit/miss observations across different fusion mutation tiers and repeats in HD786 reference standard and contrived admixtures of fragmented cell-line DNA. Each row represents a unique fusion. Detected SNV/INDEL variants are highlighted in orange and missed variants in gray.

<https://doi.org/10.1371/journal.pone.0267389.g002>

sensitivity of 100% at 5% and 1% VAF, and 98.92% at 0.1% VAF, with individual sensitivities of 99.38% and 95.83% for SNVs and INDELS respectively (Table 1). These results confirm a 90% limit of detection (LoD₉₀) of 0.1% VAF for both SNVs and INDELS. Across the HD780 reference standards, pairwise analysis gave an overall reproducibility of 100% at 5% and 1% VAF, and 97.22% at 0.1% VAF, with no noticeable difference within or between operators, reagent lots, and sequencers (Fig 2A). Quantitative accuracy of variant calls was high, with a good degree of correlation between observed and expected VAFs (Fig 2C; R² = 0.97, y = 0.90).

To validate the clinical accuracy of LiquidHALLMARK for SNV and INDEL variant calls, 355 lung cancer specimens received for clinical testing were concurrently analyzed for 10 specific *EGFR* mutations, including exon 19 deletions, L858R, T790M, G719X, and L861Q, using

Table 1. Analytical sensitivity of the LiquidHALLMARK assay.

Mutation Tier	Variant Type	Sensitivity (95% CI)
3.6–6.5%	SNVs	100% (97.91%, 100%)
	INDELs	100% (89.28%, 100%)
	Fusions	100% (77.19%, 100%)
1.0–1.3%	SNVs	100% (97.68%, 100%)
	INDELs	100% (86.20%, 100%)
	Fusions	95.45% (78.20%, 99.77%)
0.5%	Fusions	91.67% (74.15%, 98.52%)
0.1–0.13%	SNVs	99.38% (96.59%, 99.97%)
	INDELs	95.83% (79.76%, 99.79%)

<https://doi.org/10.1371/journal.pone.0267389.t001>

a reference allele-specific PCR (AS-PCR) method [28]. LiquidHALLMARK and AS-PCR were concordant for 127 positive calls (positive percent agreement, PPA 95.49%; 95% CI, 90.51–97.92%; NGS VAF 0.03–90.43%) and 206 negative calls (negative percent agreement, NPA 92.79%; 95% CI, 88.61–95.52%), giving an overall concordance of 93.80% (95% CI, 90.80–95.87%; **Tables 2 & S2**). When segregated by *EGFR* variant type (exon 19 deletion, L858R, T790M, or G719X/L861Q), both PPA and NPA remained high (>94% and >97% respectively) for all four *EGFR* variant types (**Fig 2D & 2E**), with good agreement between quantitative calls made by both methods (**Fig 2F**). In total, four exon 19 deletion, one L858R, and one T790M variants were missed by LiquidHALLMARK at detected VAFs close to or below the LoD₉₀ (AS-PCR VAF 0.03–0.16%). LiquidHALLMARK detected three additional G719X, four exon 19 deletion, five L858R, and five T790M variants (NGS VAF 0.02–0.50%); 10 of these were detected specifically by AS-PCR melt curve analysis below the reportable range. Of the remaining seven discordant variants, six were detected close to or below the established LoD of AS-PCR, with one discordant exon 19 deletion (L747_P753delinsS) detected at 0.5% VAF in LiquidHALLMARK. These findings corroborate the high sensitivity and accuracy of the LiquidHALLMARK assay.

Fusion detection performance

To determine the analytical sensitivity of the assay for the detection of gene fusions, the HD786 reference standard (Horizon Discovery) and contrived admixtures of fragmented fusion positive cell-line DNA were used to probe four gene fusions (*EML4-ALK*, *SLC34A2-ROS1*, *TMPRSS2-ERG*, and *CCDC6-RET*) across 5%, 1%, and 0.5% VAF. A total of 12 fusions were tested at 5% VAF across three replicates and 24 fusions at 1% and 0.5% VAF across 12 replicates each, with 1 *EML4-ALK* fusion missed at 1% VAF and 1 *EML4-ALK* fusion and 1 *TMPRSS2-ERG* fusion missed at 0.5% VAF (**Fig 2G**), giving a sensitivity of 100%,

Table 2. Comparison of LiquidHALLMARK with a reference AS-PCR assay for the detection of EGFR L858R, exon 19 deletion, T790M, and uncommon G719X/L861Q variants.

LiquidHALLMARK	Reference assay (AS-PCR)			PPA: 95.49% (90.51–97.92%) NPA: 92.79% (88.61–95.52%) OPA: 93.80% (90.80–95.87%)
		Positive	Negative	
Positive	127	16	143	
Negative	6	206	212	
Total	133	222	355	

PPA, positive percent agreement; NPA, negative percent agreement; OPA, overall percent agreement.

<https://doi.org/10.1371/journal.pone.0267389.t002>

95.45%, and 91.67% respectively (Tables 1 & S3). Thus, the LoD_{90} was empirically determined to be 0.5% VAF for gene fusions.

Due to a lack of fusion-positive genetic material, the platform technology was further validated for the ability to detect novel fusions by analyzing 29 natural killer (NK)/T-lymphoma tissue samples for programmed death ligand-1 (PD-L1) 3'-untranslated region (3'-UTR) rearrangement positivity [29]. Of these, nine samples tested positive for PD-L1 3'-UTR rearrangement and 20 samples tested negative; all calls were orthogonally confirmed by whole-genome, targeted, or Sanger sequencing, giving an overall concordance of 100% (S4 Table). In one sample where matched plasma was available, cfDNA testing identified the same PD-L1 3'-UTR rearrangement.

Copy number alteration (CNA) detection performance

Validation of CNA detection performance was conducted using contrived admixtures of fragmented cell-line DNA from cell-lines well-established to harbor gene amplifications in *EGFR*, erb-b2 receptor tyrosine kinase 2 (*ERBB2*, also known as human epidermal growth receptor 2 (HER2)), *NRAS* proto-oncogene, GTPase (*NRAS*), and cyclin dependent kinase 6 (*CDK6*) [30–32], as well as the HD786 reference standard which harbors focal amplification in *MET* proto-oncogene, receptor tyrosine kinase (*MET*) and *MYC* proto-oncogene, bHLH transcription factor (*MYC*) at 4.5 and 9.5 copies respectively, instead of normal diploid two copies. Across 19 observations from 2.4 to 65.2 gene copies in the dilution series, all amplifications were positively detected, with a high degree of correlation between expected and observed copy numbers ($R^2 = 0.92$, $y = 0.70$; Fig 3A). Based on gene amplifications confidently detectable above the baseline range of normalized copy numbers determined from a set of selected reference plasma cfDNA samples ($n = 27$), the LoD for gene amplifications was determined to be a 1.5X increase from the normal diploid number, which is equivalent to a low-level amplification (copy number of 6) at 25% tumor fraction (TF) or a high-level amplification (copy number of 10) at 12.5% TF.

To assess the concordance between gene amplification calls in plasma cfDNA and fluorescence *in situ* hybridization (FISH) in tissue, nine matched plasma cfDNA samples from *ERBB2* (HER2) amplification-positive samples were analyzed by LiquidHALLMARK. Six

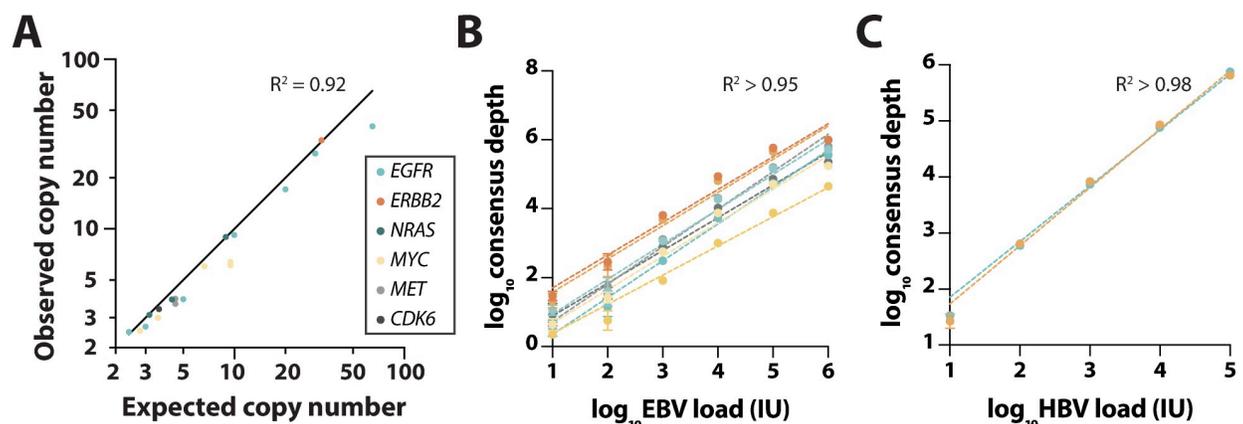


Fig 3. Validation of LiquidHALLMARK read count-based biomarkers. (A) Positive concordance of observed and expected copy numbers for amplification in six genes using HD786 reference standard and contrived admixtures of fragmented cell-line DNA. Observed consensus depth relative to (B) Epstein-Barr virus (EBV) and (C) hepatitis B virus (HBV) viral loads in plasma. Data is represented as mean \pm SD from four replicate experiments. Each color represents a unique virus-specific amplicon. Correlation between consensus depth and viral load is based on linear regression.

<https://doi.org/10.1371/journal.pone.0267389.g003>

samples were positive for *ERBB2* (HER2) amplification in plasma cfDNA, giving a positive concordance between cfDNA and tissue testing of 66.67% (S5 Table).

Detection performance for other biomarkers

In addition to detecting conventional biomarkers in SNVs, INDELS, gene fusions, and CNAs, the LiquidHALLMARK assay detects and quantifies viral loads of EBV and HBV in plasma cfDNA via the inclusion of amplicons specific for these viruses. EBV detection performance was verified by spiking plasma cfDNA with known viral loads of the external reference standard for EBV (1st WHO International Standard for Epstein-Barr Virus, NIBSC code: 09/260) spanning six orders of magnitude. Based on the dilution series performed in four replicates, a high degree of linearity could be observed for all eight EBV-specific amplicons ($R^2 > 0.95$ for all amplicons, Fig 3B). These experiments also establish an LoD of 10 IU EBV within a sequencing run, equivalent to 2 IU/mL plasma when the routine 5 mL of plasma is used for cfDNA extraction. To validate the clinical performance of the assay for EBV detection, 19 clinical cfDNA specimens were analyzed by LiquidHALLMARK and an orthogonal PCR method [33]. A total of 16 positive calls and 1 negative call were observed by both methods, giving a PPA of 94.12% and an NPA of 50.00% (S6 Table).

HBV detection performance was assessed by spiking plasma cfDNA with known viral loads of the external reference standard for HBV (4th WHO International Standard for HBV DNA for NAT, NIBSC code: 10/266) spanning six orders of magnitude. Based on the dilution series performed in four replicates, a high degree of linearity of response was observed across both HBV-specific amplicons ($R^2 > 0.98$ for both amplicons, Fig 3C).

The detection threshold for MSI was determined using 91 plasma cfDNA samples and 29 FFPE tissue samples with no known MSI, with the threshold for calling MSI set as the 95th percentile of deletion length (for each of the six MSI loci) detected across all microsatellite stable (MSI-negative) samples. The LoD for MSI was validated using dilutions of DNA from the MSI-positive cell-line RKO in a wild-type cfDNA background and confirmed to be 5% tumor fraction (S3 Fig). Due to the lack of characterized MSI-positive (MSI-low or MSI-high) cfDNA material, tumor tissue DNA was analyzed for 36 clinical samples in which microsatellite status was determined by fragment size analysis. LiquidHALLMARK and fragment size analysis were concordant for 4 positive and 31 negative calls, giving an overall concordance of 97.22% (PPA 80.00%, NPA 100%; S7 Table). In a subset of 19 cases for which matched cfDNA was available, MSI was detectable in 1 cfDNA sample by LiquidHALLMARK and 16 MSI-negative cases were also MSI-negative in their matched cfDNA (PPA 33.33%, NPA 100%; S8 Table).

Clinical validation for detection of *EGFR* alterations

To assess the clinical accuracy of LiquidHALLMARK for *EGFR* SNV and INDEL variant calls, LiquidHALLMARK was externally validated against the FDA-approved cobas *EGFR* Mutation Test v2. Fifty samples analyzed by LiquidHALLMARK, including 38 positive and 12 negative for *EGFR* variants covered by the cobas assay were submitted to the Mayo Clinic Laboratories (Rochester Main Campus, MN) for cobas testing to benchmark the concordance between LiquidHALLMARK and cobas calls. Of the 50 samples, LiquidHALLMARK and cobas were concordant for 31 positive (PPA 96.88%; 95% CI, 84.26–99.84%) and 11 negative variant calls (NPA 61.11%; 95% CI, 38.62–79.69%), giving an overall concordance of 84.00% (95% CI, 71.49–91.66%; Tables 3 & S9). PPA was high (>91%) across the four *EGFR* variant types (exon 19 deletion, L858R, T790M, and uncommon G719X/S768I/exon 20 insertion variants; Fig 4A), with 1 *EGFR* L858R detected by cobas but not LiquidHALLMARK. To resolve the

Table 3. Comparison of LiquidHALLMARK with cobas EGFR Mutation Test v2 for the detection of EGFR L858R, exon 19 deletion, T790M, and G719X/S768I/exon 20 insertion variants.

LiquidHALLMARK	Reference assay (cobas EGFR Mutation Test v2)			PPA: 96.88% (84.26–99.84%) NPA: 61.11% (38.62–79.69%) OPA: 84.00% (71.49–91.66%)
	Positive	Negative	Total	
Positive	31	7	38	
Negative	1	11	12	
Total	32	18	50	

PPA, positive percent agreement; NPA, negative percent agreement; OPA, overall percent agreement.

<https://doi.org/10.1371/journal.pone.0267389.t003>

true nucleotide identity of this potential false negative variant, plasma cfDNA from this sample was orthogonally tested by AS-PCR, which corroborated the absence of EGFR L858R. Hence, orthogonal AS-PCR validation coupled with the established 99% specificity of cobas for EGFR L858R alterations (benchmarked against tissue testing) [34] suggest that the potential false negative by LiquidHALLMARK could have been a false positive via cobas testing, supporting a 100% PPA of LiquidHALLMARK with cobas. Across the four *EGFR* variant types, NPA also remained high (>93%; Fig 4B), with LiquidHALLMARK detecting seven additional variants across 0.05–0.33% VAF, including 2 L858R, 2 exon 19 deletion, 2 uncommon (G719A and S768I), and 1 T790M variants, which are within the range of variants detectable by the cobas assay (Fig 4C & S9 Table). To resolve these potential false positive calls, plasma cfDNA from these samples were orthogonally validated by AS-PCR. Of the seven variants, six were positively detected by AS-PCR, including one exon 19 deletion (EGFR E746_A750del, detected at 0.15% VAF), one G719A (detected at 0.08% VAF), and one T790M variant (detected at 0.23% VAF) specifically detected by melt curve amplification below the reportable range of the AS-PCR assay. The remaining variant (EGFR S768I, detected at 0.12% VAF) was not covered by the AS-PCR assay and could not be orthogonally confirmed. These results are consistent with the established 25–100 copies/mL plasma [35] and reported 0.1–0.8% VAF [36] LoD of

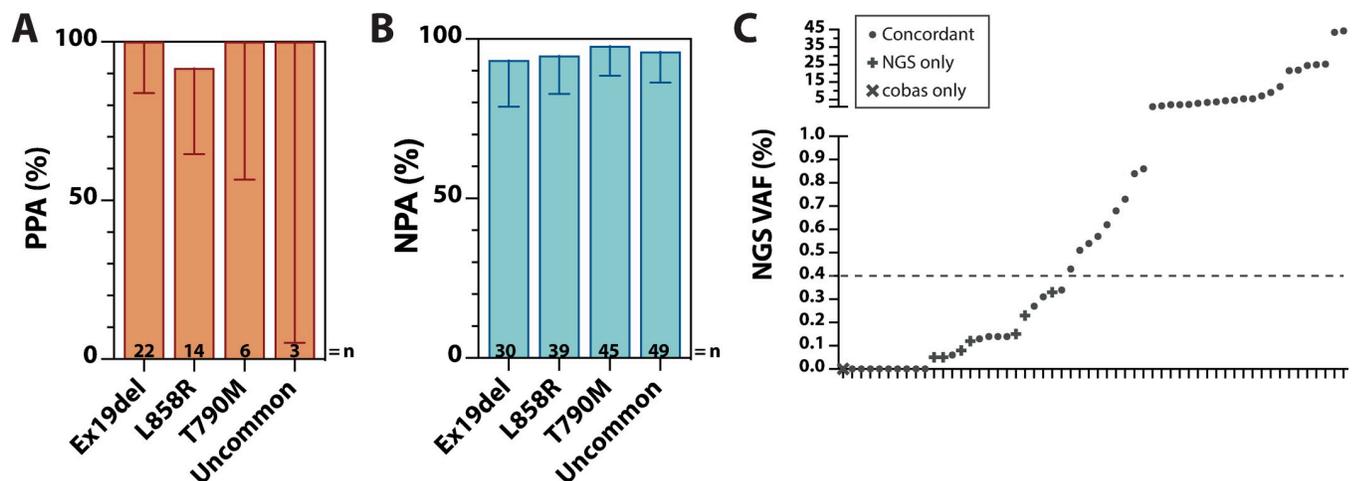


Fig 4. Clinical validation for detection of *EGFR* alterations. (A) Positive percent agreement (PPA), and (B) negative percent agreement (NPA) of LiquidHALLMARK *EGFR* variant calls relative to cobas EGFR Mutation Test v2 in 50 clinical lung cancer samples. Uncommon *EGFR* variants include G719X, S768I, and exon 20 insertions. Error bars in A and B represent 95% CI. (C) Positive and negative concordance of LiquidHALLMARK *EGFR* variant calls relative to cobas EGFR Mutation Test v2. VAF, variant allele frequency. Each data point represents a negative sample (VAF = 0%) or *EGFR* variant; •, concordant call between LiquidHALLMARK and cobas; x, detected only by cobas; +, detected only by LiquidHALLMARK (NGS). Above 0.4% VAF, both assays were concordant for all detected *EGFR* variants.

<https://doi.org/10.1371/journal.pone.0267389.g004>

the cobas assay, with all discordant variant calls falling below 0.4% VAF (Fig 4C), establishing the exceptional sensitivity of LiquidHALLMARK.

Clinical utility of the LiquidHALLMARK assay

While validation studies are crucial for examining the technical performance of an assay, insight into the performance of an assay in a real-world setting is equally, if not more informative on the utility of a liquid biopsy test in guiding clinical decisions. To that end, we examined all samples submitted to our CAP- and CLIA-accredited laboratory to elucidate real-world assay performance metrics. Within the study period from January 2018 to May 2021, a total of 1,592 samples were analyzed and reported (100% report rate), with 52.1% of the clinical volume originating from lung cancer patients (Fig 5A). While the assay is intended for advanced cancer patients to aid clinical therapeutic decision making, a significant portion ($n = 71$; 4.5%) of the clinical volume comprised either screening cases in healthy individuals ($n = 61$), or suspected cancer cases ($n = 10$). Additionally, samples from patients with localized tumors constituted 8.7% of the clinical volume. Overall, 73.6% (1120/1521) of cancer samples harbored at least one detectable genetic variant (ctDNA positive), including 40.6% of samples from localized tumors and 78.5% of samples from metastatic tumors. When non-genetic biomarkers (EBV/HBV/MSI) were included, the detection rate increased to 74.8% ($n = 1137$). While ctDNA detection was largely uniform across most tumor types, several cancers such as kidney cancer and melanoma displayed lower ctDNA detection rates (Fig 5A), consistent with previous reports on tumor type-specific ctDNA prevalence [15]. Among ctDNA-positive samples, a median of two variants were detected per sample, with a median VAF of 0.86% (range 0.01–95.30%). Additionally, more than half of ctDNA-positive cancer samples contained at least one variant $\leq 0.3\%$ VAF; this was again consistent across tumor type, highlighting the ability of the assay to sensitively detect variants at sub-1% VAFs. Across all non-cancer screening samples ($n = 61$), two likely pathogenic SNV variants (JAK2 V617F, detected at 0.81% VAF, and TP53 M246I, detected at 0.40% VAF) were observed, while no INDELS, CNAs, gene fusions, or non-genomic biomarkers, were detected, giving a per-sample clinical specificity of 96.72%, or per-base specificity of $\geq 99.9999\%$ (2 SNVs across 3.59 million bp, or 61 samples \times 58.8 kb per sample).

Given that the presence of ctDNA correlates with tumor burden [37–40], we evaluated the correlation between plasma cfDNA concentration and tumor fraction (approximated as the variant with the highest VAF observed in any given sample). Plasma cfDNA concentration was positively correlated with TF, although this varied across different tumor types; for instance, ctDNA-positive colorectal cancer samples possessed a higher median concentration (median 25.57 ng/mL plasma) compared to ctDNA-positive lung cancer samples (median 15.50 ng/mL; Fig 5B). While the median concentration of ctDNA-negative samples was lower (10.87 ng/mL, range 2.29–746.67 ng/mL) compared to that of ctDNA-positive samples (17.56 ng/mL, range 1.15–3160 ng/mL), a significant portion of ctDNA-negative samples possessed high cfDNA concentrations (Fig 5B), highlighting a potential limitation of targeted mutation-based liquid biopsy approaches in accurately determining tumor burden. Across all ctDNA-positive samples, the median TF was 2.53% (2.51% in lung cancer samples), with 20.2% of samples having a TF $\leq 0.3\%$ (19.4% in lung cancer samples; Fig 5C), emphasizing the importance of high sensitivity in liquid biopsy assays.

Clinical actionability

A cornerstone of NGS-based liquid biopsy assays is the ability to detect clinically actionable alterations not easily detected via alternative testing methods. Hence, we further analyzed the

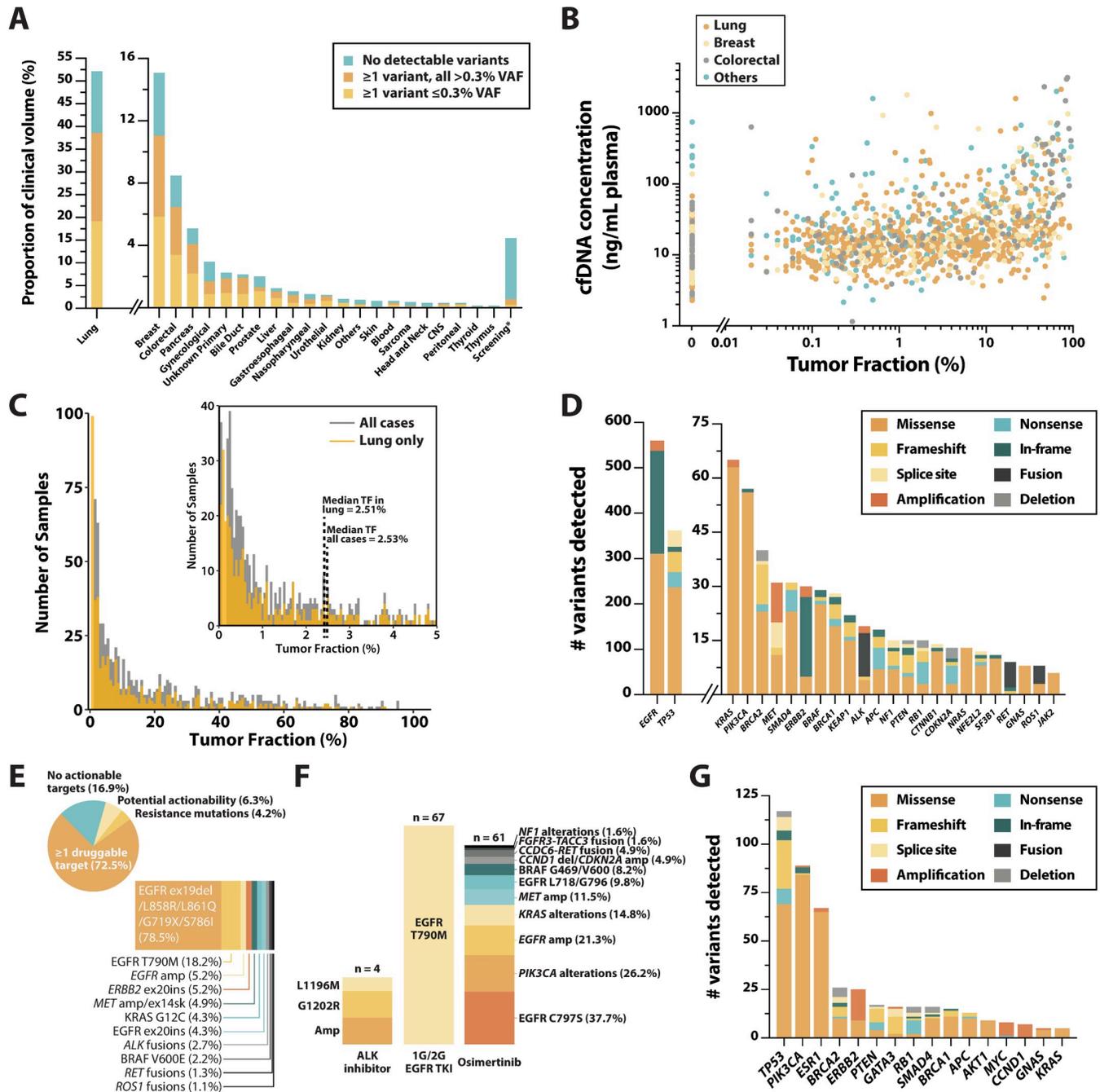


Fig 5. Clinical utility of the LiquidHALMARK assay. (A) Proportion of clinical volume (n = 1,592) by tumor origin. Samples with no detectable variants are shaded in blue, while positive samples are shaded in yellow if at least 1 variant was detected at a variant allele frequency (VAF) $\leq 0.3\%$ and orange if all variants detected were $>0.3\%$ VAF. CNS, central nervous system. * Screening includes suspected cancer cases. (B) Correlation between cell-free DNA (cfDNA) concentration and tumor fraction, approximated as the variant with the highest VAF observed in any given sample. Lung, breast, and colorectal cancer samples are highlighted in orange, yellow and gray respectively; all other tumor types are highlighted in blue. (C) Distribution of positive samples by tumor fraction. The distribution for positive lung cancer samples is shown in yellow. (D) Prevalence and variant type of the top 25 genes altered in lung cancer samples. Each color represents a different variant type. (E, top) Proportion and breakdown of druggable targets in lung cancers with guideline-approved biomarkers. (E, bottom) Prevalence and breakdown of druggable targets in lung cancers with guideline-approved biomarkers. (F) Prevalence and breakdown of alterations associated with resistance observed in lung cancer patients treated with an ALK inhibitor, 1st or 2nd generation (1G/2G) EGFR tyrosine kinase inhibitor (TKI), or osimertinib. Ex20ins, exon 20 insertion; ex14sk, exon 14 skipping; amp, amplification; del, deletion. (G) Prevalence and variant type of the top 15 genes altered in breast cancer samples. Each color represents a different variant type.

<https://doi.org/10.1371/journal.pone.0267389.g005>

prevalence of driver genes and clinically relevant variants detected within the 1,592 clinical samples. In line with known prevalence of significantly mutated genes in cancer [41,42], *TP53*, *KRAS*, *PIK3CA*, *APC*, *SMAD4*, and *PTEN* were among the most frequently altered genes across all ctDNA-positive samples (S4A Fig). In addition, *EGFR* alterations were detected in 36.1% of all ctDNA-positive samples; this was driven by the large proportion of lung cancers among clinical samples. Indeed, in ctDNA-positive lung cancer samples, the detection rate of *EGFR* alterations was 62.8%, with more than a third of detected *EGFR* alterations being in-frame exon 19 deletions (Fig 5D & S4B Fig). Significant alteration frequencies were also observed in a repertoire of clinically relevant genes including *KRAS*, *PIK3CA*, *MET*, *ERBB2*, *BRAF*, and *ALK*, as well as driver genes with potential actionability such as *NF1*, *KEAP1*, and *NFE2L2* (Fig 5D). Importantly, of the 615 ctDNA-positive lung cancer samples, 83.1% (n = 511) harbored at least one genetic biomarker possessing biological evidence for actionability, including 72.5% (n = 446) with ≥ 1 biomarker with a guideline-approved drug, and an additional 4.2% (n = 26) and 6.3% (n = 39) harboring either biomarkers predicting drug resistance or with biological evidence suggesting potential clinical actionability (both drug sensitivity and resistance) respectively (Fig 5E top). Significantly, 175 of the 511 lung cancer samples with actionable biomarkers (28.5% of all ctDNA-positive lung cancer samples) harbored an actionable biomarker at low VAF ($\leq 0.3\%$). Sensitizing *EGFR* mutations (*EGFR* L858R, exon 19 deletion, or L861Q/G719X/S786I) were by far the most prevalent druggable target, being observed in 78.5% of lung cancer samples with guideline-approved biomarkers, while *EGFR* T790M, exon 20 insertions, and amplifications, were also observed at significant frequencies (Fig 5E bottom). Beyond drug sensitivity, the prevalence of alterations that predict treatment resistance was also further analyzed. On-target *ALK* mutations [43] including *ALK* G1202R, *ALK* L1196M, and *ALK* amplification were observed in 14.8% (4/27) of *ALK* inhibitor-treated patients; *EGFR* T790M [44] was observed in 37.0% (67/181) of first- or second-generation *EGFR* TKI-treated patients, and a range of resistance mutations (in accordance with literature on osimertinib resistance [45]), including *EGFR* C797, *PIK3CA* and *KRAS* alterations, *MET* amplification, copy number changes in cell cycle genes, *FGFR3* and *RET* fusions, was observed in 37.4% (61/163) of osimertinib-treated patients (Fig 5F). Strikingly, 32.8% (22/67) of all *EGFR* T790M resistance mutations in first- or second-generation *EGFR* TKI-treated patients were detected at VAF $\leq 0.3\%$, while 42.6% (26/61) of all resistance mutations detected in osimertinib-treated patients were observed at VAF $\leq 0.3\%$. In breast cancer, clinically relevant alterations in *PIK3CA*, *ESR1*, *BRCA2*, and *ERBB2* were among the most prevalent (Figs 5G & S4C); 46.8% (59/126) of HR+ HER2- breast cancer samples harbored either an *ESR1* mutation predicting resistance to aromatase inhibitors [46], or a *PIK3CA* mutation predicting sensitivity to PI3K pathway inhibition by alpelisib [47], with 28.8% (17/59) of these samples harboring concurrent *PIK3CA* or *ESR1* mutations (S5A Fig). Additionally, loss-of-function alterations in *RBI*, predicted to contribute to CDK4/6 inhibitor resistance [48], were observed in 13.3% (6/56) of CDK4/6 inhibitor-treated HR+ HER2- breast cancer patients. In colorectal cancer patients, 54.1% of all samples harbored a mutation known to confer anti-*EGFR* antibody resistance [49,50] (S5B Fig). Across the three cancer types, actionable variants were among the most frequently observed variants, with VAFs ranging from 0.01–94.30% (Table 4). Collectively, these results demonstrate the capability of LiquidHALLMARK to guide clinical therapeutic decisions.

Molecular profiling to delineate origins of cancers of unknown primary

Broad genomic coverage also aids the molecular profiling and differential diagnosis of cancers of unknown primary (CUP). Of the 36 CUP cases analyzed, 80 unique genetic alterations

Table 4. Median variant allele fraction (VAF) of top 10 mutated genes in clinical lung, breast, and colorectal samples.

Gene	Variant	Variant Type	Median VAF % (Range)
Lung cancer, n = 615			
EGFR	Ex19del	INDEL	2.99% (0.01–90.40%)
EGFR	L858R	SNV	2.42% (0.01–73.20%)
EGFR	T790M	SNV	0.72% (0.04–67.03%)
EGFR	C797S	SNV	1.50% (0.14–63.90%)
EGFR	Amplification	CNA	2.06X (1.42–8.51X)
KRAS	G12C	SNV	3.80% (0.06–26.80%)
PIK3CA	E545K	SNV	2.90% (0.03–48.58%)
ERBB2	Y772_A775del	INDEL	0.84% (0.09–23.67%)
KRAS	G12D	SNV	1.49% (0.02–59.94%)
MET	Amplification	CNA	2.50X (1.64–6.73X)
ALK	EML4-ALK	Fusion	3.20% (0.19–54.20%)
Breast cancer, n = 176			
PIK3CA	H1047R	SNV	3.09% (0.06–61.90%)
ESR1	D538G	SNV	1.19% (0.11–68.06%)
ERBB2	Amplification	CNA	2.95X (1.40–10.09X)
PIK3CA	E545K	SNV	7.60% (0.22–41.19%)
PIK3CA	E542K	SNV	2.30% (0.11–69.50%)
AKT1	E17K	SNV	2.80% (0.25–23.11%)
ESR1	Y537S	SNV	1.46% (0.13–32.70%)
TP53	Y163S	SNV	0.36% (0.26–0.68%)
CCND1	Amplification	CNA	2.30X (2.01–3.99X)
MYC	Amplification	CNA	2.42X (1.66–5.41X)
ESR1	E380Q	SNV	2.60% (0.06–21.22%)
Colorectal cancer, n = 103			
KRAS	G12D	SNV	15.90% (0.06–76.30%)
KRAS	G13D	SNV	0.37% (0.06–27.12%)
KRAS	G12V	SNV	0.58% (0.02–34.81%)
KRAS	Q61H	SNV	3.91% (0.23–34.50%)
MYC	Amplification	CNA	2.00X (1.55–23.31X)
APC	S1465Wfs*3	INDEL	0.61% (0.27–48.41%)
TP53	R273H	SNV	4.35% (0.35–86.10%)
APC	c.835-8A>G	SNV	1.18% (0.66–52.16%)
APC	R1450*	SNV	12.82% (0.45–40.30%)
APC	R876*	SNV	2.06% (0.92–48.99%)
BRAF	V600E	SNV	6.76% (0.02–38.90%)
PIK3CA	E542K	SNV	9.05% (0.23–20.67%)
PTEN	R130*	SNV	0.17% (0.08–21.40%)
TP53	G245S	SNV	28.01% (0.99–61.59%)
TP53	Y220C	SNV	84.40% (59.89–94.90%)

<https://doi.org/10.1371/journal.pone.0267389.t004>

across 34 genes were detected in 30 cases, giving a ctDNA detection rate of 83.3% (S6 Fig). Detection of tumor type-specific alterations, including *AR*, *ESR1*, *GNAS*, and *EGFR* alterations in seven samples, as well as EBV, HBV, and MSI positivity in three samples, informed potential tumor origins. Within genes with high prevalence of mutations, tumor specific prevalence biases based on the COSMIC database [51] could be used to discriminate between tumor

types. For instance, several *TP53* alterations identified favored a large intestinal origin (R273C, R273H), while others favored a lung origin (K305*, V157F, R110P).

Cancer monitoring through alternative non-genomic biomarkers

In tumor types where the prevalence of conventional genetic biomarkers is low, such as nasopharyngeal cancers (NPC), viral load quantification acts as a useful marker of tumor burden [52]. Of the NPC samples analyzed, 64.3% (9/14) were ctDNA positive, with *TP53* alterations being the most prevalent. By contrast, 69.2% (9/13; one sample was not tested for EBV) of NPC samples were EBV positive (S7A Fig), including two ctDNA-negative samples, enabling an additional diagnostic yield of 14.3%. In liver cancer, although 85.0% (17/20) of samples were ctDNA positive and only 55.6% (10/18; 2 samples were not tested for HBV) of samples were positive for HBV (S7B Fig), one ctDNA-negative sample was HBV positive, providing an additional diagnostic yield of 5.0%. The detection of microsatellite instability also contributed to diagnostic yield; 16 samples were identified as MSI positive, including six MSI-high samples and 10 MSI-low samples. Of these, three were negative for any ctDNA-based genetic variant.

Discussion

Tissue biopsy has historically been the gold standard of cancer diagnostics, enabling the histopathological analysis of tumors, and since the advent of precision oncology, facilitates the molecular profiling of tumors to inform therapeutic decisions [53]. Yet, as the field of molecular biology advances, the limitation of tissue biopsy for molecular profiling becomes increasingly evident. Critically, the isolated snapshot provided by a single tissue biopsy is unable to wholly recapitulate a tumor, particularly given extensive evidence of intratumoral heterogeneity, which may impact therapeutic options [54–56]. This emphasizes the need for serial biopsies, an option that is often infeasible given the invasive nature of tissue biopsies and their accompanying clinical risks, potential surgical complications, and economic costs [57]. In scenarios where tumor accessibility is an issue, obtaining even a single tissue biopsy can be onerous [53]. In light of this, liquid biopsy presents a minimally invasive opportunity to interrogate tumor biology without additional risks to patients, enabling early cancer detection, prognostication, and disease monitoring through a simple blood draw. Several cfDNA-based liquid biopsy companion diagnostics currently exist; these assays probe specific molecular targets to predict drug sensitivity, for instance, *EGFR* in the cobas EGFR Mutation Test v2 and *PIK3CA* in the theascreen PIK3CA RGQ PCR Kit [10]. As the rapidly expanding repository of targeted therapies grows, the need for more comprehensive panels that probe a wider spectrum of vulnerabilities becomes increasingly imperative.

Here, we present the analytical and clinical validation of LiquidHALLMARK, an amplicon-based NGS liquid biopsy assay which interrogates 80 cancer-related genes for SNVs, INDELs, CNAs, and gene fusions, as well as additional biomarkers including oncogenic viruses (EBV and HBV) and MSI. These results underscore the exceptional sensitivity, specificity, precision, and accuracy of the LiquidHALLMARK panel. Across 674 observations in 49 unique alterations in analytical validation, the assay demonstrated highly quantitative and reproducible VAF measurements, demonstrating a false negative rate of 0.62% for SNVs and 4.17% for INDELs at 0.1% VAF, and 8.33% for fusions at 0.5% VAF, and giving an LoD_{90} of 0.1% VAF for SNVs and INDELs, and 0.5% VAF for fusions. The specificity of the assay, enhanced through error suppression via UMI sequencing coupled with statistical background error-modeling, was demonstrated through the real-world analysis of non-cancer screening patients. Across 61 non-cancer screening samples, two likely pathogenic variants, JAK2 V617F and TP53 M246I, were detected at 0.81% and 0.40% VAF respectively, demonstrating a per-base

specificity of $\geq 99.9999\%$, giving a false positive variant rate of $\leq 0.0001\%$ (1 in 1,000,000 bp). Furthermore, given that JAK2 V617F and mutations in *TP53* are two of the most common somatic alterations detected in plasma cfDNA due to clonal hematopoiesis (CH) [29], the possibility that these two mutations arose due to CH cannot be ruled out. Orthogonal validation of LiquidHALLMARK *EGFR* variant calls with a reference AS-PCR method in 355 lung cancer specimens revealed an overall concordance of 93.80% (PPA 95.49%; NPA 92.79%), highlighting the excellent accuracy of the assay. Comparison of LiquidHALLMARK against the FDA-approved cobas *EGFR* Mutation Test v2 for 50 lung cancer specimens demonstrated an overall concordance of 84.00% (PPA 96.88%; NPA 61.11%), with eight discordant calls (LiquidHALLMARK VAF 0.05–0.33%) between LiquidHALLMARK and cobas. Of these, seven were resolved by orthogonal AS-PCR testing, supporting the calls made by LiquidHALLMARK. These results exemplify the high sensitivity of LiquidHALLMARK, even at VAFs as low as 0.05%. In orthogonal comparisons of *ERBB2* (HER2) amplification and MSI calls made by LiquidHALLMARK against FISH and fragment size analysis using patient-matched tumor tissue samples, a PPA of 66.67% and 80.00% respectively was demonstrated. Although patient-matched tumor tissue and plasma samples were obtained, these samples were not temporally matched, and the treatment status of patients between collection of tumor tissue and blood is unknown. This potentially underestimates the true PPA of the assay and represents a limitation of the tissue concordance studies performed. In addition to high sensitivity, the assay exhibits uniformly high sequencing depth across the targeted regions of genome coverage, recovering a median 69.6% of unique input DNA molecules (range 42.0–98.4%), with 98.7% of calls having a unique coverage of $> 1000\times$. Several additional capabilities of the assay are highlighted in these validation studies. The ability of the platform technology to detect novel fusions was exemplified by an overall concordance of 100% against orthogonal testing methods for structural rearrangements in PD-L1 that disrupt its 3'-UTR [29]. Further, the quantitative detection of oncogenic viruses such as EBV and HBV demonstrated excellent linearity ($R^2 > 0.95$) across six orders of magnitude, down to an LoD of 2 IU/mL plasma.

When used within a real-world setting for 1,592 clinical samples, 74.8% of cancer samples harbored at least one clinically useful biomarker for drug targeting, clinical trial inclusion, or disease monitoring, including 36.2% with on-label drug recommendations, and an additional 5.7% and 8.1% harboring biomarkers associated with treatment resistance and potential actionability respectively. Importantly, the diverse spectrum of alterations across the 80 cancer-related genes covered in the panel is highlighted by the 1607 unique somatic variants identified across all clinical samples, emphasizing the necessity of broad genomic coverage. Underpinning the high sensitivity of the assay is the frequent detection of samples with TFs $\leq 0.3\%$, with variants routinely detected below the LoD; indeed, 10.2% of all reported variants were found at VAFs between 0.01% and 0.09%, with calls being supported by high variant allele read counts and a high depth of coverage. Although ctDNA detection rate was largely uniform across tumor types, several cancers including kidney cancer and melanoma displayed lower ctDNA detection rates, consistent with reported tumor type-specific ctDNA prevalence [15]. The detection rate of ctDNA also differed between metastatic and localized tumors; 40.6% of samples from localized tumors harbored at least one detectable genetic variant, compared to 78.5% of samples from metastatic tumors. This is representative of the differential extent of tumor shedding between localized and metastatic tumors, with higher quantities of ctDNA being present in patients experiencing progressive or advanced stage disease [37–40]. Indeed, we noted a positive correlation between plasma cfDNA concentration and tumor fraction, and further observed a lower median TF of 0.49% in ctDNA-positive samples from localized tumors compared to a median TF of 2.96% in ctDNA-positive samples from metastatic tumors. Taken together, these parameters enable the use of the LiquidHALLMARK assay in

disease monitoring, including during MRD, even though the assay is intended for patients with advanced or metastatic tumors for the purposes of guiding clinical therapeutic decisions. To that end, the assay frequently identified actionable biomarkers which directly impact therapeutic decisions. In lung cancer, the assay identified an actionable biomarker in 83.1% of ctDNA-positive cases, including 28.5% of cases where the actionable biomarker was detected at a low VAF of $\leq 0.3\%$. Significantly, nearly a third of first- or second-generation EGFR TKI-associated EGFR T790M resistance mutations, and more than two-fifths of osimertinib-associated resistance mutations, were detected at low VAFs $\leq 0.3\%$. These findings reinforce the necessity of high sensitivity in liquid biopsies, particularly in lung cancers where such actionable findings can directly impact treatment outcomes in patients. Besides the sensitive detection of SNVs and INDELS, gene fusions in *ALK*, *RET*, or *ROS1* were detected in 2.9% of ctDNA-positive lung cancer cases, exemplifying the clinical applicability of an amplicon-based assay for the detection of novel fusions. In addition, the inclusion of less commonly-targeted oncogenic biomarkers in LiquidHALLMARK enhances diagnostic yield. For instance, EBV clearance in ctDNA has been investigated as a surrogate for clinical outcomes in the only prospective, randomized, interventional study in the MRD setting [58] and an EBV qPCR assay has also been used as a blood-based primary screening test for nasopharyngeal cancer [59], while MSI is an actionable biomarker that is routinely tested in tumor tissue but is currently limited in liquid biopsy assays. EBV and HBV detection enabled a further diagnostic yield of 14.3% and 5.0% respectively in nasopharyngeal and liver cancers, while MSI detection provided on-label drug recommendations in six samples and additional diagnostic yield in three ctDNA-negative samples.

Collectively, these results establish comparable analytical performance of an amplicon-based liquid biopsy assay with existing hybrid capture-based NGS liquid biopsy assays [15,16,24]. Although conventional whole-exome amplicon-based NGS has been associated with lower uniformity of depth [60], this limitation is easily circumvented in targeted liquid biopsy panels by rational design, particularly given that cfDNA is fragmented in a non-random fashion [61–63]. Indeed, LiquidHALLMARK demonstrates highly uniform sequencing depth, with an average 97.1% of calls having a depth no less than 0.2X of the mean panel coverage and 79.0% of calls having no less than half the mean panel coverage. This, coupled with the unmatched specificity of PCR combined with UMI and statistical error-modeling, affords the LiquidHALLMARK assay exquisite sensitivity, even at low cfDNA input amounts of 20 ng (containing approximately 6000 haploid genome equivalents), ensuring a high assay robustness (100% report rate in 1,592 consecutive clinical samples) and reliability of test results, and highlighting the value and utility of amplicon-based methods in monitoring oncology treatment sensitivity and resistance. Additional studies to demonstrate clinical validity and utility are ongoing.

Materials and methods

Sample collection and processing

Two 10 mL tubes of whole blood were collected for each patient in Streck Cell-Free DNA BCT (Streck), shipped to Lucence Diagnostics at ambient temperature, and processed within 6–48 h of sample collection. Upon sample receipt, samples were accessioned and both 10 mL tubes of whole blood per patient were fractionated to isolate plasma via centrifugation at 1,600 g for 10 min at 4°C. Isolated plasma was clarified via further centrifugation at 16,000 g for 10 min at 4°C before being aliquoted and stored at -20°C or processed immediately for cfDNA extraction.

cfDNA extraction, library preparation, and sequencing

Cell-free DNA (cfDNA) extraction, library preparation, and sequencing were performed in a CLIA-certified, CAP-accredited laboratory (Lucence Diagnostics). Briefly, cfDNA was extracted from 4–8 mL plasma (from one tube of whole blood) using the QIAamp Circulating Nucleic Acid Kit (Qiagen) and quantified on a Qubit® 2.0 fluorometer (Thermo Fisher), yielding a median of 14.13 ng cfDNA/mL plasma (range 1.15–3160 ng/mL plasma). For each sample, 10–20 ng of extracted cfDNA was converted to an Illumina-compatible sequencing amplicon library, first via limited PCR amplification using random oligonucleotide barcodes (IDT), then enriched with sample-specific barcodes and Illumina sequencing adapters. For the detection of fusions, 7.5–15 ng of extracted cfDNA was barcoded via limited PCR amplification before undergoing A-tailing and single-ended ligation of partial Illumina sequencing adapters, followed by enrichment with sample-specific barcodes and Illumina sequencing adapters. Libraries were purified using AMPure XP beads (Beckman Coulter) and quantified using the KAPA Library Quantification Kit (Roche) before being pooled for sequencing by synthesis on a NextSeq 550 (Illumina) to generate paired-end reads (300 cycles).

Bioinformatics processing and variant calling

Illumina binary base call sequencing files were demultiplexed and converted to fastq file format using the bcl2fastq v2.20 software (Illumina) before being processed with an in-house bioinformatics pipeline (Lucence Diagnostics) for base quality filtering (Phred quality score <15), adapter trimming, and consensus UMI clustering. Consensus reads were then aligned to the hg19 reference genome using BWA-MEM v0.7.17 [64]. For SNV and INDEL detection, consensus reads were compared against position- and sequence run-specific noise level profiles trained using a curated selection of 110 past samples run through the in-house pipeline to estimate the probability of noise for each variant in a sample. Position-specific *p*-value cutoffs were used to discriminate between true and noise variants. All true variants found within coding sequences and splice sites, as well as pathogenic variants within promoter regions based on ClinVar classification were reported [65]. For fusion detection, paired-end reads were mapped to the reference genome and alignment information including directionality and soft-clipped sequences were analyzed to identify potential fusion events. To detect CNAs, probe-level consensus coverage was self-normalized to a set of invariant genomic positions and compared against the normalized coverage distribution constructed from baseline samples before being aggregated to the gene level (z-score) [66]. MSI status was determined by comparing deletion lengths at six repetitive loci (NR27, NR24, BAT25, MONO27, BAT26, and NR21) against a reference baseline; samples were considered MSI-low if one of six markers deviated from the baseline and MSI-high if two or more markers deviated from the baseline. Viral loads of HBV and EBV were quantified based on consensus read coverage using a standard curve method, generated using NIBSC EBV and HBV standards. Alterations were categorized as having strong biological evidence for actionability if the therapeutic indications were FDA-approved, recommended by NCCN or other professional guidelines, or supported by results from well-powered phase III clinical trials with expert consensus. Alterations with actionability evidence in multiple small, published studies or case reports were categorized as having potential clinical actionability.

Contrived sample characteristics

SNVs and INDELS. To determine SNV and INDEL detection performance, the reference standard sets HD780 and Tru-Q, as well as the HD786 reference standard, were used (Horizon Discovery). The HD780 Multiplex I cfDNA Reference Standard Set contains cell-line-derived

DNA fragmented to an average of 160 bp, and harbors eight variants, including six SNVs and two INDELS, in four genes (*EGFR*, *KRAS*, *NRAS*, and *PIK3CA*) at 5%, 1%, and 0.1%, and 0% VAFs. The Tru-Q Reference Standard set contains six standards, covering 38 SNVs and two INDELS across 19 genes at 5%, 1%, and 0% VAF, as well as four additional endogenous variants at VAF $\geq 8\%$. To generate the Tru-Q 0.1% VAF standard, Tru-Q7 (1% VAF standard) was diluted tenfold in Tru-Q0 (wild-type standard). The HD786 Structural Multiplex cfDNA Reference Standard contains six SNVs at 5–5.6% VAF, four INDELS at 5.3–5.6% VAF, one INDEL at 2.5% VAF, and two endogenous SNVs $\geq 16\%$ VAF. All alterations are verified by ddPCR. A full list of alterations covered by the three reference standards is found in [S1 Table](#).

Fusions. In addition to the SNVs and INDELS described above, the HD786 reference standard also harbors a *SLC34A1-ROS1* fusion at 5.6% and *CCDC6-RET* fusion at 5.0%. To generate additional fusion-positive materials, DNA from the *TMPRSS2-ERG* fusion heterozygous cell-line VCaP, the HD664 *EML4-ALK* Reference Standard, 50% (Horizon Discovery), and the HapMap cell-line NA24129 were fragmented to 200 bp by acoustic shearing on an ME220 Focused-ultrasonicator (Covaris). Admixtures of the three cell-lines in 1:1:8, 1:1:48, and 1:1:98 ratios were performed to generate *TMPRSS2-ERG* and *EML4-ALK* fusion at 5%, 1%, and 0.5% VAF.

CNAs. For CNA detection performance, cell-line DNA from MCF7 (Lot #60540387), HCC2218 (Lot #64097142), and HCC827 (Lot #64216187; all obtained from ATCC) and the HapMap cell-line NA24143 were used following acoustic shearing to 200 bp on an ME220 Focused-ultrasonicator. The HCC827 cell-line harbors a 32.6-fold increase in *EGFR* copy number over the diploid number, as verified by the manufacturers via ddPCR. Fragmented HCC2218 DNA was mixed with fragmented NA24143 DNA to generate cell-line DNA admixtures containing *EGFR* copy numbers of 30, 20, 10, 4, 3, and 2.4. The HCC2218 cell-line harbors a 16.32-fold increase in *ERBB2* (HER2) copy number [30], and the MCF7 cell-line harbors an estimated 4.46-fold increase in *NRAS* copy number, 3.34-fold increase in *MYC* copy number, and a 1.80-fold increase in *CDK6* copy number [31,32]. Fragmented MCF7 DNA was mixed with fragmented NA24143 DNA to generate cell-line DNA admixtures containing 4.3 copies of *NRAS* and 3.6 copies of *MYC*, and 3.2 copies of *NRAS* and 2.8 copies of *MYC*. The HD786 reference standard harbors focal *MET* and *MYC* amplification at 4.5 and 9.5 copies respectively.

Oncogenic viruses. For EBV and HBV detection, EBV reference standard (1st WHO International Standard for Epstein-Barr Virus, NIBSC code: 09/260) and HBV reference standard (4th WHO International Standard for HBV DNA for NAT, NIBSC code: 10/266) were each spiked into 1 mL plasma at viral loads spanning 0 to 1,000,000 IU before cfDNA was extracted as previously described.

MSI. For MSI detection, admixtures of the MSI-positive fragmented (200 bp) RKO cell-line DNA with plasma cfDNA in 1:99, 1:95, and 1:1 ratios were performed to generate MSI-positive DNA at 1%, 5%, and 50% TF.

Limit of detection

The lower limit of detection is defined as the lowest VAF at which an alteration can be consistently measured $\geq 90\%$ of the time (LoD₉₀), and was determined empirically based on all assay replicates run on all reference standard materials (HD780, HD786, Tru-Q, and cell-line DNA admixtures) used for analytical validation. The LoD was determined as VAFs for SNVs, INDELS, and fusions, copy number fold-changes for CNAs, and IU/mL plasma for EBV and HBV viral loads.

Per-base specificity

Analytical specificity for SNVs and INDELS was assessed across all base positions interrogated by the panel using 61 non-cancer screening cfDNA samples. Analytical specificity was calculated as the number of alterations detected across a cumulative target region of 3.59 million bp (61 samples x 58.8 kb per sample)

Clinical sample characteristics

Clinical samples used for validation were collected at the Concord Hospital, National University Hospital, and the National Cancer Centre, Singapore under the respective institutional review board (IRB) approvals. Investigations were carried out according to the principles expressed in the Declaration of Helsinki. The 1,592 consecutive clinical samples included in the observational analyses were received as routine clinical samples and underwent standard processing and reporting at a CAP- and CLIA-accredited laboratory (Lucence Diagnostics). Samples collected between January 2018 and May 2021 were used; all patients provided written informed consent, all data was de-identified and exemption from IRB review for study using archived clinical samples and data was provided by Parkway Independent Ethics Committee (PIEC)—Reference PIEC/2022/007. These data span three versions of the LiquidHALLMARK assay (80-gene, 61-gene, and 51-gene panels), with the inclusion of a greater number of targeted genes and genomic regions without changing the underlying molecular principle of the assay. SNVs and INDELS were analyzed in 49–77 genes, CNAs were analyzed in 12–18 genes, and structural rearrangements were analyzed in 3–10 genes, depending on the assay version used (see [S1 Fig](#)).

Clinical and orthogonal validation

The reference AS-PCR method [28] used for orthogonal *EGFR* variant detection was performed in-house using variant-specific PCR primer pairs with wild-type- or mutant-specific fluorescent probes for 10 pre-determined *EGFR* variants. The limit of detection (LoD) for the reference AS-PCR method is 0.05% VAF for *EGFR* exon 19 deletions (E746_A750del, L747_P753delinsS, L747_A750delinsP, and L747_T751del) and L861Q, 0.1% VAF for *EGFR* G719A, G719S, T790M, and L858R, and 0.5% for *EGFR* G719C. For external validation of *EGFR* variants, 2 mL of plasma cfDNA from deidentified clinical samples were submitted to the Mayo Clinic Laboratories (Rochester, MN) for cobas *EGFR* Mutation Test v2. Orthogonal EBV detection was performed by PCR as previously described [33].

Contingency tables were constructed to calculate positive percent agreement (PPA), negative percent agreement (NPA), and overall concordance using standard methods [67]; 95% confidence intervals (95% CI) were calculated using a hybrid Wilson/Brown method [68].

Supporting information

S1 Fig. Genome coverage of the LiquidHALLMARK assay. Genes highlighted in black have near full (>98%) exonic coverage. *TERT* (highlighted in blue) is targeted only in the promoter region. SNV, single nucleotide variant; INDEL insertion/deletion; CNA, copy number alteration. Genes not covered in the 61-gene and 51-gene panels are indicated by * and # respectively. (EPS)

S2 Fig. Statistical background error-modeling enables further error suppression of UMI-corrected sequencing data. Output of all variants (excluding known SNPs) detected in cfDNA following sequencing (top) before and (bottom) after noise filtering based on statistical

error-modelling. In this actual sample, a total of 209 variants are detected even with UMI-based error suppression, owing to the ultra-high depth of sequencing. Following noise filtering, only 19 true variants confidently detected above background noise remain.

(EPS)

S3 Fig. Validation of MSI detection. Observed deletion length at the (A) NR27, (B) NR24, (C) BAT25, (D) MONO27, (E) BAT26, and (F) NR21 locus in plasma cfDNA (0%) and 1%, 5%, and 50% admixtures of MSI-positive RKO cell-line DNA in plasma cfDNA, representing the tumor fraction (% TF). VAF, variant allele frequency. Broken lines represent the 95th percentile of deletion length in 91 MSI-negative plasma cfDNA and 29 MSI-negative FFPE tissue samples.

(EPS)

S4 Fig. Prevalence of the top 10 genes altered in (A) all ctDNA-positive samples, (B) ctDNA-positive lung cancers, and (C) ctDNA-positive breast cancers.

(EPS)

S5 Fig. Prevalence of resistance-associated alterations in (A) HR+ HER2- breast cancer and (B) colorectal cancer. Amp, amplification; EGFR ECD, EGFR extracellular domain.

(EPS)

S6 Fig. Prevalence and variant type of all alterations detected in cancers of unknown primary (CUP). Each color represents a different variant type.

(EPS)

S7 Fig. Prevalence of biomarkers altered in (A) nasopharyngeal and (B) liver cancer samples. Epstein-Barr virus (EBV) and hepatitis B virus (HBV) are highlighted in red.

(EPS)

S1 Table. Analytical validation of SNV/INDELS using Horizon Discovery reference standards HD780, HD786, and Tru-Q.

(XLSX)

S2 Table. EGFR variants detected by LiquidHALLMARK and a reference AS-PCR assay.

(XLSX)

S3 Table. Analytical validation of fusions using Horizon Discovery reference standards HD786 and HD664, and VCaP cell-line.

(XLSX)

S4 Table. PD-L1 structural variants detected by LiquidHALLMARK and a reference assay.

(XLSX)

S5 Table. ERBB2 copy number gains detected by LiquidHALLMARK in cfDNA and by ERBB2 FISH.

(XLSX)

S6 Table. Plasma EBV DNA detected by LiquidHALLMARK and a reference PCR assay.

(XLSX)

S7 Table. Microsatellite instability detected by LiquidHALLMARK and fragment size analysis.

(XLSX)

S8 Table. Orthogonal comparison of LiquidHALLMARK cfDNA and matched tissue testing for microsatellite instability.

(XLSX)

S9 Table. EGFR variants detected by LiquidHALLMARK and cobas EGFR Mutation Test v2.

(XLSX)

S1 File. Minimal data set for figures.

(PDF)

Acknowledgments

The authors would like to thank all staff at Lucence Diagnostics who have contributed to the development and validation of LiquidHALLMARK.

Author Contributions

Conceptualization: Jonathan Poh, Kao Chin Ngeow, Kian-Hin Tan, Yukti Choudhury.

Data curation: Jonathan Poh, Michelle Pek, Jing Shan Lim, Yukti Choudhury.

Formal analysis: Jonathan Poh, Kao Chin Ngeow, Kian-Hin Tan, Yukti Choudhury.

Investigation: Jonathan Poh, Kao Chin Ngeow, Michelle Pek, Kian-Hin Tan, Jing Shan Lim, Yukti Choudhury.

Project administration: Yukti Choudhury.

Resources: Choon Kiat Ong, Jing Quan Lim, Soon Thye Lim, Chwee Ming Lim, Boon Cher Goh.

Software: Hao Chen.

Supervision: Jonathan Poh, Kao Chin Ngeow, Michelle Pek, Yukti Choudhury.

Validation: Jonathan Poh, Kao Chin Ngeow, Kian-Hin Tan.

Visualization: Jonathan Poh, Yukti Choudhury.

Writing – original draft: Jonathan Poh, Yukti Choudhury.

Writing – review & editing: Jonathan Poh, Kao Chin Ngeow, Michelle Pek, Kian-Hin Tan, Jing Shan Lim, Hao Chen, Choon Kiat Ong, Jing Quan Lim, Soon Thye Lim, Chwee Ming Lim, Boon Cher Goh, Yukti Choudhury.

References

1. Basik M, Aguilar-Mahecha A, Rousseau C, Diaz Z, Tejpar S, Spatz A, et al. Biopsies: Next-generation biospecimens for tailoring therapy. *Nature Reviews Clinical Oncology*. 2013; 10: 437–450. <https://doi.org/10.1038/nrclinonc.2013.101> PMID: 23799370
2. Al-Kateb H, Nguyen TDT, Steger-May K, Pfeifer JD. Identification of major factors associated with failed clinical molecular oncology testing performed by next generation sequencing (NGS). *Molecular Oncology*. 2015; 9: 1737–1743. <https://doi.org/10.1016/j.molonc.2015.05.004> PMID: 26071350
3. Hagemann IS, Devarakonda S, Lockwood CM, Spencer DH, Guebert K, Bredemeyer AJ, et al. Clinical next-generation sequencing in patients with non-small cell lung cancer. *Cancer*. 2015; 121: 631–639. <https://doi.org/10.1002/cncr.29089> PMID: 25345567
4. Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, et al. Liquid biopsies come of age: Towards implementation of circulating tumour DNA. *Nature Reviews Cancer*. 2017; 17: 223–238. <https://doi.org/10.1038/nrc.2017.7> PMID: 28233803
5. Bettgowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Science Translational Medicine*. 2014; 6. <https://doi.org/10.1126/scitranslmed.3007094> PMID: 24553385

6. Network NCC. Clinical practice guidelines in oncology. Non–small-cell lung cancer. Version 3.2020. Updated 2/11/2020. https://www.nccn.org/professionals/physician_gls/pdf; 2020.
7. Yuan M, Huang LL, Chen JH, Wu J, Xu Q. The emerging treatment landscape of targeted therapy in non-small-cell lung cancer. *Signal Transduction and Targeted Therapy*. 2019. <https://doi.org/10.1038/s41392-019-0099-9> PMID: 31871778
8. 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. <https://doi.org/10.1016/j.bdq.2019.100087> PMID: 30923679
9. Cescon DW, Bratman S V., Chan SM, Siu LL. Circulating tumor DNA and liquid biopsy in oncology. *Nature Cancer*. 2020; 1: 276–290. <https://doi.org/10.1038/s43018-020-0043-5> PMID: 35122035
10. Ou S-HI, Nagasaka M, Zhu VW. Liquid Biopsy to Identify Actionable Genomic Alterations. *American Society of Clinical Oncology Educational Book*. 2018; 38: 978–997. https://doi.org/10.1200/edbk_199765 PMID: 30231331
11. FDA. FDA Approves First Liquid Biopsy Next-Generation Sequencing Companion Diagnostic Test | FDA. In: FDA New Release: Press Announcements [Internet]. 2020. Available: <https://www.fda.gov/news-events/press-announcements/fda-approves-first-liquid-biopsy-next-generation-sequencing-companion-diagnostic-test>.
12. FDA. FDA approves liquid biopsy NGS companion diagnostic test for multiple cancers and biomarkers. FDA. 2020.
13. Aggarwal C, Thompson JC, Black TA, Katz SI, Fan R, Yee SS, et al. Clinical Implications of Plasma-Based Genotyping with the Delivery of Personalized Therapy in Metastatic Non-Small Cell Lung Cancer. *JAMA Oncology*. 2019; 5: 173–180. <https://doi.org/10.1001/jamaoncol.2018.4305> PMID: 30325992
14. Li BT, Janku F, Jung B, Hou C, Madwani K, Alden R, et al. Ultra-deep next-generation sequencing of plasma cell-free DNA in patients with advanced lung cancers: Results from the actionable genome consortium. *Annals of Oncology*. 2019; 30: 597–603. <https://doi.org/10.1093/annonc/mdz046> PMID: 30891595
15. Odegaard JI, Vincent JJ, Mortimer S, Vowles J V., Ulrich BC, Banks KC, et al. Validation of a plasma-based comprehensive cancer genotyping assay utilizing orthogonal tissue- and plasma-based methodologies. *Clinical Cancer Research*. 2018; 24: 3539–3549. <https://doi.org/10.1158/1078-0432.CCR-17-3831> PMID: 29691297
16. Clark TA, Chung JH, Kennedy M, Hughes JD, Chennagiri N, Lieber DS, et al. Analytical Validation of a Hybrid Capture–Based Next-Generation Sequencing Clinical Assay for Genomic Profiling of Cell-Free Circulating Tumor DNA. *Journal of Molecular Diagnostics*. 2018; 20: 686–702. <https://doi.org/10.1016/j.jmoldx.2018.05.004> PMID: 29936259
17. Fox EJ, Reid-Bayliss KS. Accuracy of Next Generation Sequencing Platforms. *Journal of Next Generation Sequencing & Applications*. 2014; 01. <https://doi.org/10.4172/jngsa.1000106> PMID: 25699289
18. Salk JJ, Schmitt MW, Loeb LA. Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nature Reviews Genetics*. 2018. pp. 269–285. <https://doi.org/10.1038/nrg.2017.117> PMID: 29576615
19. Newman AM, Lovejoy AF, Klass DM, Kurtz DM, Chabon JJ, Scherer F, et al. Integrated digital error suppression for improved detection of circulating tumor DNA. *Nature Biotechnology*. 2016; 34: 547–555. <https://doi.org/10.1038/nbt.3520> PMID: 27018799
20. Stetson D, Ahmed A, Xu X, Nuttall BRB, Lubinski TJ, Johnson JH, et al. Orthogonal Comparison of Four Plasma NGS Tests With Tumor Suggests Technical Factors are a Major Source of Assay Discordance. *JCO Precision Oncology*. 2019; 1–9. <https://doi.org/10.1200/PO.18.00191> PMID: 35100678
21. Meador CB, Oxnard GR. Effective cancer genotyping—many means to one end. *Clinical Cancer Research*. 2019; 25: 4583–4585. <https://doi.org/10.1158/1078-0432.CCR-19-1233> PMID: 31138589
22. Toppmeyer DL, Press MF. Testing considerations for phosphatidylinositol-3-kinase catalytic subunit alpha as an emerging biomarker in advanced breast cancer. *Cancer Medicine*. 2020. pp. 6463–6472. <https://doi.org/10.1002/cam4.3278> PMID: 32697890
23. Anedda A, López-Bernardo E, Acosta-Iborra B, Saadeh Suleiman M, Landázuri MO, Cadenas S. The transcription factor Nrf2 promotes survival by enhancing the expression of uncoupling protein 3 under conditions of oxidative stress. *Free Radical Biology and Medicine*. 2013; 61: 395–407. <https://doi.org/10.1016/j.freeradbiomed.2013.04.007> PMID: 23597505
24. Woodhouse R, Li M, Hughes J, Delfosse D, Skoletsky J, Ma P, et al. Clinical and analytical validation of foundation one liquid CDx, a novel 324-Gene cfDNA-based comprehensive genomic profiling assay for cancers of solid tumor origin. *PLoS ONE*. 2020; 15: e0237802. <https://doi.org/10.1371/journal.pone.0237802> PMID: 32976510

25. Gale D, Lawson ARJ, Howarth K, Madi M, Durham B, Smalley S, et al. Development of a highly sensitive liquid biopsy platform to detect clinically-relevant cancer mutations at low allele fractions in cellfree DNA. *PLoS ONE*. 2018; 13: e0194630. <https://doi.org/10.1371/journal.pone.0194630> PMID: 29547634
26. Williamson JB, Solano L, Yuki A, Burkhart VD, Chitwood J, Cao R, et al. Analytical validation of the OncoPrint Pan-Cancer Cell-Free Assay in a CLIA- and CAP-regulated laboratory for detection of solid tumor-derived variants in blood plasma. *Journal of Clinical Oncology*. 2019; 37: e14614–e14614. https://doi.org/10.1200/jco.2019.37.15_suppl.e14614
27. Heitzer E, Haque IS, Roberts CES, Speicher MR. Current and future perspectives of liquid biopsies in genomics-driven oncology. *Nature Reviews Genetics*. 2019; 20: 71–88. <https://doi.org/10.1038/s41576-018-0071-5> PMID: 30410101
28. Shi J, Mong J, Chin TM, Lim YH, Tan WL, Toh CK, et al. EGFR mutation detection in plasma cell-free DNA correlates with clinical outcomes in non-small cell lung cancer. *Annals of Oncology*. 2017; 28: v476–v477. <https://doi.org/10.1093/annonc/mdx380.040>
29. Lim JQ, Huang D, Tang T, Tan D, Laurensia Y, Peng RJ, et al. Whole-genome sequencing identifies responders to Pembrolizumab in relapse/refractory natural-killer/T cell lymphoma. *Leukemia*. 2020; 34: 3413–3419. <https://doi.org/10.1038/s41375-020-1000-0> PMID: 32753688
30. Lanman RB, Mortimer SA, Zill OA, Sebisano D, Lopez R, Blau S, et al. Analytical and clinical validation of a digital sequencing panel for quantitative, highly accurate evaluation of cell-free circulating tumor DNA. *PLoS ONE*. 2015; 10: e0140712. <https://doi.org/10.1371/journal.pone.0140712> PMID: 26474073
31. Reinhold WC, Sunshine M, Liu H, Varma S, Kohn KW, Morris J, et al. CellMiner: A web-based suite of genomic and pharmacologic tools to explore transcript and drug patterns in the NCI-60 cell line set. *Cancer Research*. 2012; 72: 3499–3511. <https://doi.org/10.1158/0008-5472.CAN-12-1370> PMID: 22802077
32. Shankavaram UT, Varma S, Kane D, Sunshine M, Chary KK, Reinhold WC, et al. CellMiner: A relational database and query tool for the NCI-60 cancer cell lines. *BMC Genomics*. 2009; 10: 277. <https://doi.org/10.1186/1471-2164-10-277> PMID: 19549304
33. Vo JH, Nei WL, Hu M, Phyo WM, Wang F, Fong KW, et al. Comparison of Circulating Tumour Cells and Circulating Cell-Free Epstein-Barr Virus DNA in Patients with Nasopharyngeal Carcinoma Undergoing Radiotherapy. *Scientific Reports*. 2016; 6: 1–9. <https://doi.org/10.1038/s41598-016-0001-8> PMID: 28442746
34. Gray JE, Okamoto I, Sriuranpong V, Vansteenkiste J, Imamura F, Lee JS, et al. Tissue and plasma EGFR mutation analysis in the FLAURA trial: Osimertinib versus Comparator EGFR tyrosine kinase inhibitor as first-line treatment in patients with EGFR-mutated advanced non-small cell lung cancer. *Clinical Cancer Research*. 2019; 25: 6644–6652. <https://doi.org/10.1158/1078-0432.CCR-19-1126> PMID: 31439584
35. Cobas E. Mutation Test v2. FDA website. 2017. <https://doi.org/10.1080/14737159.2017.1288568> PMID: 28129709
36. Kim Y, Shin S, Lee KA. A Comparative Study for Detection of EGFR Mutations in Plasma Cell-Free DNA in Korean Clinical Diagnostic Laboratories. *BioMed Research International*. 2018; 2018. <https://doi.org/10.1155/2018/7392419> PMID: 29854785
37. Oxnard GR, Paweletz CP, Kuang Y, Mach SL, O'Connell A, Messineo MM, et al. Noninvasive detection of response and resistance in egfrmutant lung cancer using quantitative next-generation genotyping of cell-free plasma DNA. *Clinical Cancer Research*. 2014; 20: 1698–1705. <https://doi.org/10.1158/1078-0432.CCR-13-2482> PMID: 24429876
38. Lee Y, Park S, Kim WS, Lee JC, Jang SJ, Choi J, et al. Correlation between progression-free survival, tumor burden, and circulating tumor DNA in the initial diagnosis of advanced-stage EGFR-mutated non-small cell lung cancer. *Thoracic Cancer*. 2018; 9: 1104–1110. <https://doi.org/10.1111/1759-7714.12793> PMID: 29989342
39. Namløs HM, Boye K, Mishkin SJ, Barøy T, Lorenz S, Bjerkehagen B, et al. Noninvasive detection of ctDNA reveals intratumor heterogeneity and is associated with tumor burden in gastrointestinal stromal tumor. *Molecular Cancer Therapeutics*. 2018; 17: 2473–2480. <https://doi.org/10.1158/1535-7163.MCT-18-0174> PMID: 30097488
40. Marczyński GT, Laus AC, dos Reis MB, Reis RM, Vazquez V de L. Circulating tumor DNA (ctDNA) detection is associated with shorter progression-free survival in advanced melanoma patients. *Scientific Reports*. 2020; 10: 1–11. <https://doi.org/10.1038/s41598-019-56847-4> PMID: 31913322
41. Priestley P, Baber J, Lolkema MP, Steeghs N, de Bruijn E, Shale C, et al. Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature*. 2019; 575: 210–216. <https://doi.org/10.1038/s41586-019-1689-y> PMID: 31645765

42. The I, of Whole TP-CA, Consortium G. Pan-cancer analysis of whole genomes. *Nature*. 2020; 578: 82. <https://doi.org/10.1038/s41586-020-1969-6> PMID: 32025007
43. Gainor J, Dardaei L, Yoda S, Friboulet L, Leischler I, Katayama R, et al. Molecular mechanisms of resistance to first- and second-generation ALK inhibitors in ALK-rearranged lung cancer. *European Journal of Cancer*. 2016; 69: S138. [https://doi.org/10.1016/s0959-8049\(16\)33009-x](https://doi.org/10.1016/s0959-8049(16)33009-x)
44. Denis MG, Vallée A, Théoleyre S. EGFR T790M resistance mutation in non small-cell lung carcinoma. *Clinica Chimica Acta*. 2015; 444: 81–85. <https://doi.org/10.1016/j.cca.2015.01.039> PMID: 25668228
45. Piotrowska Z, Isozaki H, Lennerz JK, Gainor JF, Lennes IT, Zhu VW, et al. Landscape of acquired resistance to osimertinib in EGFR-mutant NSCLC and clinical validation of combined EGFR and RET inhibition with osimertinib and BLU-667 for acquired RET fusion. *Cancer Discovery*. 2018; 8: 1529–1539. <https://doi.org/10.1158/2159-8290.CD-18-1022> PMID: 30257958
46. Jeselsohn R, Buchwalter G, De Angelis C, Brown M, Schiff R. ESR1 mutations—a mechanism for acquired endocrine resistance in breast cancer. *Nature Reviews Clinical Oncology*. 2015; 12: 573–583. <https://doi.org/10.1038/nrclinonc.2015.117> PMID: 26122181
47. André F, Ciruelos E, Rubovszky G, Campone M, Loibl S, Rugo HS, et al. Alpelisib for PIK3CA -Mutated, Hormone Receptor–Positive Advanced Breast Cancer. *New England Journal of Medicine*. 2019; 380: 1929–1940. <https://doi.org/10.1056/NEJMoa1813904> PMID: 31091374
48. O’leary B, Cutts RJ, Liu Y, Hrebien S, Huang X, Fenwick K, et al. The genetic landscape and clonal evolution of breast cancer resistance to palbociclib plus fulvestrant in the PALOMA-3 trial. *Cancer Discovery*. 2018; 8: 1390–1403. <https://doi.org/10.1158/2159-8290.CD-18-0264> PMID: 30206110
49. Zhao B, Wang L, Qiu H, Zhang M, Sun L, Peng P, et al. Mechanisms of resistance to anti-EGFR therapy in colorectal cancer. *Oncotarget*. 2017; 8: 3980–4000. <https://doi.org/10.18632/oncotarget.14012> PMID: 28002810
50. Brand TM, Iida M, Wheeler DL. Molecular mechanisms of resistance to the EGFR monoclonal antibody cetuximab. *Cancer Biology and Therapy*. 2011; 11: 777–792. <https://doi.org/10.4161/cbt.11.9.15050> PMID: 21293176
51. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: The Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Research*. 2019; 47: D941–D947. <https://doi.org/10.1093/nar/gky1015> PMID: 30371878
52. Fung SYH, Lam JWK, Chan KCA. Clinical utility of circulating Epstein-Barr virus DNA analysis for the management of nasopharyngeal carcinoma. *Chinese Clinical Oncology*. 2016; 5: 18. <https://doi.org/10.21037/cco.2016.03.07> PMID: 27121878
53. Pessoa LS, Heringer M, Ferrer VP. ctDNA as a cancer biomarker: A broad overview. *Critical Reviews in Oncology/Hematology*. 2020; 155: 103109. <https://doi.org/10.1016/j.critrevonc.2020.103109> PMID: 33049662
54. Alizadeh AA, Aranda V, Bardelli A, Blanpain C, Bock C, Borowski C, et al. Toward understanding and exploiting tumor heterogeneity. *Nature Medicine*. 2015; 21: 846–853. <https://doi.org/10.1038/nm.3915> PMID: 26248267
55. Swanton C. Intratumor heterogeneity: Evolution through space and time. *Cancer Research*. 2012; 72: 4875–4882. <https://doi.org/10.1158/0008-5472.CAN-12-2217> PMID: 23002210
56. Gerlinger M, Catto JW, Orntoft TF, Real FX, Zwarthoff EC, Swanton C. Intratumour heterogeneity in urologic cancers: From molecular evidence to clinical implications. *European Urology*. 2015; 67: 729–737. <https://doi.org/10.1016/j.eururo.2014.04.014> PMID: 24836153
57. Crowley E, Di Nicolantonio F, Loupakis F, Bardelli A. Liquid biopsy: monitoring cancer-genetics in the blood. *Nature reviews Clinical oncology*. 2013; 10: 472. <https://doi.org/10.1038/nrclinonc.2013.110> PMID: 23836314
58. Chan ATC, Hui EP, Ngan RKC, Tung SY, Cheng ACK, Ng WT, et al. Analysis of plasma Epstein-Barr virus DNA in nasopharyngeal cancer after chemoradiation to identify high-risk patients for adjuvant chemotherapy: A randomized controlled trial. *Journal of Clinical Oncology*. 2018. pp. 3091–3100. <https://doi.org/10.1200/JCO.2018.77.7847> PMID: 29989858
59. Chan KCA, Woo JKS, King A, Zee BCY, Lam WKJ, Chan SL, et al. Analysis of Plasma Epstein–Barr Virus DNA to Screen for Nasopharyngeal Cancer. *New England Journal of Medicine*. 2017; 377: 513–522. <https://doi.org/10.1056/NEJMoa1701717> PMID: 28792880
60. Samorodnitsky E, Jewell BM, Hagopian R, Miya J, Wing MR, Lyon E, et al. Evaluation of Hybridization Capture Versus Amplicon-Based Methods for Whole-Exome Sequencing. *Human Mutation*. 2015; 36: 903–914. <https://doi.org/10.1002/humu.22825> PMID: 26110913
61. Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. Cell-free DNA comprises an in vivo nucleosome footprint that informs its tissues-of-origin. *Transplantation*. 2016; 100: 698. <https://doi.org/10.1097/TP.0000000000001176>

62. Jiang P, Sun K, Tong YK, Cheng SH, Cheng THT, Heung MMS, et al. Preferred end coordinates and somatic variants as signatures of circulating tumor DNA associated with hepatocellular carcinoma. *Proceedings of the National Academy of Sciences of the United States of America*. 2018; 115: E10925–E10933. <https://doi.org/10.1073/pnas.1814616115> PMID: 30373822
63. Chandrananda D, Thorne NP, Bahlo M. High-resolution characterization of sequence signatures due to non-random cleavage of cell-free DNA. *BMC Medical Genomics*. 2015; 8: 29. <https://doi.org/10.1186/s12920-015-0107-z> PMID: 26081108
64. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:13033997. 2013.
65. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: Improving access to variant interpretations and supporting evidence. *Nucleic Acids Research*. 2018; 46: D1062–D1067. <https://doi.org/10.1093/nar/gkx1153> PMID: 29165669
66. Eijkelenboom A, Tops BBJ, van den Berg A, van den Brule AJC, Dinjens WNM, Dubbink HJ, et al. Recommendations for the clinical interpretation and reporting of copy number gains using gene panel NGS analysis in routine diagnostics. *Virchows Archiv*. 2019. <https://doi.org/10.1007/s00428-019-02555-3> PMID: 30888490
67. Godsey JH, Silvestro A, Barrett JC, Bramlett K, Chudova D, Deras I, et al. Generic Protocols for the Analytical Validation of Next-Generation Sequencing-Based ctDNA Assays: A Joint Consensus Recommendation of the BloodPAC's Analytical Variables Working Group. *Clinical chemistry*. 2020; 66: 1156–1166. <https://doi.org/10.1093/clinchem/hvaa164> PMID: 32870995
68. Brown LD, Cai TT, Das Gupta A. Interval estimation for a binomial proportion. *Statistical Science*. 2001; 16: 101–117. <https://doi.org/10.1214/ss/1009213286>