

Tumor fraction in cell-free DNA as a biomarker in prostate cancer

Atish D. Choudhury,^{1,2,3} Lillian Werner,^{1,2} Edoardo Francini,^{1,4} Xiao X. Wei,^{1,2} Gavin Ha,^{1,3} Samuel S. Freeman,³ Justin Rhoades,³ Sarah C. Reed,³ Gregory Gydush,³ Denisse Rotem,³ Christopher Lo,³ Mary-Ellen Taplin,^{1,2} Lauren C. Harshman,^{1,2} Zhenwei Zhang,¹ Edward P. O'Connor,¹ Daniel G. Stover,⁵ Heather A. Parsons,^{1,2,3} Gad Getz,^{3,6} Matthew Meyerson,^{1,2,3} J. Christopher Love,^{3,7} William C. Hahn,^{1,2,3} and Viktor A. Adalsteinsson^{3,7}

¹Dana-Farber Cancer Institute, Boston, Massachusetts, USA. ²Harvard Medical School, Boston, Massachusetts, USA. ³Eli and Edythe L. Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA. ⁴Sapienza University of Rome, Rome, Italy. ⁵The Ohio State University, Columbus, Ohio, USA. ⁶Massachusetts General Hospital, Boston, Massachusetts, USA. ⁷Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

BACKGROUND. Tumor content in circulating cell-free DNA (cfDNA) is a promising biomarker, but longitudinal dynamics of tumor-derived and non-tumor-derived cfDNA through multiple courses of therapy have not been well described.

METHODS. CfDNA from 663 plasma samples from 140 patients with castration-resistant prostate cancer (CRPC) was subject to sparse whole genome sequencing. Tumor fraction (TFx) estimated using the computational tool ichorCNA was correlated with clinical features and responses to therapy.

RESULTS. TFx associated with the number of bone metastases (median TFx = 0.014 with no bone metastases, 0.047 with 1–3 bone metastases, 0.190 for 4+ bone metastases; $P < 0.0001$) and with visceral metastases ($P < 0.0001$). In multivariable analysis, TFx remained associated with metastasis location ($P = 0.042$); TFx was positively correlated with alkaline phosphatase ($P = 0.0227$) and negatively correlated with hemoglobin (Hgb) ($P < 0.001$), but it was not correlated with prostate specific antigen (PSA) ($P = 0.75$). Tumor-derived and non-tumor-derived cfDNA track together and do not increase with generalized tissue damage from chemotherapy or radiation at the time scales examined. All new treatments that led to $\geq 30\%$ PSA decline at 6 weeks were associated with TFx decline when baseline TFx was $> 7\%$; however, TFx in patients being subsequently maintained on secondary hormonal therapy was quite dynamic.

CONCLUSION. TFx correlates with clinical features associated with overall survival in CRPC, and TFx decline is a promising biomarker for initial therapeutic response.

TRIAL REGISTRATION. Dana-Farber/Harvard Cancer Center (DF/HCC) protocol no. 18-135.

FUNDING. Wong Family Award in Translational Oncology, Dana Farber Cancer Institute Medical Oncology grant, Gerstner Family Foundation, Janssen Pharmaceuticals Inc., and Koch Institute Support (core) grant P30-CA14051 from the National Cancer Institute (NCI).

Conflict of interest: ADC has pending research funding with Bayer with salary support. AR has consulting role and equity with Celsius Pharmaceuticals. GG receives research funding from IBM and Pharmacyclics and has filed patent applications related to cancer genomics (US20150178445A1, US20150197785A1). MM receives research funding from Bayer; had equity interest in Foundation Medicine (former, partially sold to Roche); had consulting or advisory role with Foundation Medicine (former) and OrigiMed; and received patent royalties on intellectual property from Dana-Farber Cancer Institute licensed to Labcorp (#7,294,468; 7,964,349; 8,105,769; 8,465,916; 9,035,036). JCL received research funding from Janssen. VAA, GH, and SSF filed a patent (WO2017161175A1) on methods applied in this study.

License: Copyright 2018, American Society for Clinical Investigation.

Submitted: May 7, 2018

Accepted: October 2, 2018

Published: November 2, 2018

Reference information:

JCI Insight. 2018;3(21):e122109.

<https://doi.org/10.1172/jci.insight.122109>.

insight.122109.

Introduction

As treatment options for metastatic prostate cancer expand, choosing the optimal agent, timing, sequencing, and combinations to maximize clinical benefit has become increasingly challenging. Treating patients with ineffective therapy exposes them to unnecessary toxicity while allowing symptomatic progression of their cancer. Conversely, early discontinuation of an agent that is providing some degree of clinical benefit relative to other available modalities or no treatment can also be detrimental. The Prostate Cancer Clinical Trials Working Group 3 (PCWG3) guidelines (1) address challenges in efficacy assessment posed by standard bone scintigraphy, CT, and prostate specific antigen (PSA) assessments and place increased emphasis on outcomes that reflect patient benefit, even after initial progression on therapy or after discontinuation of

therapy. Among the recommendations from the PCWG3 is direct biologic characterization of the tumor at the time a new treatment is being considered, including blood-based diagnostics such as circulating nucleic acids, to improve understanding of disease biology and identify potential predictive molecular biomarkers.

Circulating cell-free DNA (cfDNA) derived from patient plasma is attractive as a noninvasive method for genetic characterization of tumor burden, given its high concordance with matched metastatic biopsies (2, 3). In addition, both the total quantity of cfDNA in the circulation and estimates of the tumor-derived contribution to cfDNA (tumor fraction; TFX) have been proposed as prognostic biomarkers (4) and indicators of response and resistance to therapy (5). One study suggested an association between the total cfDNA concentration in the circulation with clinical outcomes after taxane-based chemotherapy in metastatic castration-resistant prostate cancer (mCRPC) (6), but other studies did not demonstrate an independent correlation on multivariable analysis between baseline cfDNA concentration and outcomes for abiraterone/enzalutamide (7, 8), taxane chemotherapy (9), or olaparib (5). This may be because total cfDNA concentration measured in the blood is not a reliable biomarker for tumor burden and is subject to significant technical and biological variability, sources of which include reagents and conditions related to DNA isolation, as well as biological features that are poorly understood (e.g., levels of DNase and anti-dsDNA antibodies in the circulation, or upregulation or downregulation of mechanisms of cfDNA clearance) (10).

As such, we hypothesized that the fraction of cfDNA derived from tumor cells (i.e., TFX) would be a more reliable biomarker of disease dynamics. TFX is promising as a biomarker of disease burden and therapeutic response, since — unlike circulating protein tumor markers — this metric is agnostic to cancer type. Among the clinical applications of such a biomarker would be the disambiguation of radiographic flare or pseudoprogression responses after starting new agents (particularly hormonal or immune therapies), or the early measurement of treatment response to allow earlier intensification or discontinuation of ineffective therapy prior to radiographic progression.

We have previously described a method using a software tool called ichorCNA (2) to derive TFX through copy number alterations detected by sparse ($\sim 0.1\times$ coverage) whole genome sequencing (WGS) (3, 11), which we termed ultra-low pass WGS (ULP-WGS). Measuring tumor content through quantification of the presence of individual alleles (12) has demonstrated utility in assessment of response or resistance to therapy in cancers where those alleles are known clonal drivers. An advantage of ichorCNA is that it does not require prior knowledge of alterations present in that patient's cancer, while accounting for differences in clonality and copy number at each locus. Accounting for both are important features of ichorCNA as compared with quantifying individual alleles; for example, point mutations in genes such as *AR* (13), *EGFR* (14), and *ERBB2* (15) often are associated with amplifications at those loci and may occur within distinct subclones, so estimating TFX based on read count of the alternate allele could be inaccurate in these cases. We previously showed that TFX estimated by ULP-WGS using ichorCNA demonstrates close concordance with that estimated from whole exome sequencing using a different method for deriving TFX from somatic DNA alterations called ABSOLUTE (16, 17), thus validating our method for quantification. We also previously demonstrated that TFX is correlated with overall survival in metastatic triple-negative breast cancer (TNBC) (17). Here, we examine clinical correlations of TFX in patients with CRPC and assess TFX as a biomarker of tumor burden and disease dynamics in these patients.

Results

Participating patients were identified for inclusion in this study in clinical cohorts as described in the Methods section. A total of 722 plasma specimens from 164 patients with CRPC were obtained and planned for ULP-WGS — 53 banked samples from 50 patients and 669 prospectively collected samples from 114 patients (Figure 1). Of the banked samples, 17 were low yield and 3 failed library construction (primarily due to inadequate cfDNA extraction from samples where only 1 ml plasma was available), leaving 33 banked specimens from 31 patients available for study. Of the 669 prospectively collected samples, 15 were low yield and 3 failed library construction, and another 19 had poor quality sequencing (as suggested by median absolute deviation [MAD] score > 0.2 , see Methods). Two of the patient identification (ID) codes could not be linked to their medical record, leaving 630 prospectively collected samples from 109 patients available for study. In our previous study, there was no statistically significant difference in cfDNA yield and TFX between banked and prospectively collected specimens (17), and in this current study, there were no obvious differences in these parameters, fragment length, or sequencing quality between the banked and prospectively collected specimens. Thus, we combined these sample sets for a total of 663 samples from

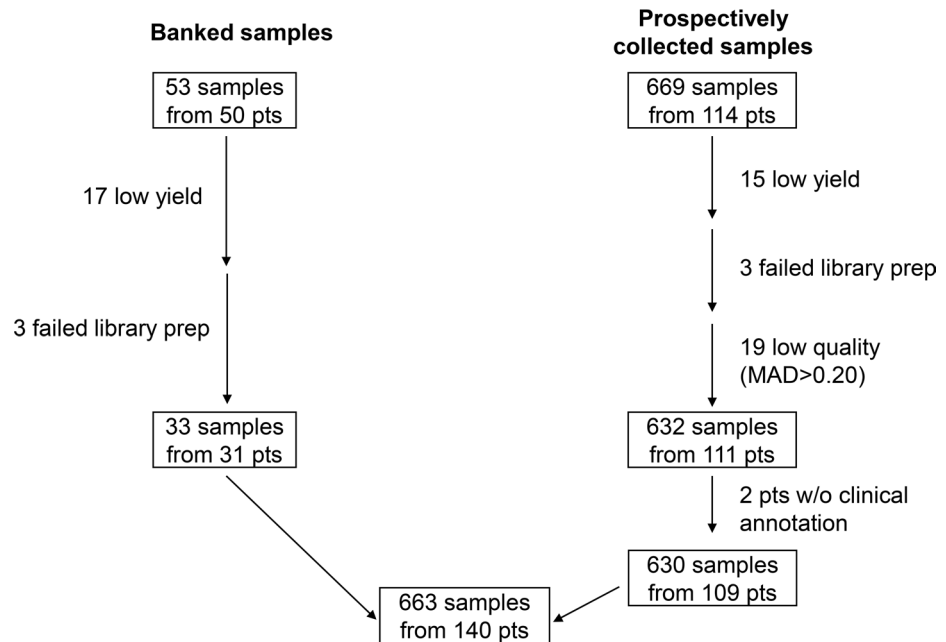


Figure 1. Schema of the clinical cohort. Diagram depicting reasons for excluding samples from the 722 total collected specimens (from 164 patients) to yield 663 samples (from 140 patients) amenable to analysis

140 patients for this analysis (median, 3 samples/patient; range, 1–20). The number of samples per patient, the current (or most recent) treatment at the time of first cfDNA collection, and the number of treatment switches or additions during longitudinal monitoring are summarized in Table 1.

Correlation of TFX with clinical parameters. To better understand the role of TFX as a clinical biomarker, we sought to determine the relationship of TFX with PSA and clinical features associated with overall survival in patients with mCRPC. A recently described multivariable analysis demonstrated that independent prognostic features in mCRPC patients treated with docetaxel were Eastern Cooperative Oncology Group (ECOG) performance status, alkaline phosphatase, hemoglobin, lactate dehydrogenase (LDH), and number of metastases (18). For our analysis, we focused on the time point during longitudinal monitoring when TFX was measured highest as a uniform way to account for the heterogeneity in the number of samples per patient. TFX positively correlated with PSA ($r = 0.41$, $P < 0.0001$) and alkaline phosphatase ($r = 0.47$, $P < 0.0001$). In contrast, there was a negative correlation between hemoglobin ($r = -0.49$, $P < 0.0001$) and TFX (Figure 2, A–C). LDH is not routinely checked at our institution, so TFX was unable to be correlated with this parameter.

Next, we examined the correlation of TFX with sites of metastasis, as determined through conventional CT imaging ($n = 138$), and number of bone metastases visualized by Technetium-99m skeletal scintigraphy ($n = 140$). The presence or absence of metastases involving lymph nodes did not correlate with TFX ($P = 0.74$). However, TFX was significantly associated with the presence (Figure 2D) and number (Figure 2E) of bone metastases — median TFX was 0.014 with no bone metastases, 0.047 with 1–3 bone metastases, and 0.190 for 4+ bone metastases; $P < 0.0001$. TFX was also associated with the site of metastasis — median TFX was 0.34 with distant visceral metastases, 0.077 with bone metastases (without visceral metastases), 0.018 with lymph node–only metastases, and 0.009 with no visualized metastases ($P < 0.0001$) (Figure 2F and Table 2).

We included variables that were significantly associated with TFX on univariate analyses into a multivariable model to assess for whether the associations remained. In the model, the location of metastasis, hemoglobin, and alkaline phosphatase were still associated with TFX (Table 3). Patients with distant metastasis (lung, liver, and brain) were likely to have a higher TFX compared with patients with bone, nodal, or no metastasis. Lower levels of hemoglobin and higher levels of alkaline phosphatase were associated with higher TFX. PSA was not independently associated with TFX in this multivariable model.

Correlation of TFX with response to therapy. We next examined changes in TFX over longitudinal monitoring in 50 patients meeting criteria for this analysis, as described in the Methods. Changes in

Table 1. Description of the patient cohort by number of patients meeting the specified criteria

Number of samples per pt	<i>n</i>	Treatment at first cfDNA draw	<i>n</i>	Treatment Switches/Additions ^A	<i>n</i>
1	51	LHRH-A alone	10	Nilutamide	4
2	18	Bicalutamide	17	Sipuleucel-T	2
3	16	Estradiol/leukine	1	Abiraterone	21
4	8	Nilutamide	6	Enzalutamide	10
5–10	23	Sipuleucel-T	3	+investigational agent	14
11–20	24	Abiraterone	24	Ketoconazole	1
		+investigational agent	8	Docetaxel	10
Prior chemo for HSPC ^B	16	Enzalutamide	25	+carboplatin	3
Prior chemo for CRPC ^B	37	+investigational agent	10	Cabazitaxel	7
		Ketoconazole	2	Radium-223	4
		Docetaxel	9	Olaparib	4
		+investigational agent	1	Mitoxantrone	1
		+carboplatin	3	Investigational agent	3
		Cabazitaxel	5	Palliative RT	9
		Paclitaxel	1		
		Radium-223	5		
		Olaparib	2		
		Best Supportive Care	5		
		Investigational Agent	3		

^AAfter first cfDNA draw, with a plasma sample drawn for cfDNA analysis subsequent to the treatment switch/addition. An individual patient can have multiple treatment changes over the course of monitoring. ^BPrior to first cfDNA draw.

TFx and PSA in 3 patients who responded and then became resistant to cytotoxic chemotherapy, as suggested by PSA changes, are depicted in the left panels of Figure 3. Patient 35 had a rising PSA while being treated with estradiol/leukine and then responded — and subsequently became resistant — to docetaxel chemotherapy. Patient 82 initially responded and subsequently became resistant to enzalutamide in combination with an investigational agent; the patient then responded and subsequently became resistant to cabazitaxel and then did not experience a response to the combination of docetaxel and carboplatin. Patient 85 responded and subsequently became resistant to cabazitaxel, but the patient did not respond to the combination of docetaxel and carboplatin or mitoxantrone.

In all 3 patients, changes in TFx mirrored changes in PSA, suggesting that these changes likely reflect changes in disease burden with response and resistance to therapy. Indeed, all new treatments that led to $\geq 30\%$ PSA decline at ≥ 6 weeks were associated with a decline in TFx from baseline ($n = 16$, median TFx change, -92.2% ; range, -70.3% to -100%) when baseline TFx was $>7\%$. TFx of $>7\%$ was chosen as a threshold for this analysis, as fluctuating values up to ~ 0.07 were seen in patients during periods of disease stability (Supplemental Figure 1; supplemental material available online with this article; <https://doi.org/10.1172/jci.insight.122109DS1>).

We observed that the magnitude of change in TFx was not always concordant with the degree of change in PSA, as some patients had a decline in TFx by an order of magnitude or more but had only modest decreases in PSA. This suggests that TFx and PSA are not interchangeable biomarkers and represent different features of tumor burden and/or biology.

Longitudinal monitoring of non-tumor-derived cfDNA. Accurate estimation of the fraction of cfDNA in the circulation derived from cancerous tissues allows us to estimate the total quantity of tumor-derived cfDNA per 1 ml of plasma, which can be expressed as $\text{TFx} \times \text{total cfDNA yield}$ (ng cfDNA/ml of plasma). Likewise, the total yield of cfDNA derived from noncancerous tissues can be expressed as $(1 - \text{TFx}) \times \text{total cfDNA yield}$ (ng cfDNA/ml of plasma). The total cfDNA yield, the tumor-derived cfDNA yield, and the non-tumor-derived cfDNA yield for patients 35, 82, and 85 are shown in the right panels of Figure 3. There appears to be more variability in the measurements of total cfDNA yield and tumor-derived cfDNA yield, as compared with calculated TFx in longitudinal monitoring, in consonance with our previous findings in TNBC (17). In addition, changes in TFx appear to track

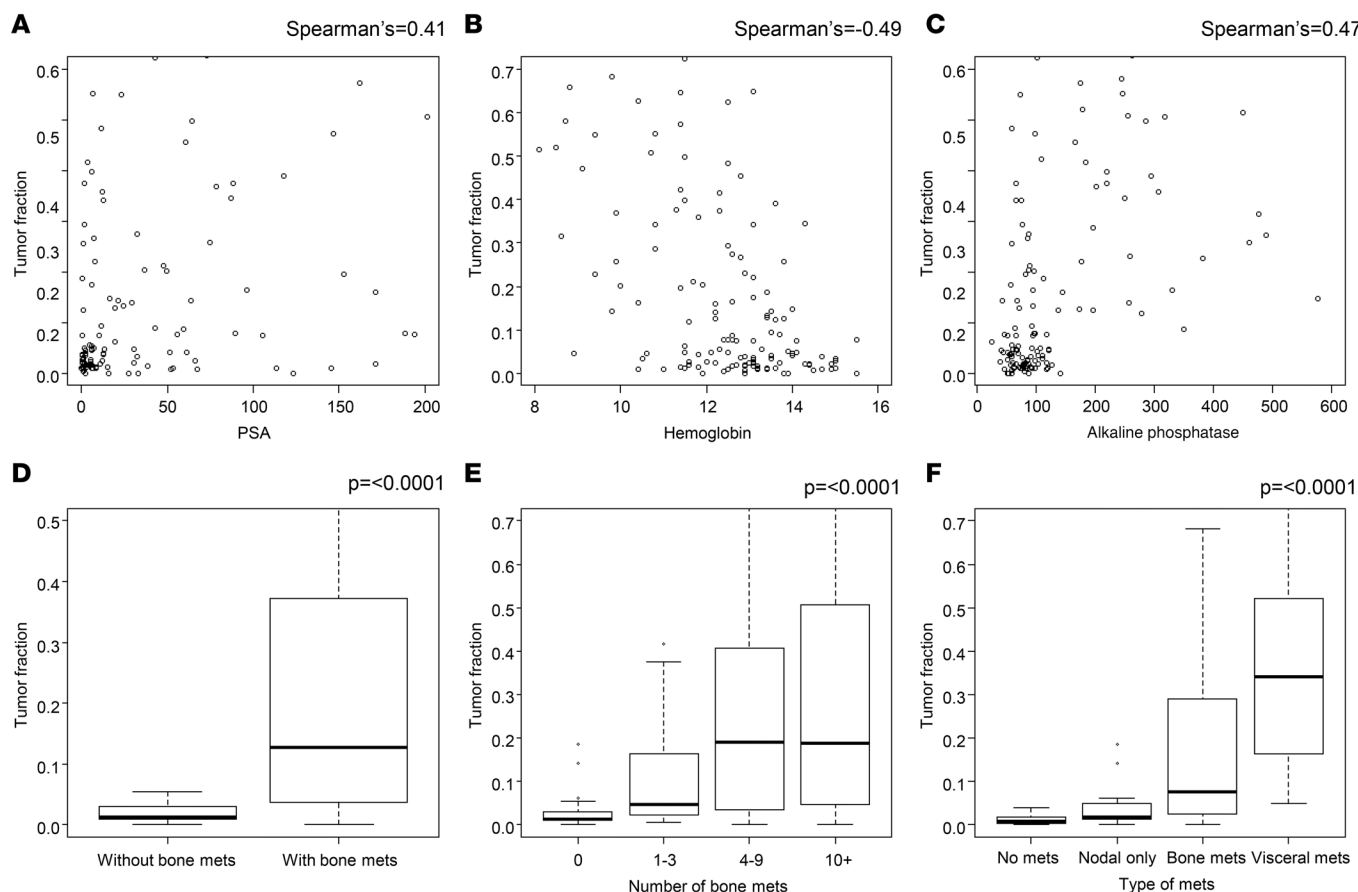


Figure 2. Correlation of Tfx with clinical features. (A–C) Scatter plots depicting the relationship between Tfx at the time point during longitudinal monitoring when the highest Tfx value was seen with PSA (A), hemoglobin (B), and alkaline phosphatase (C) at that time point. Each dot represents an individual patient ($n = 140$). Spearman correlation coefficients are noted in the corner of each plot. (D–F) Box plots depicting the relationship between Tfx at the time point during longitudinal monitoring when the highest Tfx value was calculated with presence or absence of bony metastases by skeletal scintigraphy ($n = 140$) (D), number of bony metastases detected by skeletal scintigraphy ($n = 140$) (E), and site(s) of metastasis detected by bone scintigraphy and/or CT imaging at that time point ($n = 138$) (F). “Nodal only” denotes macrometastatic disease involving lymph nodes but not distant bony or visceral sites. “Bone only” denotes metastatic disease involving bone (+/- lymph nodes) but not distant visceral sites. “Visceral mets” denotes involvement of distant visceral sites (liver, lung, brain, adrenal gland) and does not include extension of local disease (i.e., to bladder/ureter or rectum.) Each patient is represented by a single value within the box plots in each figure (i.e., no patient is represented more than once). Two-sided P values per Wilcoxon rank sum test.

more closely with PSA changes than changes in total cfDNA yield and tumor-derived cfDNA yield, suggesting that Tfx is likely a better biomarker of disease dynamics than these other parameters. (The complete data set for Tfx, total cfDNA yield, tumor cfDNA yield, and normal cfDNA yield, as well as PSA over time for the 50 patients included in this analysis, is included in Supplemental File 1; correlation of total cfDNA yield with clinical parameters as in Figure 2 is depicted in Supplemental Figure 3). Interestingly, the total yield of cfDNA derived from both tumor and noncancerous tissues tracked similarly with Tfx. This suggests that the quantity of DNA in the circulation derived from noncancerous tissues is not static and correlates with changes in tumor burden.

One might expect that taxane chemotherapy would lead to ongoing tissue damage to noncancerous cells in the body that would be detected as an increase in non-tumor-derived cfDNA while on treatment. However, such an increase is not seen in these 3 patients over the time points examined. We cannot rule out temporary increases a few hours or days after each infusion that were not captured during longitudinal monitoring; however, the dosing schedule of these agents every 3 weeks suggests that they should continue to have some biological activity by the time the next dose is due.

Similarly, one might expect that radiation therapy, both external beam radiation therapy (EBRT) and Radium-223, would lead to ongoing damage to noncancerous tissues that would lead to an increase in non-tumor-derived cfDNA. Three patients treated with radiation are depicted in Figure 4. Patient 20 was treated with EBRT to the cervical and thoracic spine, followed by treatment with Radium-223; at day 406,

Table 2. Associations between tumor fraction and location of metastasis in univariate analysis

	<i>n</i>	Median	Q1	Q3	<i>P</i> value
Bone mets					<0.0001
No	25	0.014	0.01	0.03	
Yes	115	0.13	0.04	0.38	
Nodal mets					0.74
No	67	0.08	0.02	0.35	
Yes	71	0.08	0.02	0.29	
Liver mets					<0.0001
No	125	0.05	0.02	0.26	
Yes	13	0.37	0.20	0.66	
Number of bone mets					<0.0001
0	25	0.014	0.01	0.031	
1–3	33	0.047	0.022	0.164	
4–9	20	0.19	0.034	0.407	
10+	62	0.19	0.047	0.507	
Types of mets					<0.0001
No mets	7	0.009	0	0.028	
Nodal only	17	0.018	0.012	0.05	
Bone (no visceral)	95	0.077	0.025	0.29	
Distant visceral	21	0.34	0.16	0.52	

Mets, metastasis; Q1, quartile 1; Q3, quartile 3.

there was a high yield of cfDNA derived from both cancerous and noncancerous tissues, but the subsequent time points showed a return of cfDNA yield to baseline, despite ongoing treatment with Radium-223. Patient 19 had a plasma specimen collected for cfDNA while he was undergoing EBRT to the prostate bed — cfDNA from this time point showed a notable increase in TFX and tumor-derived cfDNA yield but only a slight increase in non-tumor-derived cfDNA. Patient 83 had cfDNA collected soon after completion of EBRT to the spine, and cfDNA from this time point demonstrated no increase in cfDNA from tumor-derived or non-tumor-derived tissue.

TFx in patients maintained on secondary hormonal therapy. As in the clinical trials demonstrating a survival benefit to abiraterone (19) and enzalutamide (20), it is standard practice for patients receiving these agents to remain on therapy until clinical or radiographic progression, even when PSA is rising. Three patients who remained on enzalutamide or abiraterone for >100 days after initial PSA rise on therapy are depicted in Figure 5. Unlike the patients in Figure 3 and Figure 4, where changes in TFX mirrored PSA closely, here we see fluctuations in TFX that do not correlate with PSA changes. These changes are not artifactually related to variability in cfDNA yield, as the cfDNA yield from tumor and noncancerous tissues depicted in the right panels do not obviously correlate with the TFX changes. These cases were manually reviewed to confirm accuracy of results generated by ichorCNA, and visual inspection of genome-wide copy number plots showed increases and decreases in the amplitude of detected copy number changes concordant with TFX changes as calculated by ichorCNA (Supplemental Figure 2). Taken together, we believe that these TFX changes represent bursts of increased tumor-derived cfDNA in the circulation during a period of consistently rising PSA but radiographic stability.

Discussion

Here, we present results of longitudinal monitoring of TFX in cfDNA from 663 prospectively collected and banked plasma samples from 140 patients with CRPC. We have described the correlation of TFX with clinical parameters associated with overall survival in mCRPC, and we have demonstrated that PSA declines with treatment initiation were associated with declines in TFX. We also observed that changes in TFX appear to track more consistently with PSA than changes in total cfDNA or tumor-derived cfDNA, suggesting that TFX is a more reliable biomarker of disease activity than these parameters. TFX during longitudinal monitoring on secondary hormonal therapy can fluctuate dramatically, and cfDNA derived from noncancerous tissues is also dynamic. These observations may lead to mechanistic insights into processes that induce increases in cfDNA derived from both cancerous and noncancerous tissues in the bloodstream and have therapeutic implications.

Table 3. Associations between tumor fraction and location of metastasis, PSA, Hgb, and Alk phos in multivariable model

Variable	Estimate	Error	P value
Intercept	0.6620	0.1493	
Mets location			0.042
No mets	0	.	
Nodal	0.0265	0.08145	
Bone	0.1065	0.07062	
Visceral	0.1926	0.08239	
PSA	-3.92E-6	0.000013	0.7561
Hgb	-0.04876	0.009776	<0.0001
alkaline phosphatase	0.000144	0.000063	0.0227

Mets, metastasis.

TFx was positively correlated with PSA and alkaline phosphatase in our study, as has been observed in a previous report (21); however, negative correlation with hemoglobin was not seen in this previous report. Potential reasons for this discrepancy include different methods for estimating TFx, as well as different methods to assess for correlation; in our study, we correlated hemoglobin with TFx as a continuous variable, whereas in the prior study, hemoglobin was correlated with TFx as binned into 3 levels (undetectable, 2%–30%, and >30%). In our multivariable analysis, TFx positively correlated with alkaline phosphatase and negatively correlated with hemoglobin, but it was not significantly correlated with PSA. These results suggest that the absolute PSA value is not independently related to disease burden across patients, which is expected, given that aggressive variant prostate cancers with predilection to visceral metastases (22), as well as cancers that dedifferentiate with resistance to therapy (23), often produce relatively low levels of PSA. This finding is also consistent with a recent multivariable analysis that demonstrated that alkaline phosphatase, hemoglobin, and number of metastases were independent predictors of overall survival in patients with mCRPC, while PSA was not (18). This may be because hemoglobin and alkaline phosphatase are parameters that are more consistently and predictably affected by organ involvement (primarily bone and liver) across patients with similar disease burden. While our heterogeneous patient cohort with a limited number of patients with >2 samples analyzed did not allow a rigorous statistical analysis of TFx as an independent prognostic or predictive biomarker, we demonstrated that TFx correlated with many of the same features that are prognostic for overall survival in a similar population.

Our findings demonstrate that the TFx in cfDNA is higher in patients with distant visceral metastases (particularly the liver) as compared with those with lymph node or bone metastases. It is unclear whether this is due to burden of disease, acquisition of more aggressive disease biology late in cancer progression, or the anatomy and blood flow of the metastatic milieu of the visceral sites. This finding is in consonance with a previous autopsy study that demonstrated that, in a patient with metastatic neuroendocrine prostate cancer, tumor-derived DNA in the blood had higher contribution from liver metastases than metastases from other sites (24). One important translational impact of these findings is prediction of which patients (or time points from patients) are unlikely to have detectable tumor-derived cfDNA. For instance, clinical cfDNA tests are less likely to detect tumor mutations in patients with normal hemoglobin (Hgb) and alkaline phosphatase, no visceral metastases, or low-volume bone metastases or patients who are responding to therapy; however, cfDNA testing is more likely to detect tumor mutations at the time of progression in these patients.

Monitoring of cfDNA with initiation of systemic treatments for mCRPC demonstrated that hormonal therapies and chemotherapies that led to a PSA decline of $\geq 30\%$ at 6 weeks from treatment initiation also led to a decrease in TFx. This is concordant with a prior study demonstrating that TFx was lower in plasma from patients at the time of response to abiraterone compared with the time of treatment progression (25), as well as the observation that a decrease in cfDNA concentration correlated with outcome for the PARP inhibitor olaparib in the Phase II Trial of Olaparib in Patients with Advanced Castration Resistant Prostate Cancer (TOPARP-A) (5). This finding would suggest that decrease in TFx with treatment initiation carries the potential for assessment of initial treatment response in cancers without a serum biomarker and cancers

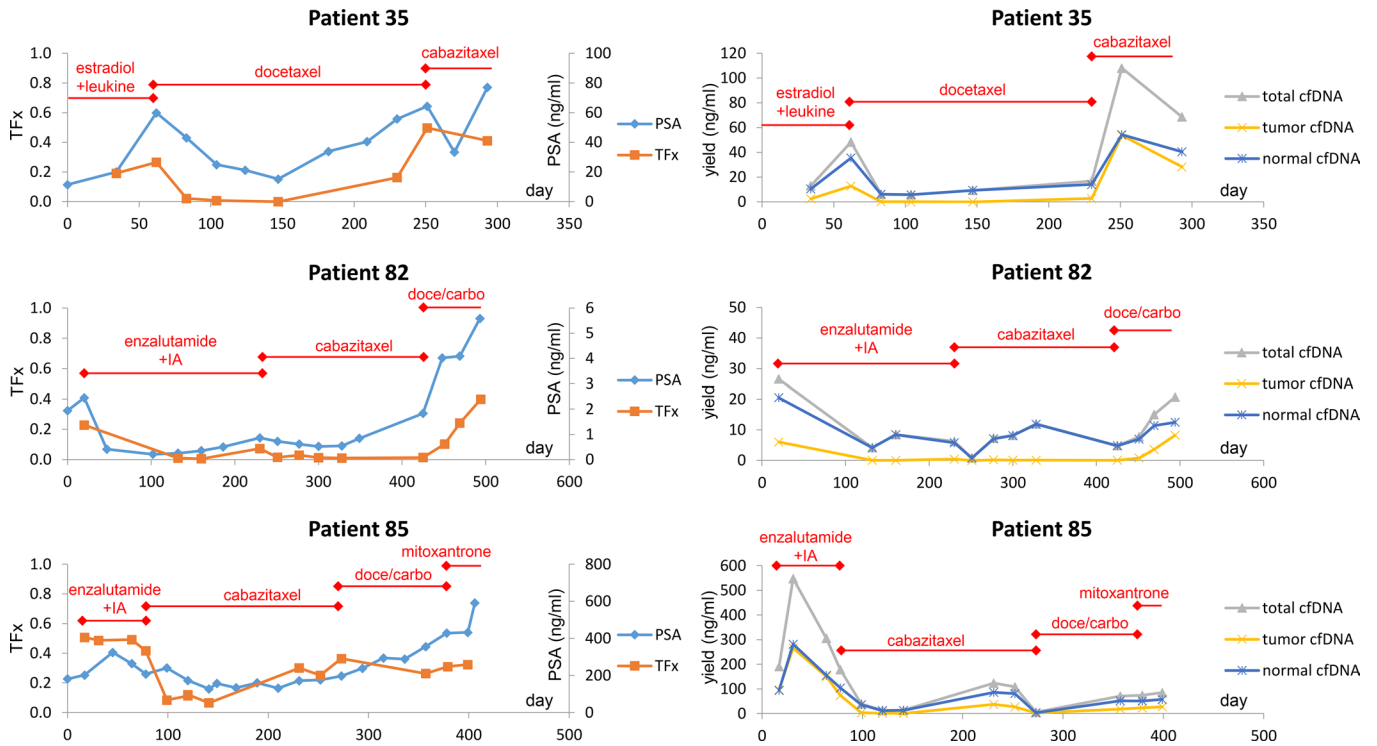


Figure 3. Tumor fraction and cfDNA yield with PSA over time in 3 patients receiving chemotherapy. Left panels: Depiction of TFx (orange markers, left axis) and PSA (blue markers, right axis) over time (measured in days) for 3 patients in the cohort treated with chemotherapy. Start and stop dates for treatments are depicted with a diamond; lines continuing to the edge of the graph without a diamond denote that the therapy was initiated prior to the time points depicted in the graph or completed after the time points depicted in the graph. IA, investigational agent. Two markers from the same time point on day 120 for patient 85 denote replicates from the same cfDNA isolate but with independent library construction and sequencing steps. Right panels: Depiction of total cfDNA yield (in ng/ml of plasma, lavender), tumor DNA yield (TFx × total cfDNA yield, orange), and noncancerous (“normal”) DNA yield ($[1 - \text{TFx}] \times \text{total cfDNA yield}$, blue) in the same 3 patients over time. Note that the scales of total cfDNA yield (y axes) are different for the 3 patients.

without measurable disease (e.g., bone-only metastases). This assay may allow for disambiguation of radiographic flare responses or pseudoprogression, for example, to hormonal therapies and immunotherapeutics. Indeed, a favorable circulating tumor DNA profile using a different assay has been reported to accurately differentiate pseudoprogression from true progression in melanoma (26). TFx could also serve as an early biomarker of disease response to allow for earlier switch or intensification of therapy. The use of TFx decline as an early biomarker of response to therapy should be validated prospectively.

In addition, dense longitudinal monitoring across this heterogeneous cohort has demonstrated unexpected findings that warrant further study. First, the total yield of cfDNA derived from both tumor and noncancerous tissues appears to track with TFx. The fact that cfDNA derived from noncancerous tissues tends to decrease and increase in response to progression of therapy suggests that the same process that leads to tumor-derived DNA in the circulation also contributes to cfDNA from these noncancerous elements. These could include tumor-associated cells (e.g., fibroblasts, endothelial cells, inflammatory infiltrate), normal cells from the metastatic niche (e.g., hepatocytes, osteoblasts/osteoclasts), or cell populations not originating in the tumor (e.g., leukocytes, other BM-derived cell populations). Detection of tissue-specific methylation patterns in cfDNA (27–29) may allow us to identify tissue of origin of these noncancerous elements.

Interestingly, DNA from noncancerous cells in the circulation does not increase in response to initiation of chemotherapy, external beam radiation, or Radium-223 at a subsequent time point 3–4 weeks from treatment start (though there may be a short-term increase that is not captured at the time scale of our longitudinal sampling). This suggests that cfDNA in the circulation is not simply associated with tissue injury, as there is certainly ongoing tissue injury from chemotherapy and radiation that would not be expected to be associated with treatment with abiraterone/enzalutamide, for example. Of note, a single blood draw was obtained in this cohort on a day when a patient was actively receiving radiation therapy, and at that time point, we did see an increase in TFx but not a pronounced increase in cfDNA derived from noncancerous tissues compared with surrounding time points.

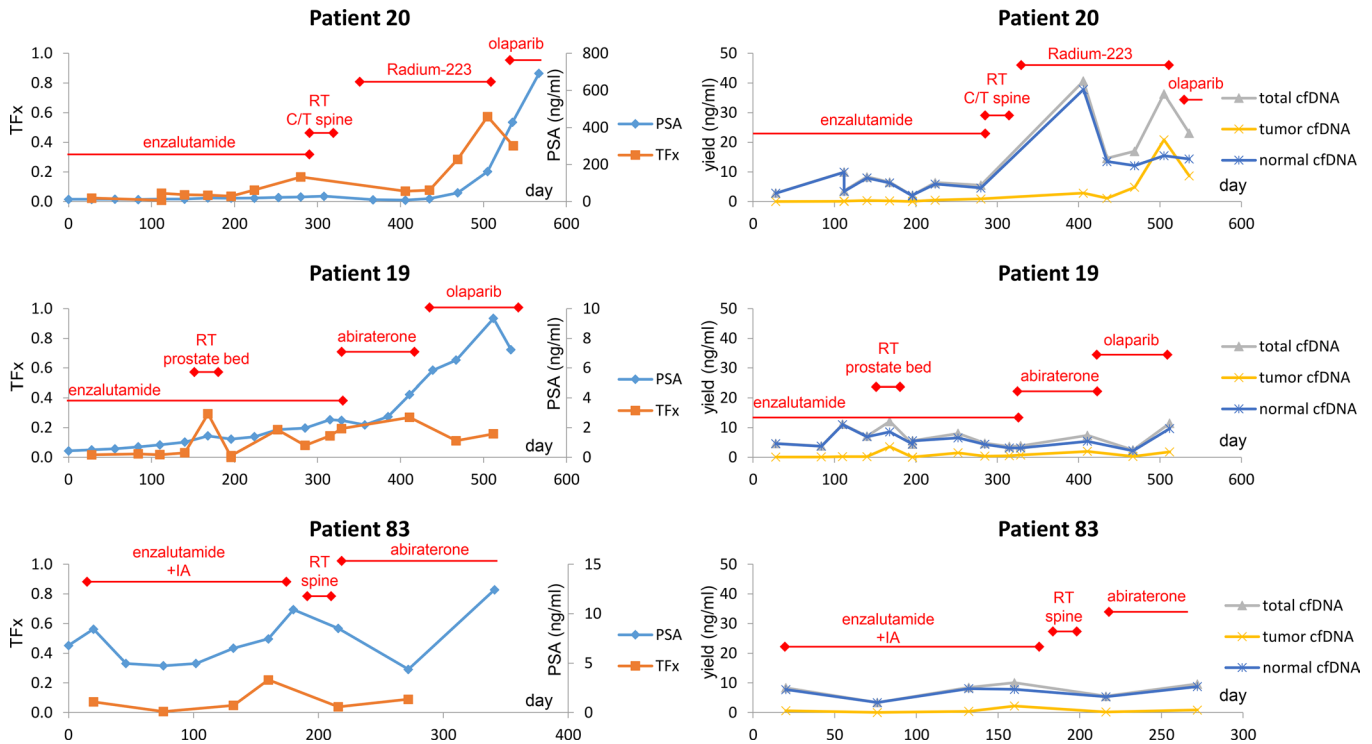


Figure 4. Tumor fraction and cfDNA yield with PSA over time in 3 patients receiving radiation therapy. Left panels: Depiction of TFx and PSA over time for 3 patients in the cohort treated with radiation therapy (RT). Note that only patient 19 had plasma collected for cfDNA on a day when he received external beam radiation therapy. Two markers from the same time point on days 112 and 196 for patient 20, and on day 196 for patient 19 denote replicates from the same blood draw but with independent cfDNA isolation, library construction, and sequencing steps. Right panels: Depiction of total cfDNA yield, tumor DNA yield, and noncancerous (“normal”) DNA yield over time in the same 3 patients.

Our data demonstrate discordance in magnitude of changes in TFx vs. PSA with response to therapy, as well as unexpected fluctuations in TFx during periods of overall disease stability for patients receiving secondary hormonal therapy. Both findings suggest that our understanding of processes leading to changes in TFx is incomplete. Indeed, other studies suggest that cfDNA in the circulation may be more related to tumor metabolism than tumor burden (30). We hypothesize that there are short-term processes, such as inflammatory or metabolic stresses, leading to bursts of cfDNA release that do not necessarily reflect increased tumor burden. As such, a rise in TFx while undergoing treatment is not a reliable indicator of disease progression and should not be used to withdraw a patient from therapy. We are actively investigating the properties of the tumor-derived and noncancerous elements present during these bursts of cfDNA release to suggest the biological underpinnings. Given the short half-life of cfDNA (estimates range from 16 minutes to several hours; ref. 31) in relation to longer-term changes in tumor dynamics that would be captured through a protein-based circulating biomarker or radiographic changes, we posit that examination of circulating cfDNA provides a window into short-term processes that would not easily be assayed through other means.

Although our study focused on using ULP-WGS of cfDNA to quantify TFx and explore its association with clinical parameters in mCRPC, ULP-WGS has also enabled noninvasive tracking of somatic copy number alterations and structural variants (2, 3, 11). For instance, serial analysis of cfDNA has revealed changes in focal amplifications in the tumors of mCRPC patients over the course of therapy (10). Previously, deep targeted sequencing of cfDNA from pretreatment plasma specimens from patients treated with abiraterone or enzalutamide revealed genetic alterations in *AR*, *TP53*, *ATM*, and *BRCA2* that were associated with rapid resistance to therapy (21). ULP-WGS has the capability of detecting certain copy number alterations (particularly losses) or signatures (e.g., tandem duplicator phenotype; ref. 32) that would not be easily detectable through targeted sequencing, and we hypothesize that a subset of these, such as copy losses of *PTEN*, *RBI*, or DNA damage repair genes, would have prognostic and predictive utility. Accordingly, further analyses of ULP-WGS data are underway from specimens obtained in the context of completed and currently accruing

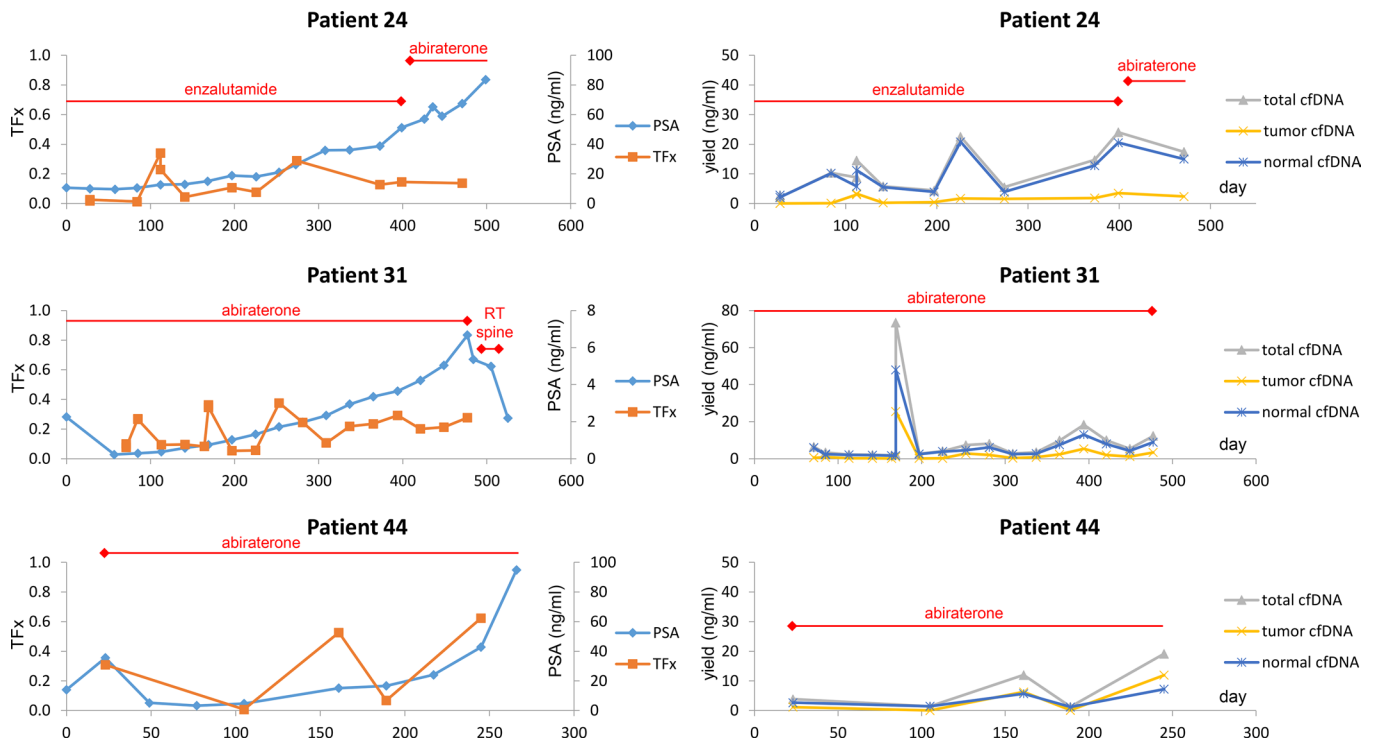


Figure 5. Tumor fraction and cfDNA yield with PSA over time in 3 patients receiving AR-targeted therapy. Left panels: Depiction of TFx and PSA over time for 3 patients in the cohort treated with abiraterone or enzalutamide for >100 days with rising PSA. Two markers from the same time point on days 28 and 112 for patient 24, and on day 85 for patient 31, denote replicates from the same blood draw but with independent cfDNA isolation, library construction, and sequencing steps. Two markers from the same time point on day 71 for patient 31 denote replicates from the same cfDNA isolate but with independent library construction and sequencing steps. Right panels: Depiction of total cfDNA yield, tumor DNA yield, and noncancerous (“normal”) DNA yield over time in the same 3 patients.

clinical trials to derive TFx and characterize copy number alterations and to correlate these features with clinical outcomes in these trials. We anticipate that ULP-WGS will be a complementary approach to targeted sequencing: clinical applications for the combination of ULP-WGS with deep targeted sequencing include (a) revealing a set of genetic features to inform prognosis and optimize initial treatment selection; (b) assessing for initial therapeutic response such that a suboptimal reduction in TFx on-treatment as determined by ULP-WGS would allow for earlier treatment intensification or discontinuation; and (c) revealing new genetic features after therapeutic resistance that would allow for optimal selection of subsequent treatments.

In conclusion, our study demonstrates the potential role of TFx as a clinically relevant biomarker that warrants further study as a prognostic feature and as an early marker of response to therapy. If validated, it would be particularly useful in cancer types without a reliable serologic biomarker. As a relatively low-cost, noninvasive test that is easily accessible by standard phlebotomy, deriving tumor fraction from ultra-low pass whole genome sequencing has the potential to be broadly disseminated to help guide clinical decision making.

Methods

Human subjects. Eligible CRPC patients were identified through the Prostate Clinical Research Information System (CRIS) database at Dana-Farber Cancer Institute (33). The CRIS system comprises data-entry software, a central data repository, collection of patient data including comprehensive follow-up of all patients, and tightly integrated security measures, as previously described (33). The cohorts accrued to this study were patients who either (a) were identified based on prospective chart review to have PSA >20 ng/ml, progressive disease based on rising PSA, and scan progression; (b) were participants in a Phase I study of crizotinib in combination with enzalutamide (DF/HCC protocol no. 14-230, NCT02207504) or a Phase Ib study of abiraterone in combination with ARN-509 (DF/HCC protocol no. 12-338, NCT01792687); (c) were eligible for metastasis biopsy after progression on enzalutamide or abiraterone through the Stand Up 2 Cancer/PCF Dream Team effort based on participation in a Phase II study of abiraterone in combination with dutasteride (DF/HCC protocol no. 10-448, NCT01393730), a Phase II trial of enzalutamide with

correlative assessment of AR signaling (DF/HCC protocol no. 13-301, NCT01942837), a Phase II trial of abiraterone without exogenous glucocorticoids (DF/HCC protocol no. 13-449, NCT02025010), or a tumor biopsy protocol to assess tissue correlates of therapeutic response (DF/HCC protocol no. 09-171); or (d) had frozen plasma banked prior to treatment with carboplatin or a taxane through a tissue and blood specimen banking protocol (DF/HCC protocol no. 01-045). Patients in all of these cohorts had castration-resistant disease, but participants in trial 13-449 were not required to have radiographically detectable metastases. Sampling of plasma from multiple time points could be done during clinical trial or standard of care treatment if the patient had consented to protocol banking protocol DF/HCC protocol no. 01-045. Medical record review was approved by the IRB per DF/HCC protocol no. 18-135.

Clinical specimens. Prospectively collected venous blood samples (10 cc) were collected in EDTA (BD Biosciences) or CellSave Preservative (Cell Search) tubes. Blood was processed to component plasma, buffy coat, and erythrocytes within 4 hours of collection through standard density gradient centrifugation. Plasma samples were subjected to an additional high-speed spin at 19,000 *g* for 10 minutes, and plasma was frozen at -80°C until further processing. Plasma specimens banked under protocol no. 01-045 were processed from blood using standard methods at the time of collection (i.e., without an additional high-speed spin that is performed for prospectively collected samples and with an uncertain gap between blood collection and plasma processing up to ~8 hours) and were subsequently frozen. Frozen aliquots of plasma were thawed at room temperature, and banked plasma specimens were subjected to high-speed spin after thawing. cfDNA was extracted from plasma (4 ml plasma from prospectively collected specimens; 1 ml from banked specimens) using the Qiagen Circulating DNA kit on the QIA Symphony liquid handling system. Quantification of extracted cfDNA was performed using the PicoGreen (Invitrogen) assay on a Hamilton STAR-line liquid handling system.

ULP-WGS. Library construction of cfDNA was performed using the Kapa HyperPrep kit with custom adapters (Integrated DNA Technologies). Generally, 3–20 ng of cfDNA input (median 5 ng, 1,000–7,000 haploid genome equivalents) was used for ULP-WGS. A Hamilton STAR-line liquid handling system was used to automate and perform this method. Constructed sequencing libraries were pooled (2 μl of each \times 96 per pool) and sequenced using 100 bp paired-end reads over 1 lane on a HiSeq2500 (Illumina) for ULP-WGS.

Analysis of ULP-WGS data. ULP-WGS of cfDNA was performed to average genome-wide fold coverage of 0.1 \times . Segment copy number and TFX were derived via ichorCNA (2). Briefly, the genome is divided into nonoverlapping windows, or bins, of 1 Mb, and aligned reads are counted based on overlap within each bin using the tools in HMMcopy Suite1 (<http://compbio.bccrc.ca/software/hmmcopy/>). Centromeres are filtered based on chromosome gap coordinates obtained from UCSC for hg19, including one 1 Mb bin up- and downstream of the gap. The read counts are then normalized to correct for GC-content and mappability biases using HMMcopy R package. \log_2 copy ratios were computed for each bin relative to ULP-WGS data from a reference panel of 27 healthy donors. The cancer patient cfDNA CNA signals is composed of an admixture between DNA fragments derived from tumor and nontumor cells so discrete copy number prediction and TFX estimation is derived using a hidden Markov model. The copy number states can be mapped to integer values corresponding to hemizygous deletions (HETD, 1 copy), copy neutral (NEUT, 2 copies), copy gain (GAIN, 3 copies), amplification (AMP, 4 copies), and high-level amplification (HLAMP, 5 copies). Of note, homozygous deletion states typically occur at smaller scales than the bin sizes used for the analysis and are therefore excluded. Further assumptions of the model are described in ref. 2. Samples were excluded if the MAD of the copy ratios (\log_2 ratio) between adjacent bins, genome-wide, was greater than 0.20, suggesting poor-quality sequence data. Genome-wide copy number plots were generated by ichorCNA. The software ichorCNA is available at <https://github.com/broadinstitute/ichorCNA>.

Data availability. Sequencing data available upon request under dbGaP accession code phs001417.

Statistics. All statistical analyses and data visualizations were performed in SAS 9.4 and R version 3.3.1. Spearman's correlation was used to summarize correlation of PSA, alkaline phosphatase, and hemoglobin levels with TFX. Wilcoxon's rank sum test was used to assess the association between TFX and metastasis site. For the multivariable analysis, a linear mixed effects model with restricted maximum likelihood (REML) method for variance was used to assess the association between location of metastasis, PSA, hemoglobin, alkaline phosphatase, and TFX. We performed 2-sided testing for all statistical tests, and we considered a nominal *P* value less than 0.05 to be significant. Analyses from longitudinal monitoring of TFX were restricted to patients with more than 1 plasma sample that passed the quality checks in Figure 1 with TFX measured at >0.07 at any time point during longitudinal monitoring. TFX of 0.07 was chosen as a threshold,

as fluctuating values up to 0.07 were seen in patients during periods of disease stability (Supplemental Figure 1). We also wanted to focus our analyses above this threshold to mitigate the effect of measurement error at or near the lower limit of detection of 0.03, given a MAD of error of ≤ 0.014 for TFX estimation (2).

Study approval. The present studies in humans were reviewed and approved by the IRB of the DF/HCC. Subjects provided written informed consent to allow the collection of tissue and blood and analysis of clinical and genetic data for research purposes, as well as data sharing through DF/HCC protocol no. 11-104 (Profile - Research on Clinically Acquired Patient Material in Cancer), no. 01-045, or no. 09-171. Medical record review was approved by the IRB per DF/HCC protocol no. 18-135.

Author contributions

ADC and VAA designed the experiments, analyzed and interpreted the results, and wrote the manuscript. ADC, XXW, and EF retrieved clinical data, and LW performed the statistical analyses. GH developed the analytical method of ichorCNA; GH, SSF, JR, and CL performed the analyses and benchmarking. LCH and MET provided clinical samples, and ZZ and EPO coordinated the acquisition and transportation of the specimens and maintained the database linking deidentified patient IDs to the medical record number. G. Gydush, SCR, and DR performed the extraction and sequencing of cfDNA from patient blood samples. LCH, MET, DGS, HAP, G. Getz, MM, JCL, and WCH provided scientific guidance and clinical interpretation.

Acknowledgments

The authors would like to first and foremost acknowledge the courageous patients and their families for their participation and contribution to this study. The authors also acknowledge the generous support from a Wong Family Award in Translational Oncology, a Dana-Farber Cancer Institute Medical Oncology grant, the Gerstner Family Foundation, Janssen Pharmaceuticals Inc., and the Koch Institute Support (core) grant P30-CA14051 from the NCI. The authors acknowledge the Stand Up To Cancer—Prostate Cancer Foundation Prostate Dream Team for its contributions to whole-exome sequencing of prostate tumor biopsies. Stand Up To Cancer is a program of the Entertainment Industry Foundation administered by the American Association for Cancer Research (SU2C-AACR-DT0712). We acknowledge the Arthur and Linda Gelb Center for Translational Research for the acquisition and annotation of clinical samples. ADC has been funded by the Prostate Cancer Foundation Young Investigator Award and the Department of Defense Physician Scientist Training Award. JCL is a Camille Dreyfus Teacher-Scholar.

Address correspondence to: Viktor A. Adalsteinsson, Eli and Edythe L. Broad Institute of MIT and Harvard, 415 Main Street, Cambridge, Massachusetts 02142, USA. Phone: 617.714.7971; Email: viktor@broadinstitute.org. Or to: Atish D. Choudhury, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, Massachusetts 02215, USA. Phone: 617.632.6328; Email: ACHOUDHURY@partners.org.

1. Scher HI, et al. Trial Design and Objectives for Castration-Resistant Prostate Cancer: Updated Recommendations From the Prostate Cancer Clinical Trials Working Group 3. *J Clin Oncol.* 2016;34(12):1402–1418.
2. Adalsteinsson VA, et al. Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. *Nat Commun.* 2017;8(1):1324.
3. Hovelson DH, et al. Rapid, ultra low coverage copy number profiling of cell-free DNA as a precision oncology screening strategy. *Oncotarget.* 2017;8(52):89848–89866.
4. Fan G, Zhang K, Yang X, Ding J, Wang Z, Li J. Prognostic value of circulating tumor DNA in patients with colon cancer: Systematic review. *PLoS One.* 2017;12(2):e0171991.
5. Goodall J, et al. Circulating Cell-Free DNA to Guide Prostate Cancer Treatment with PARP Inhibition. *Cancer Discov.* 2017;7(9):1006–1017.
6. Kienel A, Porres D, Heidenreich A, Pfister D. cfDNA as a Prognostic Marker of Response to Taxane Based Chemotherapy in Patients with Prostate Cancer. *J Urol.* 2015;194(4):966–971.
7. Conteduca V, et al. Androgen receptor gene status in plasma DNA associates with worse outcome on enzalutamide or abiraterone for castration-resistant prostate cancer: a multi-institution correlative biomarker study. *Ann Oncol.* 2017;28(7):1508–1516.
8. Gordevičius J, et al. Cell-Free DNA Modification Dynamics in Abiraterone Acetate-Treated Prostate Cancer Patients. *Clin Cancer Res.* 2018;24(14):3317–3324.
9. Haidl F, Kienel A, Pfister D, Heidenreich A. Cell-free DNA as a surrogate marker in patients with CRPCA undergoing taxane-based chemotherapy. *J Clin Oncol.* 2017;35(6_suppl):e584.
10. Elshimali YI, Khaddour H, Sarkissyan M, Wu Y, Vadgama JV. The clinical utilization of circulating cell free DNA (CCFDNA) in blood of cancer patients. *Int J Mol Sci.* 2013;14(9):18925–18958.

11. Ulz P, et al. Whole-genome plasma sequencing reveals focal amplifications as a driving force in metastatic prostate cancer. *Nat Commun.* 2016;7:12008.
12. Lee JY, et al. Longitudinal monitoring of EGFR mutations in plasma predicts outcomes of NSCLC patients treated with EGFR TKIs: Korean Lung Cancer Consortium (KLCC-12-02). *Oncotarget.* 2016;7(6):6984–6993.
13. Jernberg E, Bergh A, Wikström P. Clinical relevance of androgen receptor alterations in prostate cancer. *Endocr Connect.* 2017;6(8):R146–R161.
14. Shan L, et al. Concurrence of EGFR amplification and sensitizing mutations indicate a better survival benefit from EGFR-TKI therapy in lung adenocarcinoma patients. *Lung Cancer.* 2015;89(3):337–342.
15. Wen W, et al. Mutations in the Kinase Domain of the HER2/ERBB2 Gene Identified in a Wide Variety of Human Cancers. *J Mol Diagn.* 2015;17(5):487–495.
16. Carter SL, et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol.* 2012;30(5):413–421.
17. Stover DG, et al. Association of Cell-Free DNA Tumor Fraction and Somatic Copy Number Alterations With Survival in Metastatic Triple-Negative Breast Cancer. *J Clin Oncol.* 2018;36(6):543–553.
18. Anampa-Guzman AC, Sulca-Huamani O, Perez-Mendez R, Mendoza-Soto G, Chavez PC, Aliaga R. Prognostic factors for overall survival in patients with metastatic castration-resistant prostate cancer: Secondary analysis. *J Clin Oncol.* 2017;35(6_suppl):e597.
19. de Bono JS, et al. Abiraterone and increased survival in metastatic prostate cancer. *N Engl J Med.* 2011;364(21):1995–2005.
20. Scher HI, et al. Increased survival with enzalutamide in prostate cancer after chemotherapy. *N Engl J Med.* 2012;367(13):1187–1197.
21. Annala M, et al. Circulating Tumor DNA Genomics Correlate with Resistance to Abiraterone and Enzalutamide in Prostate Cancer. *Cancer Discov.* 2018;8(4):444–457.
22. Beltran H, et al. Aggressive variants of castration-resistant prostate cancer. *Clin Cancer Res.* 2014;20(11):2846–2850.
23. Bluemn EG, et al. Androgen Receptor Pathway-Independent Prostate Cancer Is Sustained through FGF Signaling. *Cancer Cell.* 2017;32(4):474–489.e6.
24. Beltran H, et al. Whole exome sequencing (WES) of circulating tumor DNA (ctDNA) in patients with neuroendocrine prostate cancer (NEPC) informs tumor heterogeneity. *J Clin Oncol.* 2017;35(15_suppl):5011.
25. Romanel A, et al. Plasma AR and abiraterone-resistant prostate cancer. *Sci Transl Med.* 2015;7(312):312re10.
26. Lee JH, et al. Association Between Circulating Tumor DNA and Pseudoprogression in Patients With Metastatic Melanoma Treated With Anti-Programmed Cell Death 1 Antibodies. *JAMA Oncol.* 2018;4(5):717–721.
27. Lehmann-Werman R, et al. Identification of tissue-specific cell death using methylation patterns of circulating DNA. *Proc Natl Acad Sci USA.* 2016;113(13):E1826–E1834.
28. Sun K, et al. Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments. *Proc Natl Acad Sci USA.* 2015;112(40):E5503–E5512.
29. Kang S, et al. CancerLocator: non-invasive cancer diagnosis and tissue-of-origin prediction using methylation profiles of cell-free DNA. *Genome Biol.* 2017;18(1):53.
30. Morbelli S, et al. Circulating Tumor DNA Reflects Tumor Metabolism Rather Than Tumor Burden in Chemotherapy-Naive Patients with Advanced Non-Small Cell Lung Cancer: 18F-FDG PET/CT Study. *J Nucl Med.* 2017;58(11):1764–1769.
31. Leung F, et al. Circulating Tumor DNA as a Cancer Biomarker: Fact or Fiction? *Clin Chem.* 2016;62(8):1054–1060.
32. Viswanathan SR, et al. Structural Alterations Driving Castration-Resistant Prostate Cancer Revealed by Linked-Read Genome Sequencing. *Cell.* 2018;174(2):433–447.e19.
33. Oh WK, et al. Development of an integrated prostate cancer research information system. *Clin Genitourin Cancer.* 2006;5(1):61–66.