CANCER GENOMICS

original reports

Cell-Free DNA Detection of Tumor Mutations in Heterogeneous, Localized Prostate Cancer Via Targeted, Multiregion Sequencing

Emmalyn Chen, PhD[1]; Clinton L. Cario, PhD[1]; Lancelote Leong, BA[1]; Karen Lopez, BS[2]; César P. Márquez, MD[3,4]; Patricia S. Li, BS[2]; Erica Oropeza, BS[2]; Imelda Tenggara, BS[2]; Janet Cowan, MA[2]; Jeffry P. Simko, MD, PhD[2,5]; Robin Kageyama, PhD[6]; Daniel K. Wells, PhD[6]; June M. Chan, DSc[1]; Terence Friedlander, MD[3]; Rahul Aggarwal, MD[3]; Pamela L. Paris, PhD[2]; Felix Feng, MD[2]; Peter R. Carroll, MD, MPH[2]; and John S. Witte, PhD[1,2]

abstract

PURPOSE Cell-free DNA (cfDNA) may allow for minimally invasive identification of biologically relevant genomic alterations and genetically distinct tumor subclones. Although existing biomarkers may detect localized prostate cancer, additional strategies interrogating genomic heterogeneity are necessary for identifying and monitoring aggressive disease. In this study, we aimed to evaluate whether circulating tumor DNA can detect genomic alterations present in multiple regions of localized prostate tumor tissue.

METHODS Low-pass whole-genome and targeted sequencing with a machine-learning guided 2.5-Mb targeted panel were used to identify single nucleotide variants, small insertions and deletions (indels), and copy-number alterations in cfDNA. The majority of this study focuses on the subset of 21 patients with localized disease, although 45 total individuals were evaluated, including 15 healthy controls and nine men with metastatic castration-resistant prostate cancer. Plasma cfDNA was barcoded with duplex unique molecular identifiers. For localized cases, matched tumor tissue was collected from multiple regions (one to nine samples per patient) for comparison.

RESULTS Somatic tumor variants present in heterogeneous tumor foci from patients with localized disease were detected in cfDNA, and cfDNA mutational burden was found to track with disease severity. Somatic tissue alterations were identified in cfDNA, including nonsynonymous variants in *FOXA1*, *PTEN*, *MED12*, and *ATM*. Detection of these overlapping variants was associated with seminal vesicle invasion ($P = .019$) and with the number of variants initially found in the matched tumor tissue samples ($P = .0005$).

CONCLUSION Our findings demonstrate the potential of targeted cfDNA sequencing to detect somatic tissue alterations in heterogeneous, localized prostate cancer, especially in a setting where matched tumor tissue may be unavailable (ie, active surveillance or treatment monitoring).

JCO Precis Oncol 5:710-725. © 2021 by American Society of Clinical Oncology

INTRODUCTION

Prostate cancer accounts for approximately 20% of all new cancer diagnoses in American men, and an estimated 33,330 men will die from this disease in 2020.[1] Although the majority of prostate cancers are diagnosed when the disease is still localized and are successfully treated, an estimated 20%-40% of patients undergoing radical prostatectomy (RP) will experience biochemical recurrence within 10 years.[2] Additionally, approximately 2.9% of patients develop bone metastases and 2.4% die of prostate cancer within 6 years.[3-5] This is believed in part to be because of pathologically heterogeneous and genetically multiclonal disease, which likely determines available tumor escape mechanisms that allow for tumor survival and proliferation, subsequently driving disease progression and treatment outcome.[6-9] A variety of existing biomarkers in addition to prostate-specific antigen (PSA) have been critical in predicting treatment outcomes, including nomograms that incorporate clinical and pathologic factors, parametric MRI, and molecular testing.[10-12] However, PSA is not tumor-specific and can be elevated in other conditions such as prostatitis. Additionally, tissue biopsies may miss some of the tumor and can lead to underestimation of disease grade and stage, motivating the need for additional modalities to comprehensively assess disease heterogeneity.[14] This may be particularly important if tumor tissue is unavailable during active surveillance, or during disease monitoring after surgery—or other treatments—for detection of minimal residual disease or progression.

Multiple studies have tracked the evolutionary trajectory of localized prostate cancers and have found a number of genomic factors to be predictive of poor outcomes. Specifically, genomic instability resulting from recurrent copy-number alterations in genes such as *MYC*, *NKX3-1*, and *PTEN* is prognostic for

ASSOCIATED CONTENT

Appendix

Data Sharing Statement

Data Supplement

Author affiliations and support information (if applicable) appear at the end of this article.

Accepted on March 11, 2021 and published at ascopubs.org/journal/po on April 27, 2021: DOI https://doi.org/10.1200/PO.20.00428

CONTEXT

Key Objective

The pathologically heterogeneous and genetically multiclonal nature of localized prostate cancer likely determines available tumor escape mechanisms, subsequently driving disease progression and treatment outcome. We evaluated whether cell-free DNA (cfDNA) can detect genomic alterations present in multiple regions of localized prostate tumor tissue. Both low-pass whole-genome and targeted sequencing were performed.

Knowledge Generated

Somatic tumor mutations present in tumor tissue were detected by targeted sequencing of cfDNA collected before radical prostatectomy. Circulating tumor DNA was detected in 12 of 21 patients and was associated with seminal vesicle invasion and tumor mutational burden. However, circulating tumor DNA was not detectable using low-pass whole-genome sequencing.

Relevance

Although cfDNA detected a portion of the genomic heterogeneity present in localized prostate cancer, comprehensive detection of early-stage prostate cancer remains challenging because of significant limitations in assay sensitivity. The utility of this method may be limited to those with more aggressive prostate cancer.

biochemical recurrence following surgery or radiotherapy.[15-18] Prostate tumors also harbor a large proportion of somatic single nucleotide variants that are not protein-altering and exhibit extensive intrafocal heterogeneity.[6] Although non-synonymous mutations have been found in *SPOP*, *FOXA1*, *MED12*, and *ATM*, these recurrently mutated genes are only found in $< 10\%$ of patients.[15,19] Overall, polyclonal tumors with multiple tumor populations originating from a single clone are more likely to result in adverse outcomes.[7] Although PSA is used to monitor localized prostate cancer, it is unable to detect these genomic features, which can be distributed heterogeneously across tumor cell subpopulations. However, plasma cell-free DNA (cfDNA) may provide a way to assess the genomic profile of the tumor without the use of invasive tissue biopsies.

Plasma cfDNA remains a promising tool for directly detecting tumor DNA that is shed into the bloodstream. Both droplet digital polymerase chain reaction (ddPCR) and next-generation sequencing (NGS) have been successfully used to depict clonal evolution and identify genomic alterations in the context of early detection and disease monitoring. For example, personalized multiplex-PCR next-generation sequencing of cfDNA has been used to derive tumor phylogenetic trees and characterize post-operative non–small-cell lung cancer relapse.[20] Targeted error correction sequencing with dual-index barcodes has been used to discover somatic alterations in early-stage cancers, including breast cancer, colorectal cancer, lung cancer, and ovarian cancer.[21] Detection of *BRCA2* reversion mutations in cfDNA have been associated with resistance to PARP inhibitors, allowing for monitoring of treatment resistance in patients with metastatic castration-resistant prostate cancer (mCRPC).[22] Recently, ultra-low-pass whole-genome sequencing (WGS) and targeted resequencing were unable to detect tumor fragments in cfDNA in patients with localized prostate cancer.[23] Although detection of somatic alterations in cfDNA for localized disease may be more challenging because of low disease burden, advances in sample processing, library preparation, targeted panel design, and bioinformatic tools have the potential to be used for broad yet sensitive variant detection in the cfDNA of a heterogeneous disease such as localized prostate cancer.[24]

In this study, extensive tissue sampling was used to capture tumor heterogeneity and provide a patient-specific gold standard for comparison of matched cfDNA. We performed both targeted and low-pass whole-genome sequencing of cfDNA, which was collected from blood draws taken immediately before surgery, and multiregion tumor tissue, which was collected from surgically resected prostate tissue from 21 men with localized prostate cancer, who underwent RP as primary treatment (Appendix Fig A1). This allowed for assessing the genomic heterogeneity of localized prostate cancer that is a result of clonal evolution. Next, we used a 2.5-Mb targeted panel to also evaluate the mutational burden found in 15 healthy controls and nine men with mCRPC. Importantly, the targeted panel, as previously described, was generated without a priori patient-specific tumor mutational information in an attempt to capture a wide range of potentially important mutations, as well as reflect the scenario of repeated cfDNA blood collection in a clinical setting when tumor tissue biopsy is not possible.[25] Additionally, plasma cfDNA was barcoded with duplex unique molecular identifiers (UMIs) to improve variant detection in a setting where the fraction of circulating tumor DNA can be low (Appendix).

This study was carried out in accordance to the Human Research Protection Program Institutional Review Board at UCSF (IRB 11-05226 and IRB 12-09659). All subjects provided written informed consent in accordance with the Declaration of Helsinki.

RESULTS

Plasma cfDNA Mutational Burden and Prostate Cancer

The median cfDNA variant count was 1,089 (interquartile range [IQR] = 761) for nondiseased controls, 1,843 (IQR = 605) for men with localized prostate cancer, and 5,081 (IQR = 716) for men with mCRPC. The average cfDNA variant count for men with localized prostate cancer was statistically significantly higher than those observed in nondiseased controls (Table 1; Appendix Fig A2; $P < .01$). Men with mCRPC had a statistically significantly higher cfDNA variant count than men with localized disease ($P < .0001$) or controls ($P < .0001$). Age was not correlated with cfDNA variant count within each group.

Genomic Heterogeneity in Localized Prostate Cancer

A total of 21 men had between one and nine samples collected from distinct tumor foci (71 specimens total), which were then sequenced with our targeted panel in an effort to capture regions with varying histology and identify potential clonal and subclonal mutations. Targeted sequencing identified somatic variants in all 71 prostate cancer specimens. A median of eight variants with a range from 1 to 1,091 variants (IQR = 8) were identified across all foci (Fig 1A). Although 17% of the target panel was composed of noncoding variants, 88% of the tissue variants identified were in noncoding regions.

Nonsynonymous variants were identified in 22 of the 71 tissue specimens (Fig 1B). Alterations were discovered among commonly mutated genes in localized prostate cancer, including *FOXA1*, *PTEN*, *MED12*, and *ATM*. The majority of somatic mutations observed in patients with multiple regions sequenced were private to each tumor focus, with a subset of mutations present in all regions. In one patient, the potentially clonal mutation identified in all six tissue regions was found in *ATM* (Fig 2).

Tissue samples from five patients for which there was sufficient DNA underwent WGS at 4× coverage and were found to harbor a median of 18 copy-number alterations with a range from 2 to 626 copy-number alterations (IQR = 153). Across all patients, 2%-25% of the copy-number changes were likely clonal and found in all foci for a given patient. For each patient, copy-number alterations are shown in Appendix Figure A3, with multiple tracks overlaid for different samples. Samples had a median percentage of genome altered (PGA) of 9% with a range from 0.002% to 17.8% (IQR = 4.8%). Loss of *CHD1*, *NKX3-1*, *CDKN1B*, *PTEN*, and *TP53* were found in all sequenced tissue regions for two of the five patients. Notably, *MYC* amplification was found in a subset of the regions for three patients and found to co-occur with PTEN loss in one patient. However, somatic copy-number alterations were not detected in the cfDNA of patients with localized prostate cancer.

Somatic Tumor Tissue Variants Identified in cfDNA With 2.5-Mb Targeted Panel Sequencing in Localized Prostate Cancer

Somatic tissue variants were identified in cfDNA sequenced without prior knowledge of the variants present in tumor tissue. A matched source of normal tissue was used to exclude germline variants from the analysis. Overlapping variants were found in 12 of the 21 patients, with a range of 1-62 variants and 0.2% to 13.8% tissue variants detected by cfDNA (Fig 3, Data Supplement). For the majority of the patients, overlapping variants were subclonal and found in a subset of the tumor tissue regions sequenced. Although the targeted panel was composed of 17% noncoding variants, 85% of the overlapping variants were found in intronic and intergenic regions, which is comparable to the 88% noncoding variants found in tumor tissue. Of the 21 variants found in coding regions, four were nonsynonymous missense mutations. The cfDNA variant allele frequency for the overlapping variants ranged from 0.9% to 19% (Fig 4).

TABLE 1. Clinical Characteristics of Individuals Included in the Study at Baseline

Characteristics	Healthy (n = 15)	Localized (n = 21)	mCRPC (n = 9)
Age, years			
Median ± IQR	33 ± 19	66 ± 10	63 ± 3
Range	22-63	50-74	59-75
Pathologic Gleason			
6	—	1	—
7	—	11	—
8-10	—	9	—
Pathologic stage			
Organ confined (pT2)	—	6	—
Not organ confined (pT3, pT4)	—	15	—
Extraprostatic extension (pT3a)	—	9	—
Seminal vesicle invasion (pT3b)	—	4	—
Lymph node involvement (N1)	—	6	—
PSA, ng/mL			
Median ± IQR	—	9.1 ± 12	34.5 ± 67[a]
Range	—	2-69.9	0-263
No. tissue samples per patient			
Median	—	3	—
Range	—	1-9	—
Tumor tissue variants[b]			
Median ± IQR	—	8 ± 8	—
Range	—	1-1,091	—
cfDNA variants[b]			
Median ± IQR	1,089 ± 761	1,843 ± 605	5,081 ± 716
Range	598-2,423	1,172-2,595	4,285-5,938

Abbreviations: cfDNA, cell-free DNA; IQR, interquartile range; mCRPC, metastatic castration-resistant prostate cancer; PSA, prostate-specific antigen.

[a]One man with unknown data in the cohort.

[b]Sequenced with 2.5-Mb targeted panel.

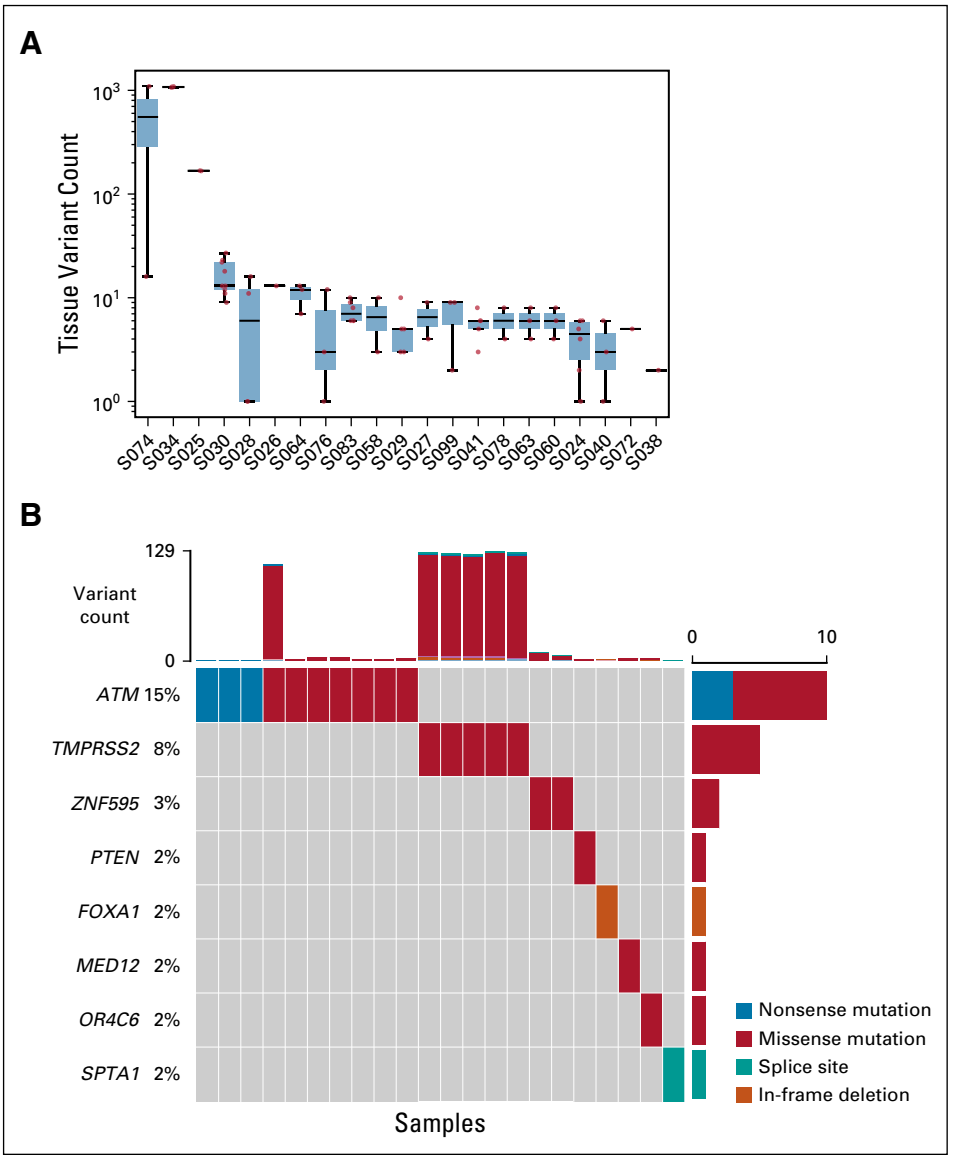


FIG 1. Localized prostate cancer tumors harbor a wide range of somatic variants both across patients and within a patient's tumor foci. (A) Distribution of variant counts for a cohort of 21 men with one to nine tumor tissue samples collected from regions of varying histology that were sequenced with a targeted panel. Patients are on the *x*-axis and each dot represents the variant count for a single tissue sample, and tissue variant counts are on the log-scaled *y*-axis. A median of eight variants with a range from 1 to 1,091 variants identified across foci. Matched whole blood was used as a source of normal for patients S025, S034, S041, and S060. (B) Nonsynonymous variants in the listed genes were detected in 22 of the 71 tissue samples. (Top) The number of variants in each sample. (Left) Gene names and percentage of samples with mutations in a given gene. (Center) Mutations colored by coding consequence. (Right) Number of mutations in a given gene.

Determinants of Somatic Tissue Variant Detection in cfDNA for Localized Prostate Cancer

Our ability to detect some of the observed somatic variants in cfDNA was positively associated with the number of variants coidentified in tumor tissue (P = .005) and with seminal vesicle invasion (P = .019). There was no clear pattern of association observed with the other clinical factors. However, the small number of patients with overlapping variants may not be sufficiently powered to detect a difference between groups (Appendix Table A1).

DISCUSSION

We found that targeted sequencing of cfDNA—without a priori patient-specific tumor mutation information—identified somatic alterations found in matched tumor tissue from multiple regions. Detection of these concordant variants was associated with seminal vesicle invasion and the number of somatic variants initially found in the tumor tissue samples, predicating its use for patients with poor prognostic factors in a localized setting. Our study demonstrates the ability of cfDNA to detect a portion of the genomic heterogeneity present in localized prostate cancer, potentially allowing for dynamic monitoring of emerging resistant subclones throughout the course of disease. However, comprehensive detection of early-stage prostate cancer with cfDNA remains elusive, with significant limitations in blood volume collection and assay sensitivity acting as barriers to implementation in the clinic.

Although tumor mutations were identified in cfDNA for 57% of patients with localized disease, the remaining patients did not have detectable overlapping variants. Interestingly, clinicopathologic factors may play an important role in the detection of somatic tissue variants in plasma cfDNA. Specifically, tissue variants were found in the cfDNA of three men with the highest diagnostic PSA levels (ranging

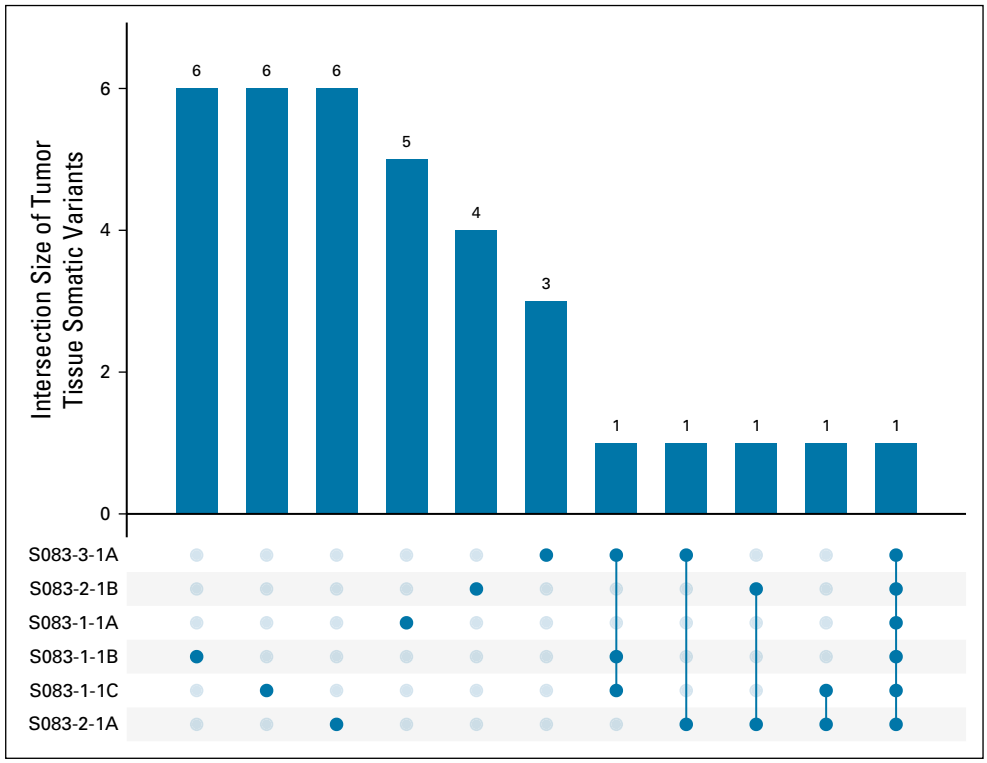


FIG 2. Within a single patient, there appears to be likely clonal and subclonal mutations. Rows are tumor tissue samples and bars show the number of variants that are either private to each sample or are shared among multiple tissue samples. Nonsynonymous tumor tissue variants were identified in several well-characterized genes, including in *ATM*, *FOXA1*, *PTEN*, and *MED12*. For this patient, the mutation shared among all six samples is found in *ATM*, which is commonly found in localized prostate cancer and known to play an important role in cell cycle regulation and maintenance of genomic integrity. Mutations in *FOXA1* have been associated with increased AR-driven transcription, whereas *PTEN* is known to play a role in cell migration and DNA repair as an effector of the PI3K signaling pathway. For patients with mCRPC, nonsynonymous variants were found in *TP53*, *CDK12*, *PTEN*, and *AR*, which have also been previously reported. mCRPC, metastatic castration-resistant prostate cancer; PI3K, phosphatidylinositol 3-kinase.

from 39 to 70 ng/mL) who also had high Gleason scores of nine. Two of these patients had metastasis to nearby lymph nodes and pathologic T4 staging, with invasion of nearby structures. Notably, two of these patients had the two highest tumor volumes measured. This suggests that patients with cfDNA detection likely have further micrometastatic disease, limiting the utility of this assay to those with more aggressive prostate cancer.

Technical considerations related to library preparation and sequencing also affect cfDNA detection. A majority of the tumor tissue variants that were not detected in cfDNA had zero read coverage, suggesting that cfDNA fragments with the variants may not have been present initially during cfDNA extraction or were present, but either not captured during the hybridization step or failed to bind to the flow cell before sequencing. Other tissue variants that were not detected in cfDNA had sufficient total coverage but insufficient UMI family coverage supporting the alternate allele (Appendix Fig A4). This remains a major limitation and demonstrates the need for rigorous assay development and validation before further evaluation of the utility of cfDNA in localized prostate cancer.

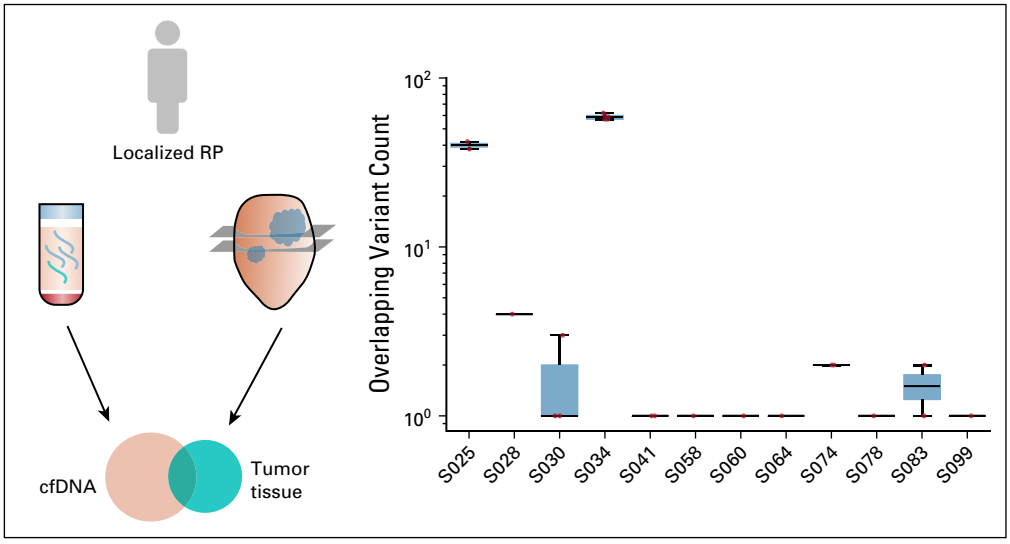


FIG 3. Somatic tissue variants were detected through targeted cfDNA sequencing without prior identification of variants present in tumor tissue. Boxplots show the distribution of variants overlapping between cfDNA and tumor tissue. Patients are on the *x*-axis and each dot represents the overlapping variant count for a single tissue sample, with counts shown on the log-scaled *y*-axis. Overlapping variants were detected in 12 of the 21 patients sequenced with the targeted panel and ranged from 1 to 62 variants. cfDNA, cell-free DNA; RP, radical prostatectomy.

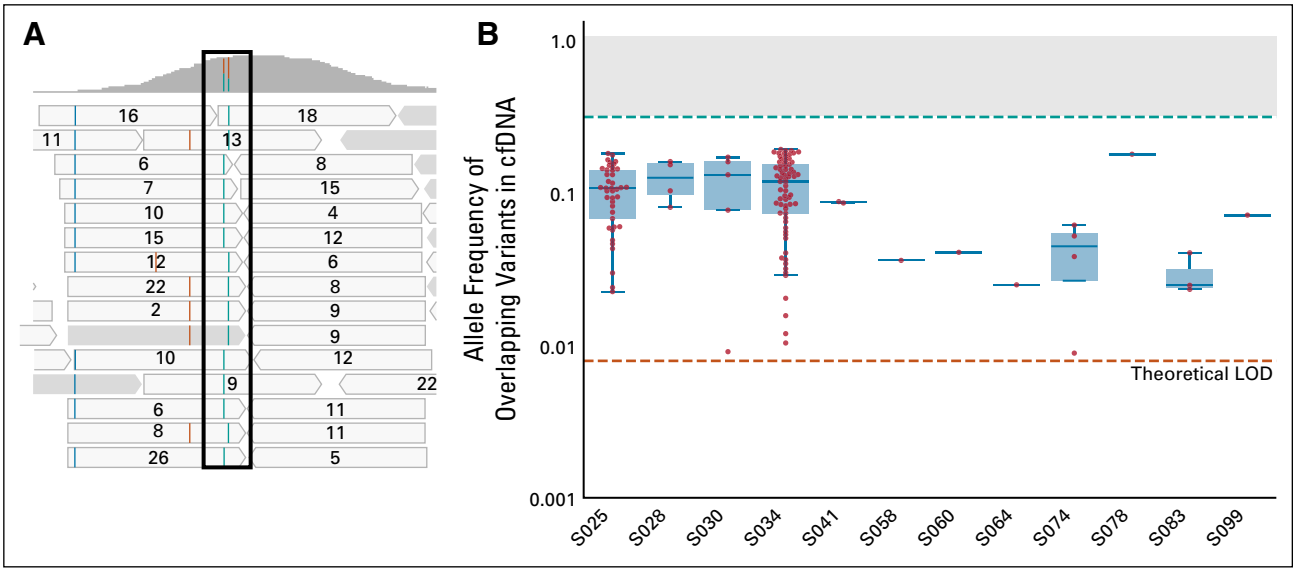


FIG 4. VAF of overlapping variants in cfDNA ranged from 0.9% to 19%. (A) Schematic of an example cfDNA variant highlighted by a vertical blue bar and the number of reads supporting the call. Numbers reflect the number of reads in the same UMI family used for consensus calling. (B) Boxplots show the allele frequency distribution for overlapping variants on a log-scaled *y*-axis. Each dot represents a variant identified in a given patient. A germline threshold is shown at 20% and a theoretical LOD at 0.8%. cfDNA, cell-free DNA; LOD, limit of detection; UMI, unique molecular identifier; VAF, variant allele frequency.

Importantly, many of the somatic tumor tissue alterations identified in this study's cohort of localized cases were also previously identified in large cohorts of patients with localized prostate cancer, providing a patient-matched gold standard for comparison with cfDNA variants detected in this study. A fraction of these alterations were found to be common across all tissue samples, with others identified in only a subset of the samples, confirming the importance of comprehensive sampling.

Although 17% of our 2.5-Mb targeted panel was composed of noncoding variants, 88% of the variants found in the tumors of men with localized disease were in noncoding regions. The impact of these specific alterations remains largely unknown, but prior studies have discovered cancer driver noncoding elements in regulatory regions (ie, promoters and enhancers) and noncoding single nucleotide variants that alter RNA secondary structures.[33-36] In a recent study, noncoding mutations were found to target cis-regulatory elements of *FOXA1* and modulate the binding of transcription factors, exposing a potential therapeutic target and highlighting the importance of mutations in untranslated regions.[36] Similar to the percentage of noncoding variants detected in the tumors of men with localized prostate cancer, 85% of the overlapping variants identified in both cfDNA and matched tumor tissue were found in noncoding regions. As a result, the design of future panels could focus on the improved detection of noncoding variants. Variants in *FOXA1* were identified in both cfDNA and tumor tissue, although found at different locations in the gene. Both likely clonal and subclonal mutations were identified, supporting the ability of cfDNA to capture somatic alterations from multiple tumor cell populations and detecting intrapatient heterogeneity.

To ensure that the tissue variants identified in cfDNA were not clonal hematopoietic mutations of indeterminate potential (CHIP), which typically accumulate during the aging process, we looked for the presence of alterations in genes commonly associated with CHIP.[37] Variants were found in *DNMT3A*, *ASXL1*, *TET2*, and *NOTCH2* in the cfDNA of healthy patients, in white blood cells from patients with localized disease, and in cfDNA from patients with localized disease, which is expected since the majority of cfDNA is derived from hematopoietic cells; however, variants overlapping between cfDNA and localized tumor tissue were not found in these genes, suggesting that the CHIP effect does not explain our findings.[38] Giving credence to the cfDNA variants found in genes previously found to be mutated in prostate cancer, 5% of the cfDNA variants found in coding regions were identified in *AR*, *ATM*, *BRCA2*, *BRAF*, *CDK12*, *CHEK2*, *IDH1*, *PIK3CA*, *MYC*, and *FOXA1*, although not detected in tumor tissue.[15] Interestingly, we observed genes that were altered in both cfDNA and tumor tissue for a given patient, albeit at a different locus.

In undertaking this study, we leveraged a number of methods to ensure broad and sensitive detection of cfDNA variants for patients with localized prostate cancer. Relatively large blood volumes, between 13 and 25 mL, were collected and centrifuged with an initial low spin at 1,900*g* followed by a high spin at 16,000*g* to remove leukocytes and cellular debris. During library preparation, UMIs were used to barcode cfDNA fragments and take advantage of

sequence complementarity of the double-stranded DNA and duplicates that arise during amplification. Importantly, the 2.5-Mb targeted panel used in this study was generated by using a classification and ranking model trained on WGS data from 550 prostate tumors in the International Cancer Genome Consortium, and included both coding and noncoding regions. This optimized the composition of the panel to capture the heterogeneity previously seen in localized prostate cancer, while limiting the panel in size to allow for higher coverage at a lower sequencing cost. Additionally, analysis of cfDNA variants was performed with matched normal samples to filter out germline and clonal hematopoiesis variants and compared with matched tumor tissue from multiple regions to confirm variant calls.

There are a number of limitations to this study that merit consideration. First, somatic tissue variants were detected in cfDNA for only a subset of our study subjects. Although factors including seminal vesicle invasion and tumor mutational burden were predictors of detection, patients with earlier stage disease may have fewer tissue variants detected in their cfDNA. The few number of variants identified in the tumor tissue biopsies remains a challenge in detection, since these variants were used to verify those identified in cfDNA for this study. These ancillary analyses investigating potential determinants of detection are also limited because of small sample sizes and are meant to be hypothesis-generating analyses. Second, the selection of a matched source of normal tissue likely affects the final set of cfDNA variants identified because of its effect on the exclusion of germline and CHIP variants during analysis. We used normal tissue from nearby seminal vesicles for 17 of 21 patients, which may have a genomic profile that is more similar to the prostate tissue than to patient-matched whole blood. Consequently, CHIP variants may remain after cfDNA variant calling, and somatic alterations arising from mosaicism, a process where mutations occur during development and propagate to a subset of tissues, may be removed. Additionally, the median age of healthy individuals was lower than those of patients with prostate cancer. Although age was not found to be correlated with cfDNA variant count in this study, previous studies have shown tumor mutational burden to increase significantly with age.[39] Future studies evaluating the cfDNA mutational burden would benefit from the selection of age-matched controls. Finally, the relatively short follow-up time for a protracted disease like localized prostate cancer means detection of relapse may be difficult. After surgery, four of the 21 patients experienced biochemical recurrence with a median follow-up time of 2.34 years and a range of 66-1,502 days.

In previous studies, features identified in the subclonal architecture of localized prostate cancer have been found to identify patients at higher risk of relapse. In one study, almost two thirds of men with localized prostate cancer had tumors that harbored multiple subclones, and these men relapsed following treatment at a much higher rate than men with monoclonal tumors.[7] In this study, among men with localized disease, all patients with more than one sequenced tissue region had tumors that harbored subclonal mutations, suggesting from previous studies that these men may have an increased risk of subsequent relapse after surgery. To this end, they may benefit from the use of cfDNA to identify alterations indicative of poor prognosis, which may help distinguish between patients who should be treated immediately from those who could remain on active surveillance.

In summary, we show that targeted sequencing of cfDNA without prior patient-specific tumor mutation information can be used to identify somatic alterations, with implications for disease monitoring and detection of emerging subclones through repeat sampling. Targeted sequencing of cfDNA molecules can detect both potentially clonal and subclonal somatic tissue variants, with clinicopathologic and technical factors influencing detection. Future studies investigating the regulatory role of noncoding somatic mutations in localized prostate cancer will help elucidate the functional impact of cfDNA detection of these types of alterations. Combined with previous studies, the importance of detecting somatic alterations using cfDNA in localized disease is developing. Our study supports the use of cfDNA in the assessment of heterogeneous, localized prostate cancer, which will be further strengthened by ongoing technological advances to enrich for tumor fragments found in cfDNA.

AFFILIATIONS

[1]Department of Epidemiology and Biostatistics, University of California, San Francisco, CA
[2]Department of Urology, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, CA
[3]Division of Hematology/Oncology, University of California, San Francisco, CA
[4]School of Medicine, Stanford University, Stanford, CA
[5]Department of Anatomic Pathology, University of California, San Francisco, CA
[6]Parker Institute for Cancer Immunotherapy, San Francisco, CA

CORRESPONDING AUTHOR

John S. Witte, MS, PhD, University of California, San Francisco, UCSF Helen Diller Research Building, 1450 Third St, San Francisco, CA 94143-0794; e-mail: JWitte@ucsf.edu.

SUPPORT

Supported by the UCSF Goldberg-Benioff Program in Cancer Translational Biology, NIH R01 CA088164, NIH R01 CA201358, the UCSF Moritz-Heyman Discovery Fellows Program, and the Steven and Christine Burd Safeway Distinguished Professorship.

DATA SHARING STATEMENT

A data sharing statement provided by the authors is available with this article at DOI https://doi.org/10.1200/po.20.00428.

AUTHOR CONTRIBUTIONS

Conception and design: Emmalyn Chen, Clinton L. Cario, Jeffry P. Simko, John S. Witte
Financial support: Emmalyn Chen, John S. Witte
Administrative support: Lancelote Leong, Peter R. Carroll, John S. Witte
Provision of study materials or patients: Lancelote Leong, Karen Lopez, César P. Márquez, Erica Oropeza, Jeffry P. Simko, June M. Chan, Terence Friedlander, Felix Feng, Peter R. Carroll, John S. Witte
Collection and assembly of data: Emmalyn Chen, Clinton L. Cario, Lancelote Leong, Karen Lopez, César P. Márquez, Patricia S. Li, Erica Oropeza, Imelda Tenggara, Janet Cowan, Jeffry P. Simko, June M. Chan, Terence Friedlander, Peter R. Carroll, John S. Witte
Data analysis and interpretation: Emmalyn Chen, Clinton L. Cario, César P. Márquez, Robin Kageyama, Daniel K. Wells, Terence Friedlander, Rahul Aggarwal, Pamela L. Paris, Felix Feng, Peter R. Carroll, John S. Witte
Manuscript writing: All authors
Final approval of manuscript: All authors
Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/po/author-center.

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians (Open Payments).

Emmalyn Chen
Consulting or Advisory Role: Avail Bio

Clinton L. Cario
Employment: Avail Bio
Stock and Other Ownership Interests: Avail Bio

Janet Cowan
Stock and Other Ownership Interests: GlaxoSmithKline, McKesson

Jeffry P. Simko
Stock and Other Ownership Interests: 3D Biopsy, Lightspeed Microscopy, JNJ, Gilead Sciences, Abbvie, Proetan BioDiagnostics
Consulting or Advisory Role: 3D Biopsy

Robin Kageyama
Employment: Gilead Sciences
Stock and Other Ownership Interests: Gilead Sciences

Daniel K. Wells
Employment: Immunai
Leadership: Immunai
Stock and Other Ownership Interests: Immunai
Consulting or Advisory Role: Immunai, Rubius Therapeutics

June M. Chan
Employment: GRAIL, Myriad Genetics
Stock and Other Ownership Interests: GRAIL
Research Funding: GRAIL, Genomic Health, GenomeDx, Myriad Genetics
Travel, Accommodations, Expenses: GRAIL, Myriad Genetics

Terence Friedlander
Leadership: Med BioGene
Honoraria: EMD Serono, AstraZeneca/MedImmune, Seattle Genetics/Astellas
Consulting or Advisory Role: AstraZeneca, Janssen, Pfizer, Abbvie, Dendreon, Dava Oncology, EMD Serono, Merck
Research Funding: Janssen, Seattle Genetics, Incyte, Bristol Myers Squibb, Neon Therapeutics, Roche/Genentech
Travel, Accommodations, Expenses: AstraZeneca/MedImmune, Genentech/Roche, Jounce Therapeutics

Rahul Aggarwal
Honoraria: Clovis Oncology
Consulting or Advisory Role: AstraZeneca, Dendreon, Advanced Accelerator Applications, Clovis Oncology, Axiom Biotechnologies
Research Funding: Zenith Epigenetics, Novartis, Xynomic Pharma, Cancer Targeted Technology, Janssen, Merck, Abbvie, Amgen, AstraZeneca, BioXCel Therapeutics

Felix Feng
Leadership: PFS Genomics
Stock and Other Ownership Interests: PFS Genomics
Honoraria: Genentech
Consulting or Advisory Role: Blue Earth Diagnostics, Celgene, Janssen Biotech, Genentech, Myovant Sciences, Roivant, Astellas Pharma, SerImmune
Research Funding: Zenith Epigenetics
Patents, Royalties, Other Intellectual Property: I helped develop a molecular signature to predict radiation resistance in breast cancer, and this signature was patented by the University of Michigan, my employer. It is in the process of being licensed to PFS Genomics, a company that I helped found

Peter R. Carroll
Stock and Other Ownership Interests: Nutcracker Therapeutics Inc
Consulting or Advisory Role: Nutcracker Therapeutics Inc, Insightec, Progenics, Francis Medical

John S. Witte
Stock and Other Ownership Interests: Avail Bio
Expert Testimony: Pfizer, Sanofi

No other potential conflicts of interest were reported.

ACKNOWLEDGMENT

The authors thank all participating patients and their families.

REFERENCES

1. Siegel RL, Miller KD, Jemal A: Cancer statistics, 2020. Cancer J Clin 70:7-30, 2020
2. Freedland SJ, Humphreys EB, Mangold LA, et al: Risk of prostate cancer–specific mortality following biochemical recurrence after radical prostatectomy. JAMA 294:433-439, 2005
3. Boorjian SA, Thompson RH, Siddiqui S, et al: Long-term outcome after radical prostatectomy for patients with lymph node positive prostate cancer in the prostate specific antigen era. J Urol 178:864-871, 2007
4. Litwin MS, Tan HJ: The diagnosis and treatment of prostate cancer: A review. JAMA 317:2532-2542, 2017

5. Cooperberg MR, Broering JM, Carroll PR: Risk assessment for prostate cancer metastasis and mortality at the time of diagnosis. J Natl Cancer Inst 101:878-887, 2009
6. Boutros PC, Fraser M, Harding NJ, et al: Spatial genomic heterogeneity within localized, multifocal prostate cancer. Nat Genet 47:736-745, 2015
7. Melijah S, Liu LY, Rubanova Y, et al: The evolutionary landscape of localized prostate cancers drives clinical aggression. Cell 173:1003-1013.e15, 2018
8. Caswell DR, Swanton C: The role of tumour heterogeneity and clonal cooperativity in metastasis, immune evasion and clinical outcome. BMC Med 15:133, 2017
9. Cooper CS, Eeles R, Wedge DC, et al: Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. Nat Genet 47:367-372, 2015
10. Lilja H, Ulmert D, Vickers AJ: Prostate-specific antigen and prostate cancer: Prediction, detection and monitoring. Nat Rev Cancer 8:268-278, 2008
11. Carroll PH, Mohler JL: NCCN guidelines updates: Prostate cancer and prostate cancer early detection. JNCCN J Natl Compr Cancer Netw 16:620-623, 2018
12. Marrone M, Potosky AL, Penson D, et al: A 22 gene-expression assay, decipher® (GenomeDx biosciences) to predict five-year risk of metastatic prostate cancer in men treated with radical prostatectomy. PLoS Curr 7:1-8, 2015
13. Dhanasekaran SM, Barrette TR, Ghosh D, et al: Delineation of prognostic biomarkers in prostate cancer. Nature 412:822-826, 2001
14. Siddiqui MM, Rais-Bahrami S, Turkbey B, et al: Comparison of MR/ultrasound fusion-guided biopsy with ultrasound-guided biopsy for the diagnosis of prostate cancer. JAMA 313:390-397, 2015
15. Cancer Genome Atlas Research Network: The molecular taxonomy of primary prostate cancer. Cell 163:1011-1025, 2015
16. Lalonde E, Ishkanian AS, Sykes J, et al. Tumour genomic and microenvironmental heterogeneity for integrated prediction of 5-year biochemical recurrence of prostate cancer: A retrospective cohort study. Lancet Oncol 15:1521-1532, 2014
17. Locke JA, Zafarana G, Ishkanian AS, et al: NKX3.1 haploinsufficiency is prognostic for prostate cancer relapse following surgery or image-guided radiotherapy. Clin Cancer Res 18:308-316, 2012
18. Price TJ, Hardingham JE, Lee CK, et al: Prognostic impact and the relevance of PTEN copy number alterations in patients with advanced colorectal cancer (CRC) receiving bevacizumab. Cancer Med 2:277-285, 2013
19. Fraser M, Sabelnykova VY, Yamaguchi TN, et al. Genomic hallmarks of localized, non-indolent prostate cancer. Nature 541:359-364, 2017
20. Abbosh C, Birkbak NJ, Wilson GA, et al: Phylogenetic ctDNA analysis depicts early stage lung cancer evolution. Nature 545:446-451, 2017
21. Phallen J, Sausen M, Adleff V, et al. Direct detection of early-stage cancers using circulating tumor DNA. Sci Transl Med 9:eaan2415, 2017
22. Quigley D, Alumkal JJ, Wyatt AW, et al: Analysis of circulating cell-free DNA identifies multi-clonal heterogeneity of BRCA2 reversion mutations associated with resistance to PARP inhibitors. Cancer Discov 7:999-1005, 2017
23. Hennigan ST, Trostel SY, Terrigino NT, et al: Low abundance of circulating tumor DNA in localized prostate cancer. JCO Precis Oncol 3:PO.19.00176, 2019
24. Kinde I, Wu J, Papadopoulos N, et al: Detection and quantification of rare mutations with massively parallel sequencing. Proc Natl Acad Sci USA 108:9530-9535, 2011
25. Cario CL, Chen E, Leong L, et al: A machine learning approach to optimizing cell-free DNA sequencing panels: With an application to prostate cancer. BMC Cancer 20:820-829, 2020
26. Cooperberg MR, Hilton JF, Carroll PR: The CAPRA-S score: A straightforward tool for improved prediction of outcomes after radical prostatectomy. Cancer 117:5039-5046, 2011
27. Pérez-Barrios C, Nieto-Alcolado I, Torrente M, et al: Comparison of methods for circulating cell-free DNA isolation using blood from cancer patients: Impact on biomarker testing. Transl Lung Cancer Res 5:665-672, 2016
28. Benjamin D, Sato T, Cibulskis K, et al: Calling somatic SNVs and indels with Mutect2. bioRxiv:1-8, 2019. doi: 10.1101/861054
29. Ramos AH, Lichtenstein L, Gupta M, et al: Oncotator: Cancer variant annotation tool. Hum Mutat 36:E2423-E2429, 2015
30. Adalsteinsson VA, Ha G, Freeman SS, et al: Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. Nat Commun 8:1324, 2017
31. Lang SH, Swift SL, White H, et al: A systematic review of the prevalence of DNA damage response gene mutations in prostate cancer. Int J Oncol 55:597-616, 2019
32. Grasso CS, Wu YM, Robinson DR, et al: The mutational landscape of lethal castration-resistant prostate cancer. Nature 487:239, 2012
33. Ashouri A, Sayin VI, Van den Eynden J, et al: Pan-cancer transcriptomic analysis associates long non-coding RNAs with key mutational driver events. Nat Commun 7:13197, 2016
34. Lanzós A, Carlevaro-Fita J, Mularoni L, et al: Discovery of cancer driver long noncoding RNAs across 1112 tumour genomes: New candidates and distinguishing features. Sci Rep 7:41544, 2017
35. Fredriksson NJ, Ny L, Nilsson JA, et al: Systematic analysis of noncoding somatic mutations and gene expression alterations across 14 tumor types. Nat Genet 46:1258-1263, 2014
36. He F, Wie R, Zhou Z, et al: Integrative analysis of somatic mutations in non-coding regions altering RNA secondary structures in cancer genomes. Sci Rep 9:8205, 2019
37. Abbosh C, Swanton C, Birkbak NJ: Clonal haematopoiesis: A source of biological noise in cell-free DNA analyses. Ann Oncol 30:358-359, 2019
38. Liu J, Chen X, Wang J, et al: Biological background of the genomic variations of cf-DNA in healthy individuals. Ann Oncol 30:464-470, 2019
39. Chalmers ZR, Connelly CF, Fabrizio D, et al: Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. Genome Med 9:34-14, 2017

APPENDIX

MATERIALS AND METHODS

Patient Cohort

A total of 45 individuals recruited between August 2015 and November 2019 were included in this study: 15 healthy donors, 21 patients with localized prostate cancer, and nine patients with metastatic castration-resistant prostate cancer (mCRPC) (Table 1). Healthy control samples were collected from volunteers at UCSF, and all patients with localized disease underwent RP at UCSF. Patients with mCRPC were included from the UCSF PROMOTE study investigating predictive markers of response. Blood samples, matched normal tissue from adjacent seminal vesicles or whole blood, and multiple tumor regions (one to nine samples per patient) were collected from patients undergoing RP. Whole peripheral blood was collected immediately before surgery for patients with localized disease and before treatment initiation for patients with mCRPC.

Clinicopathologic variables that play an important role in surgical management after prostatectomy were also collected for patients with localized disease who underwent RP, including pathologic T stage, tumor volume, and the Cancer of the Prostate Risk Assessment Postsurgical (CAPRA-S) score, which is a prediction model used to assess risk of recurrence after surgery and encompasses presurgical prostate-specific antigen level, pathologic Gleason score, presence of positive surgical margins, extracapsular extension, seminal vesicle invasion, and lymph node involvement.[26] Biochemical recurrence was defined as two consecutive prostate-specific antigen levels of > 0.2 ng/mL at least 8 weeks after surgery. Approval for this study was granted by the Human Research Protection Program Institutional Review Board at UCSF (IRB 11-05226, IRB 12-09659, and IRB 12-10340), and written informed consent was obtained from all patients.

Tissue Sample Collection

For the men with localized prostate cancer undergoing RP, multiple 3-mm punches were collected from the index lesion, regions with varying histology or Gleason grade, as well as other spatially distinct tumors from surgically resected prostates. Prostate tissue was cut into quadrants and snap-frozen in optimal cutting temperature compound, using isopentane chilled by dry ice. A cryostat was used to create 5-μm sections. Tumor locations, Gleason grade, and tumor content were verified by a pathologist's examination of hematoxylin and eosin–stained sections of prostate tissue. Tissue punches were stored at –80°C. Matched normal tissue samples were also collected from nearby seminal vesicles or peripheral whole blood when the prior was unavailable. For the patients with mCRPC, tissue samples were obtained using image-guided core needle biopsy of the metastatic lesion in the bone or soft tissue, and formalin-fixed, paraffin-embedded for histopathologic review.

Tissue Processing and Sequencing

DNA was extracted from normal and tumor tissue using the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Redwood City, CA). For each sample, genomic DNA was sheared with a Covaris M220 ultrasonicator to a target size of approximately 300 bp, and assembled into a library with Illumina TruSeq adapters. For each sample, 10-100 ng of DNA was used for targeted and whole-genome sequencing (see Data Supplement). A set of custom myBaits (Arbor Biosciences, Ann Arbor, MI) hybrid capture probes ± 175 bp and tiled 3× were designed to target mutations in a custom panel (described below) and applied by Arbor Biosciences before sequencing to a target depth of 200× on an Illumina HiSeq 2500 at the CLIA-certified laboratory of the UCSF Institute for Human Genetics Genomics Core Facility (CLIA #05D2080584). For patients with mCRPC, four of the nine men had tumor tissue samples from metastatic lesions available for sequencing as a part of the UCSF PROMOTE study (ClinicalTrials.gov identifier: NCT02735252).

Targeted Panel Design

We used a custom designed 2.5-Mb targeted panel that included 7,034 mutations identified using a Support Vector Machine classification and ranking model, as previously described.[25] Briefly, this model was trained on whole-genome sequence data from 550 primary prostate tumors from the International Cancer Genome Consortium (release 23), along with biological feature annotations (ie, Combined Annotation-Dependent Depletion and PhyloP deleterious measures, annotation in KEGG, amino acid identity, and evolutionary conservation). The resulting panel included single-point mutations as well as small < 200 bp indels in both coding (83%) and noncoding regions (17%).

cfDNA Extraction and Quantification

For healthy controls, 20 mL of whole peripheral blood was collected in PAXgene Blood ccfDNA tubes (Qiagen, Redwood City, CA). From 13 to 25 mL of whole peripheral blood was collected immediately before surgery for patients with localized prostate cancer and before treatment initiation for patients with mCRPC. Plasma was generated from whole blood samples within 2 hours for blood collected in K3EDTA tubes or within 7 days for blood collected in PAXgene Blood ccfDNA tubes (Qiagen, Redwood City, CA). We used a two-step centrifugation protocol: first centrifuging the blood at 1,900*g* for 10 minutes at 21°C, followed by centrifugation of the supernatant at 16,000*g* for 10 minutes to remove leukocytes and cellular debris. DNA was extracted from 7 to 29 mL of plasma using the Qiagen QIAamp Circulating Nucleic Acid Kit (Qiagen, Redwood City, CA), double eluted with 40 μL of Qiagen Elution Buffer, and analyzed on the Agilent 2100 Bioanalyzer with High-Sensitivity DNA Chips (Agilent Technologies, Santa Clara, CA) for assessment of sample purity, concentration, and fragment size distribution according to the manufacturer's instructions. Plasma cell-free DNA (cfDNA) concentration was determined with the Agilent 2100 Bioanalyzer Expert software and calculated across the first three peaks (between 75 and 675 bp) corresponding to the length of nucleosomal footprints and linkers derived from apoptotic cells.[27]

cfDNA Library Preparation and Sequencing

A minimum of 10 ng of cfDNA from each sample was used to prepare sequencing libraries by concentrating the cfDNA with a Zymo Clean and Concentrator Kit (Zymo Research, Irvine, CA) and tagging molecules with unique molecular identifiers (UMIs) with the ThruPLEX Tag-Seq 48S kit (Takara Bio, Mountain View, CA) before PCR amplification (7-11 cycles). The UMIs included two 6 nucleotide barcodes and two 8-11 nucleotide stems on each end of the insert, with an 8 nucleotide Sanger index on the 3′ end. Finally, samples were again analyzed on the Agilent 2100 Bioanalyzer after AMPure XP bead cleanup for quality control (Beckman Coulter, San Jose, CA). Hybrid capture with custom myBaits (Arbor Biosciences, Ann, Arbor, MI) were applied to the libraries before sequencing to a target depth of 2,500× on the Illumina HiSeq 2500. Samples were also whole-genome sequenced to a target depth of 4× (see Data Supplement).

Because of the low tumor fractions typically found in localized prostate cancer, special consideration was given to UMI-tagged cfDNA sequencing depth calculations. The average sequencing depth can be defined theoretically as LN/G, where L is the read length, N is the number of reads, and G is the haploid genome length. For sequencing with the targeted panel, the on-target hybrid capture efficiency was estimated to be 40% with 10% duplicates. The number of total reads was found by identifying the minimum number of raw reads per UMI family necessary to generate consensus reads for variant calling, where a UMI family is a set of reads constructed from both strands of the original dsDNA molecule.

Tissue Sequencing Data Analysis

Quality assessment of sequence reads was first evaluated using FastQC, which includes metrics on per base quality, GC content, sequence length distribution, and overrepresented sequences. Whole-genome and targeted sequencing data were then analyzed using the

pipelines developed by the Broad Institute of MIT and Harvard on the Terra platform with the GATK4.1 tools release.

All tissue sequencing data were preprocessed to produce analysis-ready BAM files before somatic variant calling and copy-number analysis. Raw paired-end reads (150 bp) in FASTQ format were merged and aligned to the Genome Reference Consortium Human Build 37 (GRCh37) with BWA. Bases with a Phred quality score < 20 were filtered out to remove poor-quality reads, likely because of sequencing errors. Picard tools were used to sort, index, and merge files, as well as mark and remove duplicate reads that originated from the same DNA fragment, which are nonindependent observations. Base quality scores were also recalibrated to correct for systematic errors to produce a final BAM file for further analysis.

For the samples sequenced with the targeted panel, MuTect2 was used to perform somatic variant calling on matched tumor-normal BAMs to detect single nucleotide variants and small INDELS, which used annotation files contained in the GATK bundle.[28] Variant filtering was performed to remove potential technical or germline artifacts, including cross-sample contamination. Variants were retained if the filter parameter was designated as PASS and subsequently annotated with Oncotator.[29] Manual review of the variants was performed with Integrative Genomics Viewer (IGV) to inspect variants for sequencing evidence.

Somatic copy-number alterations in tumor tissue were identified in whole-genome sequence data using GATK ModelSegments, using a panel of normals generated from whole-genome normal samples sequenced at the Broad Institute Genomics Platform. When creating a genomic intervals list to define the resolution of the analysis, bin lengths were set to 1,000 bp. Read coverage data are denoised against the panel of normals using principal component analysis, and both kernel segmentation and Markov-chain Monte Carlo are used with copy ratio and allelic counts data to group contiguous segments and make calls. Genomic instability was assessed with the percentage of genome altered metric, which was calculated by dividing by the number of base pairs affected by copy-number changes by the total length of the genome.

cfDNA Sequencing Data Analysis

Plasma cfDNA barcoded with UMIs and sequenced with the targeted panel underwent variant calling using the Curio Genomic platform, which was specifically designed for processing UMI-tagged sequences prepared with the Takara ThruPlex Tag-seq kit. Raw paired-end reads were merged and aligned to GRCh37 with BWA, and the 6nt UMIs were extracted for downstream analysis. Duplex UMI processing was enabled to group reads from both strands of the original dsDNA molecule into UMI families. Consensus reads were generated from UMI families before variant calling with Curio version 1.4.1 with the following parameter thresholds: (1) a minimum base quality Phred score of 30, corresponding to 99.9% base call accuracy; (2) a minimum of four reads in every UMI family to filter out smaller families with few reads; (3) a minimum of 75% of the reads with the same base call in a UMI family at a given position; (4) an allowable UMI hamming distance of four bases to differ across both the read and its paired-end mate; (5) a minimum nonreference coverage or number of unique UMI families supporting the variant was set to three; and (6) an allele frequency less than 20% to exclude potentially homozygous and heterozygous germline variants.

Low-pass whole-genome sequencing data were used to identify large-scale copy-number alterations and estimate the fraction of tumor in cfDNA using HMMcopy and ichorCNA.[30] Briefly, whole genomes were binned into 1 Mb windows, and a Hidden Markov Model was used to segment the copy-number profile into regions predicted to be generated by the same copy-number variant event, as well as identify copy-number alterations for each segment.

Statistical Analysis

Since cfDNA variant counts were not normally distributed ($P < .001$, Shapiro-Wilk test), we evaluated the difference across healthy and prostate cancer groups using the Mann-Whitney nonparametric test. Correlations between clinical categorical variables (ie, biochemical recurrence, seminal vesicle invasion, and Gleason score $\leq 3 + 4$ $v \geq 4 + 3$) and somatic tumor tissue variant detection in cfDNA was assessed with Fisher's exact test. The nonparametric Mann-Whitney U-test was used to evaluate differences in cfDNA variant detection for continuous variables (ie, starting amount of extracted cfDNA, number of tumor tissue or cfDNA variants, CAPRA-S score, pathologic tumor volume, and sequencing depth of coverage). All statistical analyses were performed using R version 3.6.1.

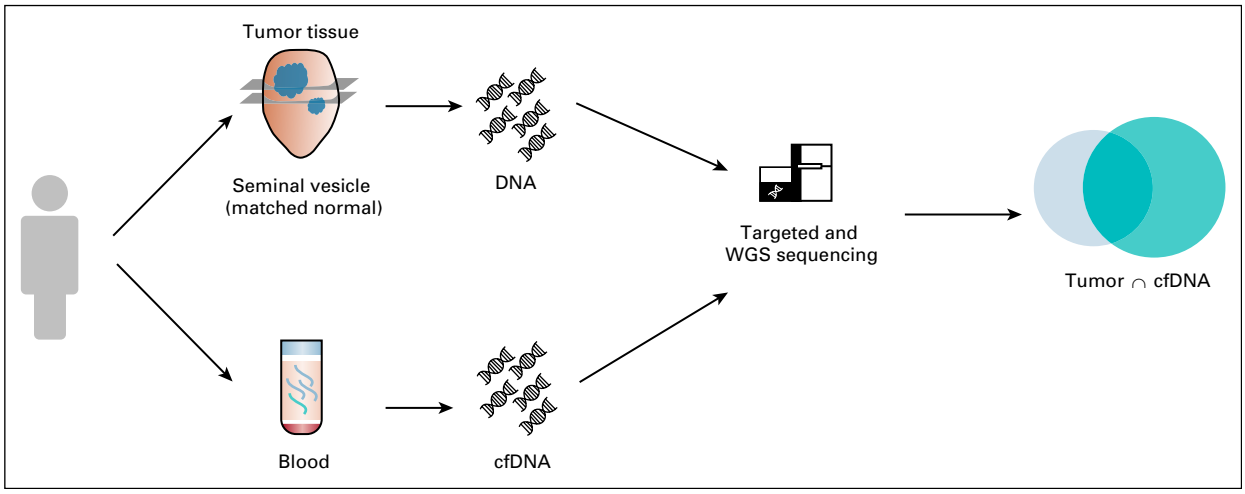


FIG A1. A total of 71 tissue specimens from surgically resected prostates were collected from 21 men. For each patient, multiple tumor tissue samples were obtained from regions with varying histology when possible, and matched normal tissue was collected from nearby seminal vesicles or peripheral whole blood when the prior was unavailable. Venous whole blood was drawn in K3EDTA or PAXgene ccfDNA tubes for all patients. DNA extraction and library preparation were then performed before targeted sequencing with a 2.5 Mb panel or WGS to assess genomic heterogeneity among localized prostate tumors. Additionally, cfDNA molecules were barcoded with UMIs to group reads from both strands of the original double-stranded DNA molecule into UMI families during variant calling. cfDNA, cell-free DNA; UMI, unique molecular identifier; WGS, whole-genome sequencing.

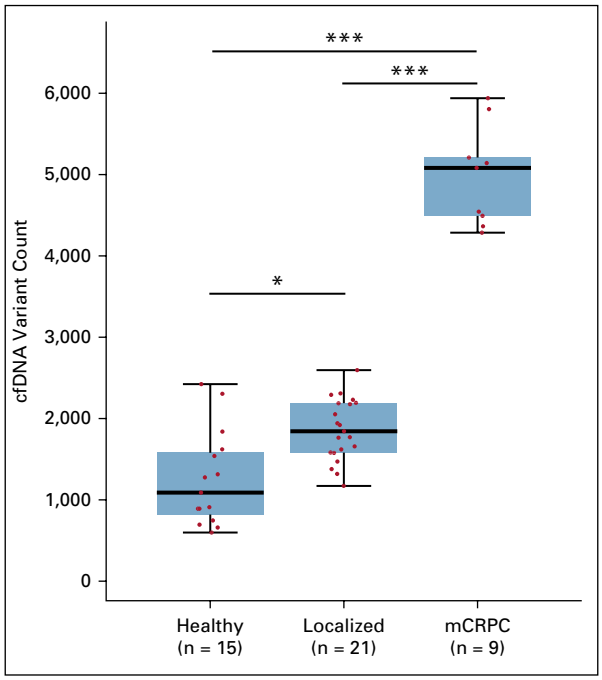


FIG A2. Plasma cfDNA mutational burden assessed by targeted sequencing increases with disease severity. Boxplots show the distribution in cfDNA variant count across healthy controls (n = 15), patients with localized disease (n = 21), and patients with mCRPC (n = 9) from blood samples collected at baseline before surgery or treatment initiation. Men with localized disease had significantly higher counts than those observed in controls ($P < .01$), and men with mCRPC had significantly higher counts compared with those found in men with localized disease ($P < .0001$) or controls ($P < .0001$). *$P < .01$, ***$P < .0001$. cfDNA, cell-free DNA; mCRPC, metastatic castration-resistant prostate cancer.

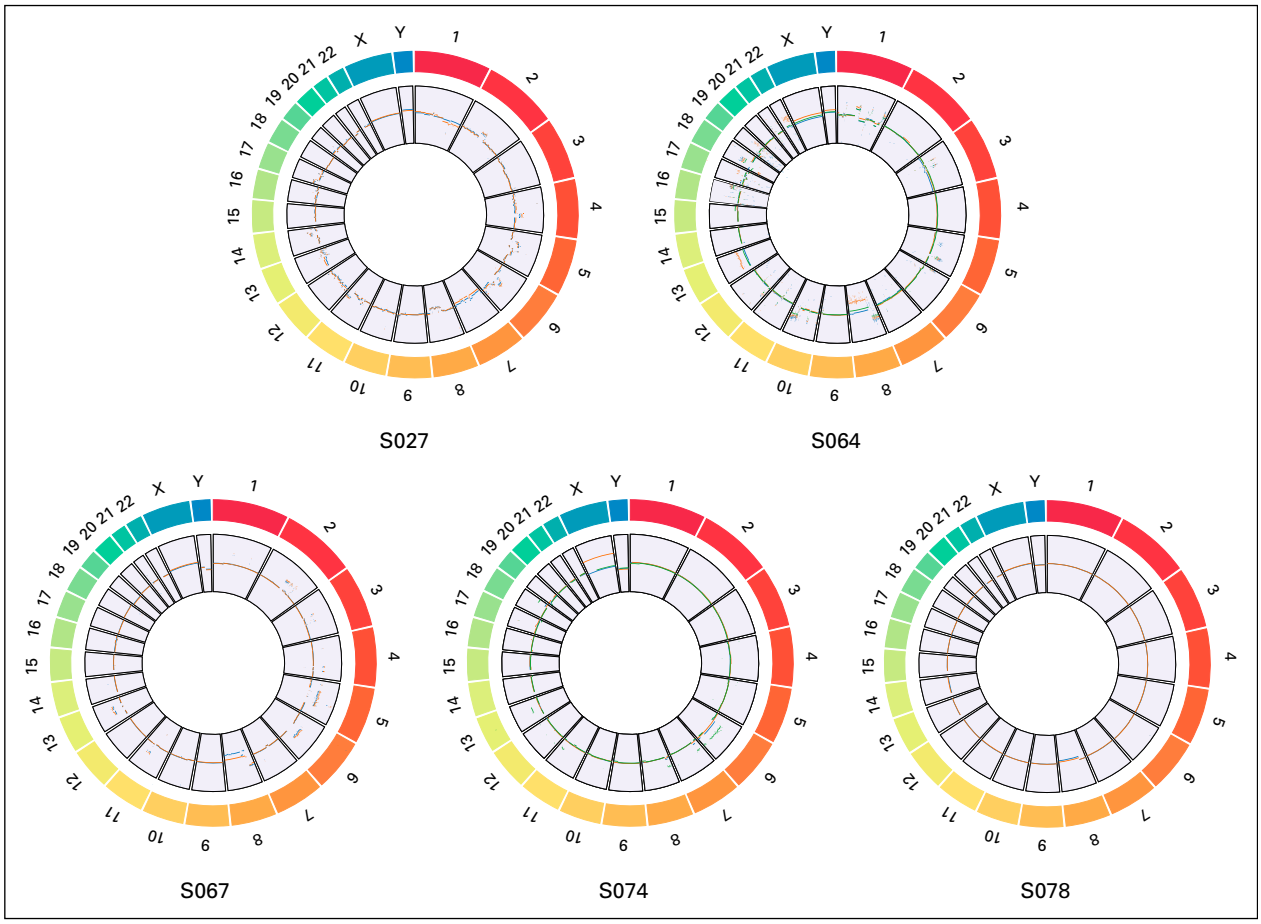
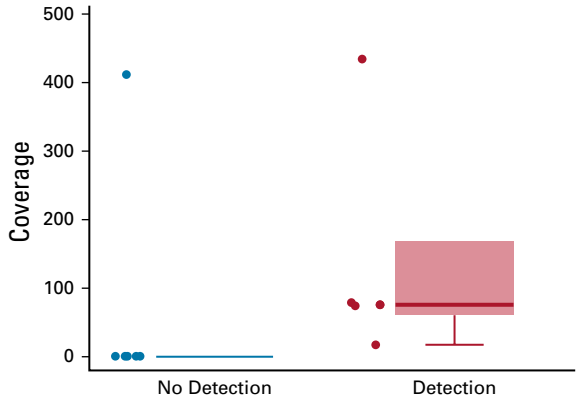


FIG A3. Low-pass whole-genome sequencing was performed for 12 tissue specimens from five patients with localized prostate cancer. Each Circos plot depicts the genomic location in the outer ring and chromosomal $\log_2$ copy number in the inner ring, with multiple samples overlaid for the same patient. Likely clonal and subclonal copy-number alterations were identified for *CHD1*, *NKX3-1*, *CDKN1B*, *MYC*, *PTEN*, and *TP53*. In one patient, one of the three tissue samples harbored both *MYC* amplification and *PTEN* loss, which is prognostic for biochemical recurrence, and 17.8% of the genome was affected by copy-number changes.



Detection in cfDNA	Chromosome	Position	Total Coverage	Variant Coverage
Yes	chr2	91775688	434	4
Yes	chrY	10037785	74	10
Yes	chrY	10037787	76	6
Yes	chrY	10037790	79	13
Yes	chr19	53357000	17	3
No	chr1	152127865	412	1
No	chr2	152127871	0	0
No	chr3	158450451	0	0
No	chr4	158812248	0	0
No	chr5	97552665	0	0
No	chr5	121739540	0	0
No	chr6	75814974	0	0
No	chr7	100678421	0	0
No	chr8	100678426	0	0

FIG A4. Total coverage and variant allele coverage in cfDNA were potential determinants of somatic tissue variant detection in cfDNA. Shown here is a subset of somatic tissue variants for a single patient and whether or not the variants were detected in cfDNA. A majority of variants that were not detected in cfDNA had zero coverage or sufficient total coverage but not enough UMI family coverage supporting the alternate allele to be called as a variant. cfDNA, cell-free DNA; UMI, unique molecular identifier.

TABLE A1. Correlation between cfDNA Variant Detection and Clinical Characteristics

Clinical Feature	No Detection in cfDNA	Detection in cfDNA	*P*
Pathologic T stage			
≤ T3a (No SVI)	9	6	.019
≥ T3b	0	6	
Gleason			
≤ 7 (3 + 4)	4	5	1.0
≥ 7 (4 + 3)	5	7	
Biochemical recurrence			
No	8	6	.087
Yes	1	6	
No. tissue samples			
< 3	4	4	.67
≥ 3	5	8	
Tumor tissue variant count			
Average	14	597	.0046
Starting amount of cfDNA, ng			
Average	18.5	26.1	.25
CAPRA-S score			
Average	4	6.6	.08
Tumor volume, cc			
Average	4	4	.61

NOTE. Fisher's exact test was used to investigate the correlation between categorical features and detection of tumor tissue variants from targeted cfDNA sequencing in patients with localized prostate cancer. Mann-Whitney *U*-test was used to assess the association between continuous clinical features and detection of tumor tissue variants from targeted cfDNA sequencing in patients with localized prostate cancer. SVI and tumor tissue variant count were significantly associated with detection in cfDNA.

Abbreviations: CAPRA-S, Cancer of the Prostate Risk Assessment Postsurgical; cfDNA, cell-free DNA; SVI, seminal vesicle invasion.