ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

Annals of Oncology 29: 1445–1453, 2018
doi:10.1093/annonc/mdy119
Published online 9 April 2018

ORIGINAL ARTICLE

Targeted methylation sequencing of plasma cell-free DNA for cancer detection and classification

L. Liu[1*†], J. M. Toung[1†], A. F. Jassowicz[2†], R. Vijayaraghavan[1], H. Kang[1], R. Zhang[1], K. M. Kruglyak[1], H. J. Huang[2], T. Hinoue[3], H. Shen[3], N. S. Salathia[1], D. S. Hong[2], A. Naing[2], V. Subbiah[2], S. A. Piha-Paul[2], M. Bibikova[1], G. Granger[1], B. Barnes[1], R. Shen[1], K. Gutekunst[1], S. Fu[2], A. M. Tsimberidou[2], C. Lu[4], C. Eng[5], S. L. Moulder[6], E. S. Kopetz[5], R. N. Amaria[7], F. Meric-Bernstam[2], P. W. Laird[3], J.-B. Fan[1] & F. Janku[2*]

[1]Illumina, Inc., San Diego; [2]Department of Investigational Cancer Therapeutics (Phase I Clinical Trials Program), The University of Texas MD Anderson Cancer Center, Houston; [3]Van Andel Research Institute, Grand Rapids; Departments of [4]Thoracic/Head and Neck Medical Oncology; [5]Gastrointestinal Medical Oncology; [6]Breast Medical Oncology; [7]Melanoma Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, USA

Correspondence to: Dr Filip Janku, Department of Investigational Cancer Therapeutics (Phase I Clinical Trials Program), The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030, USA. Tel: +1-713-563-0803; E-mail: fjanku@mdanderson.org

Dr Li Liu, Illumina Inc., 5200 Illumina Way, San Diego, CA 92122, USA. E-mail: lliu3@illumina.com

†These authors contributed equally to the work.

Background: Targeted methylation sequencing of plasma cell-free DNA (cfDNA) has a potential to expand liquid biopsies to patients with tumors without detectable oncogenic alterations, which can be potentially useful in early diagnosis.

Patients and methods: We developed a comprehensive methylation sequencing assay targeting 9223 CpG sites consistently hypermethylated according to The Cancer Genome Atlas. Next, we carried out a clinical validation of our method using plasma cfDNA samples from 78 patients with advanced colorectal cancer, non-small-cell lung cancer (NSCLC), breast cancer or melanoma and compared results with patients' outcomes.

Results: Median methylation scores in plasma cfDNA samples from patients on therapy were lower than from patients off therapy (4.74 versus 85.29; $P = 0.001$). Of 68 plasma samples from patients off therapy, methylation scores detected the presence of cancer in 57 (83.8%), and methylation-based signatures accurately classified the underlying cancer type in 45 (78.9%) of these. Methylation scores were most accurate in detecting colorectal cancer (96.3%), followed by breast cancer (91.7%), melanoma (81.8%) and NSCLC (61.1%), and most accurate in classifying the underlying cancer type in colorectal cancer (88.5%), followed by NSCLC (81.8%), breast cancer (72.7%) and melanoma (55.6%). Low methylation scores versus high were associated with longer survival (10.4 versus 4.4 months, $P < 0.001$) and longer time-to-treatment failure (2.8 versus 1.6 months, $P = 0.016$).

Conclusions: Comprehensive targeted methylation sequencing of 9223 CpG sites in plasma cfDNA from patients with common advanced cancers detects the presence of cancer and underlying cancer type with high accuracy. Methylation scores in plasma cfDNA correspond with treatment outcomes.

Key words: cell-free DNA, methylation, liquid biopsy, next-generation sequencing, cancer, plasma

Introduction

Molecular testing of plasma-derived circulating cell-free DNA (cfDNA) in cancer, also known as 'liquid biopsy', is a promising tool for minimally invasive molecular diagnostics and disease monitoring [1–3]. Molecular testing of liquid biopsies typically probes hot spot mutations in common cancer genes of known clinical relevance detected in small subsets of patients, which limits potential utility, especially for early diagnosis [4–6]. Expanding molecular testing of liquid biopsies to include epigenetic signatures can increase its diagnostic yield and utility [7]. Cancer is a genetic and epigenetic disease, and alterations in DNA methylation profiles are one of the earliest, most robust, and most frequent signatures in cancer [8]. Unlike hot spot mutations, the genome-wide distribution of numerous, densely

clustered DNA methylation alterations can enable more robust cancer detection and higher sensitivity in cancer diagnostics [9]. Furthermore, cancers originating from different tissue types may share similar genotypes but display different methylation profiles; thus, cancer-type-specific methylation signatures can potentially be used to identify cancer tissue of origin [10, 11].

In order to enable analysis of genome-wide methylation markers, we developed a next-generation sequencing (NGS) targeted methylation sequencing assay to simultaneously measure the methylation status of 9223 CpG (5′-C-phosphate-G-3′) sites known to be hypermethylated in cancer. We carried out clinical validation of this assay in plasma cfDNA samples from patients with advanced cancers and compared methylation-based results with clinical outcomes.

Materials and methods

Patients

The study enrolled patients with progressing advanced cancers who were referred to MD Anderson Cancer Center's Department of Investigational Cancer Therapeutics for experimental therapies from July 2011 to August 2016. Patients had the option of providing longitudinally collected plasma samples during the course of their therapy. The study was conducted in accordance with MD Anderson's Institutional Review Board guidelines.

Plasma collection and cfDNA isolation

Whole blood was collected in ethylenediaminetetraacetic acid-containing tubes and centrifuged within 2 h to yield plasma, which was stored at −80°C. Circulating cfDNA was extracted from 1 to 4 ml of plasma (4 ml was used whenever available) using the QIAamp Circulating Nucleic Acid kit (catalog # 55114, Qiagen, Valencia, CA) as previously described [7]. Plasma from normal individuals were purchased from BioreclamationIVT (Baltimore, MD). Concentrations of cfDNA were quantified by the 2100 Bioanalyzer System with High Sensitivity DNA Kit using the 75–250 bp range (Agilent Technologies, Santa Clara, CA).

Pan-cancer methylation panel design and optimization

The Pan-cancer methylation panel targets were particularly selected for hypermethylated CpG sites in tumor versus normal tissue based on The Cancer Genome Atlas (TCGA) data. A total of 10 888 CpG sites in 34 major cancer types and subtypes were selected. Probe sequences for targeted CpG sites were selected from the Infinium HM450 array (Illumina, San Diego, CA). Probes were individually synthesized and 5′-biotinylated at Illumina. Analysis of 20 normal plasma samples resulted in removal of 1235 (11%) sites with mean methylation levels >2% (supplementary Figure S1, available at *Annals of Oncology* online) and 430 (3.9%) sites with coverage below the fifth percentile in at least 10 samples, resulting in a total of 9223 CpG sites on the panel. Quality control was imposed after sequencing by excluding any samples with zero coverage at more than 15% of panel sites.

Targeted bisulfite sequencing library preparation and sequencing

Extracted cfDNA was bisulfite treated and purified using EZ DNA Methylation-Lightning Kit (Zymo Research). The assay input was capped at 30 ng regardless of the extraction yield. Whole genome amplification of bisulfite-converted DNA was carried out using Accel-NGS® Methyl-Seq DNA Library Kit (Swift Biosciences) following recommended conditions. Target enrichment was carried out on the whole genome bisulfite libraries to specifically pull down DNA fragments that contain target CpG sites using 5′-biotinylated capture probes. Hybrid selection was carried out using Illumina TruSight® Rapid Capture Kit. Capture Target Buffer 3 (Illumina) instead of enrichment hybridization buffer was used in the hybridization step. Following hybridization, the captured DNA fragments were amplified with 14 PCR cycles. Target capture libraries were sequenced on an Illumina HiSeq®2500 Sequencer using 2 × 100 cycle runs, with four to five samples in rapid run mode. A 10% PhiX was spiked into bisulfite sequencing libraries to increase base diversity for better sequencing quality.

Alignment

FASTQ files were generated using Illumina's bcl2fastq (v2.17.1.14) and mapped to the bi-sulfite-converted hg19 reference genome using BWA (v0.7.12; default parameters). Reads containing more than one alignment location as well as PCR duplicates flagged by Picard (v1.129) were removed from further downstream analyses.

Calculation of a sample methylation score

To account for variation in methylation profiles among healthy individuals, a z-score ($z=(x-\mu)/\sigma$), where μ and σ are the mean and standard deviation of methylation levels from a baseline set ($n=20$) of normal plasma samples was calculated for each CpG site. This z-score was transformed into a one-sided P-value [$P=1-\Phi(z)$, where $\Phi(z)$ represents the cumulative distribution function of the standard normal distribution], which normalizes the range from 0 to 1, and increases the signal-to-noise ratio. Finally, these P-values were log-transformed and weighted by a factor of $-2c$, where c represents the coverage (post-duplication removal) at each site, and aggregated into a single methylation score via Fisher's method.

Cancer-type classification

Methylation signatures for each cancer type were derived for those cancers with at least 100 tumor samples available in the TCGA database (32 cancers in total). For each cancer type, the available tumor samples were separated into training (80% of samples) and testing (20% of samples) sets. The training sets were used to identify signature CpG sites for the respective classifier. For each cancer type, the top 3% of CpG sites with the highest mean methylation levels across the training samples were included in the classifier. Any sites with mean methylation levels below 6% were excluded. This resulted in 32 classifiers for 32 cancer types respectively, each comprised of a unique set of CpG sites, although some individual CpG sites are present in multiple classifiers. For each tumor sample, 32 classification scores were calculated as the mean methylation level across all CpG sites in each of the respective classifiers. A tumor sample was classified as the cancer type with the highest classification score.

Statistical analysis

Overall survival (OS) was defined as the time from the date of sample collection to the date of death or last follow-up. The Royal Marsden Hospital (RMH) score was calculated on the basis of lactate dehydrogenase levels (greater than the upper limit of normal versus normal), albumin levels (<3.5 g/ml versus ≥3.5 g/ml), and number of metastatic sites (>2 sites versus ≤2 sites) [12]. Time-to-treatment failure (TTF) was defined as the time from the date of systemic therapy initiation to the date of removal from the treatment. Tumor responses were evaluated according to Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 [13]. The Kaplan–Meier method was used to estimate OS and TTF, and a log-rank test was used to compare OS and TTF among patient subgroups. Cox proportional hazards regression models were fit to assess the association between patient characteristics and OS. Correlations were assessed using the Spearman coefficient. All tests were two-sided, and

P values <0.05 were considered statistically significant. All statistical analyses were carried out with SPSS 23 (SPSS, Chicago, IL).

Results

Development of pan-cancer methylation profiling assay

We developed a comprehensive methylation panel targeting more than 9223 CpG sites by analyzing 32 cancer types and identifying sites that are consistently hypermethylated in tumors. Methylation microarray data from ~10 000 cancer samples from TCGA were collected for analysis. A set of cancer-type-specific hypermethylation sites was selected for each targeted cancer type, although we observed that some cancer types, such as stomach and colorectal, had higher overall methylation signatures compared with others (Figure 1A). Candidate sites were filtered by requiring low (<20%) methylation levels in normal blood cells. Annotation of the targeted CpG sites showed that 77% are located in CpG islands and that they are evenly distributed across gene regions, except for the 3′UTR, which is underrepresented (Figure 1B and C). We also observed that the targeted CpG sites are associated with known regulatory regions such as promoters (27.5%), enhancers (18.2%), and cell-type-specific regions (16.4%) (UCSC hg19; Figure 1D).

PCR-based methylation assays are limited by the number of targets that can be evaluated at one time, whereas a sequencing-based methylation assay allows for the simultaneous evaluation of the methylation status of thousands of CpG sites; however, sequencing-based assays in which enrichment is carried out before PCR amplification often do not retain sufficient diversity for a robust signal [14–16]. Our targeted methylation workflow measures methylation status of low input cfDNA while maintaining sufficient diversity. Extracted cfDNA was treated with bisulfite reagents followed by whole genome library preparation to amplify the bisulfite-converted DNA and incorporate sequencing adaptors. This amplification preserves the molecular diversity of the small amount of cfDNA in the subsequent target enrichment step. Enrichment probes were designed based on the Infinium HumanMethylation450 BeadChip in liquid hybridization—50mer probes that can capture both the methylated and unmethylated forms of bisulfite converted DNA after amplification (see Materials and methods) [17].

Next, we developed a novel algorithm to integrate the pan-cancer methylation sequencing data into a single methylation score to identify cancer samples (see Materials and methods). Briefly, z-scores were calculated for individual CpG sites; these sites were normalized by transforming to P-values; and a weighted sum of these P-values was calculated for the final sample-specific methylation score. We confirmed that this score is not biased to age or gender in healthy controls (supplementary Figure S2, available at *Annals of Oncology* online).

Performance of pan-cancer methylation profiling in cancer and healthy controls

The technical performance of this assay on plasma cfDNA was tested on samples from 66 healthy controls with males and females equally represented. The subjects ranged from 18 to 78 years in age and were evenly distributed between males and females. The per-sample cfDNA input ranged from 4.5 to 30 ng, with an average of 21.1 ng. We carried out deep sequencing of the targeted bisulfite libraries to achieve ~10 000× mean target coverage (~101× mean deduped coverage). On average, 95.3% of reads were properly aligned and 81.1% were on-target. The assay showed high uniformity of 81% (defined as the percentage of sites covered at ≥0.2× mean depth), with only 0.06% of target sites having zero coverage.

To determine the assay performance in distinguishing cancer from normal, we evaluated 37 plasma samples from 30 advanced colorectal cancer patients (several patients had plasma samples collected at multiple time points) and 66 plasma samples from healthy individuals. To calculate the methylation scores, we first carried out an outlier analysis of the 66 samples from healthy individuals (see Materials and methods). One sample was detected as an outlier and removed from further analysis. The remaining 65 normal samples were randomly divided to three groups. Group I ($n=20$) was used to establish a methylation baseline for z-score calculations. Group II ($n=20$) was used as a training set to derive the cutoff (calculated as 3 standard deviations above the mean methylation score) to separate cancer from normal. Normal samples in group III ($n=25$) and colorectal cancer samples that passed QC ($n=36$; see Materials and methods) were used to evaluate sensitivity and specificity. Methylation scores were calculated for each sample in the training and testing sets (Figure 1E). Using the cutoff derived from the training set (7.52), we correctly called 34 out of the 36 cancer samples and all 25 normal samples in the testing group (supplementary Figure S3A, available at *Annals of Oncology* online), resulting in a sensitivity of 94.4% for colorectal cancer detection and a specificity of 100%. Based on ROC analysis, the corresponding area under the curve (AUC) was 0.969 (Figure 1F).

Based on our previous observation of the high duplication rate (>10 000× raw coverage versus 101× deduped coverage), we re- carried out the same cancer/normal classification after down-sampling 30-fold to 10 million reads per sample and saw comparable performance (supplementary Figure S3B, available at *Annals of Oncology* online).

Next, we assessed whether a pan-cancer methylation panel can predict cancer tissue of origin. We identified methylation signatures for 32 cancer types, and unique sets of CpG sites for each cancer type were selected as the basis for the classification algorithm (see Materials and methods). We examined the performance of the algorithm on the TCGA tumor tissue testing sets and obtained an overall accuracy of 83.5% across all 32 cancer types (Figure 2A). We then compared the methylation profiles from colorectal cancer plasma cfDNA samples and cancer tissues in the TCGA database and found a relatively high correlation for colon and rectum adenocarcinomas ($r=0.435$ and 0.474, respectively), but not for other cancer types (mean $r=0.138$; Figure 2B). No correlation was observed between the methylation profiles of normal samples and any cancer types from TCGA (max $r=0.157$; Figure 2B). Combined, these results suggest that the methylation profiles of cancer plasma DNA reflect their tissue of origin despite the presence of normal cfDNA.

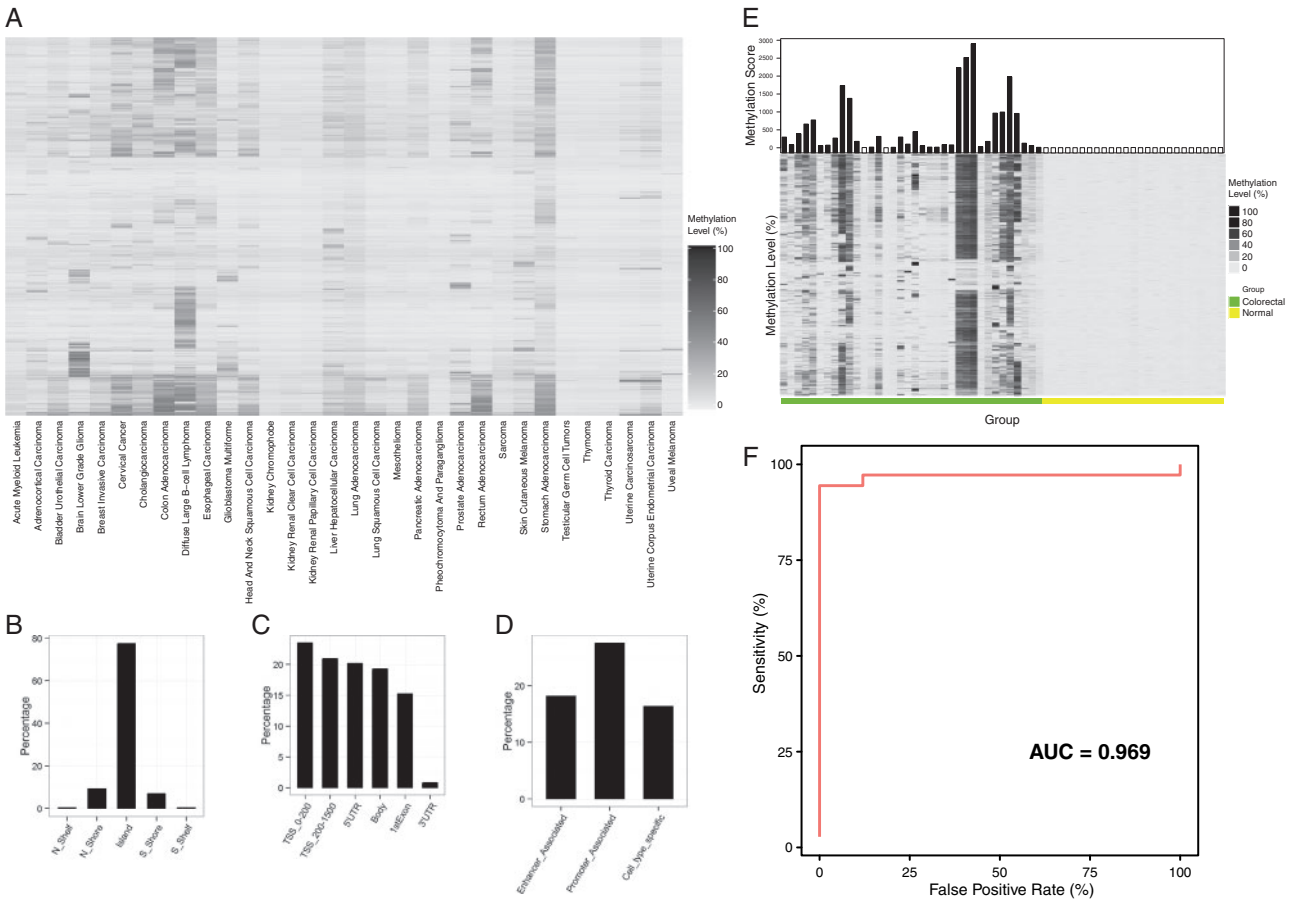

Figure 1. Pan-cancer methylation panel target sites: (A) heat map of methylation levels of 9223 CpG sites in 32 cancer types and subtypes based on the TCGA database. Distribution of targeted CpG sites relative to CpG islands (B), gene feature categories (C), and known regulatory feature groups (D) based on the UCSC Genome Browser Database. Targeted methylation profiling for cancer/normal classification: (E) heat map displaying measured methylation levels at each targeted CpG site of the methylation panel for all samples tested. The bar chart above the heat map indicates calculated sample methylation scores: black bars indicate classification of cancer and white bars indicate classification of normal. (F) ROC curve of cancer/normal classification for colorectal cancer (AUC = 0.969).

Clinical validation of pan-cancer methylation profiling in plasma cfDNA from patients with advanced cancers

In total, 78 patients with advanced colorectal cancer, non-small-cell lung cancer (NSCLC), breast cancer, or melanoma were enrolled. From those patients, we collected 82 plasma cfDNA samples comprising 68 samples from patients not receiving systemic therapy at the time of blood draw and 14 samples from patients on systemic therapy at the time of blood draw. Of the 68 samples from patients off therapy, 55 (80.9%) patients received subsequent systemic therapy and 13 (19.1%) patients did not. The median age was 61 years (range: 28–82 years). The majority of patients were female (40, 51.3%), the predominant ethnicity was white (62, 79.5%), and colorectal cancer was the most frequent cancer type (29, 37.2%), followed by NSCLC (22, 28.2%), breast cancer (15, 19.2%) and melanoma (12, 15.4%). Details are summarized in supplementary Table S1, available at *Annals of Oncology* online.

In samples from 68 patients off therapy, methylation scores correctly detected cancer in 57 (83.8%) of them. In the disease-specific analysis, methylation scores accurately detected cancer in 26 of 27 (96.3%) colorectal cancer patients, 11 of 12 (91.7%) breast cancer patients, 9 of 11 (81.8%) melanoma patients and 11 of 18 (61.1%) of NSCLC patients. In samples from those patients for whom cancer had been detected, methylation-based signatures correctly classified the underlying cancer type in 45 (76.3%) of them. Methylation-based signatures accurately classified the underlying cancer type in 23 of 26 (88.5%) colorectal cancer patients, 8 of 11 (72.7%) breast cancer patients, 5 of 9 (55.6%) melanoma patients, and 9 of 11 (81.8%) of NSCLC patients (Table 1).

In samples from the 14 patients on therapy, methylation scores accurately detected cancer in 7 (50.0%) of them. In the disease-specific analysis, methylation scores accurately detected cancer in five of the six (83.3%) colorectal cancer patients, zero of the three (0%) breast cancer patients, one of the one (100%) melanoma patients and one of the four (25.0%) of NSCLC patients. In samples from those patients for whom cancer had been detected, methylation-based signatures correctly classified the underlying cancer type in six (85.7%) of them. Methylation-based signatures accurately classified the underlying cancer type in five of the five (100%) colorectal cancer patients, one of the one (100%) melanoma patients, and zero of the one (0%) NSCLC patients (Table 1).

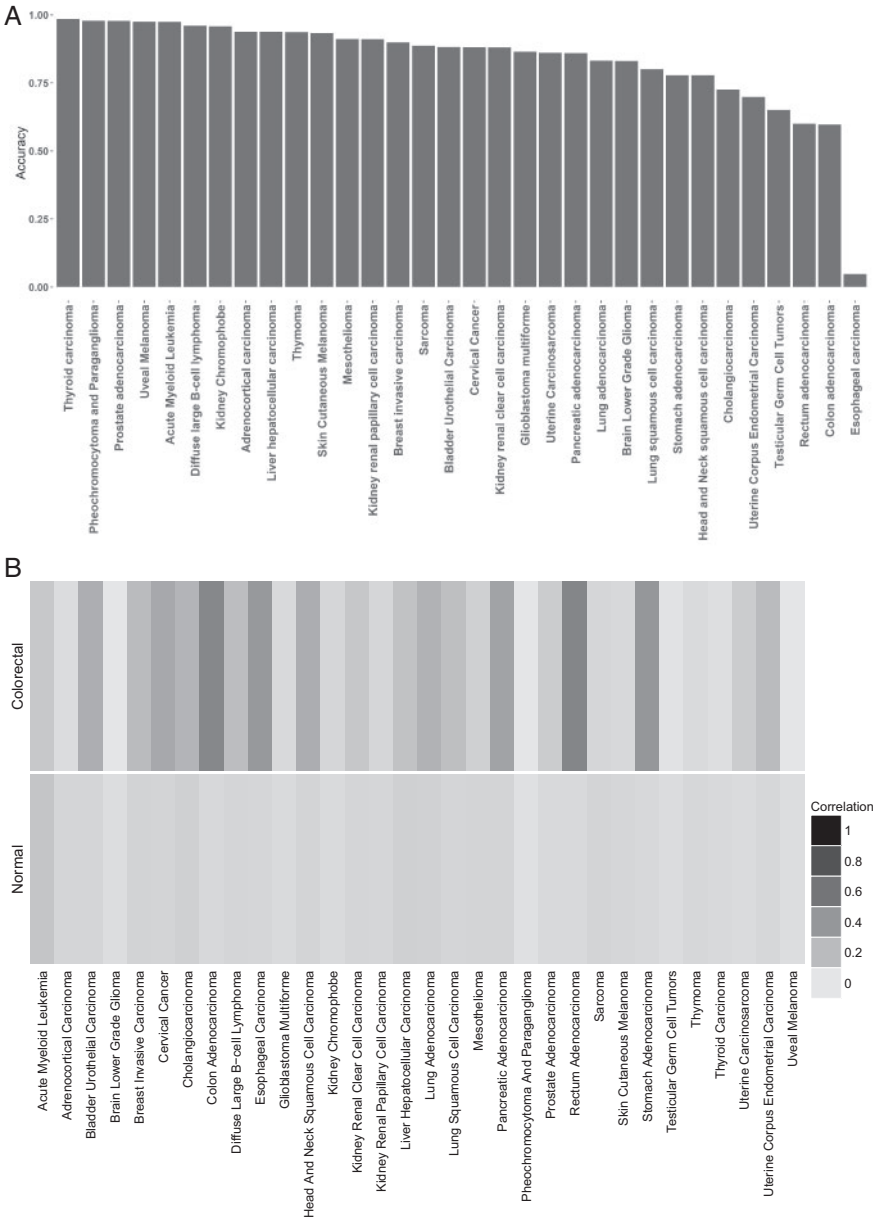

Figure 2. Cancer type classification accuracy on TCGA tumor tissue samples and correlation in methylation profiles between cancer plasma and cancer tissue samples. (A) The cancer type classification algorithm was applied on the TCGA tumor tissue testing sets. Classification accuracy was shown for 32 cancer types. (B) The correlation between methylation profiles from an individual plasma sample and each TCGA cancer type was calculated. The average correlation of each plasma cancer type (horizontal) and TCGA tissue cancer type (vertical) are displayed as a heat map. Black to white indicates correlation from strong to weak.

Finally, we compared methylation scores in plasma cfDNA from all 82 samples and demonstrated that a median methylation score in 14 samples obtained on therapy was lower than a median methylation score in 68 samples obtained from patients off therapy (mean 4.74 versus 85.29, $P = 0.001$; Figure 3A).

Pan-cancer methylation profiling of plasma cfDNA in cancer and clinical outcomes

Next, we assessed if methylation scores from patients off therapy at the time of blood draw corresponded with the length of survival. Forty-four patients with low methylation scores ($\leq$ 5% trimmed mean of 267.54) had a longer median overall survival (OS) compared with 24 patients with high methylation scores [>5% trimmed mean of 267.54; 10.4 months, 95% confidence interval (CI) 7.1–13.7 versus 4.4 months, 95% CI 3.3–5.4; $P < 0.001$; Figure 3B]. The 5% trimmed mean was selected over the median as a cutoff, in order to eliminate potential bias caused by the presence of samples with extreme methylation scores. In order to evaluate the methylation score as a possible independent prognostic factor for OS, we used the prospectively validated RMH prognostic score as described in Materials and methods. Indeed, 31 patients with RMH scores of 0–1 had longer median OS than patients with RMH scores of 2–3 (10.8 months, 95% CI 5.5–16.1 versus

Table 1. Classification accuracy using pan-methylation assay in clinical plasma cfDNA samples from patients with breast cancer, colorectal cancer, NSCLC and melanoma off or on systemic therapy at the time of collection

Plasma samples collected from patients off-therapy

		Actual class			
		Breast ($n = 12$)	**Colorectal cancer ($n = 27$)**	**NSCLC ($n = 18$)**	**Melanoma ($n = 11$)**
Predicted Class[a]	AML	2	0	0	0
	Breast	8	0	0	0
	Colorectal	1	23	0	0
	Cholangiocarcinoma	0	0	0	1
	Esophageal	0	0	0	1
	Liver	0	1	1	1
	Lung	0	0	9	0
	Lymphoma	0	1	0	0
	Melanoma	0	0	0	5
	Pancreatic	0	1	0	0
	Sarcoma	0	0	1	1
	Stomach	0	0	0	0
	Not cancer	1	1	7	2
Total		12	27	18	11
Accurate classification of cancer irrespective of cancer type		91.7% 83.8%	96.3%	61.1%	81.8%
Accurate classification of cancer type (out of set classified as cancer)		72.7% 78.9%	88.5%	81.8%	55.6%

Plasma samples collected from patients on-therapy

		Actual class			
		Breast ($n=3$)	**Colorectal cancer ($n=6$)**	**NSCLC ($n=4$)**	**Melanoma ($n=1$)**
Predicted class[a]	Breast	0	0	0	0
	Colorectal	0	5	0	0
	Lung	0	0	0	0
	Melanoma	0	0	0	1
	Stomach	0	0	1	0
	Not cancer	3	1	3	0
Accurate classification of cancer irrespective of cancer type		0% 50%	83.3%	25%	100%
Accurate classification of cancer type (out of set classified as cancer)		N/A 85.7%	100%	0%	100%

[a]Each sample was evaluated against 32 cancer type signatures. Cancer types to which no samples were assigned are not listed in this table.

4.7 months, 95% CI 3.4–5.9, $P = 0.009$, supplementary Figure S4, available at *Annals of Oncology* online). A multivariable analysis, which included methylation score (low versus high), RMH score (0–1 versus 2–3) and tumor type (colorectal cancer versus breast cancer versus NSCLC versus melanoma), demonstrated that a low methylation score is an independent prognostic factor for OS [hazard ratio (HR) 0.43, 95% CI 0.21–0.90, $P = 0.025$, Table 2].

Additionally, we assessed whether methylation scores are associated with treatment outcomes in 55 patients who went on to receive systemic therapy. Thirty-six patients with low methylation scores had a longer median TTF compared with 19 patients with high methylation scores (2.8 months, 95% CI 1.6–4, 0 versus 1.6 months, 95% CI 1.3–1.9, $P = 0.016$, Figure 3C). A multivariable analysis, which included methylation score (low versus high), RMH score (0–1 versus 2–3) and tumor type (colorectal cancer versus breast cancer versus NSCLC versus melanoma), demonstrated that a low methylation score is an independent prognostic factor for TTF (HR 0.38, 95% CI 0.16–0.88, $P = 0.025$, Table 2).

Finally, we assessed whether methylation scores are correlated with the best change in the sum of target lesions measured by RECIST 1.1 [13]. In 45 of the 55 patients with measurable disease

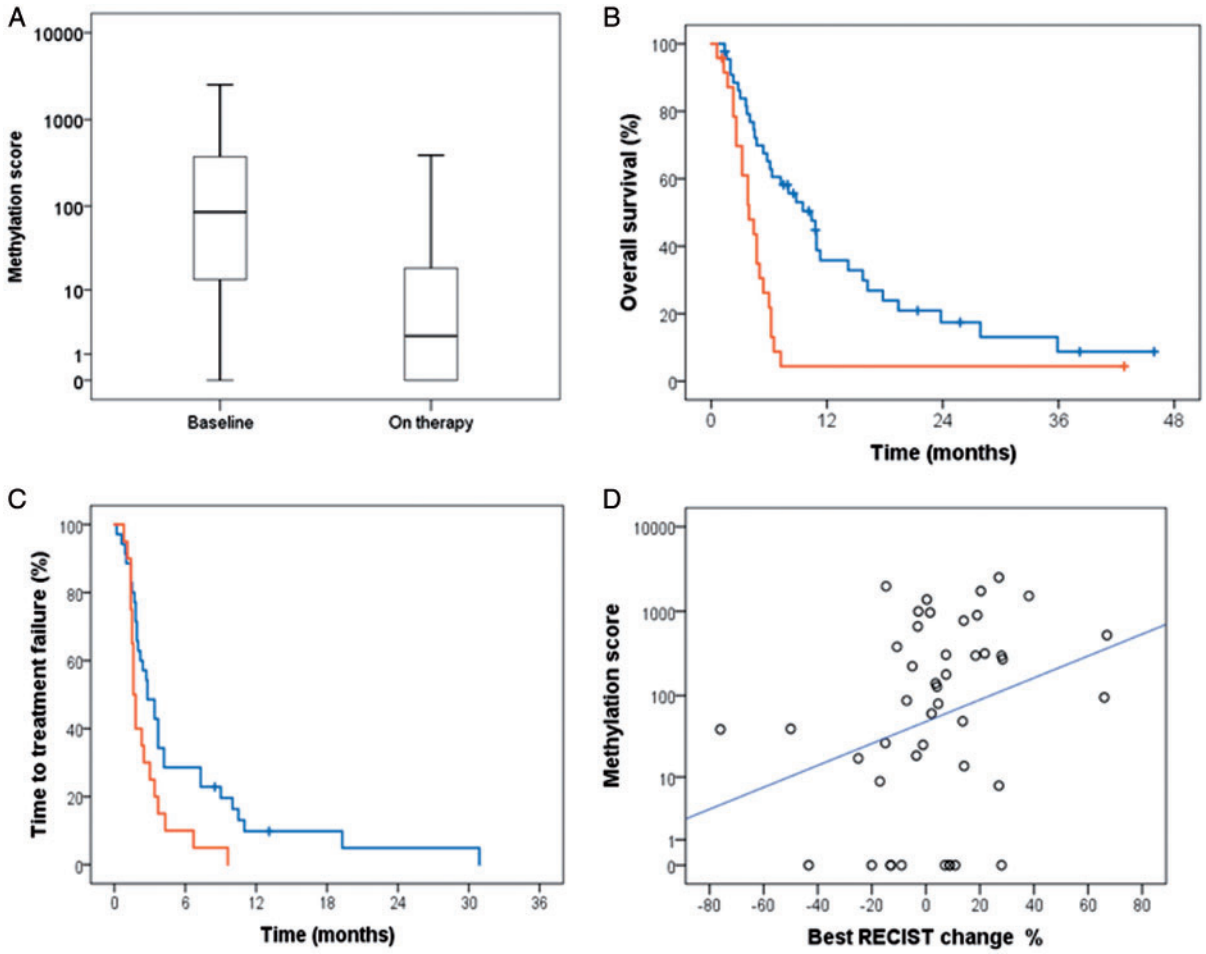

Figure 3. Clinical validation of pan-cancer methylation profiling in plasma cfDNA from patients with advanced cancers. (A) Boxplot of 82 patients including 68 patients off therapy during sample collection (baseline) and 14 patients on therapy during sample collection (on therapy) shows that patients before systemic therapy had higher methylations scores (median 85.29) compared with methylation scores for patients on therapy (median 4.7). (B) In 68 patients, whose cfDNA samples were tested and methylation scores were calculated, the median OS for patients with high methylation scores (> 267.54, red) was significantly shorter with 4.4 months (95% CI 3.3–5.4) compared with patients with low methylation scores (≤267.54, blue) who had a median OS of 10.4 months (95% CI 7.1–13.7; $P < 0.001$). (C) For 55 patients who received systemic therapy after cfDNA samples were drawn and methylation scores were calculated, patients with high methylation scores (> 267.54, red) had a shorter TTF of 1.6 months (95% CI 1.3–1.9), while in patients with low methylation scores (≤267.54, blue) TTF was significantly longer with 2.8 months (95% CI 1.6–4.0; $P = 0.016$). (D) Linear regression (blue line) between methylation score and best RECIST change of 45 patients who received systemic therapy after sample collection shows a positive correlation ($r = 0.32$, $P = 0.034$).

receiving systemic cancer therapy, methylation scores positively correlated with the percentage change in the sum of target lesions, suggesting that patients with low methylation scores are more likely to have tumor shrinkage ($r = 0.32$, $P = 0.034$, Figure 3D).

Discussion

Methylation profiling of cfDNA has been investigated in cancer diagnostics and assessment of therapeutic outcomes [18–24]. Guo et al. have shown that identifying methylation patterns from cfDNA can be used to map tissue-of-origin in lung and colorectal cancer patients [25]. We have developed a highly sensitive assay targeting more than 9000 CpG sites to determine methylation scores in plasma cfDNA from patients with advanced breast cancer, colorectal cancer, NSCLC and melanoma. Our results demonstrate that methylation scores from plasma cfDNA samples from patients with these cancers can accurately classify the presence of cancer in 83.8% of cases with 100% specificity. In addition, methylation scores from plasma cfDNA accurately predicted cancer type in 78.9% of cases (breast cancer, 72.7%; colorectal cancer, 88.5%; NSCLC, 81.8%; melanoma, 55.6%). Methylation scores were highly predictive of tumor origin in colorectal cancer, but much less in melanoma. It is plausible that our assay includes CpG sites that can perform better in cancers with strong hypermethylation signatures, such as colorectal cancer [26–28]. Another potential explanation may be differences in tumor DNA shedding into the circulation; however, small numbers of patients precludes a final conclusion. To our knowledge, our study is the first report on comprehensive targeted methylation profiling of plasma cfDNA in advanced cancer using NGS to simultaneously classify the presence of cancer and predict the underlying cancer type. Our assay covers

Table 2. Multivariable cox regression analysis of methylation score and outcomes

Multi-variable cox regression models evaluating Methylation scores and Royal Marsden Hospital (RMH) score with respect to overall survival (OS) and time-to-treatment failure (TTF)

Variable	HR	95% CI	*P* value
OS (68 patients with plasma collection at progression)			
Methylation scores (≤267.54 versus >267.54)	0.43	0.21–0. 90	0.025
RMH scores (0 or 1 versus 2 or 3)	0.55	0.29–1.07	0.079
TTF (55 patients with plasma collection at progression and subesquent systemic cancer therapy)			
Methylation scores (≤267.54 versus >267.54)	0.38	0.16–0.88	0.025
RMH scores (0 or 1 versus 2 or 3)	0.73	0.37–1.45	0.371

more than 9000 CpG sites while using a simplified workflow and commercially available reagents. Whole-genome bisulfite sequencing can provide a more comprehensive methylation profile; however, the cost may preclude advances to the clinic in the short-term [22]. On the other hand, PCR-based approaches are less expensive, though, their utility may be limited by the small number of CpG sites assessed [25].

Furthermore, previous studies with plasma-based liquid biopsies have demonstrated that the amount of mutant cfDNA for common oncogenic hot-spot mutations can be associated with progression-free survival (PFS), TTF and OS in patients with advanced cancer [1, 2, 4, 5, 7, 29–31]. Visvanathan et al. [32] demonstrated use of a 10-gene methylation panel to test serum samples from 141 women with metastatic breast cancer and demonstrated that high methylation serum scores were associated with shorter PFS and OS. In addition, blood DNA methylation changes induced by first line chemotherapy were associated with OS [33]. We demonstrated that patients with advanced breast cancer, colorectal cancer, NSCLC, or melanoma and low methylation scores (≤5% trimmed mean) in plasma cfDNA compared with those with high methylation scores (>5% trimmed mean) had a longer median OS, which was confirmed on multicovariate analysis (HR 0.43, $P = 0.025$). Additionally, our data demonstrated that patients with low methylation scores in plasma cfDNA collected before systemic therapy compared with those with high methylation scores had a longer median TTF (HR 0.38, $P = 0.025$). We also demonstrated that in patients with measurable disease, low methylation scores correlated with tumor shrinkage ($r = 0.32$, $P = 0.034$). Furthermore, we also demonstrated that samples of plasma cfDNA collected on therapy had lower median methylation scores than samples collected before therapy ($P = 0.001$).

Our study also has several limitations. First, the study was retrospective and the number of patients enrolled was limited. Second, we enrolled a patient population with four common tumor types, which were treated with diverse experimental therapies. Third, we only included patients with advanced cancers and it is plausible that sensitivity of our method could be less if directed at early stage cancers. Fourth, our assay included 9223 CpG sites and it is plausible that including more tumor-specific CpG sites might improve sensitivity, especially in some tumor types. Fifth, even though in our study methylation scores were associated with treatment outcomes such as OS, TTF and change in tumor size, the limited number of patients and tumor types assessed preclude us from concluding whether these findings are universally applicable or tumor specific.

In conclusion, comprehensive targeted methylation sequencing of over 9000 CpG sites in plasma cfDNA collected before therapy from patients with common advanced cancers predicts the presence of cancer and underlying tumor type with high accuracy, and methylation scores correspond with treatment outcomes. Further studies of methylation profiling of plasma cfDNA in cancer are warranted including titration studies to further establish the analytical sensitivity of the assay as a function of the overall number of methylated sites and methylation levels and prospective clinical trials in which therapeutic intervention is tailored on the basis of methylation profile.

Funding

This study was supported by the Sidney Kimmel Foundation for Cancer Research, the Sheikh Khalifa Al Nahyan Ben Zayed Institute for Personalized Cancer Therapy, the National Center for Advancing Translational Sciences (grant no. UL1 TR000371), and the National Institutes of Health through MD Anderson's Cancer Center Support Grant (P30 CA016672).

Disclosure

FJ has research support from Novartis, Genentech, FujiFilm Pharmaceuticals, Plexxikon, Deciphera, Symphogen, Piqur, BioMed Valley Discoveries, and Upsher Smith Laboratories; is on the Scientific Advisory Boards of Deciphera, Guardant Health, IFM Therapeutics; is a paid consultant for Immunomet and Trovagene; and has ownership interests in Trovagene. LL, RV, HK, KK, MB, GG, BB and KG are current employees and shareholders of Illumina, Inc. RZ, NS, RS, and J-BF are shareholders of Illumina, Inc. RA has research support from Merck, Bristol-Myers Squibb and Array Biopharma. All remaining authors have declared no conflicts of interest.

References

1. Polivka J Jr, Pesta M, Janku F. Testing for oncogenic molecular aberrations in cell-free DNA-based liquid biopsies in the clinic: are we there yet? Expert Rev Mol Diagn 2015; 15(12): 1631–1644.
2. Bettegowda C, Sausen M, Leary RJ et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 2014; 6: 224ra224.
3. Cai X, Janku F, Zhan Q, Fan JB. Accessing genetic information with liquid biopsies. Trends Genet 2015; 31(10): 564–575.
4. Janku F, Huang HJ, Claes B et al. BRAF mutation testing in cell-free DNA from the plasma of patients with advanced cancers using a rapid, automated molecular diagnostics system. Mol Cancer Ther 2016; 15(6): 1397–1404.

5. Janku F, Huang HJ, Fujii T et al. Multiplex KRASG12/G13 mutation testing of unamplified cell-free DNA from the plasma of patients with advanced cancers using droplet digital polymerase chain reaction. Ann Oncol 2017; 28: 642–650.
6. Diehl F, Schmidt K, Choti MA et al. Circulating mutant DNA to assess tumor dynamics. Nat Med 2008; 14(9): 985–990.
7. Janku F, Zhang S, Waters J et al. Development and validation of an ultradeep next-generation sequencing assay for testing of plasma cell-free DNA from patients with advanced cancer. Clinical Cancer Res 2017; 23(18): 5648–5656.
8. Esteller M. Epigenetics in cancer. N Engl J Med 2008; 358(11): 1148–1159.
9. Hao X, Luo H, Krawczyk M et al. DNA methylation markers for diagnosis and prognosis of common cancers. Proc Natl Acad Sci USA 2017; 114(28): 7414–7419.
10. Fernandez AF, Assenov Y, Martin-Subero JI et al. A DNA methylation fingerprint of 1628 human samples. Genome Res 2012; 22(2): 407–419.
11. Moran S, Martinez-Cardus A, Sayols S et al. Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. Lancet Oncol 2016; 17(10): 1386–1395.
12. Arkenau HT, Barriuso J, Olmos D et al. Prospective validation of a prognostic score to improve patient selection for oncology phase I trials. JCO 2009; 27(16): 2692–2696.
13. Eisenhauer EA, Therasse P, Bogaerts J et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009; 45(2): 228–247.
14. Eads CA, Danenberg KD, Kawakami K et al. MethyLight: a high-throughput assay to measure DNA methylation. Nucleic Acids Res 2000; 28(8): E32.
15. Kristensen LS, Mikeska T, Krypuy M, Dobrovic A. Sensitive Melting Analysis after Real Time-Methylation Specific PCR (SMART-MSP): high-throughput and probe-free quantitative DNA methylation detection. Nucleic Acids Res 2008; 36(7): e42.
16. Herman JG, Graff JR, Myohanen S et al. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci USA 1996; 93(18): 9821–9826.
17. Bibikova M, Barnes B, Tsan C et al. High density DNA methylation array with single CpG site resolution. Genomics 2011; 98(4): 288–295.
18. Chan KC, Jiang P, Chan CW et al. Noninvasive detection of cancer-associated genome-wide hypomethylation and copy number aberrations by plasma DNA bisulfite sequencing. Proc Natl Acad Sci USA 2013; 110(47): 18761–18768.
19. Fackler MJ, Lopez Bujanda Z, Umbricht C et al. Novel methylated biomarkers and a robust assay to detect circulating tumor DNA in metastatic breast cancer. Cancer Res 2014; 74(8): 2160–2170.
20. Mishima C, Kagara N, Matsui S et al. Promoter methylation of TRIM9 as a marker for detection of circulating tumor DNA in breast cancer patients. Springerplus 2015; 4(1): 635.
21. Sailer V, Holmes EE, Gevensleben H et al. PITX2 and PANCR DNA methylation predicts overall survival in patients with head and neck squamous cell carcinoma. Oncotarget 2016; 7(46): 75827–75838.
22. Qu X, Othus M, Davison J et al. Prognostic methylation markers for overall survival in cytogenetically normal patients with acute myeloid leukemia treated on SWOG trials. Cancer 2017; 123(13): 2472–2481.
23. van Nimwegen KJ, van Soest RA, Veltman JA et al. Is the $1000 genome as near as we think? A cost analysis of next-generation sequencing. Clin Chem 2016; 62(11): 1458–1464.
24. Lee EJ, Luo J, Wilson JM, Shi H. Analyzing the cancer methylome through targeted bisulfite sequencing. Cancer Lett 2013; 340(2): 171–178.
25. Guo S, Diep D, Plongthongkum N et al. Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA. Nat Genet 2017; 49(4): 635–642.
26. Lange CP, Campan M, Hinoue T et al. Genome-scale discovery of DNA-methylation biomarkers for blood-based detection of colorectal cancer. PLoS One 2012; 7(11): e50266.
27. Hinoue T, Weisenberger DJ, Lange CP et al. Genome-scale analysis of aberrant DNA methylation in colorectal cancer. Genome Res 2012; 22(2): 271–282.
28. Legendre C, Gooden GC, Johnson K et al. Whole-genome bisulfite sequencing of cell-free DNA identifies signature associated with metastatic breast cancer. Clin Epigenet 2015; 7(1): 100.
29. Janku F, Angenendt P, Tsimberidou AM et al. Actionable mutations in plasma cell-free DNA in patients with advanced cancers referred for experimental targeted therapies. Oncotarget 2015; 6(14): 12809–12821.
30. Fujii T, Barzi A, Sartore-Bianchi A et al. Mutation-enrichment next-generation sequencing for quantitative detection of KRAS mutations in urine cell-fre DNA from patients with advanced cancers. Clin Cancer Res 2017; 23(14): 3657–3666.
31. Santiago-Walker A, Gagnon R, Mazumdar J et al. Correlation of BRAF mutation status in circulating-free DNA and tumor and association with clinical outcome across Four BRAFi and MEKi Clinical Trials. Clin Cancer Res 2016; 22(3): 567–574.
32. Visvanathan K, Fackler MS, Zhang Z et al. Monitoring of serum DNA methylation as an early independent marker of response and survival in metastatic breast cancer: tBCRC 005 Prospective Biomarker Study. JCO 2017; 35(7): 751–758.
33. Flanagan JM, Wilson A, Koo C et al. Platinum-based chemotherapy induces methylation changes in blood DNA associated with overall survival in patients with ovarian cancer. Clin Cancer Res 2017; 23(9): 2213–2222.