Validation of Microsatellite Instability Detection Using a Comprehensive Plasma-Based Genotyping Panel


Abstract

Purpose: To analytically and clinically validate microsatellite instability (MSI) detection using cell-free DNA (cfDNA) sequencing.

Experimental Design: Pan-cancer MSI detection using Guardant360 was analytically validated according to established guidelines and clinically validated using 1,145 cfDNA samples for which tissue MSI status based on standard-of-care tissue testing was available. The landscape of cfDNA-based MSI across solid tumor types was investigated in a cohort of 28,459 clinical plasma samples. Clinical outcomes for 16 patients with cfDNA MSI-H gastric cancer treated with immunotherapy were evaluated.

Results: cfDNA MSI evaluation was shown to have high specificity, precision, and sensitivity, with a limit of detection of 0.1% tumor content. In evaluable patients, cfDNA testing accurately detected 87% (71/82) of tissue MSI-H and 99.5% of tissue microsatellite stable (863/867) for an overall accuracy of 98.4% (934/949) and a positive predictive value of 95% (71/75). Concordance of cfDNA MSI with tissue PCR and next-generation sequencing was significantly higher than IHC. Prevalence of cfDNA MSI for major cancer types was consistent with those reported for tissue. Finally, robust clinical activity of immunotherapy treatment was seen in patients with advanced gastric cancer positive for MSI by cfDNA, with 63% (10/16) of patients achieving complete or partial remission with sustained clinical benefit.

Conclusions: cfDNA-based MSI detection using Guardant360 is highly concordant with tissue-based testing, enabling highly accurate detection of MSI status concurrent with comprehensive genomic profiling and expanding access to immunotherapy for patients with advanced cancer for whom current testing practices are inadequate. See related commentary by Wang and Ajani, p. 6887

Introduction

Microsatellite instability (MSI) is a National Comprehensive Cancer Network (NCCN) clinical practice guidelines-recommended biomarker in at least nine cancer types—cervical, cholangiocarcinoma, colorectal, endometrial, esophageal and esophagogastric, gastric, ovarian, pancreatic, and prostate cancers (1–9)—due to its importance as a predictive biomarker for response to immune checkpoint blockade (ICB) as exemplified by pan-cancer approval of pembrolizumab (10, 11). Detection of MSI in a patient with advanced cancer can also alert the...
Willis et al.

Microsatellite instability (MSI) is an important biomarker predictive of response to immune checkpoint blockade (ICB) across solid cancers and may herald the presence of a hereditary cancer predisposition syndrome. However, despite being recommended by clinical practice guidelines, MSI is most often not assessed, in part due to tissue insufficiency, unavailability, or infeasibility. Noninvasive blood-based methods have been developed to detect MSI; however, these approaches are challenged by limited sensitivity. Here, we report the validation of a highly sensitive cell-free DNA (cfDNA) sequencing-based method that allows simultaneous guideline-comprehensive determination of MSI status and genomic biomarkers for all solid cancers. We further validate the clinical relevance of cfDNA-based MSI detection by reporting robust clinical activity of ICB therapy in 16 patients with MSI-H gastric cancer.

Materials and Methods
Microsatellite loci selection
The Guardant360 test is a 74-gene panel previously validated for detection of SNVs, indels, CNAs, and fusions in all guideline-recommended indications for advanced solid tumors (25, 26). The assay initially incorporated 99 putative microsatellite loci consisting of short-tandem repeats of length 7 or more, which were selected to include sites susceptible to instability across multiple tumor types, including three of the five Bethesda panel sites (BAT-25, BAT-26, and NR-21). The remaining two Bethesda sites (NR-24 and MONO-27) were not included because of extremely low mappability of the genomic intervals.

In this study, we sought to enhance the utility of a previously validated cfDNA-based genotyping test through the addition of MSI detection. Here, we describe the design and validation of MSI assessment on this platform, report its performance on the largest comprehensive liquid biopsies can provide all guideline-recommended somatic genomic biomarker information for all adult solid tumors in a single test.

Future directions
Despite recommendations across many cancer types in NCCN guidelines and associated FDA-approved treatment options, current rates of MSI testing outside of colorectal cancer and gastrointestinal carcinoma remain largely unknown. Even in colorectal cancer, where MSI testing recommendations have been in place since 2005 (17, 22), fewer than 50% of patients are tested (23), which results in missed ICB treatment opportunities and failure to recognize patients whose family members may be at increased risk for cancer. While multifactorial, such MSI undergenotyping is, in part, due to barriers associated with tissue acquisition, including difficulties locating archival diagnostic specimens or delays due to biopsy scheduling to obtain new tissue. In addition, invasive tissue acquisition procedures may be contraindicated in many heavily pretreated and/or frail patients and have the associated disadvantages of higher cost and procedural risk. Furthermore, the rapidly growing number of biomarkers and diversification of testing options creates daunting complexity for already overburdened physicians.

Cell-free circulating-tumor DNA (cfDNA) assays ("liquid biopsies") have successfully addressed such barriers in many genotyping indications by enabling minimally invasive profiling of contemporaneous tumor DNA. Liquid biopsies thus expand patient access to standard-of-care–targeted therapies, including ICBs, by identifying patients whose tumors harbor biomarkers of interest not otherwise identifiable due to tissue sampling limitations and do so more rapidly than typical tissue testing (24). Moreover, comprehensive liquid biopsies can provide all guideline-recommended somatic genomic biomarker information for all adult solid tumors in a single test.

In this study, we sought to enhance the utility of a previously validated cfDNA-based genotyping test through the addition of MSI detection. Here, we describe the design and validation of MSI assessment on this platform, report its performance on the largest cfDNA-tissue MSI validation cohort yet described (n = 1,145) and evaluate response prediction in 16 patients with advanced gastric cancer treated with ICB. We also report the MSI-H landscape more than 28,000 consecutive solid tumor patient samples tested in our Clinical Laboratory Improvement Amendment (CLIA)-certified, College of American Pathologists (CAP)-accredited, New York State Department of Health-approved laboratory.

Translational Relevance
Microsatellite instability (MSI) is an important biomarker predictive of response to immune checkpoint blockade (ICB) across solid cancers and may herald the presence of a hereditary cancer predisposition syndrome. However, despite being recommended by clinical practice guidelines, MSI is most often not assessed, in part due to tissue insufficiency, unavailability, or infeasibility. Noninvasive blood-based methods have been developed to detect MSI; however, these approaches are challenged by limited sensitivity. Here, we report the validation of a highly sensitive cell-free DNA (cfDNA) sequencing-based method that allows simultaneous guideline-comprehensive determination of MSI status and genomic biomarkers for all solid cancers. We further validate the clinical relevance of cfDNA-based MSI detection by reporting robust clinical activity of ICB therapy in 16 patients with cfDNA MSI-H gastric cancer.
Model description
MSI detection is based on integrating observed read sequences with molecular barcoding information into a single probabilistic model that compares the likelihood of observed data under PCR and sequencing noise assumptions with that under somatic MSI assumption. Each individual site with sufficient coverage is scored independently using Akaike Information Criterion (AIC; ref. 27), which is a statistical method for estimating the relative validity of different models describing how a given dataset was generated. The AIC model generates a locus score (ranging from 0 to infinity), reflecting the likelihood that observed variability at any given microsatellite locus is due to biological instability versus noise, and a locus is considered unstable if its score is above a trained threshold. The number of affected loci is calculated across the final 90 sites and the sample is called positive if the number of unstable loci (the "MSI Score") is above a trained threshold (n = 6). The thresholds for individual loci and total MSI score per sample were established using permutation-based simulations with data from healthy donor samples varying the frequencies of molecules with different repeat lengths and the error parameters at individual loci, as well as the overall number of unstable loci within a simulated sample. Through this approach, simulations were used to interrogate 100,000 combinations of microsatellite lengths and unstable locus numbers, which allows assessment of a diverse landscape of scenarios, some of which may not be represented in a nonsimulated dataset. The algorithm does not distinguish between microsatellite stable (MSS) and MSI-Low (a category defined by the observation of a single unstable Bethesda locus using PCR methods), grouping them into a single category. This is based on previous qpcr-colec-tion analysis of number of variants per sample and Fisher exact test for evidence intervals (CI) for binomial proportions were calculated using Wilson score interval with continuity correction.

Samples
MSI algorithm development and training was performed using simulated data as well as a set of 84 healthy donor samples. The clinical validation study included 1,145 archived samples [residual plasma and/or cell-free DNA (cfDNA)] collected and processed as part of routine standard-of-care clinical testing in the Guardant Health CLIA laboratory as described previously (25), or archival patient plasma samples collected in EDTA blood collection tubes. Twenty healthy donor samples were also used for the analytical specificity study. Contrived samples used in the analytical validation studies comprise cfDNA pools extracted from cell line supernatants and healthy donor plasma. cfDNA was prepared from culture supernatants as described previously (25) using the following cell lines (ATCC, Inc.): KM12, NCI-H660, HCC1419, NCI-H2228, NCI-H1650, NCI-H1648, NCI-H1975, NCI-H1993, NCI-H596, HCC78, GM12878, MCF-7. cfDNA isolated from cell line culture supernatant mimics the fragment size and mechanisms of extracellular release (29). Library conversion, and sequencing properties of patient-derived cfDNA, while also providing a renewable source of well-defined material of sufficient quantity to support the high material demands of studies such as limit of detection and precision.

Sample processing and bioinformatics analysis
cfDNA was extracted from plasma samples or cell line supernatants (QiAmp Circulating Nucleic Acid Kit, Qiagen, Inc.), and up to 30 ng of extracted cfDNA was labeled with nonrandom oligonucleotide barcodes (IDT, Inc.), followed by library preparation, hybrid capture enrichment (Agilent Technologies, Inc.), and sequencing by paired-end synthesis (NextSeq 500/500 or HiSeq 2500, Illumina, Inc.) as described previously. Bioinformatics analysis and variant detection were performed as described previously (25).

Analytical validation approach
The studies performed for analytical validation were based on established CLIA, Nex-StoCT Working Group, and Association of Molecular Pathologists/CAP guidance regarding performance characteristics and validation principles. To determine the sensitivity of the assay for MSI status, cfDNA from cell line supernatant from an MSI-H cell line derived from a patient with human colon cancer (KM12; ref. 30) was diluted with cfDNA from a MSS cell line (NCI-H660; refs. 31, 32) and tested at both standard (30 ng) and low (5 ng) cfDNA inputs. The dilution series targeted maximal variant allele fractions (max VAF) of 0.03%–2% for 5-ng input, and 0.01%–1% for 30-ng input. Targeted tumor fractions were verified using known germline variants unique to the titrant and diluent materials. Assessment of repeatability (within-run precision) and reproducibility (between-run precision) was based on clinical and contrived model samples. Six of the clinical samples for precision (three MSI-H, and three MSS) were selected with max VAF values of 1%–2%, representing approximately two to three times the predicted limit of detection (LoD) at 5 ng. MSI analytical specificity was determined by analyzing 20 healthy donor samples and 245 known MSS-contrived samples.

Clinical validation approach
Archived plasma or cfDNA from clinical samples from patients with available results from standard-of-care tissue-based MSI testing were tested using the ctDNA MSI algorithm (n = 1,145). Tissue-based MSI status was derived from IHC, PCR, or, less commonly, NGS. Clinical outcome data were extracted from patient medical records and deidentified by the treating physician.

Landscape analysis of plasma MSI status from 28,459 advanced cancer patient samples
The cohort comprised 28,459 consecutive advanced cancer patient samples tested using Guardant360 in the course of clinical care. All analyses were conducted with deidentified data and according to an IRB-approved protocol. The prevalence of MSI-H in this cohort was assessed across 16 primary tumor types: bladder carcinoma, breast carcinoma, cholangiocarcinoma, colon adenocarcinoma, colon adenocarcinoma, cancer of unknown primary, head and neck squamous cell carcinoma, hepatocellular carcinoma, lung adenocarcinoma, lung cancer not otherwise specified, lung squamous cell carcinoma, "other" cancer diagnosis, pancreatic adenocarcinoma, prostate adenocarcinoma, stomach adenocarcinoma, and uterine endometrial carcinoma.

Statistical analysis
Statistical analyses were performed using Student t test for analysis of number of variants per sample and Fisher exact test for comparison of proportions. The lower and upper limits of the 95% confidence intervals (CI) for binomial proportions were calculated using Wilson score interval with continuity correction.

Ethics
This research was conducted utilizing deidentified data as per a protocol approved by the Quorum Institutional Review Board.
Results

MSI algorithm development

Traditional challenges for ctDNA genotyping using NGS include efficient molecule capture due to low inputs and low tumor fraction in circulation (25, 26) and correction of sequencing and other technical artifacts. MSI detection presents additional challenges due to the need for (i) efficient molecular capture, sequencing, and mapping of repetitive genomic regions that accurately reflect MSI status; (ii) error correction and variant detection within repetitive regions; and (iii) differentiation of signal due to MSI from non-MSI somatic variation and the strong PCR slippage artifacts at sites typically impacted by somatic instability. Indeed, technical PCR error is typically at least an order of magnitude higher than typical sequencing error rates in homopolymeric sites, necessitating iterative site selection and optimal use of molecular barcoding to achieve relevant signal-to-noise detection ratios across a large number of candidate microsatellite sites.

While tissue sequencing panels often comprise sufficient informative microsatellite loci simply due to large panel size and longer DNA fragment lengths (13, 32), the moderate size of the ctDNA panel utilized here and short ctDNA fragment lengths require purposeful microsatellite selection and inclusion. To accomplish this, we used an iterative approach informed by literature and tissue sequencing compendia to evaluate candidate sites to provide pan-cancer MSI detection with minimal background noise. The list of candidate loci was further refined on the basis of the performance criteria referenced above using healthy donor ctDNA.

On the basis of the performance assessment in training healthy donor samples, informative loci were defined as those that were effectively captured, sequenced, and mapped and were associated with little variation within MSS samples (shown in gray in Fig. 1A). Uninformative loci either failed capture, sequencing, or mapping, resulting in inadequate molecular representation (shown in blue in Fig. 1A), or demonstrated substantial variation within MSS samples, resulting in excessive artificial signal (shown in red in Fig. 1A). Interestingly, the BAT-25, BAT-26, and NR-21 Bethesda loci utilized in traditional MSI tissue tests (16, 17) and some ctDNA panels (34) performed poorly relative to other candidates and were excluded from the final marker set (indicated by arrows in Fig. 1A).

Using this approach, 90 microsatellite loci were selected for inclusion in the final test version: 89 mononucleotide repeats and a single trinucleotide repeat, all of which comprise repeats of length 7 or above. Assessment of unique molecule coverage distribution demonstrated that 65% of these loci have coverage above 0.5× median sample coverage.

In addition to effective molecular capture and mapping, MSI detection also requires highly accurate differentiation of cancer-related signal from background noise due to sequencing and polymerase errors at the very low allele fractions at which ctDNA is typically found (25, 26, 34). Importantly, the same repetitive genomic context that makes microsatellite candidates informative for MSI detection due to polymerase slippage during in vivo cellular replication also makes them particularly susceptible to the same polymerase slippage during in vitro library preparation and sequencing, resulting in high levels of technical noise. To address this, Digital Sequencing error correction was used to define true biological insertion–deletion events at microsatellite loci at high fidelity as described previously (25, 26). Among these high background error repeats, Digital Sequencing was associated with 100-fold reduction in per-molecule sequencing error relative to standard sequencing approaches (Fig. 1B), allowing efficient and accurate reconstruction of microsatellite sequences of individual unique molecules present in the original patient blood sample. Site-specific and aggregate sample-level MSI status determination thresholds were then established using permutation-based threshold simulations of healthy donor samples. When these per-site and per-sample thresholds were combined with the effects of digital sequencing correction, the per-sample false positive rate was estimated to be approximately 10⁻²⁻³. In addition, titration simulations adjusted for the distribution of clinical inputs predicted robust MSI detection to approximately 0.2% tumor fraction, with a marked decline...
in detection efficiency thereafter. As such, samples with a circulating tumor fraction (as defined by the maximum somatic variant allele fraction) of <0.2% were considered unevaluable for MSI status.

Analytical validation studies
To assess the analytical sensitivity of MSI detection, cfDNA derived from the supernatant of the MSI-H cell line KM12 was diluted into MSS cfDNA targeting five titration points comprising 15 independently processed replicates bracketing the LoD predicted by the in silico simulations described above. Each titration series was analyzed at both 5 ng, the minimum acceptable cfDNA input, and 30 ng, the maximum and most common cfDNA input. Using probit analysis, the 95% LoD (LOD95) was calculated to be 0.4% at 5-ng input (Fig. 2A) and 0.1% at 30-ng input (Fig. 2B).

To assess analytical intermediate precision, we analyzed replications of four different contrived materials, two MSS and two MSI-H (Fig. 2C). Across 499 replicates, categorical concordance for MSI status was 100% (499/499, 95% CI, 99%–100%), with coefficients of variation for quantitative MSI score ranging from 6.3%–7.2% for MSI-H samples (Supplementary Table 1). Repeatability and input robustness were also assessed by replicate testing of MSS and MSI-H–contrived material at 5-, 10-, and 30-ng cfDNA input, which similarly demonstrated 100% concordance (27/27, 95% CI, 85%–100%, Supplementary Fig. S1; Supplementary Table S2). Clinical precision was confirmed in 72 independent patient sample replicates representing a range of MSI scores and tumor fractions processed across three independent batches, days, operators, and reagent lots, which demonstrated a qualitative concordance of 100% (72/72, 95% CI, 94%–100%) with coefficients of variation for the underlying quantitative MSI score of 2.0%–15.2% (Supplementary Table S3).

To assess analytical specificity, healthy donor plasma samples (distinct from those used in training), MSS-contrived materials, and MSS patient samples were analyzed for spurious MSI-H calls. Analytical specificity was 100% across healthy donor samples (20/20, 95% CI, 83%–100%), contrived material (245/245, 95% CI, 98%–100%), and patient samples (48/48, 95% CI, 92%–100%).

Clinical validation studies
As no orthogonal cfDNA-based method was available to use as a comparator, clinical accuracy was determined by comparing cfDNA MSI assessment to MSI status from the medical record determined using standard-of-care tissue testing (a mixture of IHC, PCR, and NGS methods) for 1,145 samples comprising 40 distinct cancer types, 15 of which had at least five representative specimens (Supplementary Fig. S2). In 949 unique evaluable patients, cfDNA detected 87% of patients reported as MSI-H (71/82, 95% CI, 77%–93%) and 99.5% of patients reported as MSS/MSI-L (863/867, 95% CI, 98.7%–99.8%) for an overall accuracy of 98.4% (934/949, 95% CI, 97.3%–99.1%), with a positive predictive value (PPV) of 95% (71/75, 95% CI, 86%–98%; Fig. 3A; Supplementary Table S4). Consistent with in silico modeling studies, MSI-H detection was rare (0/19) in samples classified as unevaluable due to low tumor fraction (Fig. 3B), which explained 57% (16/28) of the observed cfDNA-tissue discordance in the total unique patient sample set (Supplementary Table S4). For samples with tumor fractions above 1%, cfDNA PPA rose to 93% (54/58, 95% CI, 82%–98%, Supplementary Table S4D).

Interestingly, despite the high correlation between IHC and PCR tissue tests reported in the literature (22, 35), concordance between cfDNA and tissue MSI status here varied by tissue test methodology [97.4% by PCR (450/462), 98.0% by NGS (239/244), and 83.0% by IHC (93/112), Fig. 3C; Supplementary Table S5]. On further investigation, it was noted that this discordance was due to both an increased tissue IHC-positive, cfDNA-negative population (2.4% by PCR, 2.0% by NGS, and 12.5% by IHC, Fisher exact test P < 0.001 for IHC-PCR and IHC-NGS) and an increased tissue IHC-negative, cfDNA-positive population.

Figure 2.
Analytic validation of cfDNA MSI detection. Observed MSI detection rate was plotted by titration level (green dots), and probit regression was used to determine the 95% limit of detection for 5-ng (A) and 30-ng (B) cfDNA inputs. C, Sample-level MSI scores for 499 independent replicates of two MSS and two MSI-H–contrived materials run across 499 separate sequencing runs. Dashed line indicates the sample-level threshold for MSI detection.
Concordance of ctDNA MSI status with tissue testing.

Figure 3.

A, Sample-level MSI scores for 1,145 cfDNA samples categorized by tissue test result and observed tumor fraction. Dashed line indicates the sample-level threshold for MSI scores. B, Concordance result categorized by tissue test methodology. C, Descriptive statistics for the evaluable unique patient cohort are presented with absolute count and 95% CIs in parentheses. (0.2% by PCR, 0% by NGS, and 4.5% by IHC, Fisher exact test $P < 0.01$ for both comparisons). Of the 25 samples for which IHC and another tissue test results were available, 12 demonstrated IHC-ctDNA discordance. Importantly, PCR and/or NGS tissue testing supported the ctDNA NGS results rather than the tissue IHC. Each of 12 discordances, indicating that the discordance observed between ctDNA NGS and tissue IHC in this cohort was, at least, partially due to IHC-derived factors rather than solely ctDNA-tissue discordance. Together, these data suggest that ctDNA assessment may provide a valuable corroborative method for MSI assessment.

cDNA MSI status in 28,459 patient samples with consecutive advanced cancer

Although a number of studies have assessed the prevalence of MSI across different tumor types in tissue (13, 28, 36), to date there is no published landscape analysis of ctDNA MSI status across cancer types. To this end, we applied the MSI algorithm described above to 28,459 consecutive advanced cancer patient clinical samples tested in the Guardant Health Clinical Laboratory. In this cohort, 278 samples (tumor fraction median of 6.55%, range 0.09%–89%) comprising 16 different tumor types were identified as MSI-H by ctDNA, which corresponds to an overall pan-cancer prevalence of approximately 1%, similar to that previously reported for tissue (13, 28, 36). Similarly, MSI-H prevalence among tumor types also closely reflected that observed in tissue-based analyses (Fig. 4A); as expected, MSI-H was most prevalent in endometrial, colorectal, and gastric cancers, whereas other tumors such as lung, bladder, and head and neck cancers demonstrated lower prevalence. Specific exceptions to previous MSI-H prevalence estimates included marginally lower prevalence in endometrial, colorectal, and gastric cancers, and marginally higher prevalence in prostate cancer.

Given the pan-solid tumor nature of the ctDNA intended use population and immunotherapy approval for MSI-H tumors, microsatellite loci for this panel were intentionally selected to be informative of MSI status across all solid tumor types. Consistent with this design intent, analysis of sample- and locus-level MSI score distributions (Fig. 4B and C) demonstrated consistent performance across tumor types, with MSI-H samples demonstrating signal substantially above threshold. Moreover, the diagnostic yield of MSI assessment outside of the tumor types for which MSI is commonly tested was substantial; more than half of the identified cases (143/278) occurred in tumor types in which MSI testing is very uncommon (Fig. 4A) and thus identified patients that would otherwise never have been tested.

Consistent with what has been reported in tissue (13), the number of indels and SNVs (inclusive of nonsynonymous and synonymous variants) is significantly increased in MSI-H samples relative to those characterized as having MSS status (Fig. 5). Specifically, the median number of SNVs in MSI-H samples was 6.3 versus 1.4 in MSS ($\chi^2 P < 0.0001$) and the median number of indels in MSI-H samples was 2.6 versus 0.4 in MSS ($\chi^2 P < 0.0001$).

cDNA MSI status predicts immunotherapy response

The most salient utility of MSI status today is its ability to select patients for immunotherapy. Despite this and the barriers to obtaining tissue in many patients, the ability of ctDNA MSI status to predict response to immunotherapy has not been reported. To establish clinical validity for this biomarker, we present clinical outcomes for 16 patients with ctDNA MSI-H metastatic gastric cancer treated with pembrolizumab ($n = 15$) or nivolumab ($n = 1$) after the failure of standard-of-care chemotherapy in a phase II pembrolizumab trial in gastric cancer (NCT#02589496). ctDNA and tissue PCR MSI assessment in pretreatment samples was 100% concordant for MSI-H (16/16, 95% CI, 76%–100%). Ten of 16 patients achieved either complete ($n = 3$) or partial ($n = 7$) investigator-assessed objective response by RECIST 1.1 criteria, with an additional three patients with stable disease (Fig. 6A), for
an objective response rate of 63% (10/16, 95% CI, 36%–84%) and a disease control rate of 81% (13/16, 95% CI, 54%–95%), similar to responses previously reported for MSI-H patients defined by tissue testing (37). Importantly, even in this pretreated population, these responses were durable, with a median duration of treatment of 39 weeks. Indeed, patient 21, for example, experienced complete regression of disease following pembrolizumab treatment after failure of fluoropyrimidine/platinum...
Tumor mutation burden by MSI status. Number of SNVs (A) and indels (B) detected per sample categorized by MSI status across 278 MSI-H and 28,181 MSS samples.

Discussion

We have validated a novel cfDNA-based targeted NGS approach for MSI detection. By using a large panel of microsatellite loci, this approach achieved high sensitivity relative to tissue-based methods, while maintaining very high specificity. Plasma-detected prevalence of MSI-H across 16 common solid tumors was similar to published tissue-based compendia, compatible with the intended pan-cancer design of the sequencing panel and MSI detection algorithm. Furthermore, we demonstrate clinical utility by showing that MSI-H patients as detected by cfDNA benefit from ICB therapy in a manner similar to that reported for tissue-defined populations (37), expanding the availability of MSI detection to all patients regardless of tissue availability or requirement to undergo invasive tissue acquisition procedures.

This study demonstrates robust analytic performance for MSI detection on a cfDNA panel previously validated for detection of the other four variant types in all guideline-recommended indications (25). In particular, the analytic sensitivity for MSI detection in contrived samples demonstrated reproducible detection to 0.1%, congruent with previous reports of similar sensitivity for indels and SNVs (25). Importantly, this study assessed the performance of cfDNA MSI testing in 1,145 samples with orthogonal tissue MSI, which constitutes the largest cfDNA-tissue MSI concordance cohort yet described. Relative to standard-of-care tissue MSI testing for the same patients, cfDNA MSI assessment demonstrated high PPV (95%), which compares favorably to the reported PPV of 90%–92% reported for local versus central tissue-based MSI assessment (38), and high PPA (87%) in the evaluable population, which is consistent with previous studies examining concordance of plasma and tissue genotyping for other variant types (24, 25, 39, 40). Factors that can contribute to incomplete concordance may include tumor heterogeneity, differential shedding by the primary versus metastatic lesions, temporal discordance of tissue and plasma collection, and low tumor shedding by some tumors (41–45). Interestingly, a patient with gastric cancer identified as MSI-H by plasma and by pentaplex PCR in this report was previously reported to comprise discrete tumor populations of MSS and MSI-H disease as assessed by both IHC and PCR performed on tissue (41). The same study found 9% discordance for MSI-H between paired tissue biopsies in the same patient (41), highlighting the potential contribution of intratumoral heterogeneity to discordances in MSI status. Moreover, the observation of nontrivial discordance between PCR and IHC tissue methods in this report highlight the importance of accurate MSI testing, which has been reported as a primary source of ICB failure (38). Consistent with the challenges presented by tissue genotyping in advanced solid tumors, a study in metastatic non–small cell lung cancer has shown that relative to tissue, plasma-based testing increases the number of patients with successful tumor genotyping results, as well as the frequency of detecting targetable mutations (24, 39).

This study presents the first cfDNA-based landscape analysis of MSI in a large advanced pan-cancer cohort. Overall, the relative prevalence across tumor types in a set of >28,000 consecutive clinical samples are consistent with what has been reported for tissue (13, 28, 36), with only marginal differences. For example, the prevalence in colorectal cancer and endometrial cancer is lower than what has been reported for tissue (13), which most likely reflects the fact that tissue-based landscape analyses include large numbers of early-stage MSI-H tumors, which have a better prognosis (15) and are less likely to be part of the advanced cancer population tested with cfDNA. In contrast, the larger than expected prevalence of MSI-H prostate cancer is attributable to increased representation of MSI-H disease in advanced patients; two recent studies focusing on MSI status in advanced prostate cancer have shown MSI-H prevalence of 3.1% and 3.8% in that patient population (46, 47), which is similar to the 2.6% observed in this study. Unsurprisingly, given the design intent for pan-cancer MSI detection, landscape analysis did not reveal tumor type–specific patterns of microsatellite instability. This design approach allows standardization of MSI detection across tumor types, including those without available training sets; however, it does not preclude the possibility that in plasma, similar to what has been shown in tissue (28, 36), tumor-type–specific patterns could emerge with the assessment of larger numbers of representative MSI-H samples or larger numbers of microsatellite loci.

The clinical outcomes reported here are limited to gastric cancer; nevertheless, the observed objective response rate is consistent with expectations from tissue-based studies, suggesting that ICB treatment based on cfDNA MSI results should achieve expected outcomes across solid tumor types. In addition, lack of germline dMMR data prevented conclusions about familial implications for cfDNA-detected MSI; however, while cfDNA-detected MSI may raise suspicion for hereditary cancer predisposition syndromes, the test’s 87% PPA relative to tissue-
based testing precludes its use as a screening test for these conditions. Finally, treatment data were not available for the majority of patients with cfDNA-/tissue+ discordance, but it is expected that at least some have received ICB therapy based on the tumor result, which would suppress MSI-H disease and lead to lack of MSI detection by cfDNA, thereby contributing to the observed discordance. Future studies should be pursued to address these questions.

In summary, we have developed and validated a cfDNA-based targeted NGS panel that accurately assesses MSI status while also providing comprehensive tumor genotyping, allowing pan-solid tumor guideline-complete testing from a single peripheral blood draw with high sensitivity, specificity, and precision. Clinical validation using both comparison to tissue testing, population-level prevalence analyses, and the first-reported outcomes for cfDNA MSI-H patients treated with ICB therapy supported the clinical accuracy and relevance of this approach. Such simultaneous characterization of MSI status and tumor genotype from a simple peripheral blood draw has the potential to expand access to both targeted therapy and immunotherapies to all patients with advanced cancer including those for whom current tissue-based testing paradigms are inadequate.

**Disclosure of Potential Conflicts of Interest**

M.I. Lefterova holds ownership interest (including patents) in Guardant Health. D.V.T. Catenacci reports receiving speakers bureau honoraria from Guardant Health, Foundation Medicine, and Merck, and is a consultant/advisory board member for Merck and Bristol-Myers Squibb. M. Fakih reports receiving speakers bureau honoraria from Amgen, and is a consultant/advisory board member for Array, Seattle Genetics, and Amgen. D. Gavino holds ownership interest (including patents) in Guardant Health and Invitae. N. Peled reports receiving speakers bureau honoraria from and is a consultant/advisory board member for AstraZeneca, BI, Bristol-Myers Squibb, Foundation Medicine, Guardant, Lilly, MSD, Novartis, Pfizer, and Roche. R.N. Eskander reports receiving speakers bureau honoraria from AstraZeneca and Clovis Oncology, and is a consultant/advisory board member for Merck, Eisai, Pfizer, and Tesaro/GSK. G. Azzi holds ownership interest (including patents) in and reports receiving speakers bureau honoraria from Guardant Health. K.C. Banks, D.I. Chudova, and A. Talasaz hold ownership interest (including patents) in Guardant Health. R.B. Lanman is an employee of Guardant Health, Inc, holds ownership interest (including patents) in Guardant Health, Inc, Biolase, Inc, and Forward Medical, Inc, and is a consultant/advisory board member for Forward Medical, Inc and Biolase, Inc. No potential conflicts of interest were disclosed by the other authors.

**Authors’ Contributions**


---

**Figure 6.**

Clinical outcome to ICB therapy in ctDNA MSI-H patients. A, Swimmer plot of duration of ICB therapy in months. Baseline (B and C) and post-therapy (D and E). CT (B and D) and gastroendoscopy (C and E) for patient 21.
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): A. Artymenko, K. Modly, C. Barbacioru, S.R. Fairlough, D. Gavino, J.I. Odegaard

References

Study supervision: M.I. Lefterova, D.I. Chudova, A. Talasaz, J.I. Odegaard
Others (provided patient/control samples including tumor data and discussions while the assay was being developed by Guardant). The assay was not developed by the author and the experiments were not done by the author: P.M. Kasi

Acknowledgments
We would like to thank the Nationwide Cancer Genome Screening Project in Japan, SCRAM-Japan GI-SCREEN, for the valuable contributions of specimens and data for the clinical validation studies presented here.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received April 22, 2019; revised May 15, 2019; accepted July 10, 2019; published first August 5, 2019.

Clinical Cancer Research
Validation of Microsatellite Instability Detection Using a Comprehensive Plasma-Based Genotyping Panel

Jason Willis, Martina I. Lefterova, Alexander Artyomenko, et al.


**Updated version**
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-19-1324

**Supplementary Material**
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2019/10/01/1078-0432.CCR-19-1324.DC1

**Cited articles**
This article cites 34 articles, 14 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/25/23/7035.full#ref-list-1

**Citing articles**
This article has been cited by 4 HighWire-hosted articles. Access the articles at:
http://clincancerres.aacrjournals.org/content/25/23/7035.full#related-urls

**E-mail alerts**
Sign up to receive free email-alerts related to this article or journal.

**Reprints and Subscriptions**
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

**Permissions**
To request permission to re-use all or part of this article, use this link http://clincancerres.aacrjournals.org/content/25/23/7035. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.